The Pennsylvania State University

The Graduate School

Department of Industrial and Manufacturing Engineering

MISSING OBSERVATIONS AND RESTRICTIONS ON RANDOMIZATION IN NANOMANUFACTURING EXPERIMENTATION

A Thesis in

Industrial Engineering

by

Navinchandra N. Acharya

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2006

The thesis of Navinchandra N. Acharya was reviewed and approved* by the following:

Harriet Black Nembhard
Associate Professor and Bashore Career Professor of Industrial and Manufacturing Engineering
Thesis Advisor
Chair of Committee

Seong H. Kim
Assistant Professor of Chemical Engineering

Dennis K. J. Lin
Affiliate Faculty Member of Industrial and Manufacturing Engineering
University Distinguished Professor of Statistics and Supply Chain Management

Timothy W. Simpson
Professor of Mechanical and Industrial Engineering

Richard J. Koubek
Professor of Industrial Engineering
Head of the Department of Industrial and Manufacturing Engineering

*Signatures are on file in the Graduate School

ABSTRACT

In recent years, scientists have discovered various techniques for building nanostructures, but they have only just begun to investigate their properties and potential applications. Moreover, to translate scientific discoveries from the laboratory to commercial products, it is imperative to address some of the fundamental scientific barriers to nanomanufacturing, in addition to the ongoing research in the field of nanotechnology.

This thesis suggests that some of the above challenges can be addressed using experimental designs tailored to specific concerns. An instance of such a scenario is considered that involves a series of gas-phase nano-scale lubrication experiments for micro-electro-mechanical systems (MEMS) devices. Due to the physical unavailability of some of the C-6 alcohol molecules in the experiments, the experimenter is forced to deal with a design having one or more missing observations.

In this study, new Bayesian algorithms are proposed that combine information from the traditional Bayesian screening algorithm used for identifying active factors and three existing algorithms for missing observations. The criterion used for estimating the missing observations is predictive ability in addition to minimization of residual sum of squares (RSS). These new algorithms are applied to simulated data sets that resemble the setup of the nano-scale lubrication experiment assuming one and two missing observations.

The performance of the Bayesian algorithms are compared to the three existing algorithms that have minimal RSS as the only criterion with an appropriate performance measure, $PRESS_{Diff}$. A comparison of the algorithms over the different positions of the

missing observations reveals that in all the cases, the Bayesian algorithms perform significantly better than the non-Bayesian algorithms. In the study the robustness of the proposed algorithms to the initial model specified by the Bayesian screening method, various mismatches of active factors were considered. In all the mismatches considered for one and two missing observations, the results indicate that the Bayesian algorithms still outperform the respective non-Bayesian ones. Finally, judging from the studies performed, the Bayesian Complete RSS minimization algorithm seems to yield the closest estimates of the missing observations, while yielding the maximum predictive ability.

In some nanomanufacturing situations, due to the physical constraints in the process, it is impracticable to execute a full or fractional factorial experiment. In such cases, restriction on randomization is imposed and the experimenter is forced to resort to a split-plot design or some of its variants. Many processes in nanomanufacturing are conducted over a series of stages. Additionally, some of the process variables in some of the stages might be difficult or hard to change in terms of time, limited resources, or – in many cases – money.

Specifically, a polymerization process for the fabrication of nano-films is investigated, where the fabrication is carried out over three stages. To execute efficient experimentation and fully understand the intricacies at the nano-scale, split-plot designs that can be applied effectively over multiple stages are proposed with the aim of reducing the cost of experimentation, and their characteristics examined. General expressions for some of the properties of these designs and analysis are developed. As common design ranking criteria such as resolution and minimum aberration do not provide the “best”

designs in all cases, two new design optimality criteria are proposed. Catalogs of split-plot designs for three and four stages are created and ranked according to the proposed criteria.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This research aims at developing advanced statistically based experimental designs and integrating them into the modeling and production for nanomanufacturing research to yield strategic advantages by speeding the research and development cycle, and help in creating more reliable, robust, and better performing products.

ACKNOWLEDGEMENTS

First, I would like to thank God Almighty for giving me the patience, strength and courage to successfully achieve my goals in every walk of life. I thank my parents, my wife, and my siblings for their constant support, enduring love, and inspiration that keeps me motivated.

Many people have generously contributed towards this research. I would like to take this opportunity to thank my advisor, Dr. Harriet Black Nembhard who has been instrumental in the development of this research. Her exceptional style of mentoring aided me in formulating creative solutions to problems.

I would also like to thank Dr. Dennis K. J. Lin, Dr. Timothy W. Simpson, and Dr. Seong H. Kim for agreeing to be an integral part of my Ph.D. committee. I thank them for sharing their time, energy, and knowledge. I have learnt a great deal not only from their classes, but also from their comments and critiques.

I extend my sincere thanks to Dr. Kim and his research group for their participation and efforts in understanding the nano-scale lubrication experiments for MEMS devices. I also would like to thank Krishna Dronavajjala for his patience and willingness to share details of the multistage polymerization process, without which the illustration of a successful implementation of the proposed multistage experiments would not have been possible.

Finally, I owe a debt of gratitude to my colleagues in the Industrial and Manufacturing Engineering department. In particular, I would like to thank members of

the Laboratory for Quality Engineering and Systems Transitions (QUEST): Shuohui Chen, Pannapa Changpetch, and Wenny Chandra, for their unconditional support. My special thanks to Vamsi Salaka, Atul Rangarajan, Piyush Goel, Dheeban Kannan, Chen-Yang Cheng, and Parameshwaran Iyer for their camaraderie and feedback throughout the course of my PhD program.

To My Family

1. INTRODUCTION

Nanomanufacturing can be thought of as fabrication and assembly of nano-particles into devices and structures and integration of nano-particles into larger scale structures. The idea of making "nanostructures" that comprise just one or a few atoms has great appeal, both as a scientific challenge and for practical reasons. There is a lot of excellent nanoscience in the literature and in research labs, but there is a gap between it and nanotechnology. These barriers are widely recognized as problems related to viable commercial scale-up of production volumes. Some of the other technical barriers are low reliability and yield, lack of control in manufacturing processes, and too much batch-to-batch variation (repeatability and reproducibility).

In a nano-scale environment, the effects of design parameters on product characteristics usually cannot be known purely from phenomenological models and therefore, nanotechnology design often requires data collection and analysis. Often the experimentation used to build up the background of nanoscience is based on changing one key factor at a time while keeping others constant. While methodical, this approach is not efficient. Through interdisciplinary collaborations, the importance of statistically based design of experiments (DOE) to the research and development of nanomanufacturing has been identified. Drawing from these interactions as well as the literature, this research addresses two cases that closely link DOE with the needs of nanomanufacturing. These two research topics are discussed briefly below. Although the field of nanomanufacturing is the chosen context, we note that the research and development may not be exclusive to this field.

1.1 Research Topics

1.1.1 Missing Observations in Experimental Designs

In typical two-level factorial experiments, the design is balanced over every factor, which means that each factor column has an equal number of low and high levels. If all the runs in such a design can be executed, there will be no missing observations. However, in many real-world applications, not all the responses in an experiment are recorded. Such situations can arise due to many reasons: machine breakdown, illegible recording of response, damaged experimental resource, and so on. In certain nanomanufacturing situations, the experimenter is forced to deal with a design having one or more missing observations because the factor combination or molecular structure suggested by a randomized design simply does not exist.

The general approach considered in traditional algorithms is to insert fictitious values in the missing cells and proceed with analysis of variance (ANOVA) as usual, reducing the degrees of freedom for error by the number of missing observations. Iterative and non-iterative algorithms for estimating values of the missing observations follow the principle that by choosing values that minimize the residual sum of squares, one can obtain the correct least squares estimates of all estimable parameters as well as the correct residual sum of squares.

Figure 1.1 provides a flowchart to illustrate the proposed procedure to handle missing observations. Once the fractional factorial (2^{n-k}) experiment is performed, the active factors are identified using a Bayesian screening algorithm. Assuming this information to be *a priori*, estimates of the missing observations are obtained using any

of the proposed Bayesian algorithms. Once the estimates are obtained, they are inserted into the missing cells and the response is said to be “pseudo-complete”. The experiment is then analyzed by the usual ANOVA, after reducing the corresponding degrees of freedom for the error term. Thus, the significant main effects and interaction terms can be identified and its effects estimated. In addition, an appropriate model can be obtained.

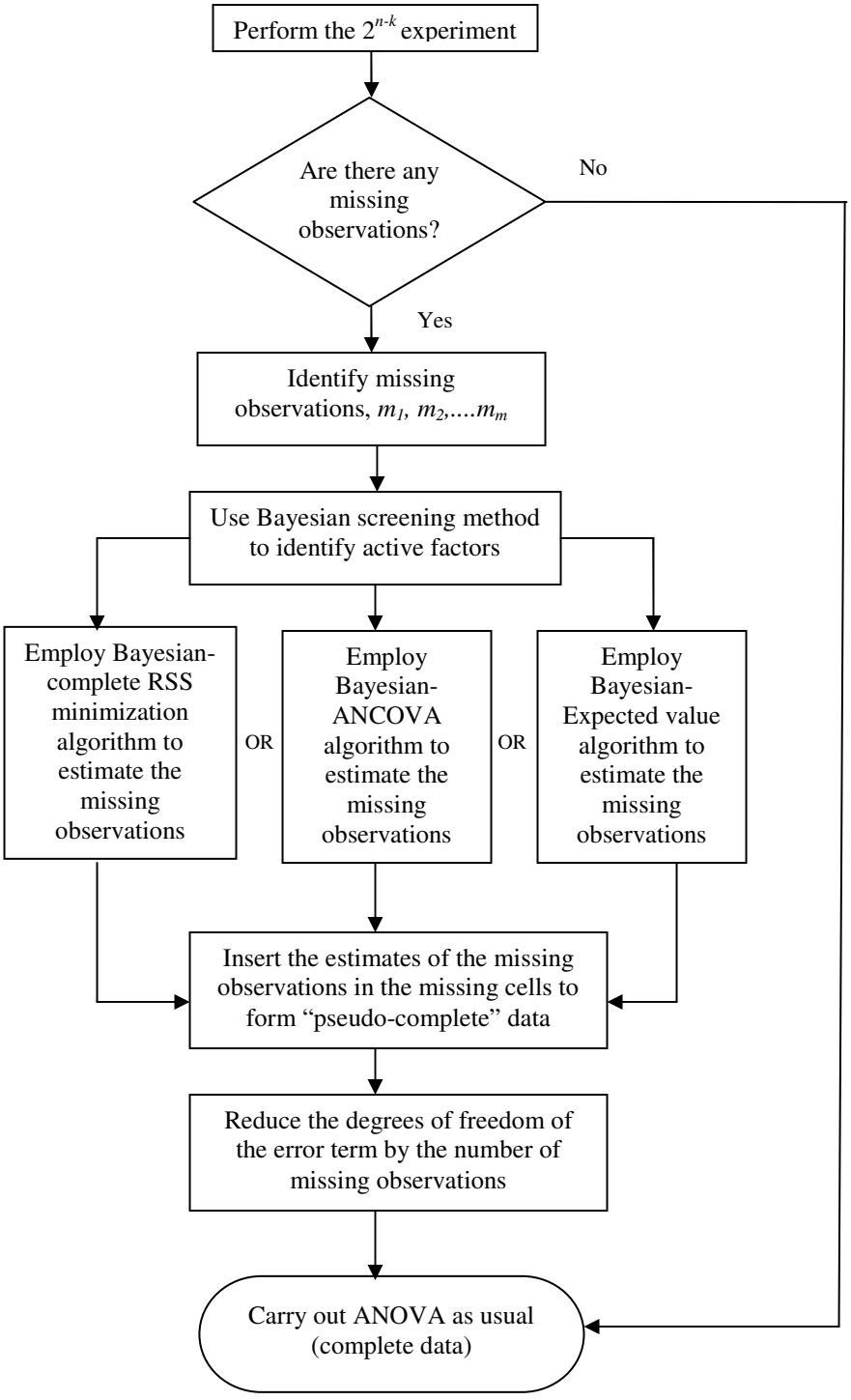

Figure 1.1: Proposed procedure for missing observations.

1.1.2 Multistage Split-Plot Experiments

Randomization is one of the key statistical principles of DOE. It refers to the concept that both the allocation of the experimental material and the order in which the individual runs or trials of the experiment are to be performed are randomly determined. However, due to the physical constraints in nanomanufacturing applications, situations often arise when it is impossible to execute a completely randomized 2^k full factorial or a 2^{k-p} fractional factorial design. Potential reasons for the inability to execute fully randomized designs may be that the manufacturing system is actually a two-stage process, it is expensive or difficult to change the levels of some of the factors, or there are physical restrictions on the process. In such cases, the design can be treated as having a split-plot structure or, more specifically, a fractional factorial split-plot (FFSP) structure. Additionally, split-plot designs result in economical experimentation by reducing the consumption of resources.

When the experiment is comprised of several stages, wherein each stage consists of multiple factors, variants of split-plot designs like strip-plot (also known as strip-block) and split-lot designs have been used recently for processes that demand restrictions on randomization. However, due to the nature of split-lot (and strip-plot), designs there is a severe limitation on degree of fractionation one could achieve. Thus, effects of interest may not be completely estimable. To counter the above limitation and to facilitate present-day experimentation in the field of nanomanufacturing, multistage fractional factorial split-plot designs are proposed. These designs require fewer runs than a split-lot design, while yielding greater information about the factors of interest.

1.2 Research Objectives

Nanotechnology is a widely emerging field and its impact is predicted to be as much as the internet (Bhushan, 2004). The overall goal of the proposed research is to address some of the main challenges in nanomanufacturing through the development of specific statistical designs. The specific objectives of the research are as follows.

a) Experimental Designs with Missing Observations:

In a typical experimental set-up of a nanomanufacturing process, there is always some probability of not being able to record all the responses. This phenomenon occurs due to the lack of proper knowledge in this developing field as newer products and processes are being discovered every day. These occurrences lead to missing observations in the design. The traditional analysis of factorial (full and fractional) designs is conducted assuming that all the responses are recorded and the design structure is orthogonal.

One specific goal of this research is to develop algorithms that are capable of analyzing a designed experiment even when one more observations are missing. These algorithms are to be developed under different criteria, such as minimum residual sum of squares and maximum predictive ability. It is also important to gain insight into the performance of the proposed algorithms over designs that are similar in structure to the motivating gas-phase nano-scale lubrication experiment. Trends in the performance of these algorithms are to be studied by varying conditions such as position of missing observation, number of missing observations, and different specification of active factors by the initial screening algorithm.

b) Experimental designs with Multiple Stages:

Many processes in nanomanufacturing are conducted over a series of stages. Additionally, some of the process variables in some of the stages might be difficult or hard to change in terms of time, limited resources, or – in many cases – money. To execute efficient experimentation and to understand the intricacies at the nano-scale, a goal of this research is to develop and investigate split-plot designs that can be applied effectively over multiple stages. Depending on the overall objective of the experimentation, a goal is to document a catalog of such designs that might be useful for situations arising in nanomanufacturing. The catalog would entail optimal designs for different scenarios and under different criteria such as minimum aberration, maximum number of clear effects, and maximum number of specific two-way interactions.

Another main objective is to study the proposed designs in detail and provide general rules on the confounding pattern and assignment of effects in different stages to the different error terms. Finally, the last objective is to look at variants of these designs. For example, a modification of the proposed design to accommodate a hard-to-change factor in the latter stages of the experiment.

1.3 Outline of the Thesis

This thesis mainly constitutes three major topics. The first part of Chapter 2 provides a literature review of the techniques used to counter the problem of missing observations. Past research on missing observations in experimental designs are presented. An overview of split-plot, strip-plot, and split-lot designs is presented next in Chapter 2. Chapter 2 also provides literature review on some of the manufacturing processes at

nano-scale that utilize statistically based experimental designs. Some of the recent trends in nanomanufacturing, challenges faced by practitioners in realizing the full potential of this technology, and the role of DOE in the nanomanufacturing sector are brought to light in Chapter 3.

New algorithms based on a Bayesian screening algorithm are proposed in Chapter 4. These algorithms are presented along with the motivating nano-scale lubrication experiment. Split-plot designs and some of its variants are discussed in Chapter 5, and a new class of designs for multistage processes is proposed. Chapter 6 draws conclusions on the proposed algorithms and designs, and a general discussion is provided as well. Finally, Chapter 6 provides the directions for further research.

2. LITERATURE REVIEW

This chapter presents a review of the literature relevant to the topics in this thesis. Specifically, it traces the root of missing observations in datasets, both large and small, and reviews existing literature on the same. It provides insight on the evolution of the missing-observation problem that has perturbed statisticians for over half a century. In addition to missing observations, this section reviews deigns in which the experimenter is forced to resort to restrictions in randomization, yielding a split-plot design. Details on the variants of split-plot designs are also reviewed.

2.1 Overview of Missing Observations

The problem of missing observation arises everyday in almost any type of study. Two instances where the presence of missing observations is prominent are in surveys and statistically designed experiments. Missing observations in surveys occur due to a multitude of reasons, some of them being refusal to answer personal questions, ambiguous questions, and time restrictions. A brief overview of some of the techniques employed to tackle missing observations in large datasets is presented in Section 2.1.1. In experimental designs, machine malfunction, faulty reading, and physical constraints on the resources are among the causes that lead to missing observations. Existing algorithms proposed to encounter missing observations in experiments are classified as iterative and non-iterative and are reviewed in Sections 2.1.2 and 2.1.3, respectively.

2.1.1 Missing Observations in Large Datasets

For a large data set, it is extremely common that not all information is present. This is the case especially in surveys in which some of the questions are related to one's attitude or behavior. In addition, Cool (2000) reports that some of the other causes of missing observations are errors in the implementation of the study, illegible recording of responses, inadmissible multiple responses to a single question, and loss of instruments.

Since mid 1970s, researchers have been considering various techniques to give a sense of completeness to the data, by either deleting or estimating the missing observations. Some of these techniques widely used are list-wise deletion, pair-wise deletion, mean substitution, regression, and more recently, multiple imputation and the EM (Expectation-Maximization) algorithm. These techniques are discussed below.

One most obvious way to deal with missing observations is to simply delete the corresponding entries and proceed with the analysis of rest of the data. The main idea behind list-wise deletion is simply to discard any observation with missing information on one or more variables. The analysis of the rest of the data can be carried out in the usual way assuming complete data. Many statistical packages including SAS® and SPSS® have this technique as the default setting for multivariate statistical procedures (Peng et al., 2002). This technique has shown to be effective by Buck (1960) for less than four variables and a small proportion of missing observations. On the contrary, Raymond and Roberts (1987) show that as the number of variables increases, an increased amount of

relevant data are discarded and might even result in complete data deletion in extreme cases.

The pair-wise deletion technique uses all the non-missing (available) observations for each variable to compute means and variances. In addition, all the available pairs of values are used to compute covariances. Correlations are thus computed using only observations that have non-missing values on both the observations of a pair. Schafer and Graham (2002) discuss the properties of the pair-wise deletion technique and provide instances in which the technique can be extremely inefficient.

An alternative to deleting portion of data would be to employ a strategy to estimate the missing values and continue with the analysis of the data. As the name suggests, the mean substitution technique estimates the value of the missing observation, for a given variable, using the overall mean of the particular variable (see Anderson et al., 1983 for further discussion on mean substitution). Tabachnick and Fidell (2001) provide an interesting variant to the mean substitution technique that estimates the value of a missing observation using the mean of the sub-group. Although mean substitution seems to be the simplest, Little and Rubin (2002) firmly advocate never using this technique as it has many disadvantages, some of them being overestimation of sample size and underestimation of variance.

In some cases, estimates of the missing observations can be obtained using a regression model. These estimates would be the predicted values derived from the regression model. This technique of estimating values of the missing observations is called regression estimation. The regression model is obtained from the non-missing observations. Thus, it uses information already existing in the data set. Kaiser (1990)

provides some variations to the regression estimation technique based on the number of variables and missing observations. Schafer and Schenker (2000) enumerate some of the disadvantages of using this technique. One major disadvantage is that correlations and covariances are inflated as the estimates are perfectly predicted from the regression model.

One of the most recent and widely used techniques by researchers and practitioners is the expectation-maximization (EM) algorithm proposed by Dempster et al. (1976). The idea behind EM algorithm can be formalized as follows:

i. Replace missing observations by estimated values.
ii. Estimate parameters.
iii. Re-estimate the values of the missing observations assuming the new parameters are correct.
iv. Re-estimate parameters iteratively until convergence.

2.1.2 Iterative Algorithms for Experimental Designs

In addition to the algorithms developed to handle missing observations in large data sets, particularly surveys, a significant amount of research has been done to counter the presence of missing observations in experimental designs. Two of the most popular iterative algorithms are the ones proposed by Healy and Westmacott (1956) and Shearer (1973). Both these iterative algorithms are based on the principle of filling in values for the missing observations so that the original balance of the experimental design is

maintained. Standard analysis, typically analysis of variance (ANOVA), can be then performed. Little and Rubin (2002) state the advantages of using this principle:

i. It is easier to specify the data structure using the terminology of experimental design.
ii. It is easier to compute necessary statistical summaries.
iii. It is easier to interpret the results of analyses since standard displays and summaries can be used.

Keeping these advantages in mind, Shearer (1973) proposed an iterative algorithm that minimizes the residual sum of squares. The proposed procedure is a two-stage iterative one and is given by

$$\begin{aligned} y_{miss}^{(i+1)} &= \sum_{j=1}^{m} b_j^{(i)} x_{miss\ j} + \frac{1}{n}\sum_{k=1}^{n} y_k^{(i)} \\ b_k^{(i)} &= \frac{1}{n}\sum_{j=1}^{n} x_{jk}\, y_j^{(i)} \qquad (k = 1, 2, ..., m) \end{aligned} \quad . \tag{2.1}$$

The second equation is just the main effect estimation procedure for a 2^m design. The first equation re-estimates the missing value using the new main effects estimates. The algorithm is initiated using any starting value $y_{miss}^{(0)}$ and the latest estimate of y_{miss} is used to re-estimate the main effects and the cycle is repeated. The properties as discussed by Shearer (1973) are that the residual sum of squares is minimized and the method reduces to a non-iterative one in the presence of a single missing observation.

A different algorithm targeted towards handling missing observations is the one proposed by Healy and Westmacott (1956). This algorithm is based on the fact that the fictitious values that are inserted are actually the expected values of the missing observations so that the residuals for the missing observations after fitting all the constants must all be zero.

The residual sum of squares due to missing observation(s) in each iteration is zero and hence the total residual sum of squares is minimized. Little and Rubin (2002) show that Healy and Westmacott's algorithm is an example of the EM algorithm with the "*M*" step corresponding to the least squares analysis on the original design and the "*E*" step corresponding to finding the expected values of the missing observation.

2.1.3 Non-Iterative Algorithms for Experimental Designs

Some of the more common non-iterative algorithms employed to address missing observations can be traced back to Allen and Wishart (1930). The authors presented a formula that estimates the value of a single missing observation. Yates (1933) extended this technique to obtain values of two or more missing observations. The approach considered in these algorithms was to insert fictitious values in the missing cells, such that the error variance obtained when unknowns are substituted is minimized, and proceed with ANOVA as usual, as if the data were complete. The degrees of freedom of the error must also be reduced by the corresponding number of missing observations. Using this approach, Yates (1933) showed that the correct least squares estimates of all estimable parameters as well as the correct residual sum of squares could be obtained. Furthermore, the contribution to the residual sum of squares in any missing cell must be

zero to avoid inflation of the residual sum of squares above the least squares value. Thus, the value to be inserted must be the least squares estimate of that cell. Over the years, many researchers have designed algorithms based on the above principle.

Rubin (1972, 1976) proposed a non-iterative algorithm for least squares estimation of missing observations. This algorithm has an advantage over iterative algorithms as it detects existence of a singular pattern of missing observations, whereas an iterative algorithm fails to produce a warning. In the case of *m* missing observations, the key idea is to find *m* values, $x_1, x_2, ..., x_m$ such that $\hat{x}_k - x_k = 0$, for $k = 1, 2, ..., m$, where $\hat{x}_k$ is the value predicted by the model in the k^{th} missing cell. Each $\hat{x}_k$ is a linear combination of $x_1, x_2, ..., x_m$ as well as the real data and is given by

$$\hat{x}_k = \sum_{i=1}^{m} b_{ki} x_i + \sum_{i=1}^{N-m} a_{ki} y_i \qquad \textbf{(2.2)}$$

where, y_i represents the observed data ($i = 1, 2, ..., N-m$). The residual in the k^{th} missing cell is given by

$$\hat{x}_k - x_k = (1 - b_{kk}) x_k - \sum_{i \neq k} b_{ki} x_i - \sum a_{ki} y_i \,. \qquad \textbf{(2.3)}$$

Rubin (1972) shows that the estimates of the missing observations can be obtained using the relationship given by

$$\boldsymbol{-\rho = XR.} \qquad \textbf{(2.4)}$$

$\boldsymbol{X}$ and $\boldsymbol{\rho}$ are ($1*m$) row vectors, $\boldsymbol{R}$ is an ($m*m$) matrix, and ρ_k is the residual in the k^{th} missing cell when all missing cells are assigned the value zero. The k^{th} row of $\boldsymbol{R}$ contains the residuals in the m missing cells when all cells are set to zero except the k^{th} missing cell, which is set to one. This algorithm is similar to the estimation of missing values using dummy variables in an analysis of covariance, proposed by Bartlett (1937) and emphasized by John and Prescott (1975). John and Prescott (1975) provide an alternate proof and illustrate the algorithm on a 2^4 full factorial experiment with two missing observations.

Draper and Stoneman (1964) proposed yet another non-iterative algorithm for factorial designs, specifically for full and fractional factorial designs. They suggest that the idea proposed by Cochran and Cox (1959) to estimate the missing observations by minimizing the sum of squares for the interaction terms that are used as error may be difficult to implement. This might be due to the experimenter's idea of the appropriate analysis and because the alias structure of the design might make prior choice of the interactions to be used as error impossible. To avoid this issue, Draper and Stoneman, and later Box (1990) employ the following algorithm:

Suppose a 2^k or a 2^{k-p} design is executed in n runs. If m observations are missing, only (n-m) effects can be estimated and once can choose which (n-m) of the original (n-1) effects are to be estimated and which m are to be sacrificed. This is the residual sum of squares, which arises from the sacrificed effects. Minimization of this residual sum of squares will provide m equations. This set of m simultaneous equations can be solved to

obtain the estimates of the m missing observations. Bias in incorrectly sacrificing a significant effect can be detected using a half-normal plot suggested by Daniel (1959).

Chauhan and Bulmahn (1993) developed a simplified formula to estimate missing observations in an $L_n(b^s)$ orthogonal array (see Taguchi, 1987), of n rows, s columns, and factors at b levels. Let Y_j be the missing observation. It is shown that the estimate of the missing observation, $\hat{Y}_j$ can be given by

$$\hat{Y}_j = \frac{b\sum l_i Y_i - (k-1)T'}{n-(b-1)k-1} \qquad \textbf{(2.5)}$$

where

l_i is the number of matches between i^{th} and j^{th} row for fixed j,

k is the number of factors, and

T' is the total of the available observations.

Some of the properties of this algorithm are investigated by Bulmahn and Chauhan (1994). They report on the performance of the estimator for missing observations in a wide range of experimental circumstances (combination of factors and number of runs).

Finally, Box and Meyer (1993) proposed a Bayesian method to identify the active factors (factors that have a significant effect on the response) in a screening experiment. Various hypotheses of factors being active can be considered using this Bayesian method, irrespective of the degree of confounding in the design. This method was initially developed to analyze different types of screening experiments such as $2^{k\text{-p}}$ fractional

factorial experiments (Montgomery, 2005) and geometric and non-geometric Plackett-Burman (Plackett and Burman, 1946) experiments. In addition, as this method can be applied to both orthogonal and non-orthogonal designs, it can be applied to situations where a fraction of the data is missing.

2.2 Overview of Randomization and Restrictions in Experimental Designs

Split-plot designs originated in the agricultural setting (see Yates, 1937; and Kempthorne, 1952), in which a single level of one treatment is applied to a relatively large plot of ground and all levels of a second treatment are applied to sub-plots within the large plot (whole plot). Recently, one of the most well known references to split-plot design is that of Taguchi's (1987) inner-outer array, forming a cross-product of two designs. The inner array usually contained the design factors and the outer array contained the environmental factors for a robust product experiment. Although, the Taguchi's inner and outer array design resembled a split-pot design, a drawback was that the experiments were conducted in completely random order, thus violating the basic principle of a split-plot design.

Bisgaard (2000) provides some detailed reasoning for not using the inner and outer array designs. One of the main disadvantages associated with these designs is that the size of the experiments often requires a substantial number of trials. Box and Jones (1992) provide alternate approaches to the inner-outer arrays. Specifically, they illustrate how split-plot designs and some of its variants can be helpful in saving runs. They provide two arrangements (a) and (b) of a full factorial split-plot design for a particular

experiment. In arrangement (a), whole plots constitute environmental factors and sub-plot constitutes the design factors. Alternatively, arrangement (b) illustrates the same experiment with the design factors now assigned to whole plots and environmental factors to sub-plots. For each of the arrangements, details about the model, error structure, and ANOVA table are provided. The relative efficiency of each design is also computed using the estimates of error variance for the designs.

Bingham and Sitter (1999) showed that even with a split-plot structure, it is possible to obtain a design that requires lot of runs. For this reason, they introduce the concept of fractional factorial split-plot (FFSP) design. These designs are obtained by combining two separate fractional factorial designs. The first design is referred to as the whole plot design, and the second as the sub-plot design. The authors emphasize that even though the design matrix for the split-plot design resembles a completely randomized design, the two designs differ in the manner in which they are executed. As the name suggests, in the completely randomized design, treatment combination of factors are applied randomly for all experimental units. On the contrary, in a split-plot design, a restriction of randomization is introduced before the treatment combinations of the sub-plot factors are applied. Catalogs of designs for different combinations of processing conditions are provided using the minimum aberration (MA) criterion proposed by Fries and Hunter (1980).

The concept of FFSP was investigated further by Bingham and Sitter (2001). The trade-off between cost of experimentation and degree of information obtained was considered. Through a real-life example, they consider the impact of restrictions on randomization, an innate characteristic of a split-plot design. It is suggested that in

selection of a FFSP design, not only does one have to consider the estimation issue captured by the MA criterion, but also the significance of the factorial effects with as much as precision as possible. Specifically, they demonstrate how issues like identifiability (ability to estimate as many of the main effects and two-way interactions) and precision (ability to detect significant effects with as much power as possible) affect the design selection.

An important application of split-plot designs, robust parameter experiments was investigated by Bingham and Sitter (2003). In a robust parameter design, primary interest is to study which control factors (*C*) have dispersion effects. The objective is to minimize process variation due to noise factors (*N*). The authors consider robust parameter designs for general factorial designs and extend the idea to split-plot designs (full and fractional factorial). They discuss the implications on design choice when robust parameter designs are run as split-plot designs and develop a catalog of such designs under the MA criterion.

Similar to the split-plot designs, strip-plot designs also originated in the agricultural settings (Yates, 1937 and Finney, 1945). As noted by Vivacqua (2003), these designs are also referred to as strip-block designs (Cochran and Cox, 1957, and Box and Jones, 1992), split-unit designs (Taguchi, 1987), or split-block designs (Hering and Wang, 1998). Strip-plot designs can be effective in situations in which the process being investigated can be separated into two distinct stages and it is possible to apply the second stage simultaneously to groups of the first stage product.

A variant of the split-plot design, strip-plot design was studied in detail by Box and Jones (1992). Using an example of cake mixes, the properties of strip-plot designs

were detailed. Configurations for the strip-plot designs were considered by Miller (1997). The author proposes an effective way for constructing strip-plot arrangements of fractional factorial designs. The procedure can be summarized in three steps as follows:

i. Identify a suitable row design.

ii. Identify a suitable column design.

iii. Select a Latin-square fraction of the product of the row and column designs.

Construction and analysis of strip-plot designs were also considered by Milliken et al. (1998). The authors considered only two-step processes, using a case study relating to an implanting-annealing process.

Strip-plot designs applied over multiple processing stages are commonly known as split-lot designs. The construction of split-lot designs were pioneered by Mee and Bates (1998). They illustrated the approach on the fabrication of integrated circuits on silicon wafers, involving a sequence of processing steps. An important aspect of such processes is that even though some processing steps are applied to individual wafers, for other steps several lots of wafers are processed simultaneously as a group. Mee and Bates (1998) present a generalized approach to construct 2^k and 2^{k-p} split-lot designs. Specific examples for up to nine processing steps are given, where each step contains a single factor. Although symmetric split-lot designs were provided, they were not verified for optimality.

Butler (2004a) extended the concept of split-lot designs and provided guidelines for construction of two-level fractional factorial split-lot designs. Examples having multiple factors in each stage were presented and split-lot designs for two, three, and four

stages were constructed using “grid representation”, first specified by Butler (2005). The author elucidates that in a split-lot experiment, settings of factors at each processing stage are used on multiple experimental units. Thus, at each processing stage, the design resembles a split-plot structure. Butler (2004a) employs the following four criteria to construct split-lot designs:

i. Minimum aberration.

ii. Main effects confounded only with the sub-lots for the stage they are in.

iii. Minimization of the number of two-way interactions that are confounded with sub-lots at each stage.

iv. Minimization of the number of alias sets that are confounded with more that one set of sub-lots.

Using “grid representation”, a catalog of fractional factorial split-lot designs are constructed based on the above four criteria. The construction covers up to four processing stages and designs having a maximum of 64 runs.

2.3 Experimental Design in Nanomanufacturing Processes

Although the use of structured experimental designs is not widespread at the nano-scale, many researchers have employed various DOE techniques. These designs are used to investigate input variables of a particular process, or obtain a model to predict conditions that would yield desired results or optimize process conditions to successfully achieve scaling-up. A brief survey of the literature for the past few years that incorporate some form of experimental design is given below.

Improvements in the rheological behavior of the nanosilica composite no-flow underfill were achieved by Sun et al. (2004) by investigating the experimental conditions of the surface treatment using silane-coupling agents with the help of a designed experiment. The objective was to obtain optimal conditions for nanosilica surface treatment to formulate nanocomposite no-flow underfill with low viscosity and good filler dispersion. Although no details on the choice of design were mentioned, it seems as though a fractional factorial design, with two baseline runs, was used for the four input variables. The significant main effects, along with an interaction term, were identified, and the input variables were set at corresponding operating conditions for further characterization.

Yong and Hahn (2005) examined the dependence of flexural properties of certain nanocomposites on dispersants and coupling agents. The authors first used a full factorial design as a screening experiment to identify potentially significant effects. Once the effects were identified, Response Surface Methodology was employed using a central composite design to determine the optimal dosage of the coupling agent and dispersant so that the flexural strength could be maximized. To test a possible curvature effect in the model, five center points and four axial points were also used. Once the optimal settings were obtained, confirmation runs were performed to validate the analysis.

In addition to processes at the nano-scale, characterization of some processes at the micro-scale is achieved with the help of statistical tools. For instance, Ji et al. (2006) have investigated behavior of isotropic etching in inductively coupled plasma etching tool. Their work bears application in microneedle fabrication for drug and gene delivery. Specifically, they study the behavior of four variables over five different output variables.

A 2^4 full factorial design was executed, and the five responses measured on each experimental unit. Optimal operating conditions of the input variables were determined for each response considered. Finally, as the fabrication time is a function of the five responses, optimal settings were recommended to achieve short fabrication time. Four additional (confirmation) runs were also carried out to verify the etching results.

A framework for statistical design and analysis of experiments on gate poly-silicon (basic component, consisting of a transistor, in an integrated circuit) critical dimension was presented by Park (2004). The framework aims to study the variation in the critical dimension by estimating the variances components and testing uniformity on the critical dimensions. The motivating factor for this work comes from the need to control gate poly-silicon critical dimension in order to achieve high net-die-per-wafer yield in a semiconductor wafer fabrication process. A two-stage nested design was used to estimate hierarchical variance components according to run-to-run, wafer-to-wafer, and die-to-die production unit changes. A factorial design assuming a reduced model (with the insignificant effects omitted) was then used to study the on-chip uniformity and on-wafer uniformity.

Another instance of the use of DOE is given in Kharbas et al. (2003). They present the influence of process conditions and nano-fillers on the microstructure and mechanical properties of a nanocomposite and its neat-resin counterpart. Two L16 Taguchi orthogonal array fractional factorial designs were executed, one each for neat resin and nanocomposite. Five process parameters were varied over four levels each in the designs. Fractionation was chosen to ensure than maximum number of factors of interest could be identified. The presence of nano-clay, a nucleating agent, was also

statistically tested and found to be significant. Thus, through DOE, optimal process conditions that would result in desirable cell size and density, thus better mechanical properties, were attained.

In addition to obtaining optimal settings of input variables to yield desired output levels, DOE has also been used in instances where certain process variables are believed to affect output but are difficult to control. These factors are termed as “noise” or “external” variables. A robust process or product is said to be one in which the variation due to the “noise” factors is minimized. One such robustness study was considered by Jagarkal et al. (2004). The objective of the study was to improve the fatigue life of solder joints of Printed Wiring Board (PWB) level electronic package. A full factorial design and a central composite design are used to study the effects of the four input (design and noise) variables.

A similar robustness study was also conducted by Dewey et al. (2000) for modeling micro-eletro-mechanical systems (MEMS) to minimize effects of device parametric variability on overall performance. Taguchi designs were used effectively in identifying factors most influential to the process output and determining the settings of the parameters that yield both an acceptable performance metric and minimize variations. In other words, sensitivity to external variables is reduced by appropriately choosing levels of controllable factors. Along similar lines, Ren et al. (2001) propose a three-step technique for quality optimization of MEMS devices using Taguchi methods. Specifically, Taguchi designs are employed in the final step of the proposed technique to reduce deviations in device performance due to parameter variations induced by fluctuations in the fabrication process. Taguchi designs were also employed by Ahmed et

al. (2006) to optimize a time-modulated chemical vapor deposition (TMCVD) process. Time, money, and effort were saved by adopting the experimental design approach in the deposition process of nanocrystalline diamond coating.

In summary, a common trend seen is the application of structured experimental designs in investigating the behavior of processed being considered. Some designs are more advanced than others. Furthermore, in many other studies, there is an indication of some structure in the manner in which experiments were conducted, but no details mentioned. For instance, Grisolia et al. (2005) studied two annealing conditions by varying process parameters such as time, temperature, and ambient atmosphere. The experimental details hint the use of a multi-way ANOVA, but no discussion is provided. For more such examples, see Jancar and Suvorov (2006), Zhang et al. (2005), and Boal et al. (2006).

3. NANOMANUFACTURING ISSUES

3.1 Advances in Nanomanufacturing

Nanotechnology is the art and science of building structures atom by atom and molecule by molecule, thus providing an excellent opportunity for a bottom-up manufacturing approach. The Royal Society (2003) defines nanotechnology as the design, characterization, production, and application of structures, devices and systems by controlling shape and size at a nanometer scale. Nanomanufacturing promises to increase quality, productivity, and efficiency of existing technologies. In addition, manufacturing at the nano-scale has potential to accelerate commercialization of products and benefit industries such as semiconductor, pharmaceutical, and automotive. Realizing this potential of nanomanufacturing, there has been a tremendous increase in the amount of investments in nanotechnology research over the past few years by industrial giants like Hewlett-Packard, Intel, General Motors, and Motorola.

Doumanidis (2004) provides some illustrations of industries benefiting from nano-scale manufacturing. One such industry that can benefit, according to Doumanidis (2004), is the materials, textile, and chemical industry. Using the principles of nanotechnology, advanced materials such as nanostructured polymers, nanotubes, and non-woven fabrics from electrospun nanofibers can be developed. A second major application area is that of aerospace. Efforts are directed towards manufacturing nanostructured alloys and composite materials for structural elements of aircraft. In

addition to the above potential impacts, the author also provides several case studies in which nanotechnology has been successfully applied in several industries.

Nanotechnology is defined by Bhushan (2004) as “any technology performed on a nano-scale having applications in the real world”. There is a popular belief that nanotechnology is likely to be the next industrial revolution, creating breakthroughs in areas like manufacturing, electronics, medicine, and biotechnology. One of the most promising devices at the nano-scale is nano-electro-mechanical-systems (NEMS). NEMS devices can be thought of as mechanical components on the nano scale that have the ability to measure small displacements and forces at a molecular scale. Generally, electromechanical systems have two principal components, *viz.* a mechanical element and transducers. The movement of the mechanical element is considered to be the output of the electromechanical system. An important characteristic of NEMS is that they dissipate very little energy, due to which they are extremely sensitive to external damping mechanisms, crucial to building many sensors.

While NEMS have tremendous potential, its progress is hindered due to issues like reliability in adhesion, wear, and contamination arising from its large surface area to volume ratio. Bhushan (2004) advocates a multidisciplinary, system-oriented approach to manufacturing NEMS that function reliably. The author suggests that such an approach can be only achieved through cross-fertilization of ideas from different disciplines and the systematic flow of information and people among research groups.

3.2 Challenges in Nanomanufacturing

As shown in Figure 3.1, nanomanufacturing is still in the conceptual stage of development, especially as compared to other mature technologies such as internet and semiconductors (see Busnaina et al., 2003). Researchers are currently looking at various techniques to manufacture devices at the nano-scale.

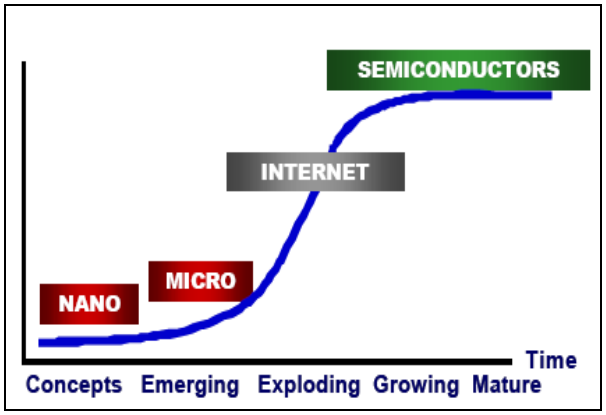

Figure 3.1: Current state of nanomanufacturing.

Two of the most common techniques of fabricating nanostructures are commonly known as top-down and bottom-up approaches. The top-down approach has been the only approach employed until recent times. It involves molding or etching materials into smaller components and has traditionally been used in making parts for computers and electronics. More formally, the top-down approach can be defined as reduction in structure sizes of microscopic elements to the nanometer scale by applying specific machining and etching techniques like lithography, etching, and grinding (see Anderson, 2005). This approach has been known to offer more reliability although producing more waste than bottom-up approaches.

On the other hand, bottom-up nanomanufacturing combines several atoms or molecules to create new devices. Although this idea is only recently being explored, the origin dates back to Feynman (1960). The bottom-up approach can be defined as the controlled assembly of atomic and molecular aggregates into larger systems. Some examples of such approaches are self-assembly and molecular fabrication. This approach provides a crucial building block to create devices and structures atom-by-atom or molecule-by-molecule. Currently, processes employing bottom-up approach are been investigated by researchers and are considered as the state-of-the-art in fabrication of structures at the nano-scale.

Scientific properties of the atoms combined to create a structure are still greatly unknown. As every atom being considered can potentially have a significant impact on the nanostructure, high-precision manufacturing techniques need to be considered at the nano-scale. Although the bottom-up approach has been applied with reasonable success, its application has been limited due to a host of reasons. A workshop conducted by the National Science Foundation (NSF) (see Busnaina et al., 2003) recognizes that one of the foremost challenges that nanomanufacturing faces is that of scaling-up. Scaling-up can be defined as the translation of the research discoveries in the laboratory to commercially available products.

In order to effectively scale-up nanotechnology to achieve mass production, several key issues need to be addressed. Some of these key issues have been recognized as low reliability and yield for nano-scale devices, repeatability and reproducibility in yielding a particular product, and lack of control of the nanomanufacturing processes. As

this technology is still in its discovery stage, there is a tremendous amount of experimentation occurring every day.

Consider a self-assembly process of creating patterns of wiring and transistor circuitry, where the structures and devices are fabricated via layer-by-layer self-assembly of molecules and atoms. These processes are typically conducted over multiple stages. Depending on the expertise in the area, the experimenter suggests a particular combination and constitution of elements to obtain a desirable output (usually width of the pattern).

Often, the first experiment does not provide desired results and hence has to be repeated. This process might take several attempts before one can get a desired output. Moreover, even when a desired output is obtained, repeatability and reproducibility are always a concern. This can be attributed to the lack of comprehensive knowledge, not only about the scientific properties, but also the process conditions. For example, a change in room temperature or humidity can cause the same experimental setup to produce different sets of results. In addition, as only the final result is observed, the failure of an experiment cannot be attributed to any particular stage.

Furthermore, Doumanidis (2002) suggests that a combination of theoretical and experimental methodologies is appropriate to address the key issues in successful scaling-up of these nanomanufacturing processes. In particular, the author suggests that predictable fabrication with repeatable quality requires fundamental understanding, modeling, and control of phenomena during processing. It is imperative to establish a set of conditions for controlling a process to obtain high reproducibility.

This research proposes that, in addition to the growth in scientific knowledge, statistically based designs can be extremely helpful in expediting the translation of nanotechnology to robust products that can be mass produced. These designs are not only useful in comprehending the effects of design factors on the final result of the process but also help in studying the effect of environmental factors such as temperature and humidity on the process. Thorough understanding of such effects will eventually result in more robust processes that can be repeated under similar operating conditions. Nembhard et al. (2005) illustrate how statistically-based design of experiments can reveal otherwise unknown relationships between process parameters.

As seen in Section 2.3, an increasing number of researchers are considering the use of structured DOE to investigate characteristics of processes and minimize variations due to external variables. However, many researchers studying the process lack expertise in designing and analyzing complex experiments and often times simple designs such as a full or a fractional factorial design is used. Similarly, there are many processes considered in the past literature for which an advanced design could yield equal or more information, while utilizing the same or fewer resources. These instances strongly advocate the use of DOE techniques, both basic and advanced to effectively investigate the processes and make them robust to uncontrollable factors.

For instance, a study conducted by Basumallick et al. (2003) investigates the effects of various factors on the synthesis of nanocomposites. Three process variables, initial temperature, heating rate, and weight percent metal to be liberated by non-isothermal reduction, were varied in accordance to a 2^3 full factorial design, and the responses (fractional conversions of $NiCl_2$ to Ni and $CoCl_2$ to Co) measured. However, a

second set of factors involving final temperature, amount of dextrose, and volume ratio of ethyl alcohol solution were kept constant throughout the experiment, as they were believed not to affect the responses. Although the statistical experimental design was used successfully to predict future values, given as set of conditions, the interaction effects involving the two sets of factors were completely ignored. One possible reason for the ignorance could be the time involved in setting the levels of the second set of variables. As shown in Chapter 5, the use of a split-plot design could potentially yield much more information, thus making the prediction ever more accurate.

Moreover, much of the work in nanomanufacturing utilizes one-factor-at-a-time (OFAT) approach to study the effects of the process parameters on the output (see Chen et al., 2006; Unalan and Chhowalla, 2005; Kim and Shahinpoor, 2003; and Huang et al., 2006 for some examples). The biggest downfall of the OFAT approach is the complete ignorance of the interaction effects. Thus, an inherent assumption being made is that the interaction effects between these variables are inconsequential and can lead to severe bias in any conclusions drawn on the relation between the input and output variables.

In summary, nanomanufacturing research could potentially benefit from well-planned experimentation. In addition, with the inclusion of the external or environmental factors, sensitivity of the design variables can be examined, and could lead to significant increase in reproducibility and repeatability. Two such cases that demand the use of statistical experimentation are presented in Chapters 4 and 5.

4. MISSING OBSERVATIONS IN 2^k AND 2^{n-k} DESIGNS

In a research environment, often times the behavior of the process variables are understood with the help of a statistical design. The designs advocate a structured method of testing appropriate combinations of the process variables of interest. When a fractional factorial design is used, it is possible to use any of the fractions of the design so that all the combinations of variables yield a response. However, obtaining such a fraction of a design is sometimes an impossible task. Hence, appropriate measures have to be taken to account for the missing observations. Such an instance in the investigation of a lubrication process is presented in the next section. The remainder of the chapter presents three new algorithms for the estimation of missing observations. Some theoretical relations between the algorithms are also discussed. Performance of the algorithms over different conditions is also examined.

4.1 Motivating Example: Nano-scale Lubrication

We consider a set of gas-phase nano-scale lubrication experiments for the study of tribological behaviors of C-6 alcohol molecules (alcohol molecules containing six carbon atoms). Lubrication problems in MEMS devices are typically solved using solid phase lubrication (e.g., Sullivan et al., 2001) or self-assembled monolayers (e.g., Ashurst et al., 2001). However, these lubrication techniques have limitations particularly in their application to sidewalls and buried surfaces (refer to Figure 4.1).

Hence, a gas-phase lubrication approach is proposed by Strawhecker et al. (2006). The essence of the approach is to use adsorption equilibrium of gas-phase molecules for continuous formation of liquid films on the working device surface for lubrication and anti-stiction. Upon adsorption on substrate surfaces, these molecules form a nano-scale lubrication and anti-stiction film. Since the molecular structures of alcohols only exist in certain combinations, there will be missing values in the experimental design.

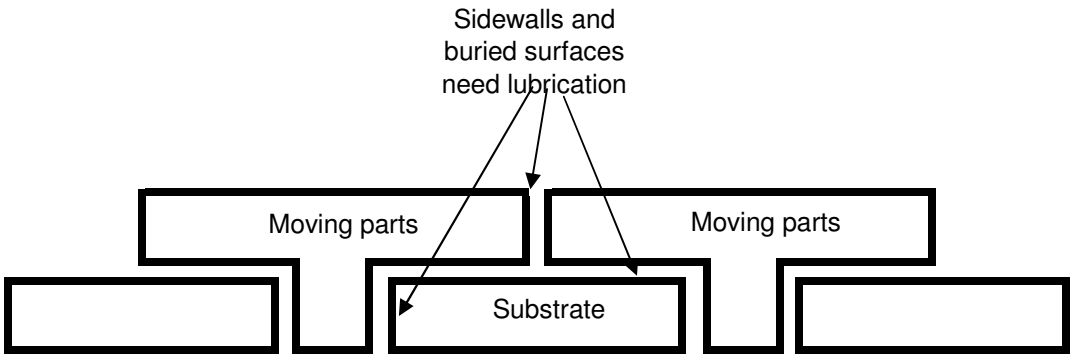

Figure 4.1: Nano-scale lubrication for MEMS devices.

Strawhecker et al. (2006) investigate the vapor phase lubrication for MEMS surfaces. More specifically, they report on the effect of saturated alcohol vapors including ethanol, n-propanol, n-butanol, and n-penatanol on the tribological properties of a model MEMS surface: the native silicon oxide surface formed on a silicon wafer.

We want to design an experiment to extend the study to understand the effect of molecular structure in C-6 alcohol molecules. These molecules have four main structural characteristics of interest listed as factors *A*, *B*, *C*, and *D* in Table 4.1. A full factorial

design to explore these characteristics would involve $2^4 = 16$ treatment combinations. Table 4.1 highlights only those molecules that are physically possible. For example, there exists no molecule that is linear, non-cyclic, alken (containing a double bond), and a secondary alcohol (corresponding to treatment combination *ab*).

Table 4.1: Factors of Interest and Available Molecules

A "+" linear "-" branched	**B** "+" non "-" cyclic	**C** "+" alkly "-" alken	**D** "+" primary alcohol "-" secondary OR tertiary alcohol	**Treatment**	**Molecule**
-	-	-	-	*(l)*	
+	-	-	-	*a*	Phenol, Cyclohexenal
-	+	-	-	*b*	1-penten-3-ol, 4-methyl
+	+	-	-	*ab*	
-	-	+	-	*c*	methylcyclopentanol
+	-	+	-	*ac*	cyclohexanol
-	+	+	-	*bc*	2,3-methyl-2-butanol
+	+	+	-	*abc*	3-hexanol
-	-	-	+	*d*	
+	-	-	+	*ad*	
-	+	-	+	*bd*	cis-3-hexen-1-ol
+	+	-	+	*abd*	trans-3-hexen-1-ol
-	-	+	+	*cd*	
+	-	+	+	*acd*	
-	+	+	+	*bcd*	3-ethyl-1-butanol
+	+	+	+	*abcd*	n - hexanol

Due to this physical unavailability of some of the C-6 alcohol molecules, if this 2^4 experiment was to be carried out, only 10 observations could be recorded and we would be left with 6 missing observations (over one-third missing). To reduce the proportion of missing observations, we considered a 2^{4-1} design. In this case, even the most efficient

combination of factor arrangement and generator choice yields at least two missing observations out of a total of eight runs.

4.2 Estimating Missing Observations

For the given data set with missing observations, the objective is to model the response as a function of the independent variables. A model not only helps in studying the optimal settings for the process, but aides in predicting values for a particular combination of a molecule. A number of techniques can be adopted to meet the above objective. One of the most familiar techniques is least squares regression (e.g. ordinary linear regression, stepwise regression, best subsets regression, etc.). However, as noted by Chipman (1998) and Samset and Tyssedal (1998), traditional approaches like least squares do not account for any uncertainty in the parameters of the model and the chosen model itself. There are several examples in the literature that show that, for a given data set, more than one model can be fit that yields approximately same values of the selection criteria. The problem of model uncertainty is even more acute in the case of a small data set (Rajagopal and del Castillo, 2005). Chipman (1998) states that the ignorance of model and parameter uncertainty can lead to unrealistic improvements and perhaps even sub-optimal performance. The author illustrates the perils of ignoring the uncertainties through an example. When parameter uncertainty is not accounted for by a model, the variability in the parameter estimates is transmitted to the variation in the response variable.

The above uncertainties can be accounted for through a Bayesian approach to model building. This approach involves the calculation of the posterior probabilities of the response based on assumed models. Consider a linear model represented by

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{4.1}$$

where

$\boldsymbol{y}$ is the vector of responses,

$\boldsymbol{X}$ is the matrix of independent variables,

$\boldsymbol{\beta}$ is the matrix of coefficients, and

$\boldsymbol{\varepsilon}$ are the errors assumed to be i.i.d. N(0, σ^2).

Let f_i be the number of terms (out of k factors) for a model M_i.

Let π_j (j = 1,.., k) be the probability of effects being in a model.

Model uncertainty is accounted for through the specification of prior probabilities (π_j) for a model. If the probabilities are assumed to be independent, then the prior probability of a model M_i can be written as

$$P(M_i) = \pi^{t_i} (1-\pi)^{k-t_i}. \tag{4.2}$$

In many applications, the researcher conducting the study has some past belief about the significance of the variables being studied. Scientific knowledge of the process under consideration usually provides the probabilities of effects being active. In situations where no past beliefs can be entertained, an alternate technique known as model averaging can be used. In this approach, a weighted average of the posterior predictive

densities of all models is computed, using the posterior probabilities of the models as weights (Hoeting et al., 1999; Raftery et al., 1997).

To study the incorporation of parameter uncertainty in the model, let us look at the Bayesian approach in more detail. Using Bayes' theorem, the posterior probability of model M_i, given the data is

$$P(M_i \mid \boldsymbol{y}) = \frac{P(M_i)P(\boldsymbol{y} \mid M_i)}{\sum_{h=0}^{m} P(M_h)P(\boldsymbol{y} \mid M_h)} \tag{4.3}$$

where

$$P(\boldsymbol{y} \mid M_i) = \int_{\sigma^2} \int_{\beta_i} P(\boldsymbol{y} \mid M_i, \sigma^2, \beta_i) P(\sigma^2, \beta_i \mid M_i)\, d\beta_i, \sigma^2. \tag{4.4}$$

For a more detailed derivation, see (Box and Meyer, 1993; Rajagopal and Del Castillo, 2005). The posterior probability of the model M_i can then be written as

$$P(M_i \mid \boldsymbol{y}) = C\left(\frac{\pi}{1-\pi}\right)^{f_i} \gamma^{-t_i} \frac{|\boldsymbol{X_0}'\boldsymbol{X_0}|^{1/2}}{|\boldsymbol{\Gamma_i} + \boldsymbol{X_i}'\boldsymbol{X_i}|^{1/2}} \left(\frac{S(\hat{\boldsymbol{\beta}}_i) + \hat{\boldsymbol{\beta}}_i'\boldsymbol{\Gamma_i}\hat{\boldsymbol{\beta}}_i}{S(\hat{\boldsymbol{\beta}}_0)}\right)^{-(n-1)/2} \tag{4.5}$$

where

C is the normalization constant, which forces all probabilities to sum to one,

γ is the scale factor relating the magnitude of real effects to noise,

$\boldsymbol{I}_{ti}$ is the identity matrix having dimension $t_i \times ti$, and

$$\boldsymbol{\Gamma}_i = \frac{1}{\gamma^2}\begin{bmatrix} 0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{t_i} \end{bmatrix} \quad .$$

The parameter estimates for the model M_i can be calculated as

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{\Gamma}_i + \boldsymbol{X}_i' \boldsymbol{X}_i)^{-1} \boldsymbol{X}_i' \boldsymbol{y} \quad . \tag{4.6}$$

The assumptions for the above calculations are that all elements of $\boldsymbol{\beta}_i$, except the constant term, are independent ~$N(0, \gamma^2\sigma^2)$. The magnitude of effect relative to the experimental noise is captured through parameter γ. The constant term and $log(\sigma^2)$ are assumed to have non-informative priors. Thus, the posterior on $\boldsymbol{\beta}_i$, given σ, captures the parameter uncertainty for a given model (see Appendix of (Chipman, 1998) for more details).

When the assumption on the distribution of $\boldsymbol{\beta}_i$ is relaxed and a non-informative prior is used, the parameter estimates from the Bayesian approach and the classical least squares approach are the same (Gelman et al., 1995). When the assumption is relaxed, the terms γ and Γ_i would not exist and the parameter estimates for the model M_i can be written as

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{X}_i' \boldsymbol{X}_i)^{-1} \boldsymbol{X}_i' \boldsymbol{y} \quad . \tag{4.7}$$

These parameter estimates are the same results obtained from the traditional least squares regression (Kutner et al., 2005). Finally, the marginal posterior probability, P_j is

the sum of the posterior probabilities of all models in which factor *j* is active and is given by

$$P_j = \sum_{M_i:\text{factor j active}} P(M_i \mid \boldsymbol{y}) \quad \textbf{(4.8)}$$

Factor *j* is said to be active only when the value of P_j is large. Box and Meyer (1993) show that this method can identify plausible explanations missed by conventional analysis (ANOVA using least squares approach). The foremost advantage of this method is that it can be applied to analyze any experiment, irrespective of the structure of the design. Furthermore, it can handle other irregularities such as missing observations.

One of the major limitations of this method is that it is most appropriate only for screening experiments, where the idea is to identify factors to be explored through subsequent experiments. If the objective of the experiment is to screen out potentially insignificant variables, then estimation might not be necessary. The Bayesian approach would work well to meet this objective.

However, when the objective is to obtain optimal settings of the variables and study the effects of noise variables on the design variables (interaction effects), the Bayesian approach might not yield desired results. Consider a 2^{4-1} design with defining relation $I = ABCD$. The corresponding design matrix is shown in Table 4.2.

Table 4.2: 2^{4-1} Design with $I = ABCD$

Run	A	B	C	D	AB	AC	AD
1	-1	-1	-1	-1	1	1	1
2	1	-1	-1	1	-1	-1	1
3	-1	1	-1	1	-1	1	-1
4	1	1	-1	-1	1	-1	-1
5	-1	-1	1	1	1	-1	-1
6	1	-1	1	-1	-1	1	-1
7	-1	1	1	-1	-1	-1	1
8	1	1	1	1	1	1	1

If runs 1 and 2 were missing, the resulting ($\boldsymbol{X'X}$) matrix would not be positive definite and hence, not invertible. Some effect(s) need to be sacrificed for the inverse to exist. Goh (1997) and Draper and Stoneman (1964) advocate the use of assumptions that set some effects to be negligible and thereby yielding estimates of missing observations. Nembhard et al. (2006) have shown certain fractional factorial designs in which this technique fails. The next three sections present efficient algorithms that impute estimates of missing observations and estimate effects of interest.

4.3 Finding Active Factors: Bayesian Screening Algorithm

The Bayesian screening algorithm proposed by Box and Meyer (1993) suggests consideration of a set of models (usually an exhaustive set), $M_0, M_1,...,M_m$. Each model M_i has an associated vector of parameters $\boldsymbol{\theta_i}$ so that the sampling distribution of data $\boldsymbol{y}$, given the model M_i is described by the probability density $P(\boldsymbol{y}|M_i,\boldsymbol{\theta_i})$. The prior probability of model M_i is given by $P(M_i)$. The prior probability density of $\boldsymbol{\theta_i}$ is $P(\boldsymbol{\theta}_i \mid M_i)$. The posterior probability of model M_i, given the data is

$$P(M_i \mid \boldsymbol{y}) = \frac{P(M_i)P(\boldsymbol{y} \mid M_i)}{\sum_{h=0}^{m} P(M_h)P(\boldsymbol{y} \mid M_h)} \tag{4.9}$$

where

$$P(\boldsymbol{y} \mid M_i) = \int_{\boldsymbol{\theta}_i} P(\boldsymbol{y} \mid M_i, \boldsymbol{\theta}_i)P(\boldsymbol{\theta}_i \mid M_i)\, d\boldsymbol{\theta}_i . \tag{4.10}$$

To simplify calculations, Box and Meyer adopt an alternate approach and show that the posterior probability of the model M_i can be written as

$$P(M_i \mid \boldsymbol{y}) = C\left(\frac{\pi}{1-\pi}\right)^{f_i} \gamma^{-t_i} \frac{|\boldsymbol{X_0}'\boldsymbol{X_0}|^{1/2}}{|\boldsymbol{\Gamma_i} + \boldsymbol{X_i}'\boldsymbol{X_i}|^{1/2}} \left(\frac{S(\hat{\boldsymbol{\beta}}_i) + \hat{\boldsymbol{\beta}}_i'\boldsymbol{\Gamma_i}\hat{\boldsymbol{\beta}}_i}{S(\hat{\boldsymbol{\beta}}_0)}\right)^{-(n-1)/2} \tag{4.11}$$

where

π is the prior probability that any one factor is active,

$\boldsymbol{X_i}$ is the design matrix including a column for the mean,

t_i is the number of effects in a model excluding the mean,

$\boldsymbol{\beta_i}$ is the vector of true effects under M_i,

C is the normalization constant, which forces all probabilities to sum to one,

γ is the scale factor relating the magnitude of real effects to noise, and

$\boldsymbol{I_{ti}}$ is the identity matrix of dimension $t_i * t_i$.

Equation 4.11 can be solved using Equations 4.12, 4.13, and 4.14.

$$\boldsymbol{\Gamma}_i = \frac{1}{\gamma^2}\begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{t_i} \end{bmatrix}. \tag{4.12}$$

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{\Gamma}_i + \boldsymbol{X}_i' \boldsymbol{X}_i)^{-1} \boldsymbol{X}_i' \boldsymbol{y}. \tag{4.13}$$

$$S(\hat{\boldsymbol{\beta}}_i) = (\boldsymbol{y} - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_i)'(\boldsymbol{y} - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}_i). \tag{4.14}$$

The marginal posterior probability, P_j is the sum of the posterior probabilities of all models in which factor j is active and is given by

$$P_j = \sum_{M_i:\text{factor j active}} P(M_i \mid \boldsymbol{y}). \tag{4.15}$$

Factor j is said to be active only when the value of P_j is large. Box and Meyer (1993) show that this method can identify plausible explanations missed by conventional analysis (ANOVA). The foremost advantage of this method is that it can be applied to analyze any experiment, irrespective of the structure of the design. Furthermore, it can handle other irregularities such as missing observations.

One of the major limitations of this method is that it is most appropriate only for screening experiments, where the idea is to identify factors to be explored through subsequent experiments. In the presence of missing observations, if the objective of the experiment is to obtain a model, like in the case of the C-6 alcohol molecules, this Bayesian method might fail (if certain observations are missing) as it is unable to estimate the effects (coefficients) of the model. Finally, the interaction effects are deemed to be significant or not solely based on the interaction plots. In the case of even a single missing observation, severe bias might be introduced into the interaction plot, which may hamper the identification of the significant interaction effects. New algorithms are proposed that overcome the limitations of the existing algorithms. The proposed algorithms extend the principle of the algorithms presented in Chapter 2 by including the information provided by the Bayesian screening algorithm. The next three sections explain the principle and limitations of the existing algorithms. Bayesian algorithms, which are extensions to the existing algorithms (by incorporating the information from the screening algorithm) are also presented.

4.4 Bayesian Complete RSS Minimization Algorithm

Shearer (1973) suggests an iterative algorithm to be used on full and fractional factorial designs in the case of missing observation(s). Rather than just minimizing a part of the residual sum of squares where the residual is estimated by the highest order interaction as in Draper and Stoneman (1964), the complete RSS minimization algorithm minimizes the complete residual sum of squares. It is a two-stage iterative algorithm given by

$$y_{miss}^{(i+1)} = \sum_{j=1}^{m} b_j^{(i)} x_{\text{miss } j} + \frac{1}{n}\sum_{k=1}^{n} y_k^{(i)} \qquad \textbf{(4.16)}$$

where

$$b_k^{(i)} = \frac{1}{n}\sum_{j=1}^{n} x_{jk}\, y_j^{(i)} \qquad (k = 1,2,.....,m). \qquad \textbf{(4.17)}$$

Equation 4.16 re-estimates the missing value using the new main effects estimates and Equation 4.17 is just the main effect estimation procedure for a 2^m design. The principle behind the algorithm is that the estimate of the missing observation be found such that the residual sum of squares due to the missing observation is zero. As the total residual sum of squares is comprised of the residual sum of squares due to the existing and missing observations, if the residual sum of squares due to the missing observations is zero, the total residual sum of squares is minimized.

As the estimate of the missing observation is obtained using the information about the existing observations and an assumed model (that includes only the main effects), the residual sum of squares is minimized for the particular assumed model. Although the residual sum of squares is minimized, the predictive ability of the model may have decreased since there may be some insignificant terms in the model assumed to estimate the missing observation(s) (Kutner et al., 2005). The Bayesian complete RSS minimization algorithm is given by

$$y_{miss}^{(i+1)} = \frac{1}{n}\sum_{k=1}^{n} y_k^{(i)} + \sum_{j \in Bayesian-specified}^{m} b_j^{(i)}\, x_{miss\ j} + \sum_{u,v \in Bayesian-specified}^{m} b_{uv}^{(i)}\, x_{miss\ u} x_{miss\ v} \qquad \textbf{(4.18)}$$

$$b_k^{(i)} = \frac{1}{n}\sum_{j=1}^{n} x_{jk}\, y_j^{(i)} \qquad (k = 1, 2, ..., m) \tag{4.19}$$

$$b_{uv}^{(i)} = \frac{1}{n}\sum_{p=1}^{n} x_{pu} x_{pv}\, y_p^{(i)} \qquad (u, v = 1, 2, ..., m) \tag{4.20}$$

where

b_{uv} is the effect of interaction uv.

Equation 4.18 re-estimates the missing observation using the Bayesian-specified terms calculated iteratively from Equations 4.19 and 4.20. The proposed Bayesian complete RSS minimization algorithm combines information about the significant terms in the model from the initial Bayesian screening algorithm and the complete RSS minimization algorithm. The Bayesian complete RSS minimization algorithm estimates the missing observation(s) using a model that only includes terms specified by the Bayesian analysis. This ensures that the predictive ability of the model is maximized, as non-significant terms are not included in the model.

4.5 Bayesian ANCOVA Algorithm using Dummy Variables

John and Prescott (1975) present an algorithm that makes use of dummy variables in order to estimate missing observation(s). It considers a linear model given by

$$\boldsymbol{y = Xa + Zb + e} \qquad \textbf{(4.21)}$$

where

$\boldsymbol{y}$ is the vector of observations,

$\boldsymbol{X}$ is the design matrix,

$\boldsymbol{Z}$ is the matrix of dummy variables,

e is the vector of errors, and

$\boldsymbol{a}$, $\boldsymbol{b}$ are parameter vectors.

This algorithm is based on Bartlett's (1937) algorithm and is extremely easy to use with very little computation. This algorithm is non-iterative with every missing observation in $\boldsymbol{y}$ given by zero. The dummy matrix $\boldsymbol{Z}$ has a column for each missing observation such that the i^{th} column $\boldsymbol{z_i}$ has a 1 corresponding to the i^{th} missing value and 0 everywhere else. The principle behind this algorithm is to estimate the missing observation(s) by carrying out a standard analysis of covariance using a dummy covariate for each of the *m* missing values, a model for which is given by

$$\boldsymbol{z_i = Xa^* + e} \qquad \textbf{(4.22)}$$

where

$\boldsymbol{a^*}$ is a parameter vector,

$\boldsymbol{r_i}$ is the vector of residuals after fitting the model in Equation 4.22, and

$\boldsymbol{R}$ is the matrix of residuals and is given by ($\boldsymbol{r_1, r_2,.....r_m}$).

John and Prescott show that the vector of estimates of the missing observations, $\hat{\boldsymbol{\theta}}$ is given by

$$\begin{aligned}\hat{\boldsymbol{\theta}} = -\hat{\boldsymbol{b}} &= -(\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}'\boldsymbol{y} \\ &= -(\boldsymbol{Z}'\boldsymbol{R})^{-1}\boldsymbol{Z}'\boldsymbol{r}_y\end{aligned} \tag{4.23}$$

where $\boldsymbol{r}_y$ is the vector of residuals after fitting the model $\boldsymbol{y} = \boldsymbol{Xa} + \boldsymbol{e}$. The i^{th} column of $\boldsymbol{Z}'\boldsymbol{R}$, $\boldsymbol{Z}'\boldsymbol{r}_i$ consists of the elements of $\boldsymbol{r}_i$, for those cells containing missing values. $\boldsymbol{Z}'\boldsymbol{r}_y$ is the vector containing the corresponding residuals in $\boldsymbol{r}_y$.

The residuals, $\boldsymbol{r}_i$ and $\boldsymbol{r}_y$, are obtained through models that are assumed to have main effects and first-order interactions only. One of the main disadvantages of assuming such models is that there is a tendency to over-fit and hence in some cases, the residuals are zero. It is seen from Equations 4.22 and 4.23 that if any residual is zero, the estimate of the missing observation(s) is also zero, thus yielding inaccurate results.

This drawback can be overcome by improving the predictive ability of the assumed models. One way to improve the predictive ability is to use the *a priori* information from Bayesian screening algorithm (Section 4.3). The Bayesian ANCOVA algorithm estimates the value(s) of the missing observation(s) non-iteratively through Equations. 4.24 and 4.25 given by

$$\hat{\boldsymbol{\theta}} = -(\boldsymbol{Z}'\boldsymbol{R})^{-1}\boldsymbol{Z}'\boldsymbol{r}_{y*} \tag{4.24}$$

$$\boldsymbol{z}_{i^*} = \boldsymbol{Xa}^* + \boldsymbol{e} \tag{4.25}$$

where $\boldsymbol{r_{y*}}$ is the vector of residuals obtained from fitting the model $\boldsymbol{y^* = Xa + e}$. Furthermore, $\boldsymbol{r_{i*}}$ is the vector of residuals obtained by fitting a model $\boldsymbol{z_{i*} = Xa^* + e}$. The highlight of these models is that $\boldsymbol{y^*}$ and $\boldsymbol{z_{i*}}$ include only the terms specified by the Bayesian screening algorithm described in Section 4.3. As a result, the predictive ability of the models is maximized by adding only the significant terms in the models. Residuals obtained from these models would be more accurate (Kutner et al., 2005) and hence, from Equation 4.24, the estimates of the missing observations would be more accurate. In addition, the contribution of the missing observation(s) towards the total residual sum of squares is still zero and the residual sum of squares is minimized.

4.6 Bayesian Expected Value Algorithm

Healy and Westmacott (1956) describe a general iterative algorithm that is applicable to any analysis in which least squares estimates are derived. The theory behind the algorithm is that the fictitious values are actually the expected values of the missing observations so that the residuals for the missing observations after fitting all the constants must all be zero. The procedure is as follows:

i. Insert a guessed value for the missing observation.
ii. Determine the residuals assuming a model containing main effects and first-order terms.
iii. Subtract the value of the residual from the guessed value of the missing observation.

iv. This replaces each guessed value by the corresponding expected value derived from the analysis based on the guessed values.
v. This process is repeated as often as necessary.
vi. When the residuals for the missing observations are small enough, the required fictitious values have been attained and the analysis of variance can be completed in the usual way.

The residual sum of squares due to missing observation(s) in each iteration is zero, and hence the total residual sum of squares is minimized. However, if non-significant terms are included in the model that is used to determine the estimate(s) of the missing observation(s), the predictive ability of the model will deteriorate. The additional terms just explain more variability and do not contribute towards the predictive ability of the model.

In addition to minimizing the total residual sum of squares, if an additional criterion could be added - that of increasing prediction ability - the residuals would be more accurate and in turn, the estimates would be more accurate. Increased prediction ability for a model is achieved by including only the terms in the model those are statistically significant (Kutner et al., 2005 and Montgomery, 2005). Therefore, any *a priori* information pertaining to the significant terms in the model would be highly beneficial in improving the predictive ability of the model. This information can be obtained from the Bayesian screening algorithm. The iterative steps proposed in the Bayesian Expected value algorithm are given by Equations 4.26 through 4.29.

$$y_{miss}^{(i)} = a \tag{4.26}$$

$$\boldsymbol{\beta}_j^{(i)} = (\boldsymbol{X}_j^{(i)\prime}\boldsymbol{X}_j^{(i)})^{-1}\boldsymbol{X}_j^{(i)\prime}\boldsymbol{Y}^{(i)} \qquad (\forall\ j \in \text{Bayesian} - \text{specified terms}) \tag{4.27}$$

$$r_{miss}^{(i)} = y_{miss}^{(i)} - \beta_0^{(i)} - \sum_{j \in \text{Bayesian-specified term}} \beta_j^{(i)} x_j^{(i)} \tag{4.28}$$

$$y_{miss}^{(i+1)} = y_{miss}^{(i)} - r_{miss}^{(i)} \tag{4.29}$$

where

a is the inital assumed value,

$\boldsymbol{\beta}_j^{(i)}$ is the estimate of the effect of term j in i^{th} iteration,

$\boldsymbol{X}_j^{(i)}$ is the column of the design matrix $\boldsymbol{X}$ corresponding to the j^{th} term,

$\boldsymbol{Y}^{(i)} = (y_1, y_2, ..., y_{miss}^{(i)}, ...y_n)'$, and

$r_{miss}^{(i)}$ is the residual for missing observation in i^{th} iteration.

The Bayesian expected value algorithm starts with an initial assumed value. The effects in Equation 4.27 are calculated using columns of the design matrix, $\boldsymbol{X}_\mathbf{j}$ corresponding to the terms specified by the Bayesian screening method. Once the effects have been obtained, the residual for the missing observation(s) can be obtained from Equation 4.28. These residuals are then used (from Equation 4.29) to obtain newer estimate(s) of the missing observation(s). This process is repeated until convergence and the latest estimate(s) of the missing observation(s) is treated as the final estimate(s).

Healy and Westmacott (1956) provide a proof to show that the residual sum of squares is minimized. This minimization holds true for the proposed Bayesian expected value algorithm as well because in every iteration the value of the residual is subtracted from the guessed value. Thus, the total residual sum of squares is not artificially inflated above its least squares value.

4.7 Properties of Bayesian Algorithms

To further understand missing observations, the properties of the proposed Bayesian algorithms are explored.

4.7.1 Equality for a single missing observation

The relationship between the values obtained by the Bayesian Complete RSS Minimization and the Bayesian ANCOVA algorithms in the presence of a single missing observation is explained by the following theorem.

Theorem 1:

If one observation is missing in an *n*-run design, then

$$\hat{y}_{RSS} = \hat{y}_{ANCOVA} \tag{4.30}$$

where

$\hat{y}_{RSS}$ is the estimate of the missing observation obtained by using the Bayesian Complete RSS minimization algorithm, and

$\hat{y}_{ANCOVA}$ is the estimate of the missing observation obtained by using the Bayesian ANCOVA algorithm.

Proof:

Let *n* be the number of runs for the design and *m* be the number of effects in the model specified by the Bayesian screening algorithm. Let $\hat{y}_{RSS}^{(i)}$ be the estimate in the i^{th} iteration and *Y* represent the limit of the sequence $\hat{y}_{RSS}^{(i)}$. If $p = \frac{n-m-1}{n}$, then Shearer (1973) shows that

$$\frac{\hat{y}_{RSS}^{(i)} - \hat{y}_{RSS}^{(i-1)}}{Y - \hat{y}_{RSS}^{(0)}} = p(1-p)^{i-1}.$$

This means that each estimate reduces the distance to *Y* by *p*. the estimate of the missing observation is then shown to be

$$Y = \frac{1}{p}(\hat{y}_{RSS}^{(1)} - \hat{y}_{RSS}^{(0)}) + \hat{y}_{RSS}^{(0)}. \quad \textbf{(4.31)}$$

In Equation 4.31, $\hat{y}_{RSS}^{(1)}$ needs to be calculated using an initial arbitrary value for $\hat{y}_{RSS}^{(0)}$. For ease of calculation, let us assume

$$\hat{y}_{RSS}^{(0)} = 0 \ . \quad \textbf{(4.32)}$$

Therefore,

$$Y = \frac{(\hat{y}_{RSS}^{(1)})}{p}. \quad \textbf{(4.33)}$$

Estimates for missing observations using Bayesian ANCOVA are given by Equation 4.23. Hence, the estimate in the case where only one observation is missing is given by

$$\hat{y}_{ANCOVA} = -(r_{z_1})^{-1} r_y = -\frac{r_y}{r_{z_1}} \tag{4.34}$$

where

r_{z_1} is the residual obtained by fitting a model, having terms specified by the Bayesian screening algorithm, to a covariate vector containing 1 corresponding to the missing observation and 0 everywhere else, and

r_y is the residual obtained by fitting a model, having terms specified by the Bayesian screening algorithm, to the response having 0 corresponding to the missing observation.

Hence,

$$r_y = 0 - (\beta_0 \sum_{j=1}^{m} \beta_j x_j). \tag{4.35}$$

In addition, Equation 4.18 can be re-written as

$$\hat{y}_{RSS}^{(i+1)} = \hat{y}_{RSS}^{(i)} + \frac{1}{n} \sum_{\forall k \neq miss} y_k^{(i)} + \sum_{\forall j=m} b_j^{(i)} x_{miss\ j}\ . \tag{4.36}$$

For the first iteration,

$$\hat{y}_{RSS}^{(1)} = \hat{y}_{RSS}^{(0)} + \frac{1}{n} \sum_{\forall k \neq miss} y_k^{(0)} + \sum_{\forall j=m} b_j^{(0)} x_{miss\ j}\ . \tag{4.37}$$

From Equation 4.32 and by comparing Equations 4.35 and 4.37, it can be seen that

$$r_y = -\hat{y}_{RSS}^{(1)} . \qquad \textbf{(4.38)}$$

Due to the result in Equation 4.38, Theorem 1 is true iff $p = r_{z_1}$.

In order to calculate r_{z_1} , β_i's need to be calculated. For a particular missing observation, the corresponding row is a combination of –1 and +1. If the cell corresponding to the missing observation and an effect k is –1, then $\beta_j = -1/n$. On the other hand, If the cell corresponding to the missing observation and effect k is +1, then $\beta_j = 1/n$. Furthermore, as only one cell has the value 1 and others 0, $\beta_j = 1/n$. In calculating r_{z_1} , the β_i's are multiplied by the sign of the corresponding effects. Hence,

$$\begin{aligned} r_{z_1} &= 1-(\beta_0 \pm \beta_1(\pm 1) \pm \beta_2(\pm 1) ..... \pm \beta_m(\pm 1)) \\ &= 1-\left(\frac{1}{n}+\sum_{j=1}^{m}\beta_j\right) \\ &= 1-\left(\frac{1}{n}+\sum_{j=1}^{m}\frac{1}{n}\right) \\ &= 1-\left(\frac{1}{n}+\frac{m}{n}\right) \\ &= \frac{n-m-1}{n} \end{aligned} \quad . \qquad \textbf{(4.39)}$$

Therefore, from Equation 4.39, and as p = [$(n - m - 1)/n$],

$$r_{z_1} = p . \qquad \textbf{(4.40)}$$

Hence, the estimates obtained from Bayesian Complete RSS minimization and Bayesian ANCOVA algorithms are exactly the same in the case of a single missing observation.

4.7.2 Equality of Iterative Algorithms

For any number of missing observations in a data set, the relationship between the Bayesian Complete RSS Minimization and the Bayesian Expected Value algorithms is explained by the following theorem.

Theorem 2:

If m observations are missing in an n-run design such that $0 < m < n$, then

$$\hat{y}_{RSS} = \hat{y}_{EV} \tag{4.41}$$

where

$\hat{y}_{RSS}$ is the estimate of the missing observation obtained by using the Bayesian Complete RSS minimization algorithm, and

$\hat{y}_{EV}$ is the estimate of the missing observation obtained by using the Bayesian Expected Value algorithm.

Proof:

The two equations for obtaining the estimates of missing observations using Bayesian Expect Value algorithm are given in Section 4.6 and are as follows:

$$r_{miss}^{(i)} = y_{miss}^{(i)} - \beta_0^{(i)} - \sum_{j \in \text{Bayesian-specified terms}} \beta_j^{(i)} \quad x_j^{(i)} \tag{4.42}$$

$$y_{miss}^{(i+1)} = y_{miss}^{(i)} - r_{miss}^{(i)} \quad . \tag{4.43}$$

Notice that $\hat{y}_{EV}$ is the limit of the sequence $y_{miss}^{(i)}$. Re-writing Equation 4.43 and substituting in Equation 4.42, we get

$$y_{miss}^{(i)} - y_{miss}^{(i+1)} = y_{miss}^{(i)} - \beta_0^{(i)} - \sum_{j \in \text{Bayesian-specified terms}} \beta_j^{(i)} \quad x_j^{(i)} \quad . \tag{4.44}$$

Therefore,

$$y_{miss}^{(i+1)} = \beta_0^{(i)} + \sum_{j \in \text{Bayesian-specified terms}} \beta_j^{(i)} \quad x_j^{(i)} \quad . \tag{4.45}$$

The first term in the right-hand side of Equation 4.45 is the average of all the observations in the i^{th} iteration and hence can be written as

$$\beta_0^{(i)} = \frac{1}{n} \sum_{k=1}^{n} y_k^{(i)} \, . \tag{4.46}$$

The second term in the right-hand side of Equation 4.45 represents parameter estimates for all effects (main and interaction) specified by the Bayesian screening algorithm and can be written as

$$\sum_{j \in \text{Bayesian-specified terms}} \beta_j^{(i)} \quad x_j^{(i)} = \sum_{j \in \text{Bayesian-specified}}^{m} b_j^{(i)} \quad x_{miss\ j} + \sum_{u,v \in \text{Bayesian-specified}}^{m} b_{uv}^{(i)} \quad x_{miss\ u} x_{miss\ v} \quad . \tag{4.47}$$

Hence, from Equations 4.46 and 4.47, we get

$$\beta_0^{(i)} + \sum_{j \in \text{Bayesian-specified terms}} \beta_j^{(i)} \; x_j^{(i)} = \frac{1}{n} \sum_{k=1}^{n} y_k^{(i)} \sum_{j \in \text{Bayesian-specified}}^{m} b_j^{(i)} \; x_{miss\,j} + \sum_{u,v \in \text{Bayesian-specified}}^{m} b_{uv}^{(i)} \; x_{miss\,u} x_{miss\,v} \; . \quad \textbf{(4.48)}$$

It is easy to see that the left-hand side and right-hand side represent the estimates obtained from Bayesian Expected Value and Bayesian Complete RSS minimization algorithms. Hence,

$$\hat{y}_{RSS} = \hat{y}_{EV} \; .$$

4.8 Case Study: Performance of Bayesian Algorithms in a Nano-scale Lubrication Process

We return now to the gas-phase lubrication problem presented in Section 4.1. Since the process is still in the discovery phase, we simulate the application of the six algorithms that we have presented – Bayesian and non-Bayesian versions of complete RSS minimization, Bayesian and non-Bayesian versions of ANCOVA, and Bayesian and non-Bayesian versions of expected value – on the process. This case study leads to an understanding of how the algorithms perform.

4.8.1 Performance Criterion

In each of the algorithms, the estimate of the missing observation is obtained using the respective algorithm. This estimate is plugged into the response, and the "pseudo-complete" data is analyzed as usual after decreasing the degrees of freedom of the error term by the number of missing observations. From this analysis, significant terms are identified, and a model for the data is obtained. These models differ in the inclusion of significant effects and their corresponding parameter estimates. In all, models obtained through the non-Bayesian algorithms need to be compared to their respective Bayesian counterparts.

Since the models obtained may be used to predict future values, it is important to evaluate the predictive ability of these models so that an appropriate algorithm can be suggested. In order to rank these algorithms, the prediction sum of squares (*PRESS*) criterion can be used. Some of the reasons for choosing the *PRESS* statistic are small sample size, unbiased measure of the model's predictive ability, and deletion of one observation at a time. It is well known (e.g., see Kutner et al., 2005) that the *PRESS* criterion is a measure of how well the use of the fitted values for a subset model can predict the observed responses, Y_i. It is obtained by deleting the i^{th} case from the data set, estimating the regression function for the subset model from the remaining (n-1) cases, and then using the fitted regression function to obtain the predicted value, $\hat{Y}_{i(i)}$ for the i^{th} case. The first subscript, i in $\hat{Y}_{i(i)}$ indicates that it is a predicted value for the i^{th} case and the second subscript i indicates that the i^{th} case was omitted when the regression function was fitted. Properties of the *PRESS* statistic such as expected value and relationship

between ordinary residual and the deleted residual are studied by Liu et al. (1999). The *PRESS* criterion over all observations can be defined as

$$PRESS = \sum_{i=1}^{n} (Y_i - \hat{Y}_{i(i)})^2 \ . \qquad \textbf{(4.49)}$$

Estimates of the missing observations are obtained through the six Bayesian and non-Bayesian algorithms and imputed into the data set. Thus, there are six different data sets. In the case of a single missing observation, all observations except the estimate of the missing observation are the same. Hence, any difference in the *PRESS* statistics for the data sets can be attributed to the estimates of the missing observations. Let the *PRESS* statistic for the data sets corresponding to the Bayesian algorithms be denoted by $PRESS_{B1}$, $PRESS_{B2}$, and $PRESS_{B3}$. Similarly, let $PRESS_{NB1}$, $PRESS_{NB2}$, and $PRESS_{NB3}$ denote the statistics corresponding to the non-Bayesian algorithms. Consider two data sets obtained using a non-Bayesian algorithm and its Bayesian counterpart. If the first observation is assumed missing in an eight-run design,

$$\begin{aligned} PRESS_B = {} & (Y_1 - \hat{Y}_{1(1)})_B^{\,2} + (Y_2 - \hat{Y}_{2(2)})_B^{\,2} + (Y_3 - \hat{Y}_{3(3)})_B^{\,2} + (Y_4 - \hat{Y}_{4(4)})_B^{\,2} + \\ & (Y_5 - \hat{Y}_{5(5)})_B^{\,2} + (Y_6 - \hat{Y}_{6(6)})_B^{\,2} + (Y_7 - \hat{Y}_{7(7)})_B^{\,2} + (Y_8 - \hat{Y}_{8(8)})_B^{\,2} \end{aligned} \ . \qquad \textbf{(4.50)}$$

$$\begin{aligned} PRESS_{NB} = {} & (Y_1 - \hat{Y}_{1(1)})_{NB}^{\,2} + (Y_2 - \hat{Y}_{2(2)})_{NB}^{\,2} + (Y_3 - \hat{Y}_{3(3)})_{NB}^{\,2} + (Y_4 - \hat{Y}_{4(4)})_{NB}^{\,2} + \\ & (Y_5 - \hat{Y}_{5(5)})_{NB}^{\,2} + (Y_6 - \hat{Y}_{6(6)})_{NB}^{\,2} + (Y_7 - \hat{Y}_{7(7)})_{NB}^{\,2} + (Y_8 - \hat{Y}_{8(8)})_{NB}^{\,2} \end{aligned} \ . \qquad \textbf{(4.51)}$$

As $\hat{Y}_{i(i)}$ is obtained by omitting the i^{th} observation and as all observations other than the estimate of the missing observation are equal,

$$(Y_1 - \hat{Y}_{1(1)})_B^{\,2} = (Y_1 - \hat{Y}_{1(1)})_{NB}^{\,2} \qquad \textbf{(4.52)}$$

$$PRESS_{Diff} = PRESS_{NB} - PRESS_B = \sum_{i=2}^{8} (Y_i - \hat{Y}_{i(i)})_{NB}^{\,2} - (Y_i - \hat{Y}_{i(i)})_B^{\,2}. \qquad \textbf{(4.53)}$$

This result can be extended to a general case of n observations with the j^{th} observation being missing.

$$PRESS_{Diff} = PRESS_{NB} - PRESS_B = \sum_{i \neq j}^{n} (Y_i - \hat{Y}_{i(i)})_{NB}^{\,2} - (Y_i - \hat{Y}_{i(i)})_B^{\,2}. \qquad \textbf{(4.54)}$$

As a lower value of the *PRESS* criterion represents better predictive ability for a model, a positive value of $PRESS_{Diff}$ indicates that the model obtained using the Bayesian algorithm to estimate the missing observation has better predictive ability than the model obtained by using the non-Bayesian algorithm. $PRESS_{Diff}$ could be calculated in a similar fashion when there are two missing observations in the data set.

In addition to using $PRESS_{Diff}$ as the performance metric, the estimates of the missing observations can be used to calculate the degree of bias induced by the Bayesian and non-Bayesian algorithms. This degree of bias can simply be calculated by taking the difference between the estimate of the missing observation and the actual value assumed

to be missing and diving by the actual value. In the case of more than one missing observation, the average of the biases is considered.

4.8.2 Effect of the Position of Missing Observation on Performance

For the sake of illustration, suppose that the process engineers think that factors *A* and *D* (linear or branched molecule and the type of alcohol, respectively) might have a significant effect on the response (friction coefficient). To gain insight on the behavior of the positions of missing observations, the Bayesian and non-Bayesian algorithms were employed on three data. The data were generated from a model having two main effects (*A* and *D*) and an interaction effect between them (*AD*) significant. The error terms of the data are assumed to follow $N(0, \sigma_\varepsilon^2)$. These data are labeled as data 1, data 2, and data 3, and are given in Table 4.3. We consider the alternate fraction of the nano-scale lubrication experiment ($I = -ABCD$).

Table 4.3: Three data sets for an eight-run experiment

Run	*A*	*B*	*C*	*D*	data 1	data 2	data 3
y_1	-1	-1	-1	1	33	25	41
y_2	1	-1	-1	-1	29	125	22
y_3	-1	1	-1	-1	66	149	37
y_4	1	1	-1	1	37	158	110
y_5	-1	-1	1	-1	71	141	43
y_6	1	-1	1	1	31	160	123
y_7	-1	1	1	1	28	17	34
y_8	1	1	1	-1	29	134	27
σ =					17.6	58.4	39.0

4.8.2.1 One missing observation

In a 2^{4-1} design with eight runs, there are eight different positions where a single observation could be missing. The positions are represented as Run 1 through Run 8. In this section, the non-Bayesian and Bayesian algorithms are applied to data 1, data 2, and data 3 over all the possible positions. Table 4.4 shows a comparison of the Bayesian and non-Bayesian algorithms for data 1. Similarly, Table 4.5 and Table 4.6 represent the comparison of algorithms using data 2 and data 3, respectively. From Theorems 1 and 2 in Section 4.7, it can be said that for a single missing observation, the three Bayesian algorithms yield the same estimate. These estimates are listed under "Bayesian Algorithms" in Table 4.4, Table 4.5, and Table 4.6. Furthermore, as the Complete RSS Minimization algorithm yields the same estimate as the Expected Value algorithm, only estimates from the former algorithm are presented. It is important to note that the active factors specified by the Bayesian screening algorithm are the same as the true significant effects (*A*, *D*, and *AD*).

Table 4.4: Comparison of Bayesian and non-Bayesian Algorithms for One Missing Observation in Data 1

Position of missing observation	Algorithm	True Significant effects = effects specified by Bayesian screening algorithm	% bias	$PRESS_{Diff}$
Run 1	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	76.76	1839.82
	Bayesian Algorithms	*A, D, AD*	15.15	
Run 2	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	110.34	2048
	Bayesian Algorithms	*A, D, AD*	0.00	
Run 3	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	32.32	1386.54
	Bayesian Algorithms	*A, D, AD*	7.58	
Run 4	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	97.30	1800
	Bayesian Algorithms	*A, D, AD*	16.22	
Run 5	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	50.70	1922
	Bayesian Algorithms	*A, D, AD*	7.04	
Run 6	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	68.81	1493.86
	Bayesian Algorithms	*A, D, AD*	19.35	
Run 7	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	114.29	1458
	Bayesian Algorithms	*A, D, AD*	17.86	
Run 8	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	87.34	1283.22
	Bayesian Algorithms	*A, D, AD*	0.00	

Table 4.5: Comparison of Bayesian and non-Bayesian Algorithms for One Missing Observation in Data 2

Position of missing observation	Algorithm	True Significant effects = effects specified by Bayesian screening algorithm	% bias	$PRESS_{Diff}$
Run 1	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	364.00	19602
	Bayesian Algorithms	*A, D, AD*	32.00	
Run 2	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	90.94	21911.6
	Bayesian Algorithms	*A, D, AD*	7.20	
Run 3	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	73.15	20402
	Bayesian Algorithms	*A, D, AD*	5.37	
Run 4	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	60.55	19078.9
	Bayesian Algorithms	*A, D, AD*	13.48	
Run 5	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	67.85	21494.9
	Bayesian Algorithms	*A, D, AD*	5.67	
Run 6	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	68.13	22898
	Bayesian Algorithms	*A, D, AD*	1.25	
Run 7	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	668.65	22332.3
	Bayesian Algorithms	*A, D, AD*	47.06	
Run 8	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	67.91	20000
	Bayesian Algorithms	*A, D, AD*	6.72	

Table 4.6: Comparison of Bayesian and non-Bayesian Algorithms for One Missing Observation in Data 3

Position of missing observation	Algorithm	True Significant effects = effects specified by Bayesian screening algorithm	% bias	$PRESS_{Diff}$
Run 1	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	134.15	7688
	Bayesian Algorithms	*A, D, AD*	17.07	
Run 2	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	322.73	8712
	Bayesian Algorithms	*A, D, AD*	22.73	
Run 3	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	182.89	10854.5
	Bayesian Algorithms	*A, D, AD*	16.22	
Run 4	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	53.03	10175.9
	Bayesian Algorithms	*A, D, AD*	11.82	
Run 5	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	135.65	5476.86
	Bayesian Algorithms	*A, D, AD*	13.95	
Run 6	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	55.02	5977.62
	Bayesian Algorithms	*A, D, AD*	10.57	
Run 7	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	208.82	8192
	Bayesian Algorithms	*A, D, AD*	20.59	
Run 8	ANCOVA	*A, D, AD*	--	--
	Complete RSS min	*A, D, AD*	203.70	7200
	Bayesian Algorithms	*A, D, AD*	18.52	

For one missing observation, a potential problem in the non-Bayesian ANCOVA algorithm is observed. Due to the small size of the data set, this algorithm can yield a zero residual due to over-fitting. Hence, no estimate of the missing observation is obtained. For each position of the missing observation, $PRESS_{Diff}$ is calculated as the difference between the *PRESS* values for the non-Bayesian algorithm ($PRESS_{NB}$) and Bayesian algorithm ($PRESS_B$). It is observed that in each of the eight positions for the missing observation in the three data sets, $PRESS_{Diff}$ has a positive value. This seems to

suggest that the performance of the Bayesian algorithms is superior to the non-Bayesian counterparts.

In addition to the $PRESS_{Diff}$ criterion, the behavior of degree of bias for all positions in the data sets are graphically illustrated in Figure 4.2 (data 1), Figure 4.3 (data 2), and Figure 4.4 (data 3). For each data set, the degree of bias in the estimated values through the non-Bayesian and Bayesian algorithms are compared over all positions.

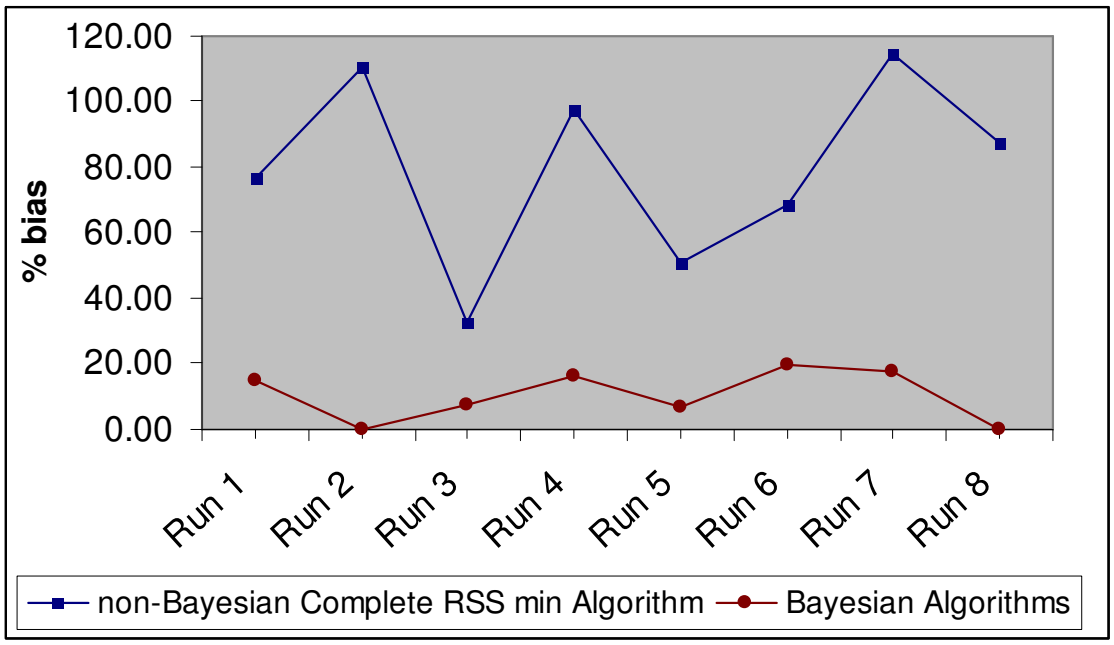

Figure 4.2: Bias for one missing observation over all positions in data 1.

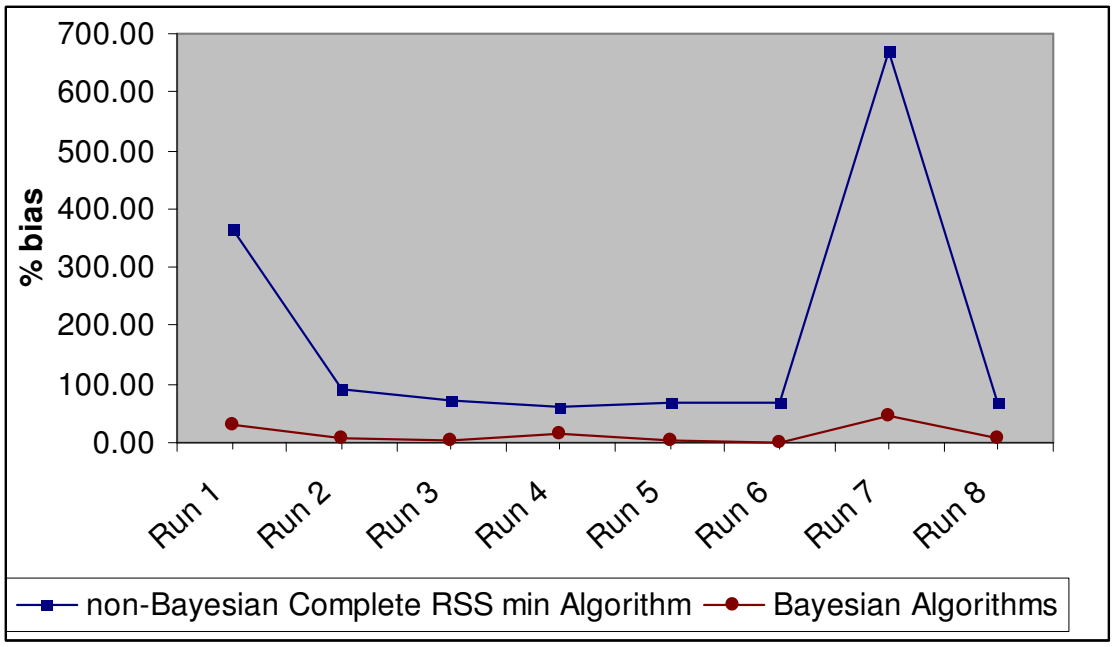

Figure 4.3: Bias for one missing observation over all positions in data 2.

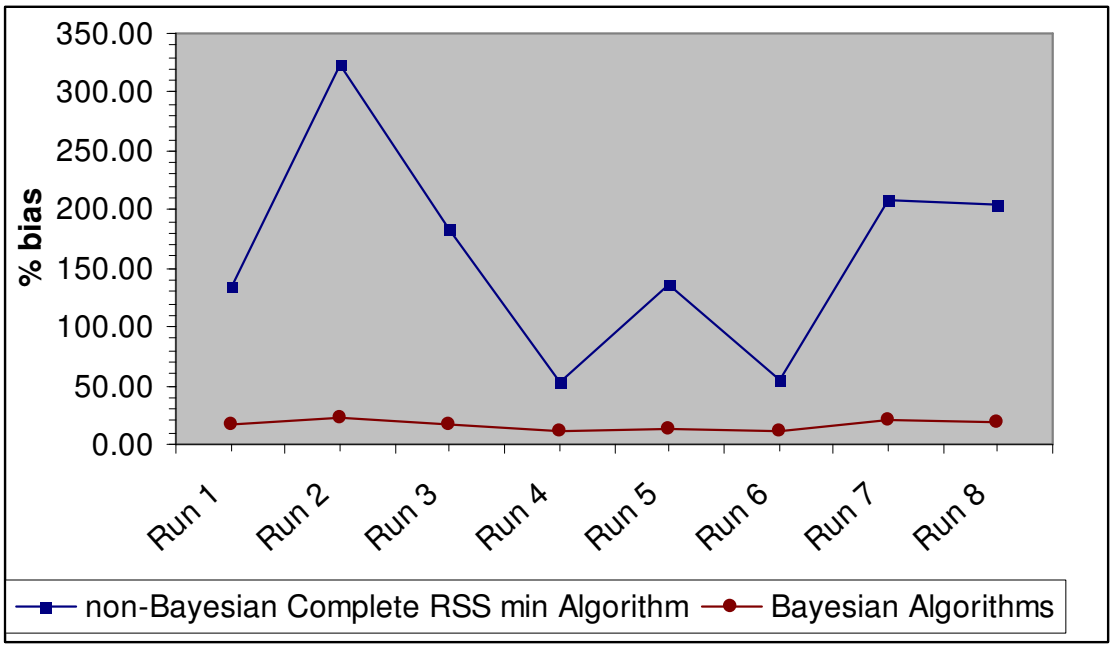

Figure 4.4: Bias for one missing observation over all positions in data 3.

In each case, it is observed that the bias due to the estimate of the missing observation is considerably less in the Bayesian algorithm than its corresponding non-Bayesian algorithm. This observation validates the claim from using the $PRESS_{Diff}$ criterion. Additionally, the variation in the estimates from the Bayesian algorithms is much less. This is observed from the standard deviations of the estimates from the Bayesian algorithms (10.4 for data 1, 14.84 for data 2, and 16.43 for data 3) and those from the non-Bayesian algorithms (79.73 for data 1, 182.65 for data 2, and 162 for data 3).

4.8.2.2 Two missing observations

In this section, we consider two missing observations in a 2^{4-1} design with eight runs. Two observations in an eight run design can be missing in ${}^{8}C_{2} = 28$ ways. Rather than compare the algorithms over all 28 pairs, let us look at the geometric representation of an eight-run design, given in Figure 4.5. The design can be represented as a cube with each run corresponding to a corner. The corners are numbered from 1 - 8 and correspond to treatment combinations (*I*), *ad*, *bd*, *ab*, *cd*, *ac*, *bc*, and *abcd*, respectively. According to the geometry of a cube, there would be 12 edges of faces, 12 diagonals of faces, and 4 lines joining opposite corners of opposite faces. These 28 edges, diagonals, and lines constitute all possible pairs of missing observations in an eight-run design. Runs (1 , 2); (1 , 5); (1 , 3); (2 , 6); (2 , 4); (3 , 4); (4 , 8); (6 , 8); (7 , 8); (5 , 6); (5 , 7); and (3 , 7) can be categorized as pairs corresponding to the edges. The pair (1, 2) is chosen as a sample representing this category. Runs (1, 6); (2, 5); (4, 6); (2, 8); (6, 7); (5, 8); (3, 5); (1, 7); (1, 4); (2, 3); (3, 8); and (4, 7) can be categorized as pairs corresponding to the diagonals of

the faces. From this category, pair (1, 6) is chosen as a sample for calculation. Finally, runs (1, 8); (2, 7); (3, 6); and (4, 5) can be categorized as lines connecting opposite corners of opposite faces and pair (4, 5) is selected as a sample.

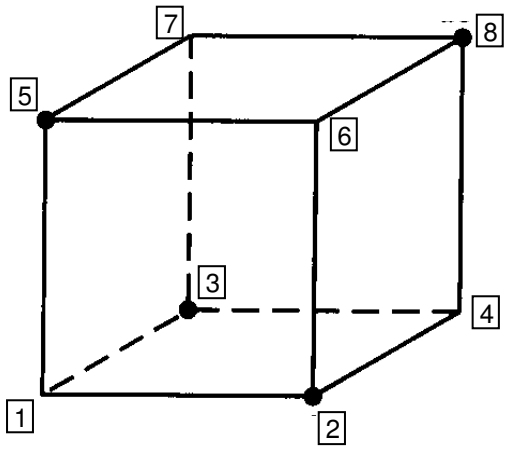

Figure 4.5: Geometric representation of the 2^{4-1} design.

First, the Bayesian screening algorithm is applied to identify the potential active effects (details of the calculations are given in Appendix 4A). The Bayesian and the non-Bayesian algorithms are applied to the three data sets mentioned in Table 4.3, having effects *A*, *D*, and *AD* significant. Two Bayesian algorithms (Bayesian ANCOVA and Bayesian Complete RSS minimization) are compared to their corresponding non-Bayesian versions when runs (1, 2); (1, 6); and (4, 5) are assumed to be missing. The results when the algorithms are applied to data 1 are shown in Table 4.7. Similarly, results from data 2 and data 3 are given in Table 4.8 and Table 4.9, respectively. Since Theorem 1 proves Bayesian Expected Value algorithm to yield same estimates of missing observations as Bayesian Complete RSS Minimization algorithm, the former can be omitted from the study. It should be noted that effects specified by the Bayesian

screening algorithm as active are the same as the true significant effects (when the observations are not assumed to be missing). The calculations for Bayesian Complete RSS Minimizaion and Bayesian ANCOVA algorithms are given in Appendix 4B and Appendix 4C, respectively.

Table 4.7: Comparison of Bayesian and non-Bayesian Algorithms for Two Missing Observations in Data 1

Position of missing observations	Algorithm	True Significant effects = Effects specified by Bayesian screening algorithm	Average % bias	$PRESS_{Diff}$
Runs 1 and 2	ANCOVA	*A, D, AD*	100.00	3249.78
	Bayesian ANCOVA	*A, D, AD*	8.14	
	Complete RSS min	*A, D, AD*	139.81	8281
	Bayesian Complete RSS min	*A, D, AD*	7.58	
Runs 1 and 6	ANCOVA	*A, D, AD*	100.00	4306
	Bayesian ANCOVA	*A, D, AD*	17.25	
	Complete RSS min	*A, D, AD*	54.45	2141
	Bayesian Complete RSS min	*A, D, AD*	17.25	
Runs 4 and 5	ANCOVA	*A, D, AD*	100.00	10361.62
	Bayesian ANCOVA	*A, D, AD*	19.85	
	Complete RSS min	*A, D, AD*	27.71	1935.76
	Bayesian Complete RSS min	*A, D, AD*	11.63	

Table 4.8: Comparison of Bayesian and non-Bayesian Algorithms for Two Missing Observations in Data 2

Position of missing observations	Algorithm	True Significant effects = Effects specified by Bayesian screening algorithm	Average % bias	$PRESS_{Diff}$
Runs 1 and 2	ANCOVA	*A, D, AD*	100.00	33448
	Bayesian ANCOVA	*A, D, AD*	28.00	
	Complete RSS min	*A, D, AD*	354.80	93636
	Bayesian Complete RSS min	*A, D, AD*	19.60	
Runs 1 and 6	ANCOVA	*A, D, AD*	100.00	50506
	Bayesian ANCOVA	*A, D, AD*	16.63	
	Complete RSS min	*A, D, AD*	150.66	24625
	Bayesian Complete RSS min	*A, D, AD*	16.63	
Runs 4 and 5	ANCOVA	*A, D, AD*	100.00	95575.06
	Bayesian ANCOVA	*A, D, AD*	4.77	
	Complete RSS min	*A, D, AD*	32.10	11200
	Bayesian Complete RSS min	*A, D, AD*	3.47	

Table 4.9: Comparison of Bayesian and non-Bayesian Algorithms for Two Missing Observations in Data 3

Position of missing observations	Algorithm	True Significant effects = Effects specified by Bayesian screening algorithm	Average % bias	$PRESS_{Diff}$
Runs 1 and 2	ANCOVA	*A, D, AD*	100.00	3759.14
	Bayesian ANCOVA	*A, D, AD*	25.20	
	Complete RSS min	*A, D, AD*	336.34	36481
	Bayesian Complete RSS min	*A, D, AD*	19.90	
Runs 1 and 6	ANCOVA	*A, D, AD*	100.00	26512
	Bayesian ANCOVA	*A, D, AD*	13.82	
	Complete RSS min	*A, D, AD*	67.07	7397
	Bayesian Complete RSS min	*A, D, AD*	13.82	
Runs 4 and 5	ANCOVA	*A, D, AD*	100.00	31352.06
	Bayesian ANCOVA	*A, D, AD*	32.27	
	Complete RSS min	*A, D, AD*	47.18	4630.3
	Bayesian Complete RSS min	*A, D, AD*	12.89	

$PRESS_{Diff}$ values are calculated by subtracting the *PRESS* values corresponding to Bayesian algorithms from those of non-Bayesian algorithms. From the above tables, it is observed that the $PRESS_{Diff}$ values are very large positive numbers, indicating much better performance of the Bayesian algorithms over the non-Bayesian ones. In order to gain insight on the two Bayesian algorithms, degree of bias for the data sets are graphed and represented in Figure 4.2 (data 1), Figure 4.3 (data 2), and Figure 4.4 (data 3).

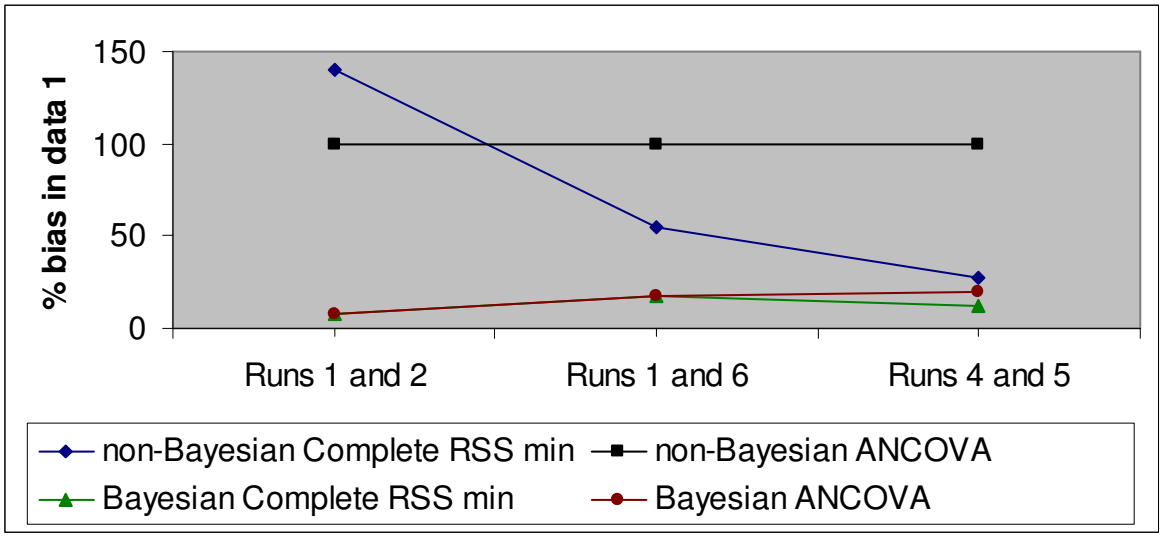

Figure 4.6: Bias for two missing observations over three pairs of positions in data 1.

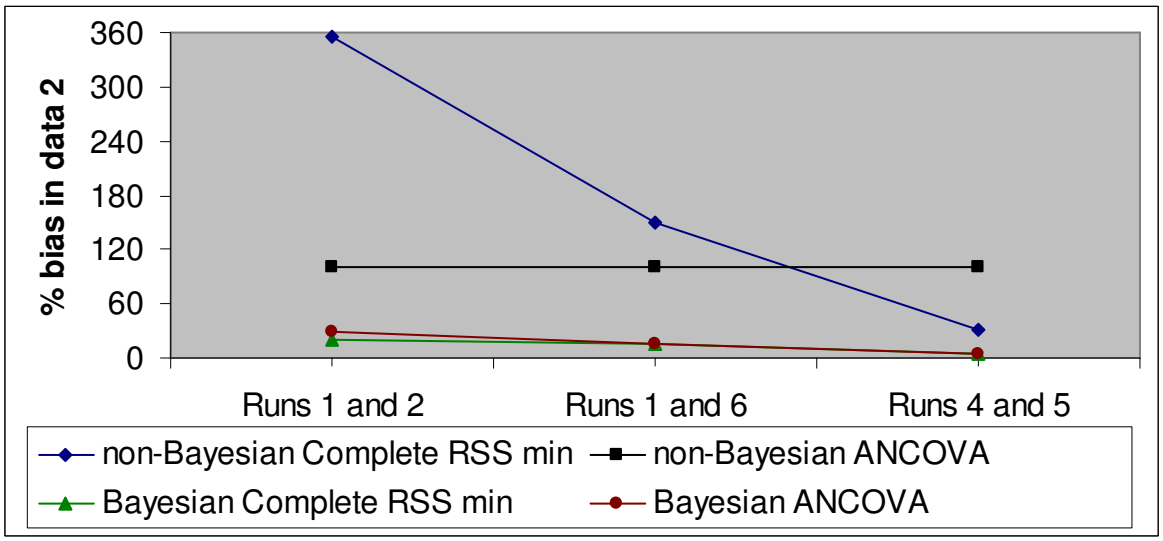

Figure 4.7: Bias for two missing observations over three pairs of positions in data 2.

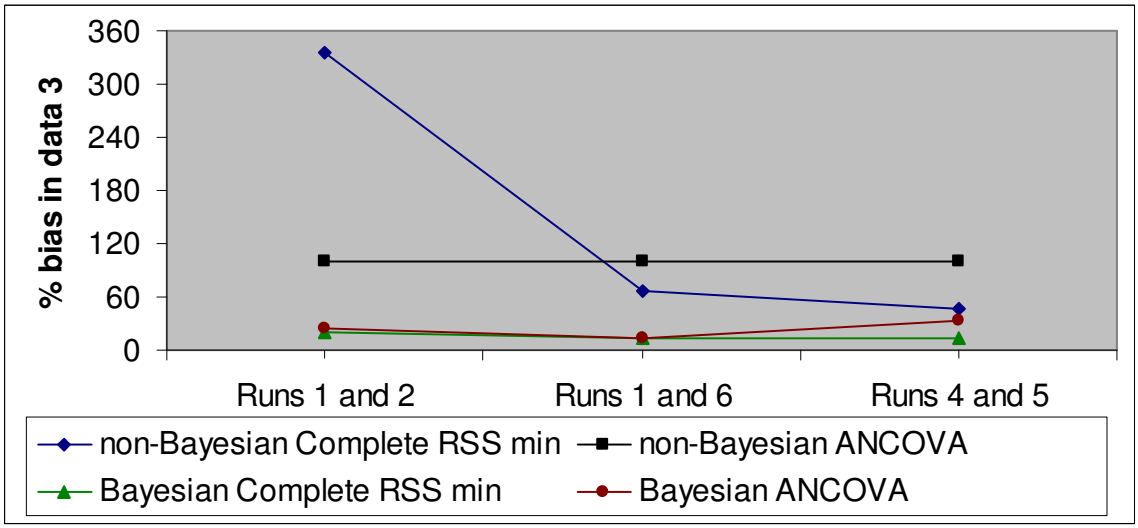

Figure 4.8: Bias for two missing observations over three pairs of positions in data 3.

In all the data sets and for the given pairs of missing observations, it is observed that either of the Bayesian algorithms has less bias than any of the non-Bayesian algorithms. This is the same trend observed in the case of a single missing observation. When the non-Bayesian ANCOVA algorithm is employed, the residuals obtained by fitting a model using main effects and interaction terms are zero. As the estimates obtained are a direct function of the residuals, when the residuals are zero due to over-fitting, no estimates are obtained. This is reflected in the bias for the non-Bayesian ANCOVA algorithm, which is 100%.

Figures 4.6 – 4.8 suggest that a better comparison of the Bayesian algorithms could be made. It seems that the estimates obtained by Bayesian Complete RSS Minimization algorithm have equal or less bias than those obtained by Bayesian ANCOVA algorithm. The only case that yields equal bias is when runs 1 and 6 are

missing in data 1, 2, or 3. In all the other cases (position of missing observation and data set), Bayesian Complete RSS Minimization has yields lesser bias than Bayesian ANCOVA does. Finally, a different pair of missing observation has the lowest bias in each of the three data sets considered, thus providing no pattern. For instance, in data 1, Bayesian Complete RSS Minimization provides the least bias when runs 1 and 2 are missing, but in data 2, the least bias corresponds to missing runs 4 and 5.

4.8.3 Effect of Incorrect Specification of Active Factors on Performance

The Bayesian Complete RSS minimization and Bayesian ANCOVA algorithms incorporate information provided on potential active factors from the Bayesian screening algorithm. In the above three data sets considered, the screening algorithm identified effects *A*, *D*, and *AD* as active, both for one and two missing observations. These effects are the "true" effects as the data were generated assuming the same effects to be significant. Hence, it is natural to examine the results of the Bayesian algorithms when there is a mismatch between the "true" effects and those specified by the screening algorithm. As the true effects consist of two main effects and an interaction term, the following effects are assumed to be active and correspond to three cases of mismatch:

- First main effect and the interaction term (*A* and *AD*)
- Second main effect and the interaction term (*D* and *AD*)
- Only the main effects (*A* and *D*)

These mismatches are considered by each of the Bayesian algorithms to be the active effects identified from the screening algorithm. Estimates are then obtained for missing observations using data 1 from Table 4.3. The results in the presence of the above mismatches for one and two missing observations are presented and examined in the subsequent sections.

4.8.3.1 One missing observation

Three cases of mismatch were examined assuming one missing observation over all possible positions (eight) in data 1. It has already been established that all Bayesian algorithms yield the same estimate in the presence of a single missing observation, and hence, results from a single Bayesian algorithm are reported. Details for the same along with the percent bias and $PRESS_{Diff}$ are shown in Table 4.10. Non-Bayesian algorithms do not consider any information on the active factors, and hence no mismatch occurs. In Table 4.10, $PRESS_{Diff}$ is calculated as difference between PRESS using a complete data set (no missing observations) and *PRESS* using the estimate of the assumed missing observation. The value for the *PRESS* statistic when no observations are missing is 172. Hence, the more negative the $PRESS_{Diff}$, the worse the algorithm performs.

Table 4.10: Incorrect Specification for One Missing Observation

Position of missing observation	Algorithm	True Significant effects	Significant effects specified by Bayesian screening algorithm	% bias	$PRESS_{Diff}$
Run 1	Bayesian Algorithms	*A, D, AD*	*A, AD*	27.878788	-353.28
	Bayesian Algorithms	*A, D, AD*	*D, AD*	55.757576	-309.12
	Bayesian Algorithms	*A, D, AD*	*A, D*	40	-238
Run 2	Bayesian Algorithms	*A, D, AD*	*A, AD*	45.517241	-348.48
	Bayesian Algorithms	*A, D, AD*	*D, AD*	49.655172	-414.72
	Bayesian Algorithms	*A, D, AD*	*A, D*	59.310345	-591.68
Run 3	Bayesian Algorithms	*A, D, AD*	*A, AD*	13.939394	-363.28
	Bayesian Algorithms	*A, D, AD*	*D, AD*	15.757576	-424.32
	Bayesian Algorithms	*A, D, AD*	*A, D*	20	-612.48
Run 4	Bayesian Algorithms	*A, D, AD*	*A, AD*	22.702703	-342.72
	Bayesian Algorithms	*A, D, AD*	*D, AD*	25.945946	-414.72
	Bayesian Algorithms	*A, D, AD*	*A, D*	59.459459	-440
Run 5	Bayesian Algorithms	*A, D, AD*	*A, AD*	24.225352	-247.68
	Bayesian Algorithms	*A, D, AD*	*D, AD*	25.915493	-309.12
	Bayesian Algorithms	*A, D, AD*	*A, D*	29.859155	-474.88
Run 6	Bayesian Algorithms	*A, D, AD*	*A, AD*	58.064516	-216
	Bayesian Algorithms	*A, D, AD*	*D, AD*	61.935484	-276.48
	Bayesian Algorithms	*A, D, AD*	*A, D*	40	-599.12
Run 7	Bayesian Algorithms	*A, D, AD*	*A, AD*	61.428571	-247.68
	Bayesian Algorithms	*A, D, AD*	*D, AD*	37.142857	-424.32
	Bayesian Algorithms	*A, D, AD*	*A, D*	75.714286	-474.88
Run 8	Bayesian Algorithms	*A, D, AD*	*A, AD*	45.517241	-348.48
	Bayesian Algorithms	*A, D, AD*	*D, AD*	49.655172	-414.72
	Bayesian Algorithms	*A, D, AD*	*A, D*	59.310345	-591.68

From Table 4.10, it is seen that none of the Bayesian algorithms perform even close to the case with no missing observations. To study the trend of the calculated

performance measures, the values are graphed over all runs and are represented in Figure 4.9 (percent bias) and Figure 4.10 ($PRESS_{Diff}$).

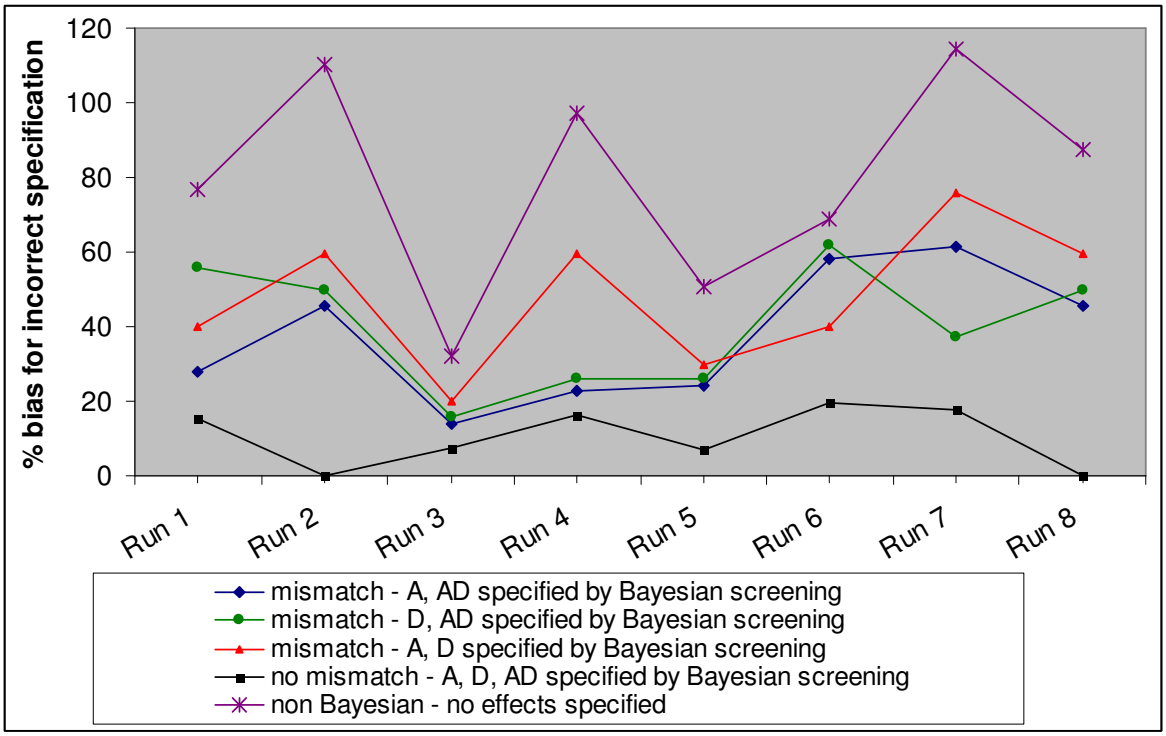

Figure 4.9: Bias in incorrect specification of active factors for one missing observation over all positions in data 1.

Figure 4.9 seems to suggest that in the presence of a mismatch, the percent bias is greater than the case where there is no mismatch (active effects from screening algorithm are the same as "true" effects). However, no case of mismatch results in greater bias than the bias due to non-Bayesian algorithms. In addition, no particular case of mismatch results in less bias over all the positions of missing observations. Hence, no general

conclusions can be drawn on the performance of a particular type of mismatch. It seems intuitive that the level of significance of an effect (included as a mismatch) would dictate the bias to a certain extent.

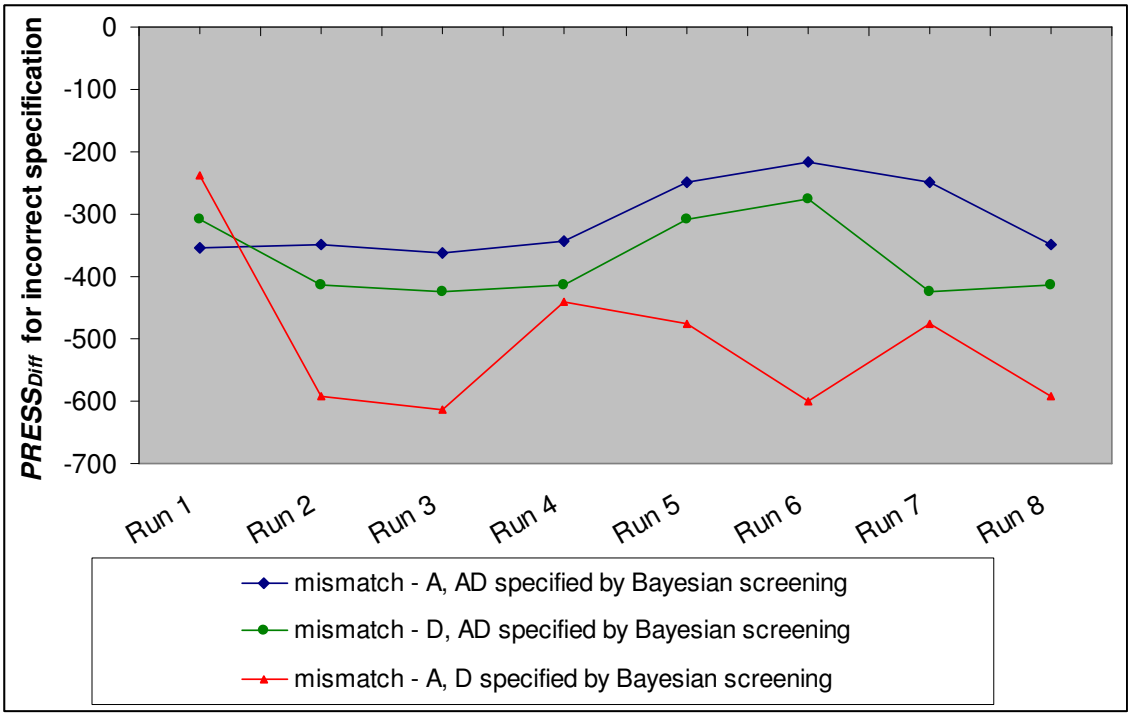

Figure 4.10: *PRESS*$_{Diff}$ in incorrect specification of active factors for one missing observation over all positions in data 1

4.8.3.2 Two missing observations

In the case of two missing observations, the same pairs of observations from Section 4.8.2 are assumed to be missing. The mismatches considered and the calculations for the percent bias and $PRESS_{Diff}$ are the same as the previous section. The results using Bayesian ANCOVA and Bayesian Complete RSS minimization algorithms are presented in Table 4.11. As before, the more negative the value of $PRESS_{Diff,}$ of a particular case, the worse is the performance of the corresponding algorithm. It is seen that the values of $PRESS_{Diff}$ are much lower than in the corresponding case of a single missing observation. Thus, the performance of the Bayesian algorithms seems to decline as the number of missing observations increase.

The values of the percent bias for the mismatches over the assumed positions of missing observations for Bayesian ANCOVA and Bayesian Complete RSS Minimization algorithms are graphically represented in Figure 4.11 and Figure 4.12, respectively.

Table 4.11: Incorrect Specification for Two Missing Observations

Position of missing observation	Algorithm	True Significant effects	Significant effects specified by Bayesian screening algorithm	Average % bias	$PRESS_{Diff}$
Runs 1 and 2	Bayesian ANCOVA	*A, D, AD*	*A, AD*	30.74504	-509.8
	Bayesian ANCOVA	*A, D, AD*	*D, AD*	33.38662	-303.71
	Bayesian ANCOVA	*A, D, AD*	*A, D*	41.15726	-856.72
	Bayesian Complete RSS min	*A, D, AD*	*A, AD*	27.29676	-423.16
	Bayesian Complete RSS min	*A, D, AD*	*D, AD*	43.75862	-449.08
	Bayesian Complete RSS min	*A, D, AD*	*A, D*	41.56426	-870.5
Runs 1 and 6	Bayesian ANCOVA	*A, D, AD*	*A, AD*	35.97849	-347.16
	Bayesian ANCOVA	*A, D, AD*	*D, AD*	49.06549	-291.72
	Bayesian ANCOVA	*A, D, AD*	*A, D*	33.32747	-923.88
	Bayesian Complete RSS min	*A, D, AD*	*A, AD*	35.97849	-347.16
	Bayesian Complete RSS min	*A, D, AD*	*D, AD*	49.06549	-291.72
	Bayesian Complete RSS min	*A, D, AD*	*A, D*	33.32747	-923.88
Runs 4 and 5	Bayesian ANCOVA	*A, D, AD*	*A, AD*	18.37381	-377.08
	Bayesian ANCOVA	*A, D, AD*	*D, AD*	21.86905	-623
	Bayesian ANCOVA	*A, D, AD*	*A, D*	41.54739	-1088.48
	Bayesian Complete RSS min	*A, D, AD*	*A, AD*	20.07994	-454
	Bayesian Complete RSS min	*A, D, AD*	*D, AD*	18.7027	-391.48
	Bayesian Complete RSS min	*A, D, AD*	*A, D*	39.03007	-588.76

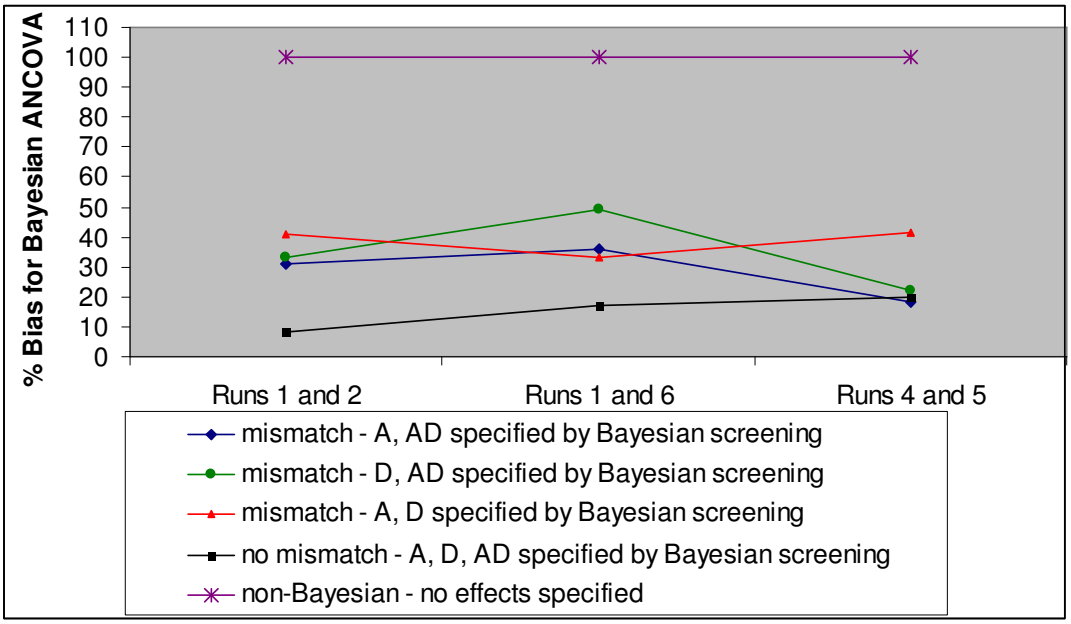

Figure 4.11: Bias in incorrect specification of active factors for two missing observations using Bayesian ANCOVA algorithm.

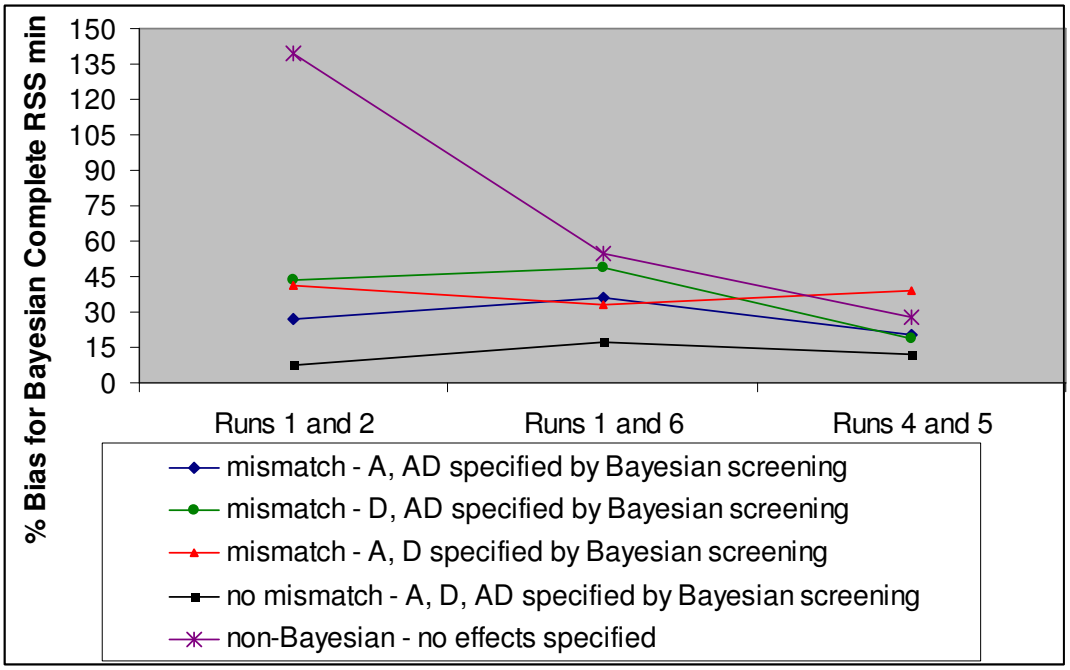

Figure 4.12: Bias in incorrect specification of active factors for two missing observations using Bayesian Complete RSS Minimization algorithm.

From Figure 4.11, it is easy to see that in the case of mismatches, Bayesian ANCOVA algorithm yields estimates that have much less bias as compared to non-Bayesian algorithm (the non-Bayesian algorithm yields 100% bias due to over fitting). In addition, bias in all cases seems to be closer to the case with no mismatch. Additionally, when runs 4 and 5 are assumed missing, the biases resulting from two of the three mismatches are almost the same as the bias due to the case with no mismatch.

Similarly, when Bayesian Complete RSS Minimization algorithm is used, any mismatch seems to increase the bias over that of the case with no mismatch. In two out of the three pairs of positions assumed to be missing, bias due to any mismatch seems lesser than bias due to the non-Bayesian algorithm. In the case of the third mismatch (*A* and *D* specified by screening algorithm), the bias seems to be greater than the bias due to the non-Bayesian algorithm. However, this difference is not large enough to cause concerns and can be ignored. Thus, it seems that the Bayesian algorithms outperform the Bayesian ones, even in the case of mismatches and hence seem robust to the screening algorithm.

When a fractional factorial design is constructed for the lubrication example with $D = -ABC$, no molecules would exist corresponding to runs 1 and 6. From performance evaluations, Bayesian Complete RSS minimization algorithm seems to yield estimates that are closer to the assumed value. Once the estimates are obtained, the characteristics of the molecules could be studied using regular analysis techniques such as ANOVA and conditions for desired film adhesion obtained.

Appendix 4A Calculations for Bayesian Screening Algorithm

Calculations shown are for two missing observations (Runs 1 and 2 assumed to be missing for data 1)

```
% 2 missing observation in a 2^(4-1) experiment
% Runs 1 and 2 missing
% using equations from Meyer and Box (1993) "Finding the active factors in fractionated screening experiments"
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

gamma = 2;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

pi = 0.5;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

n = 6;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

t_0 = 0;
t_1 = 1;
t_2 = 1;
t_3 = 1;
t_4 = 1;
t_5 = 3;
t_6 = 3;
t_7 = 3;
t_8 = 3;
t_9 = 3;
t_10 = 3;
t_11 = 7;
t_12 = 7;
t_13 = 7;
t_14 = 7;
t_15 = 15;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

f_1 = 1;
f_2 = 1;
f_3 = 1;
f_4 = 1;
f_5 = 2;
f_6 = 2;
f_7 = 2;
f_8 = 2;
f_9 = 2;
```

```
f_10 = 2;
f_11 = 3;
f_12 = 3;
f_13 = 3;
f_14 = 3;
f_15 = 4;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

E_0 = (1/(gamma^2))*(0);

E_1 = (1/(gamma^2))*[0      0;
                     0      1];

E_2 = (1/(gamma^2))*[0      0;
                     0      1];

E_3 = (1/(gamma^2))*[0      0;
                     0      1];

E_4 = (1/(gamma^2))*[0      0;
                     0      1];

E_5 = (1/(gamma^2))*[0      0       0       0;
                     0      1       0       0;
                     0      0       1       0;
                     0      0       0       1];

E_6 = (1/(gamma^2))*[0      0       0       0;
                     0      1       0       0;
                     0      0       1       0;
                     0      0       0       1];

E_7 = (1/(gamma^2))*[0      0       0       0;
                    0       1       0       0;
                    0       0       1       0;
                    0       0       0       1];

E_8 = (1/(gamma^2))*[0      0       0       0;
                    0       1       0       0;
                    0       0       1       0;
                    0       0       0       1];

E_9 = (1/(gamma^2))*[0      0       0       0;
                     0      1       0       0;
                     0      0       1       0;
                     0      0       0       1];

E_10 = (1/(gamma^2))*[0     0       0       0;
                      0     1       0       0;
                      0     0       1       0;
                      0     0       0       1];

E_11 = (1/(gamma^2))*[0     0       0       0       0       0       0       0;
                     0      1       0       0       0       0       0       0;
                     0      0       1       0       0       0       0       0;
                     0      0       0       1       0       0       0       0;
                     0      0       0       0       1       0       0       0;
                     0      0       0       0       0       1       0       0;
                     0      0       0       0       0       0       1       0;
                     0      0       0       0       0       0       0       1];
```

```
E_12 = (1/(gamma^2))*[0   0   0   0   0   0   0   0;
                      0   1   0   0   0   0   0   0;
                      0   0   1   0   0   0   0   0;
                      0   0   0   1   0   0   0   0;
                      0   0   0   0   1   0   0   0;
                      0   0   0   0   0   1   0   0;
                      0   0   0   0   0   0   1   0;
                      0   0   0   0   0   0   0   1];

E_13 = (1/(gamma^2))*[0   0   0   0   0   0   0   0;
                      0   1   0   0   0   0   0   0;
                      0   0   1   0   0   0   0   0;
                      0   0   0   1   0   0   0   0;
                      0   0   0   0   1   0   0   0;
                      0   0   0   0   0   1   0   0;
                      0   0   0   0   0   0   1   0;
                      0   0   0   0   0   0   0   1];

E_14 = (1/(gamma^2))*[0   0   0   0   0   0   0   0;
                      0   1   0   0   0   0   0   0;
                      0   0   1   0   0   0   0   0;
                      0   0   0   1   0   0   0   0;
                      0   0   0   0   1   0   0   0;
                      0   0   0   0   0   1   0   0;
                      0   0   0   0   0   0   1   0;
                      0   0   0   0   0   0   0   1];

E_15 = (1/(gamma^2))*[0   0   0   0   0   0   0   0   0   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   1   0   0   0   0   0   0   0
       0   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   1   0   0   0   0   0
       0   0   0   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   1   0   0   0
       0   0   0   0   0   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   1   0
       0   0   0   0   0   0
                      0   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
       1   0   0   0   0   0
                      0   0   0   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   1   0   0   0
                      0   0   0   0   0   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   1   0
                      0   0   0   0   0   0   0   0   0   0
       0   0   0   0   0   1];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
Y = [66;
    37;
    71;
    31;
    28;
    29];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
X_0 =   [1;
        1;
        1;
        1;
        1;
        1];

X_1 =  [1        -1;
        1         1;
        1        -1;
        1         1;
        1        -1;
        1         1];

X_2 =  [1         1;
        1         1;
        1        -1;
        1        -1;
        1         1;
        1         1];

X_3 =  [1        -1;
        1        -1;
        1         1;
        1         1;
        1         1;
        1         1];

X_4 =  [1         1;
        1        -1;
        1         1;
        1        -1;
        1        -1;
        1         1];

X_5 =   [1       -1        1        -1;
         1        1        1         1;
         1       -1       -1         1;
         1        1       -1        -1;
         1       -1        1        -1;
         1        1        1         1];

X_6 =   [1       -1       -1         1;
         1        1       -1        -1;
         1       -1        1        -1;
         1        1        1         1;
         1       -1        1        -1;
         1        1        1         1];

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];

X_8 =   [1          1         -1        -1;
          1         1         -1        -1;
          1         -1        1         -1;
          1         -1        1         -1;
          1         1         1         1;
          1         1         1         1];

X_9 =   [1          1         1         1;
          1         1         -1        -1;
          1         -1        1         -1;
          1         -1        -1        1;
          1         1         -1        -1;
          1         1         1         1];

X_10 =   [1         -1        1         -1;
           1        -1        -1        1;
           1        1         1         1;
           1        1         -1        -1;
           1        1         -1        -1;
           1        1         1         1];

X_11 =  [1          -1        1         -1        -1        1         -1        1;
           1        1         1         -1        1         -1        -1        -1;
           1        -1        -1        1         1         -1        -1        1;
           1        1         -1        1         -1        1         -1        -1;
           1        -1        1         1         -1        -1        1         -1;
           1        1         1         1         1         1         1         1];

X_12 =   [1         -1        1         1         -1        -1        1         -1;
           1        1         1         -1        1         -1        -1        -1;
           1        -1        -1        1         1         -1        -1        1;
           1        1         -1        -1        -1        -1        1         1;
           1        -1        1         -1        -1        1         -1        1;
           1        1         1         1         1         1         1         1];

X_13 =   [1         -1        -1        1         1         -1        -1        -1;
           1        1         -1        -1        -1        -1        1         -1;
           1        -1        1         1         -1        -1        1         1;
           1        1         1         -1        1         -1        -1        1;
           1        -1        1         -1        -1        1         -1        1;
           1        1         1         1         1         1         1         1];

X_14 =  [1          1         -1        1         -1        1         -1        -1;
           1        1         -1        -1        -1        -1        1         1;
           1        -1        1         1         -1        -1        1         -1;
           1        -1        1         -1        -1        1         -1        1;
           1        1         1         -1        1         -1        -1        -1;
           1        1         1         1         1         1         1         1];

X_15 = [1         -1        1         -1        1         -1        1         -1        -1        1         -1
          1        -1        -1        1         1;
      1   1        1         -1        -1        1         -1        -1        -1        -1        1         -1
          -1       1         1         1;
      1   -1       -1        1         1         1         -1        -1        -1        -1        1         1
          1        -1        -1        1;
      1   1        -1        1         -1        -1        1         -1        -1        1         -1        -1
          1        1         -1        1;
      1   -1       1         1         -1        -1        -1        1         1         -1        -1        -1
          1        -1        1         1;
```

```
        1    1           1           1           1           1           1           1           1           1           1           1
             1           1           1           1];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

beta_0 = ((inv(E_0+(X_0'*X_0)))*(X_0'*Y));

beta_1 = ((inv(E_1+(X_1'*X_1)))*(X_1'*Y));

beta_2 = ((inv(E_2+(X_2'*X_2)))*(X_2'*Y));

beta_3 = ((inv(E_3+(X_3'*X_3)))*(X_3'*Y));

beta_4 = ((inv(E_4+(X_4'*X_4)))*(X_4'*Y));

beta_5 = ((inv(E_5+(X_5'*X_5)))*(X_5'*Y));

beta_6 = ((inv(E_6+(X_6'*X_6)))*(X_6'*Y));

beta_7 = ((inv(E_7+(X_7'*X_7)))*(X_7'*Y));

beta_8 = ((inv(E_8+(X_8'*X_8)))*(X_8'*Y));

beta_9 = ((inv(E_9+(X_9'*X_9)))*(X_9'*Y));

beta_10 = ((inv(E_10+(X_10'*X_10)))*(X_10'*Y));

beta_11 = ((inv(E_11+(X_11'*X_11)))*(X_11'*Y));

beta_12 = ((inv(E_12+(X_12'*X_12)))*(X_12'*Y));

beta_13 = ((inv(E_13+(X_13'*X_13)))*(X_13'*Y));

beta_14 = ((inv(E_14+(X_14'*X_14)))*(X_14'*Y));

beta_15 = ((inv(E_15+(X_15'*X_15)))*(X_15'*Y));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

S_beta_0 = ((Y-(X_0*beta_0))'*(Y-(X_0*beta_0)));

S_beta_1 = ((Y-(X_1*beta_1))'*(Y-(X_1*beta_1)));

S_beta_2 = ((Y-(X_2*beta_2))'*(Y-(X_2*beta_2)));

S_beta_3 = ((Y-(X_3*beta_3))'*(Y-(X_3*beta_3)));

S_beta_4 = ((Y-(X_4*beta_4))'*(Y-(X_4*beta_4)));

S_beta_5 = ((Y-(X_5*beta_5))'*(Y-(X_5*beta_5)));

S_beta_6 = ((Y-(X_6*beta_6))'*(Y-(X_6*beta_6)));

S_beta_7 = ((Y-(X_7*beta_7))'*(Y-(X_7*beta_7)));

S_beta_8 = ((Y-(X_8*beta_8))'*(Y-(X_8*beta_8)));

S_beta_9 = ((Y-(X_9*beta_9))'*(Y-(X_9*beta_9)));

S_beta_10 = ((Y-(X_10*beta_10))'*(Y-(X_10*beta_10)));
```

```
S_beta_11 = ((Y-(X_11*beta_11))'*(Y-(X_11*beta_11)));
S_beta_12 = ((Y-(X_12*beta_12))'*(Y-(X_12*beta_12)));
S_beta_13 = ((Y-(X_13*beta_13))'*(Y-(X_13*beta_13)));
S_beta_14 = ((Y-(X_14*beta_14))'*(Y-(X_14*beta_14)));
S_beta_15 = ((Y-(X_15*beta_15))'*(Y-(X_15*beta_15)));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
first_part_1 = ((pi/(1-pi))^f_1)*(gamma^(-t_1));
second_part_1 = ((det(X_0'*X_0))^0.5)/((det(E_1+(X_1'*X_1)))^0.5);
third_part_1 = ((S_beta_1+(beta_1'*E_1*beta_1))/(S_beta_0))^(-(n-1)/2);
posterior_prop_1 = first_part_1 * second_part_1 * third_part_1

first_part_2 = ((pi/(1-pi))^f_2)*(gamma^(-t_2));
second_part_2 = ((det(X_0'*X_0))^0.5)/((det(E_2+(X_2'*X_2)))^0.5);
third_part_2 = ((S_beta_2+(beta_2'*E_2*beta_2))/(S_beta_0))^(-(n-1)/2);
posterior_prop_2 = first_part_2 * second_part_2 * third_part_2

first_part_3 = ((pi/(1-pi))^f_3)*(gamma^(-t_3));
second_part_3 = ((det(X_0'*X_0))^0.5)/((det(E_3+(X_3'*X_3)))^0.5);
third_part_3 = ((S_beta_3+(beta_3'*E_3*beta_3))/(S_beta_0))^(-(n-1)/2);
posterior_prop_3 = first_part_3 * second_part_3 * third_part_3

first_part_4 = ((pi/(1-pi))^f_4)*(gamma^(-t_4));
second_part_4 = ((det(X_0'*X_0))^0.5)/((det(E_4+(X_4'*X_4)))^0.5);
third_part_4 = ((S_beta_4+(beta_4'*E_4*beta_4))/(S_beta_0))^(-(n-1)/2);
posterior_prop_4 = first_part_4 * second_part_4 * third_part_4

first_part_5 = ((pi/(1-pi))^f_5)*(gamma^(-t_5));
second_part_5 = ((det(X_0'*X_0))^0.5)/((det(E_5+(X_5'*X_5)))^0.5);
third_part_5 = ((S_beta_5+(beta_5'*E_5*beta_5))/(S_beta_0))^(-(n-1)/2);
posterior_prop_5 = first_part_5 * second_part_5 * third_part_5

first_part_6 = ((pi/(1-pi))^f_6)*(gamma^(-t_6));
second_part_6 = ((det(X_0'*X_0))^0.5)/((det(E_6+(X_6'*X_6)))^0.5);
third_part_6 = ((S_beta_6+(beta_6'*E_6*beta_6))/(S_beta_0))^(-(n-1)/2);
```

```
posterior_prop_6 = first_part_6 * second_part_6 * third_part_6


first_part_7 = ((pi/(1-pi))^f_7)*(gamma^(-t_7));
second_part_7 = ((det(X_0'*X_0))^0.5)/((det(E_7+(X_7'*X_7)))^0.5);
third_part_7 = ((S_beta_7+(beta_7'*E_7*beta_7))/(S_beta_0))^(-(n-1)/2);
posterior_prop_7 = first_part_7 * second_part_7 * third_part_7


first_part_8 = ((pi/(1-pi))^f_8)*(gamma^(-t_8));
second_part_8 = ((det(X_0'*X_0))^0.5)/((det(E_8+(X_8'*X_8)))^0.5);
third_part_8 = ((S_beta_8+(beta_8'*E_8*beta_8))/(S_beta_0))^(-(n-1)/2);
posterior_prop_8 = first_part_8 * second_part_8 * third_part_8


first_part_9 = ((pi/(1-pi))^f_9)*(gamma^(-t_9));
second_part_9 = ((det(X_0'*X_0))^0.5)/((det(E_9+(X_9'*X_9)))^0.5);
third_part_9 = ((S_beta_9+(beta_9'*E_9*beta_9))/(S_beta_0))^(-(n-1)/2);
posterior_prop_9 = first_part_9 * second_part_9 * third_part_9

first_part_10 = ((pi/(1-pi))^f_10)*(gamma^(-t_10));
second_part_10 = ((det(X_0'*X_0))^0.5)/((det(E_10+(X_10'*X_10)))^0.5);
third_part_10 = ((S_beta_10+(beta_10'*E_10*beta_10))/(S_beta_0))^(-(n-1)/2);
posterior_prop_10 = first_part_10 * second_part_10 * third_part_10


first_part_11 = ((pi/(1-pi))^f_11)*(gamma^(-t_11));
second_part_11 = ((det(X_0'*X_0))^0.5)/((det(E_11+(X_11'*X_11)))^0.5);
third_part_11 = ((S_beta_11+(beta_11'*E_11*beta_11))/(S_beta_0))^(-(n-1)/2);
posterior_prop_11 = first_part_11 * second_part_11 * third_part_11


first_part_12 = ((pi/(1-pi))^f_12)*(gamma^(-t_12));
second_part_12 = ((det(X_0'*X_0))^0.5)/((det(E_12+(X_12'*X_12)))^0.5);
third_part_12 = ((S_beta_12+(beta_12'*E_12*beta_12))/(S_beta_0))^(-(n-1)/2);
posterior_prop_12 = first_part_12 * second_part_12 * third_part_12
```

```
first_part_13 = ((pi/(1-pi))^f_13)*(gamma^(-t_13));

second_part_13 = ((det(X_0'*X_0))^0.5)/((det(E_13+(X_13'*X_13)))^0.5);

third_part_13 = ((S_beta_13+(beta_13'*E_13*beta_13))/(S_beta_0))^(-(n-1)/2);

posterior_prop_13 = first_part_13 * second_part_13 * third_part_13


first_part_14 = ((pi/(1-pi))^f_14)*(gamma^(-t_14));

second_part_14 = ((det(X_0'*X_0))^0.5)/((det(E_14+(X_14'*X_14)))^0.5);

third_part_14 = ((S_beta_14+(beta_14'*E_14*beta_14))/(S_beta_0))^(-(n-1)/2);

posterior_prop_14 = first_part_14 * second_part_14 * third_part_14


first_part_15 = ((pi/(1-pi))^f_15)*(gamma^(-t_15));

second_part_15 = ((det(X_0'*X_0))^0.5)/((det(E_15+(X_15'*X_15)))^0.5);

third_part_15 = ((S_beta_15+(beta_15'*E_15*beta_15))/(S_beta_0))^(-(n-1)/2);

posterior_prop_15 = first_part_15 * second_part_15 * third_part_15

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

posterior_1 = posterior_prop_1*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_2 = posterior_prop_2*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_3 = posterior_prop_3*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_4 = posterior_prop_4*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_5 = posterior_prop_5*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_6 = posterior_prop_6*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_7 = posterior_prop_7*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

posterior_8 = posterior_prop_8*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))
```

```
  posterior_9 = posterior_prop_9*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

  posterior_10 = posterior_prop_10*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

  posterior_11 = posterior_prop_11*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

  posterior_12 = posterior_prop_12*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

  posterior_13 = posterior_prop_13*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

  posterior_14 = posterior_prop_14*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

  posterior_15 = posterior_prop_15*(1/(posterior_prop_1 + posterior_prop_2 + posterior_prop_3 + posterior_prop_4 +
posterior_prop_5 + posterior_prop_6 + posterior_prop_7 + posterior_prop_8 + posterior_prop_9 + posterior_prop_10 +
posterior_prop_11 + posterior_prop_12 + posterior_prop_13 + posterior_prop_14 + posterior_prop_15))

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Appendix 4B Calculations for Bayesian Complete RSS Minimization Algorithm

Calculations shown for two missing observations (Runs 1 and 2 assumed to be missing in data 1)

Missing observations (initial values)

Run	A	B	C	D	*Treatment*	Response
1	-1	-1	-1	1	*d*	27
2	1	-1	-1	-1	*a*	27
3	-1	1	-1	-1	*bc*	66
4	1	1	-1	1	*abd*	37
5	-1	-1	1	-1	*c*	71
6	1	-1	1	1	*acd*	31
7	-1	1	1	1	*bcd*	28
8	1	1	1	-1	*abc*	29

Iteration	y_4	y_7
0	27	27
1	27.5	28
2	27.75	28.5
3	27.875	28.75
4	27.9375	28.875
5		
6		
7		
8		
9		
10		

AD
-1
-1
1
1
1
1
-1
-1

Iteration 0

y_1 = 27

y_6 = 27

b_1 = -8.5
b_2 = 0.5
b_3 = 0.25
b_4 = -8.75
b_12 =
b_13 =
b_14 = 11.75

Iteration 1

y_3 = 27.5

y_7 = 28

b_1 = -8.4375
b_2 =
b_3 =
b_4 = -8.8125
b_12 =
b_13 =
b_14 = 11.5625

Appendix 4C Calculations for Bayesian ANCOVA Algorithm

Calculations shown for two missing observations (Runs 1 and 2 assumed to be missing in data 1)

Std Order	A	B	C	D	*T'ment*	y	z_1	z_2	AB	AC	AD
1	-1	-1	-1	1	*d*	0	1	0	1	1	-1
2	1	-1	-1	-1	*a*	0	0	1	-1	-1	-1
3	-1	1	-1	-1	*b*	66	0	0	-1	1	1
4	1	1	-1	1	*abd*	37	0	0	1	-1	1
5	-1	-1	1	-1	*c*	71	0	0	1	-1	1
6	1	-1	1	1	*acd*	31	0	0	-1	1	1
7	-1	1	1	1	*bcd*	28	0	0	-1	-1	-1
8	1	1	1	-1	*abc*	29	0	0	1	1	-1

Using y:

Effect_I = 32.75
Effect_A = -17
Effect_B = 14.5
Effect_C = 14
Effect_D = -17.5
Effect_AB = 3
Effect_AC = -2.5
Effect_AD = 37

Residual_1 = -14 **using terms in the model as specified by Bayesian analysis**
Residual_2 = -14.5

Using z_1:

Effect_I = 0.125
Effect_A = -0.25
Effect_B = -0.25
Effect_C = -0.25
Effect_D = 0.25
Effect_AB = 0.25
Effect_AC = 0.25
Effect_AD = -0.25

Residual_1 = 0.5 **using terms in the model as specified by Bayesian analysis**
Residual_2 = 0

Using z_2:

Effect_I = 0.125
Effect_A = 0.25
Effect_B = -0.25
Effect_C = -0.25
Effect_D = 0.25
Effect_AB = -0.25
Effect_AC = -0.25
Effect_AD = -0.25

Residual_1 = -0.25 **using terms in the model as specified by Bayesian analysis**
Residual_2 = 0.75

using terms in the model as specified by Bayesian analysis

$$\begin{vmatrix} -0.5 & 0 \\ 0.25 & -0.75 \end{vmatrix} \quad \begin{vmatrix} -14 \\ -14.5 \end{vmatrix}$$

$$\begin{vmatrix} -2 & 0 \\ -0.666666667 & -1.333333 \end{vmatrix}$$

$$\begin{vmatrix} 28 \\ 28.66666667 \end{vmatrix}$$

5. MULTISTAGE FRACTIONAL FACTORIAL SPLIT-PLOT DESIGNS

In some nanomanufacturing situations, due to the physical constraints in the process, it is impracticable to execute a full or fractional factorial experiment. In such cases, restriction on randomization is imposed and the experimenter is forced to resort to a split-plot design or some of its variants. Many processes in nanomanufacturing are conducted over a series of stages. Additionally, some of the process variables in some of the stages might be difficult or hard to change in terms of time, limited resources, or – in many cases – money. Specifically, a polymerization process for the fabrication of nano-films is studied, where the fabrication is carried out over three stages. To execute efficient experimentation and fully understand the intricacies at the nano-scale, split-plot designs that can be applied effectively over multiple stages are proposed and their characteristics examined. General expressions for some of the properties of these designs and their analysis are developed. As common design ranking criteria such as resolution and minimum aberration do not provide the "best" designs in all cases, two new design optimality criteria are proposed. Catalogs of split-plot designs for three and four stages are created and ranked according to the proposed criteria.

5.1 Motivating Example: Nano-Scale Polymerization Process

Nanotechnology is the science of building structures atom by atom and molecule by molecule, thus providing an excellent opportunity for a bottom-up manufacturing

approach. Two of the most common techniques of fabricating nanostructures are commonly known as top-down and bottom-up approaches.

The motivating example for this research originated from one such bottom-up approach involving fabrication of a thin film using a surface initiated polymerization process (Dronavajjala et al., 2006). For this process to be used commercially, in the manufacture of resists for instance, the kinetics needs to be well understood to achieve reproducibility, precision, and control. In order to apply a working chemistry in fabricating devices, an excellent control of the kinetics of polymerization process is required. The polymerization process is carried out over three stages as illustrated in Figure 5.1.

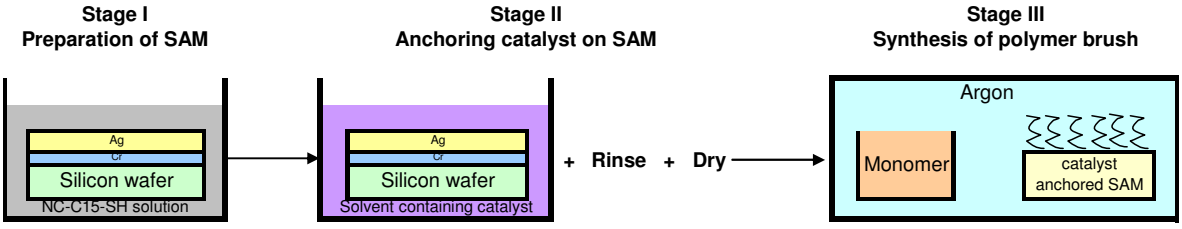

Figure 5.1: Three stage polymerization process.

Stage I. *Preparation of Self-Assembled Monolayer (SAM)*

SAMs are highly ordered quasi-two-dimensional structures formed by adsorption of appropriate precursors from solution into a solid substrate. As gold is relatively inert, has affinity to silicon, and is known to form stable monolayers (Bain et al., 1989), a gold substrate is chosen for the study. A two-inch silicon wafer is pre-coated with 10nm of

chromium to improve adhesion between the gold and silicon substrate. The gold substrates are prepared by thermal evaporation of 200nm of gold onto the wafer. The SAMs are formed by immersion of the evaporated substrates into a 1mM 16-mercapto-hexa-decane-nitrile (NC-C15-SH) solution in absolute ethanol at room temperature for 24 hours. After immersion, the substrates are rinsed with absolute ethanol and dried with nitrogen.

Stage II. *Anchoring Catalyst on the SAM*

The catalyst used in this study is a palladium complex ($[Pd(CH_3CH_2CN)_4](BF_4)_2$). This compound has been found to be an effective room temperature catalyst for the polymerization of styrene and its derivatives. The SAM formed substrate is immersed in a nitromethane solution containing the Pd complex at room temperature for 12 hours. The Pd complex reacts with the tail group (-CN) of the SAM and attaches itself to the substrate through ligand exchange. Once the catalyst anchors itself to the SAM, the substrate is rinsed with an appropriate solvent in order to wash the excess catalyst from the surface. It is imperative that the rinsing solvent not affect the synthesis of the polymer, and hence, the catalyst-anchored-SAM has to be dried. Therefore, after rinsing, the substrate is treated with a blast of high-purity nitrogen gas for a specified time, and all traces of the solvent removed.

Stage III. *Synthesis of polymer brush*

The catalyst acts as an initiator during the synthesis. The catalyst-anchored substrate is exposed to Argon gas, which is saturated with a monomer. Initiation proceeds by the

reaction of monomer and initiator, in the presence of Argon, to produce a growing species. The growing species propagates with addition of a new monomer molecule to form a longer chain. Polymers are this continuously formed as monomer reacts with the initiator. The rate of propagation should be greater than the rate of chain transfer for growth of the polymer thin film. One end of the polymer chain is tethered to a heterogeneous support (on the catalyst-anchored SAM) and the other end is growing away from the surface. The film thickness is dependent upon the length of polymer chain. Therefore, polymerization reaction should be done at conditions conducive for polymer growth where propagation is faster than that of chain transfer.

The thickness of the polymer brush at the end of the synthesis (third processing stage) is measured using an ellipsometry technique and is treated as the response.

5.2 Multistage Split-Plot (MSSP) Designs

Before constructing a design for the three-stage process described in the previous section, let us consider a simple two-stage experiment. If eight factors were identified for the investigation, a full factorial design would require $2^8 = 256$ runs, which would mean 256 different setting changes for each factor. A fractional factorial design would reduce the number of setting changes in general, but would not help if a particular factor were more difficult to change than the others were. In addition, if some factors were believed to be more significant than others were, the fractional factorial design would not provide increased precision in determining the significance of those particular effects (all effects would be tested for significance with equal precision). A better design in this case would be a split-lot design. Split-lot designs offer the advantage of running economical

experiments by consuming fewer resources while still yielding the same information as a fractional factorial experiment. However, due to the nature of the design, fractionation in a split-lot design is limited as the generating columns in a particular stage can only be a combination of the factors within the same stage. Additionally, in a split-lot design, experimental units are re-grouped in every stage to form new groups. Hence, an implicit assumption being made is that before experimental units can be processed in a stage, processing at all preceding stages need to be completed. However, in experimental setups like the polymerization process discussed in Section 5.1, it is important that some experimental units be treated over all stages before others as there might be additional "unwanted" reactions occurring if the films are not formed on the substrate.

To overcome the shortcomings of the fractional factorial and split-lot designs, a repeated measures designs could be employed in such circumstances. These designs are used when multiple measurements are taken on some of the factors under consideration. Under certain conditions, a repeated measures deign (Kutner et al., 2005) can also be viewed as a split-plot design. Let us consider a simple example to shed some light on this claim.

Example 1

Consider a simple design having three factors (p, q, and r) in which repeated measures are taken only on two (p and q) of the three factors. Following the usual terminology, factor r is known as the "within" factor, and factors p and q are referred to as the "between" factors. A fourth factor, d denotes the experimental unit. If two levels of each factor are to be investigated, the design could be as illustrated in Table 5.1.

Table 5.1: Repeated Measures Design

		r		
p	***q***	-	+	
-	-			d_1
	+			d_2
+	-			d_3
	+			d_4

Experimental units are randomly assigned to four levels of factors *p* and *q*. Factor *r* is then assigned randomly to the experimental units. An experimental unit assigned to a particular combination of factor levels is typically referred to as an "individual", following the nomenclature. The term "individual" arises from the application of repeated measures design to psychological or behavioral studies, where multiple measurements are taken on the same person (also referred to as the subject).

Comparisons between factor *p* (or *q*) level means involve differences between groups as well as differences associated with the two factor *p* (or *q*) levels. Moreover, comparisons between factor *r* level means at the same levels of factors *p* and *q* involve differences associated with the two factor *r* levels.

The following assumptions are necessary for a repeated measures design:

- *Normality*: The probability distribution of the errors follows a normal distribution with mean zero.
- *Independence*: The errors are independently distributed.
- *Constant variance*: The probability distribution of each error has a constant variance.

- *Sphericity*: Any two observations, Y_{ij} and $Y_{ij'}$ for a given subject are correlated in the same fashion for all subjects.
- *Compound symmetry*: Any two observations from different subjects (between factors) are independent.

The model (assuming no interaction between subjects and treatments) is

$$Y_{ijkl} = \mu + p_i + q_j + pq_{ij} + d(pq)_{k(ij)} + r_l + pr_{il} + qr_{jl} + pqr_{ijk} + dr(pq)_{kl(ij)} + \varepsilon_{ijkl} \qquad \textbf{(5.1)}$$

where
μ is a constant,
$d(pq)_{k(ij)} \sim N(0, \sigma_d^2), i.i.d,$
$\varepsilon_{ijkl} \sim N(0, \sigma_\varepsilon^2), i.i.d,$
$\sigma\{Y_{ijkl}, Y_{ijkl'}\} = \sigma_d^2,\ l \neq l'$; representing the sphericity assumption, and
$\sigma\{Y_{ijkl}, Y_{i'j'k'l'}\} = 0,\ i \neq i',\ j \neq j',\ k \neq' k'$; representing the compound symmetry assumption.

As differences between subjects are treated separately from differences within subjects, small consistent differences among the subjects can be detected. This increases the power of significance tests. Furthermore, the reduction in the error component of the model represents a direct increase in economy and power.

Close examination of the table of expected mean squares reveals that "between individuals" effects, *p, q,* and *pq* are tested against σ_d^2 and "within individuals" effects, *r*, *pr*, *qr*, and *pqr* are tested against σ^2.

The use of repeated measures design has been confined due to some limitations. A major disadvantage of repeated measures design is that measurements taken on an individual over adjacent periods of time might be correlated, which is not accounted for in the model. For instance, individuals might tend to give higher ratings towards the end

of a period than at the beginning, or vice-versa. Another scenario is when there exists a carry-over effect from a preceding observation. The only way to discount this effect is to randomize the order of treatments on individuals. A sphericity assumption would then be more reasonable.

However, in certain instances, it is impossible to randomize the treatment order. For instance, when measurements are to be taken on an individual in intervals of five minutes, it is impossible to randomize - hence, the limitation of repeated measures design in behavioral and life sciences studies. A more reasonable setting where these designs can be applied is in manufacturing situations where experimental units replace individuals and measurements are not taken over time.

A repeated measures design, under certain conditions, can be also thought of as a split-plot design. One such condition is when individuals in the repeated measures design can be replaced by experimental units. The sphericity assumption, then, would be much more reasonable as experimental units tend to be correlated in a constant fashion. Furthermore, the "between individual" factor becomes the whole plot factor and the "within individuals" factor resembles the sub-plot factor. As in any split-plot design, observations from different whole plots are assumed to be independent, the compound symmetry assumption from the repeated measures design is satisfied. Hence, in Example 1, factors p and q represent whole plot factors, and r represents the sub-plot factor. In general, a non-repeated ("between") factor in a repeated measures design is represented by whole plot factors and repeated ("within") factors represented by sub-plot factors in a split-plot design. An additional condition necessary to balance the orthogonal structure in multi-factor two-level split-plot designs is that the number of experimental units should

be equal to the number of runs in the sub-plot. The split-plot design for the example is shown in Table 5.2.

Table 5.2: Equivalent Split-Plot Design

p	*q*	*r*	
-	-	-	d_1
		+	$d_{1'}$
	+	-	d_2
		+	$d_{2'}$
+	-	-	d_3
		+	$d_{3'}$
	+	-	d_4
		+	$d_{4'}$

In this design, individuals (d_i) from the repeated measures design are replaced by two (d_i and $d_{i'}$) experimental units. An appropriate model for the split-plot design applied to Example 1 is

$$Y_{ijkl} = \mu + p_i + q_j + pq_{ij} + WP_E + r_k + pr_{ik} + qr_{jk} + pqr_{ijk} + SP_E \qquad \textbf{(5.2)}$$

where

μ is a constant,

$WP_E \sim N(0, \sigma^2_{WP}), i.i.d,$

$SP_E \sim N(0, \sigma^2_{SP}), i.i.d,$

$\sigma\{Y_{ijkl}, Y_{ijkl'}\} = \sigma^2_{WP}$, $l \neq l'$; representing the sphericity assumption, and

$\sigma\{Y_{ijkl}, Y_{i'j'k'l'}\} = 0$, $i \neq i'$, $j \neq j'$, $k \neq' k'$; representing the compund symmetry assumption.

The expected mean squares of the split-plot design reveals that effects *p, q,* and *pq* are tested for significance against WP_E, and effects *r, pr, qr,* and *pqr* are tested against SP_E.

In summary, split-plot designs are a special case of repeated measures designs. For a repeated measures design to be considered a split-plot, the following must be true:

i. There must exist at least one "between" factor.

ii. There must exist at least one "within" factor.

iii. Individuals can be replaced by single or multiple experimental units.

iv. The assignment order of the experimental units within "between" factor(s) must be completely randomized.

v. Sphericity assumption must be satisfied.

vi. Compound symmetry assumption must be satisfied.

A multistage split-plot (MSSP) design can be thought of as having a single whole plot and a subsequent series of sub-plots. The MSSP design is equivalent to a repeated measures design with multiple stages of "within" and "between" factors. In the next sections, various theoretical properties and characteristics of MSSP and multistage fractional factorial split-plot (MSFFSP) designs are investigated, and general expressions for *m* stages would be provided.

5.3 Design Considerations

In this section, some properties of the MSSP design are investigated through some examples. General expressions are derived from these examples, wherever appropriate.

5.3.1 Linear Model and Error Structure

Based on the randomization principle, a linear model for a split-plot design was derived by Hinklemann and Kempthorne (1994). As the principle of randomization is the same for MSSP designs, the linear model for the same need not be derived, but simply extended to fit the case for a *m*-stage MSSP.

Let

$$x^{(i)} = S_{i\text{-}1}(1 + W + \sum_{j=2}^{i-1} x^{(j)}), \tag{5.3}$$

$$y^{(m)} = \mu + r + W + WP_E + \sum_{i=2}^{m} [x^{(i)} + SP_{(i-1)_E}], \tag{5.4}$$

$$SP_{i_E} = rW \prod_{j=1}^{i} S_j, \tag{5.5}$$

where

S_i is an effect in stage i,

W is an effect in stage 1 (whole plot),

$y^{(m)}$ is response measured at the end of m^{th} stage,

μ is the overall mean,

r is the random effect of the replicate,

WP_E is the whole plot (stage 1) error ~$N(0, \sigma^2_{WP_E})$, and

SP_{i_E} is the error of stage i ~$N(0, \sigma^2_{SP_{i_E}})$.

Equations 5.3, 5.4, and 5.5 represent the general expression of the linear model for an *m*-stage MSSP design. The usual normality, independence, and constant variance assumptions are assumed for the model. Unlike a split-lot design, it is observed that the number of error terms equals the number of stages.

Example 2:

Consider a four-stage MSSP design involving factors *A*, *B*, *C*, and *D* in stages 1, 2, 3, and 4, respectively. Following the general expression, a linear model for a four-stage design is represented by

$$y^{(4)} = \mu + r + A + WP_E + x^{(2)} + SP_{1_E} + x^{(3)} + SP_{2_E} + x^{(4)} + SP_{3_E} \quad \textbf{(5.6)}$$

where

$$x^{(2)} = B + AB,$$
$$x^{(3)} = C + AC + BC + ABC, \text{ and}$$
$$x^{(4)} = D + AD + CD + ACD + BCD + ABCD + BD + ABD.$$

Therefore, by substitution, we get

$$y^{(4)} = \mu + r + A + WP_E + B + AB + SP_{1_E} + C + AC + BC + ABC + SP_{2_E} + D + AD + CD + ACD + BCD + ABCD + BD + ABD + SP_{3_E} \quad \textbf{(5.7)}$$

5.3.2 Significance Tests for Effects in MSFFSP Designs

As there are many error strata in a MSSP design, it is important to know which contrasts are to be tested against which error term in order to test for significance. A

straightforward but tedious way to do this is to derive the variance of the contrasts based on the model.

Example 3

Consider a simple unreplicated $2^2 \times 2^2 \times 2^2$ MSSP design, with factors *A* and *B* in stage 1, *P* and *Q* in stage 2, and *M* and *N* in stage 3. As this is a three-stage experiment, there would be three error terms, corresponding to the three stages. The model for the design can be represented as

$$y_{ijk} = f(x_{ijk}) + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk} \qquad i, j, k = 1..4 \tag{5.8}$$

where

i, j, and *k* represent the number of runs in stages 1, 2, and 3, respectively.

$f(x_{ijk})$ is the deterministic part of the model and contains the effects of stage 1 (St_1), stage 2 (St_2), and stage 3 (St_3), given by

$$f(x_{ijk}) = \mu + (St_1)_k + (St_2)_j + (St_3)_i. \tag{5.9}$$

ε_k is $i.i.d., \sim N(0, \sigma^2_{\mathrm{WP}})$, and represents the stage 1 error.

ε_{jk} is $i.i.d., \sim N(0, \sigma^2_{\mathrm{S_1}})$, and represents the stage 2 error.

ε_{ijk} is $i.i.d., \sim N(0, \sigma^2_{\mathrm{S_2}})$, and represents the stage 3 error.

The model terms for each of the 64 runs of the design, excluding the deterministic part is shown in Table 5.3.

Table 5.3: Model Terms

Run	*A*	*B*	*P*	*Q*	*M*	*N*	y_{ijk}	Error term
1	-	-	-	-	-	-	y_{111}	$\varepsilon_1 + \varepsilon_{11} + \varepsilon_{111}$
2	-	-	-	-	+	-	y_{211}	$\varepsilon_1 + \varepsilon_{11} + \varepsilon_{112}$
3	-	-	-	-	-	+	y_{311}	$\varepsilon_1 + \varepsilon_{11} + \varepsilon_{113}$
4	-	-	-	-	+	+	y_{411}	$\varepsilon_1 + \varepsilon_{11} + \varepsilon_{114}$
5	-	-	+	-	-	-	y_{121}	$\varepsilon_1 + \varepsilon_{12} + \varepsilon_{121}$
6	-	-	+	-	+	-	y_{221}	$\varepsilon_1 + \varepsilon_{12} + \varepsilon_{122}$
7	-	-	+	-	-	+	y_{321}	$\varepsilon_1 + \varepsilon_{12} + \varepsilon_{123}$
8	-	-	+	-	+	+	y_{421}	$\varepsilon_1 + \varepsilon_{12} + \varepsilon_{124}$
9	-	-	-	+	-	-	y_{131}	$\varepsilon_1 + \varepsilon_{13} + \varepsilon_{131}$
10	-	-	-	+	+	-	y_{231}	$\varepsilon_1 + \varepsilon_{13} + \varepsilon_{132}$
11	-	-	-	+	-	+	y_{331}	$\varepsilon_1 + \varepsilon_{13} + \varepsilon_{133}$
12	-	-	-	+	+	+	y_{431}	$\varepsilon_1 + \varepsilon_{13} + \varepsilon_{134}$
13	-	-	+	+	-	-	y_{141}	$\varepsilon_1 + \varepsilon_{14} + \varepsilon_{141}$
14	-	-	+	+	+	-	y_{241}	$\varepsilon_1 + \varepsilon_{14} + \varepsilon_{142}$
15	-	-	+	+	-	+	y_{341}	$\varepsilon_1 + \varepsilon_{14} + \varepsilon_{143}$
16	-	-	+	+	+	+	y_{441}	$\varepsilon_1 + \varepsilon_{14} + \varepsilon_{144}$
17	+	-	-	-	-	-	y_{112}	$\varepsilon_2 + \varepsilon_{21} + \varepsilon_{211}$
18	+	-	-	-	+	-	y_{212}	$\varepsilon_2 + \varepsilon_{21} + \varepsilon_{212}$
19	+	-	-	-	-	+	y_{312}	$\varepsilon_2 + \varepsilon_{21} + \varepsilon_{213}$
20	+	-	-	-	+	+	y_{412}	$\varepsilon_2 + \varepsilon_{21} + \varepsilon_{214}$
21	+	-	+	-	-	-	y_{122}	$\varepsilon_2 + \varepsilon_{22} + \varepsilon_{221}$
22	+	-	+	-	+	-	y_{222}	$\varepsilon_2 + \varepsilon_{22} + \varepsilon_{222}$
23	+	-	+	-	-	+	y_{322}	$\varepsilon_2 + \varepsilon_{22} + \varepsilon_{223}$
24	+	-	+	-	+	+	y_{422}	$\varepsilon_2 + \varepsilon_{22} + \varepsilon_{224}$
25	+	-	-	+	-	-	y_{132}	$\varepsilon_2 + \varepsilon_{23} + \varepsilon_{231}$
26	+	-	-	+	+	-	y_{232}	$\varepsilon_2 + \varepsilon_{23} + \varepsilon_{232}$
27	+	-	-	+	-	+	y_{332}	$\varepsilon_2 + \varepsilon_{23} + \varepsilon_{233}$
28	+	-	-	+	+	+	y_{432}	$\varepsilon_2 + \varepsilon_{23} + \varepsilon_{234}$
29	+	-	+	+	-	-	y_{142}	$\varepsilon_2 + \varepsilon_{24} + \varepsilon_{241}$
30	+	-	+	+	+	-	y_{242}	$\varepsilon_2 + \varepsilon_{24} + \varepsilon_{242}$
31	+	-	+	+	-	+	y_{342}	$\varepsilon_2 + \varepsilon_{24} + \varepsilon_{243}$
32	+	-	+	+	+	+	y_{442}	$\varepsilon_2 + \varepsilon_{24} + \varepsilon_{244}$
33	-	+	-	-	-	-	y_{113}	$\varepsilon_3 + \varepsilon_{31} + \varepsilon_{311}$
34	-	+	-	-	+	-	y_{213}	$\varepsilon_3 + \varepsilon_{31} + \varepsilon_{312}$
35	-	+	-	-	-	+	y_{313}	$\varepsilon_3 + \varepsilon_{31} + \varepsilon_{313}$
36	-	+	-	-	+	+	y_{413}	$\varepsilon_3 + \varepsilon_{31} + \varepsilon_{314}$
37	-	+	+	-	-	-	y_{123}	$\varepsilon_3 + \varepsilon_{32} + \varepsilon_{321}$
38	-	+	+	-	+	-	y_{223}	$\varepsilon_3 + \varepsilon_{32} + \varepsilon_{322}$
39	-	+	+	-	-	+	y_{323}	$\varepsilon_3 + \varepsilon_{32} + \varepsilon_{323}$
40	-	+	+	-	+	+	y_{423}	$\varepsilon_3 + \varepsilon_{32} + \varepsilon_{324}$
41	-	+	-	+	-	-	y_{133}	$\varepsilon_3 + \varepsilon_{33} + \varepsilon_{331}$
42	-	+	-	+	+	-	y_{233}	$\varepsilon_3 + \varepsilon_{33} + \varepsilon_{332}$
43	-	+	-	+	-	+	y_{333}	$\varepsilon_3 + \varepsilon_{33} + \varepsilon_{333}$
44	-	+	-	+	+	+	y_{433}	$\varepsilon_3 + \varepsilon_{33} + \varepsilon_{334}$
45	-	+	+	+	-	-	y_{143}	$\varepsilon_3 + \varepsilon_{34} + \varepsilon_{341}$
46	-	+	+	+	+	-	y_{243}	$\varepsilon_3 + \varepsilon_{34} + \varepsilon_{342}$
47	-	+	+	+	-	+	y_{343}	$\varepsilon_3 + \varepsilon_{34} + \varepsilon_{343}$
48	-	+	+	+	+	+	y_{443}	$\varepsilon_3 + \varepsilon_{34} + \varepsilon_{344}$
49	+	+	-	-	-	-	y_{114}	$\varepsilon_4 + \varepsilon_{41} + \varepsilon_{441}$
50	+	+	-	-	+	-	y_{214}	$\varepsilon_4 + \varepsilon_{41} + \varepsilon_{442}$
51	+	+	-	-	-	+	y_{314}	$\varepsilon_4 + \varepsilon_{41} + \varepsilon_{443}$
52	+	+	-	-	+	+	y_{414}	$\varepsilon_4 + \varepsilon_{41} + \varepsilon_{444}$
53	+	+	+	-	-	-	y_{124}	$\varepsilon_4 + \varepsilon_{42} + \varepsilon_{421}$
54	+	+	+	-	+	-	y_{224}	$\varepsilon_4 + \varepsilon_{42} + \varepsilon_{422}$
55	+	+	+	-	-	+	y_{324}	$\varepsilon_4 + \varepsilon_{42} + \varepsilon_{423}$
56	+	+	+	-	+	+	y_{424}	$\varepsilon_4 + \varepsilon_{42} + \varepsilon_{424}$
57	+	+	-	+	-	-	y_{134}	$\varepsilon_4 + \varepsilon_{43} + \varepsilon_{431}$
58	+	+	-	+	+	-	y_{234}	$\varepsilon_4 + \varepsilon_{43} + \varepsilon_{432}$
59	+	+	-	+	-	+	y_{334}	$\varepsilon_4 + \varepsilon_{43} + \varepsilon_{433}$
60	+	+	-	+	+	+	y_{434}	$\varepsilon_4 + \varepsilon_{43} + \varepsilon_{434}$
61	+	+	+	+	-	-	y_{144}	$\varepsilon_4 + \varepsilon_{44} + \varepsilon_{441}$
62	+	+	+	+	+	-	y_{244}	$\varepsilon_4 + \varepsilon_{44} + \varepsilon_{442}$
63	+	+	+	+	-	+	y_{344}	$\varepsilon_4 + \varepsilon_{44} + \varepsilon_{443}$
64	+	+	+	+	+	+	y_{444}	$\varepsilon_4 + \varepsilon_{44} + \varepsilon_{444}$

The estimate of any term in the model can be calculated as usual, i.e. the difference between the summation of all runs that have a "+" sign for the effect and the summation of all runs that have a "-" sign for that effect in the design matrix. For example, estimate of effect *A* (stage 1) in Example 3 is calculated from Table 5.3 and is given by

$$\begin{aligned}\hat{A} = \frac{1}{32}[&-16\varepsilon_1 - 4\sum_{j=1}^{4}\varepsilon_{1j} - \sum_{j=1}^{4}\varepsilon_{11j} - \sum_{j=1}^{4}\varepsilon_{12j} - \sum_{j=1}^{4}\varepsilon_{13j} - \sum_{j=1}^{4}\varepsilon_{14j} \\ &+16\varepsilon_2 + 4\sum_{j=1}^{4}\varepsilon_{2j} + \sum_{j=1}^{4}\varepsilon_{21j} + \sum_{j=1}^{4}\varepsilon_{22j} + \sum_{j=1}^{4}\varepsilon_{23j} + \sum_{j=1}^{4}\varepsilon_{24j} \\ &-16\varepsilon_3 - 4\sum_{j=1}^{4}\varepsilon_{3j} - \sum_{j=1}^{4}\varepsilon_{31j} - \sum_{j=1}^{4}\varepsilon_{32j} - \sum_{j=1}^{4}\varepsilon_{33j} - \sum_{j=1}^{4}\varepsilon_{34j} \\ &+16\varepsilon_4 + 4\sum_{j=1}^{4}\varepsilon_{4j} + \sum_{j=1}^{4}\varepsilon_{41j} + \sum_{j=1}^{4}\varepsilon_{42j} + \sum_{j=1}^{4}\varepsilon_{43j} + \sum_{j=1}^{4}\varepsilon_{44j}\end{aligned} \quad . \qquad \textbf{(5.10)}$$

Applying the variance operator, we get

$$\begin{aligned}\mathrm{Var}(\hat{A}) &= \frac{1}{32^2}[4\times 256\sigma_W^2 + 4\times 64\sigma_{S_1}^2 + 4\times 4\sigma_{S_2}^2 + 4\times 4\sigma_{S_2}^2 + 4\times 4\sigma_{S_2}^2 + 4\times 4\sigma_{S_2}^2] \\ &= \sigma_W^2 + \frac{1}{4}\sigma_{S_1}^2 + \frac{1}{16}\sigma_{S_2}^2\end{aligned} \quad . \qquad \textbf{(5.11)}$$

It can be seen that the estimate of an effect in stage 1 involves all the error terms of the model. Similarly, the estimate of an effect in stage 2 (say effect *P*) is given by

$$\hat{P} = \frac{1}{32}[-4\varepsilon_1 - 4\varepsilon_{11} - \sum_{j=1}^{4}\varepsilon_{11j} + 4\varepsilon_1 + 4\varepsilon_{12} + \sum_{j=1}^{4}\varepsilon_{12j} - 4\varepsilon_1 - 4\varepsilon_{13} - \sum_{j=1}^{4}\varepsilon_{13j} + 4\varepsilon_1 + 4\varepsilon_{14} + \sum_{j=1}^{4}\varepsilon_{14j}$$
$$- 4\varepsilon_2 - 4\varepsilon_{21} - \sum_{j=1}^{4}\varepsilon_{21j} + 4\varepsilon_2 + 4\varepsilon_{22} + \sum_{j=1}^{4}\varepsilon_{22j} - 4\varepsilon_2 - 4\varepsilon_{23} - \sum_{j=1}^{4}\varepsilon_{23j} + 4\varepsilon_2 + 4\varepsilon_{24} + \sum_{j=1}^{4}\varepsilon_{24j}$$
$$- 4\varepsilon_3 - 4\varepsilon_{31} - \sum_{j=1}^{4}\varepsilon_{31j} + 4\varepsilon_3 + 4\varepsilon_{32} + \sum_{j=1}^{4}\varepsilon_{32j} - 4\varepsilon_3 - 4\varepsilon_{33} - \sum_{j=1}^{4}\varepsilon_{33j} + 4\varepsilon_3 + 4\varepsilon_{34} + \sum_{j=1}^{4}\varepsilon_{34j}$$
$$- 4\varepsilon_4 - 4\varepsilon_{41} - \sum_{j=1}^{4}\varepsilon_{41j} + 4\varepsilon_4 + 4\varepsilon_{42} + \sum_{j=1}^{4}\varepsilon_{42j} - 4\varepsilon_4 - 4\varepsilon_{43} - \sum_{j=1}^{4}\varepsilon_{43j} + 4\varepsilon_4 + 4\varepsilon_{44} + \sum_{j=1}^{4}\varepsilon_{44j} \quad . \textbf{(5.12)}$$

Applying the variance operator, we get

$$\begin{aligned} \mathrm{Var}(\hat{P}) &= \frac{1}{32^2}[8\times16\sigma_{S_1}^2 + 8\times4\sigma_{S_2}^2 + 8\times16\sigma_{S_1}^2 + 8\times4\sigma_{S_2}^2] \\ &= \frac{1}{4}\sigma_{S_1}^2 + \frac{1}{16}\sigma_{S_2}^2 \end{aligned} \quad . \qquad \textbf{(5.13)}$$

Similarly,

$$\mathrm{Var}(\hat{M}) = \frac{1}{32^2}[64\sigma_{S_2}^2] = \frac{1}{16}\sigma_{S_2}^2 . \qquad \textbf{(5.14)}$$

Thus, it is observed that the estimate of an effect from a particular stage includes all of the error terms in that particular stage and all of the stages following it. Furthermore, Table 5.4 represents all effects that have the same variance estimates. It can be seen that all main effects in a stage, interactions between the factors in the stage ("pure" interactions), and the interactions of the main effects and "pure" interactions with all effects from the previous stages have the same estimate of variance.

Table 5.4: Contrasts having Same Variance Estimates

$\varepsilon_{k}, \varepsilon_{jk}$, and ε_{ijk}	ε_{jk} and ε_{ijk}	ε_{ijk}	
A	*P*	*M*	*AQM*
B	*Q*	*N*	*AQN*
AB	*PQ*	*MN*	*AQMN*
	AP	*AM*	*APQM*
	AQ	*AN*	*APQN*
	APQ	*AMN*	*APQMN*
	BP	*BM*	*BPM*
	BQ	*BN*	*BPN*
	BPQ	*BMN*	*BPMN*
	ABP	*ABM*	*BQM*
	ABQ	*ABN*	*BQN*
	ABPQ	*ABMN*	*BQMN*
		PM	*BPQM*
		PN	*BPQN*
		PMN	*BPQMN*
		QM	*ABPM*
		QN	*ABPN*
		QMN	*ABPMN*
		PQM	*ABQM*
		PQN	*ABQN*
		PQMN	*ABQMN*
		APM	*ABPQM*
		APN	*ABPQN*
		APMN	*ABPQMN*

The expected value of the stage 1 error consists of variances from stage 1, stage 2, and stage 3 error terms. Similarly, the expected value of the stage 2 error consists of variances from stage 2, and stage 3 error terms and finally, the expected value of the stage 3 error consists of variance from stage 3 error term only. Comparing this to the estimates of variances for the effects shown in Table 5.4, it is seen that effects consisting of ε_k, ε_{jk}, and ε_{ijk} should be tested against stage 1 error, effects containing ε_{jk}, and ε_{ijk} should be tested against stage 2 error, and effects containing ε_{ijk} should be tested against stage 3 error.

Example 4

If an additional factor has to be tested in each of the three stages of Example 3, and the number of runs (64) needs to be the same, fractionation of a certain degree needs to be considered. A MSFFSP design could be considered in this case. Let the additional factors be *C* (stage 1), *R* (stage 2), and *O* (stage 3). The design can be represented by $2^{k_1-p_1} \times 2^{k_2-p_2} \times 2^{k_3-p_3}$, where p_i represents the number of generators in stage *i*, and (k_i-p_i) is the number of runs in stage *i*. Let *N* be the number of total runs in the design and St_i be an effect in stage *i*. Following Equations 5.11, 5.13, and 5.14, the estimate of variance for a three-stage fractional factorial split-plot design can be generalized as follows:

$$\begin{aligned}\mathrm{Var}(\hat{St}_1) &= Var\left\{\frac{2}{N}\left(2^{k_2-p_2}\times 2^{k_3-p_3}\sum_{j=1}^{2^{k_1-p_1}}\pm\varepsilon_j + 2^{k_3-p_3}\sum_{i=1}^{2^{k_1-p_1}}\sum_{j=1}^{2^{k_2-p_2}}\pm\varepsilon_{ij} + \sum_{i=1}^{2^{k_1-p_1}}\sum_{j=1}^{2^{k_2-p_2}}\sum_{k=1}^{2^{k_3-p_3}}\pm\varepsilon_{ijk}\right)\right\} \\ &= \frac{4}{N^2}\begin{bmatrix} 2^{k_1-p_1}\times 2^{2(k_2-p_2)}\times 2^{2(k_3-p_3)}\sigma^2_{St_1} + 2^{k_1-p_1}\times 2^{k_2-p_2}\times 2^{2(k_3-p_3)}\sigma^2_{St_2} + \\ 2^{k_1-p_1}\times 2^{k_2-p_2}\times 2^{k_3-p_3}\sigma^2_{St_3}\end{bmatrix},\end{aligned}$$

$$\mathrm{Var}(\hat{St}_1) = \frac{4}{N}\left[2^{k_2-p_2}\times 2^{k_3-p_3}\sigma^2_{St_1} + 2^{k_3-p_3}\sigma^2_{St_2} + \sigma^2_{St_3}\right]. \quad \textbf{(5.15)}$$

$$\mathrm{Var}(\hat{St}_2) = Var\left\{\frac{2}{N}\left(2^{k_3-p_3}\sum_{i=1}^{2^{k_1-p_1}}\sum_{j=1}^{2^{k_2-p_2}}\pm\varepsilon_{ij} + \sum_{i=1}^{2^{k_1-p_1}}\sum_{j=1}^{2^{k_2-p_2}}\sum_{k=1}^{2^{k_3-p_3}}\pm\varepsilon_{ijk}\right)\right\},$$

$$\mathrm{Var}(\hat{St}_2) = \frac{4}{N}\left[2^{k_3-p_3}\sigma^2_{St_2} + \sigma^2_{St_3}\right]. \quad \textbf{(5.16)}$$

$$\mathrm{Var}(\hat{St}_3) = Var\left\{\frac{2}{N}\left(\sum_{i=1}^{2^{k_1-p_1}}\sum_{j=1}^{2^{k_2-p_2}}\sum_{k=1}^{2^{k_3-p_3}}\pm\varepsilon_{ijk}\right)\right\} = \frac{4}{N}\left[\sigma^2_{St_3}\right]. \quad \textbf{(5.17)}$$

Extending this result to an *m*-stage MSFFSP (*m*-MSFFSP) design, a general expression for the variance of an effect in stage *i* can be calculated by

$$\mathrm{Var}(\hat{St}_i) = \frac{4}{N}\left\{\left(\sum_{j=1}^{m-1} 2^{k_{j+1}-p_{j+1}} \times 2^{k_{j+2}-p_{j+2}} \times \ldots\ldots \times 2^{k_m-p_m}\,\sigma^2_{St_j}\right) + \sigma^2_{St_m}\right\};\quad i = 1,2,..,m \qquad \textbf{(5.18)}$$

where

St_i is an effect in stage *i*, and

$\sigma^2_{St_i}$ is the variance of the error term in stage *i*.

Close examination of the variance of estimates reveals that effects in latter stages are estimated with greater precision than effects from preceding stages as the variance of those contrasts are smaller. For instance, main effect *P* is estimated with higher precision than main effect *A*, main effect *M* is estimated with higher precision that main effects *A* and *P*, and so on. This result is of particular interest in constructing appropriate designs if prior knowledge of "important" effects is available.

Assignment of contrasts to appropriate error terms is a bit more challenging in a MSFFSP design. If confounding within plots for the design in Example 4 is considered, then the defining relation corresponding to maximum resolution is

$$I = ABC = PQR = MNO = ABCPQR = ABCMNO = PQRMNO = ABCPQRMNO. \qquad \textbf{(5.19)}$$

The estimate of variance for factor *A* would suggest that the main effect *A* be tested against stage 1 error and the effect *AMNO* be tested against stage 3 error. The alias

structure for the design up to fourth order interactions is provided in Table 5.5. From the table and the defining relation, it is seen that the confounding structure is

$$\boldsymbol{A} + BC + APQR + \boldsymbol{AMNO} + BCPQR + BCMNO + APQRMNO + BCPQRMNO.$$

Table 5.5: Alias Structure for Example 4

A + BC + APQR + AMNO	*AQM + BCQM + APRM + AQNO*
B + AC + BPQR + BMNO	*AQN + BCQN + APRN + AQMO*
C + AB + CPQR + CMNO	*AQO + BCQO + APRO + AQMN*
P + ABCP + QR + PMNO	*ARM + BCRM + APQM + ARNO*
Q + ABCQ + PR + QMNO	*ARN + BCRN + APQN + ARMO*
R + ABCR + PQ + RMNO	*ARO + BCRO + APQO + ARMN*
M + ABCM + PQRM + NO	*APM + BCPM + AQRM + APNO*
N + ABCN + PQRN + MO	*APN + BCPN + AQRN + APMO*
O + ABCO + PQRO + MN	*APO + BCPO + AQRO + APMN*
AP + BCP + AQR + BCQR	*BPM + ACPM + BQRM + BPNO*
AQ + BCQ + APR + BCPR	*BPN + ACPN + BQRN + BPMO*
AR + BCR + APQ + BCPQ	*BPO + ACPO + BQRO + BPMN*
BP + ACP + BQR + ACQR	*BQM + ACQM + BPRM + BQNO*
BQ + ACQ + BPR + ACPR	*BQN + ACQN + BPRN + BQMO*
BR + ACR + BPQ + ACPQ	*BQO + ACQO + BPRO + BQMN*
CP + ABP + CQR + ABQR	*BRM + ACRM + BPQM + BRNO*
CQ + ABQ + CPR + ABPR	*BRN + ACRN + BPQN + BRMO*
CR + ABR + CPQ + ABPQ	*BRO + ACRO + BPQO + BRMN*
PM + QRM + PNO + QRNO	*CPM + ABPM + CQRM + CPNO*
PN + QRN + PMO + QRMO	*CPN + ABPN + CQRN + CPMO*
PO + QRO + PMN + QRMN	*CPO + ABPO + CQRO + CPMN*
QM + PRM + QNO + PRNO	*CQM + ABQM + CPRM + CQNO*
QN + PRN + QMO + PRMO	*CQN + ABQN + CPRN + CQMO*
QO + PRO + QMN + PRMN	*CQO + ABQO + CPRO + CQMN*
RM + PQM + RNO + PQNO	*CRM + ABRM + CPQM + CRNO*
RN + PQN + RMO + PQMO	*CRN + ABRN + CPQN + CRMO*
RO + PQO + RMN + PQMN	*CRO + ABRO + CPQO + CRMN*
AM + BCM + ANO + BCNO	
AN + BCN + AMO + BCMO	
AO + BCO + AMN + BCMN	
BM + ACM + BNO + ACNO	
BN + ACN + BMO + ACMO	
BO + ACO + BMN + ACMN	
CM + ABM + CNO + ABNO	
CN + ABN + CMO + ABMO	
CO + ABO + CMN + ABMN	

Hence, the quandary of whether the above alias chain is to be tested against stage 1 or stage 3 error term needs to be addressed. This calls for rules that prioritize certain effects to be tested against appropriate error terms in the presence of complex aliasing. Similar rules were first developed for a basic split-plot design by Bisgaard (2000) and summarized by Bingham and Sitter (2001). Bisgaard (2000) considered a $2^{k-p} \times 2^{q-r}$ split plot design having $(2^{k-p} \times 2^{q-r} - 1)$ linear contrasts. The following rule was developed to assign contrasts to appropriate error terms.

Rule: The $(2^{k-p} - 1)$ contrasts generated by multiplying out the $(k - p)$ basic generators for the whole plot design in all possible ways and their aliases are tested against the whole plot error. The remaining 2^{k-p} $(2^{q-r} - 1)$ contrasts are tested against the sub-plot error.

However, there is still a need to extend these rules for an *m*-MSFFSP design and is addressed here. The basis of these rules is that "pure" interactions get precedence over other interaction terms and "pure" interactions of stage *i* gets precedence over another "pure" interaction term in stage *k*, such that $k > i$. Specific rules for the three stage MSFFSP design of Example 4 are as follows.

i. Stage 1 main effects and interactions involving only stage 1 factors (and their aliases) are tested against stage 1 error.

   e.g., main effects *A*, *B*, and *C* are tested against stage 1 error.

ii. Stage 2 main effects or interactions that are aliased with stage 1 main effects or pure stage 1 interactions are tested against stage 1 error.

e.g., none in Example 4.

iii. Stage 3 main effects or interactions that are aliased with stage 1 main effects or pure stage 1 interactions are tested against stage 1 error.

e.g., none in Example 4.

iv. Stage 3 main effects or interactions that are aliased with stage 2 main effects or pure stage 2 interactions are tested against stage 2 error.

e.g., none in Example 4.

v. Stage 2 main effects and interactions involving at least one stage 2 factor and no stage 3 factors that are not aliased with either of the following effects are tested against stage 2 error.

- stage 1 main effects
- pure stage 1 interactions

e.g., $ABCQR$ (= P), $ABCPR$ (= R), $ABCPQ$ (= R), AP, AQ, AR, BP, BQ, BR, CP, CQ, and CR are tested against stage 2 error.

vi. Stage 3 main effects and interactions involving at least one stage 3 factor that are not aliased with any of the following effects are tested against stage 3 error.

- stage 1 main effects
- pure stage 1 interactions
- stage 2 main effects
- pure stage 2 interactions
- interaction terms involving stages 1 and 2

e.g., *M*, *N*, *O*, *PM*, *PN*, *PO*, *QM*, *QN*, *QO*, *RM*, *RN*, *RO*, *AM*, *AN*, *AO*, *BM*, *BN*, *BO*, *CM*, *CN*, *CO*, *AQM*, *AQN*, *AQO*, *ARM*, *ARN*, *ARO*, *APM*, *APN*, *APO*, *BPM*, *BPN*, *BPO*, *BQM*, *BQN*, *BQO*, *BRM*, *BRN*, *BRO*, *CPM*, *CPN*, *CPO*, *CQM*, *CQN*, *CQO*, *CRM*, *CRN*, and *CRO* are tested against stage 3 error.

These rules can be generalized for the case of an *m*-MSFFSP design as follows.

Rule 1. Stage 1 main effects and interactions involving only stage 1 factors (and their aliases) are tested against stage 1 error (St_{1_E}).

Rule 2. Stage *i* main effects or interactions that are aliased with stage *j* main effects or pure stage *j* interactions (for every $i = 2,..., m$, $j = 1,.., i\text{-}1$) are tested against the error term of stage *j* (St_{j_E}).

Rule 3. Stage *i* main effects and interactions involving at least one stage *i* factor and no stage *k* factors that are aliased with stage *j* main effects, pure stage *j* interactions, or interactions involving stage *j* and stage (*j*-1) factors are tested against the stage *i* error term (St_{i_E}), where $k > i$ and for every $i = 2, ... , m$; $j = 1,...., i\text{-}1$.

When these rules are applied to an *m*-MSFFSP design, regardless of the type on confounding (confounding within plots or split-plot confounding), it is observed that the number of contrasts to be tested against the error term in stage *i* is represented by

$$C_i = 2^{(k_1 - p_1)} \times 2^{(k_2 - p_2)} \times ... \times 2^{(k_{-1i} - p_{i-1})} [2^{(k_i - p_i)} - 1],$$

$$C_i = \left[\prod_{j=1}^{i} 2^{(k_j - p_j)}\right] \times \left[2^{(k_i - p_i)} - 1\right]. \qquad \textbf{(5.20)}$$

These contrasts correspond to the sum of the following contrasts

a) Contrasts generated by multiplying out the $(k_i - p_i)$ basic generators in all possible ways, and

b) Contrasts in (a) multiplied by similar contrasts from previous (1, 2,..., i-1) stages.

For Example 4, the number of contrasts to be tested against stage 1 error is given by

$$\begin{aligned} C_1 &= 2^{k_1 - p_1} - 1 \\ &= 2^{3-1} - 1 \quad . \\ &= 3 \end{aligned}$$

These three contrasts correspond to the three effects mentioned under rule (i) for Example 4. Similarly, the number of contrasts to be tested against stage 2 error is given by

$$\begin{aligned} C_2 &= 2^{k_1 - p_1}(2^{k_2 - p_2} - 1) \\ &= 2^{3-1}(2^{3-1} - 1) \quad . \\ &= 12 \end{aligned}$$

These twelve contrasts correspond to the three effects mentioned under rule (v). Finally, the number of contrasts to be tested against stage 3 error is given by

$$
\begin{aligned}
C_3 &= 2^{k_1-p_1} \times 2^{k_2-p_2}(2^{k_3-p_3}-1) \\
&= 2^{3-1} \times 2^{3-1}(2^{3-1}-1) \\
&= 48
\end{aligned}
\quad .
$$

These forty-eight contrasts correspond to the three effects mentioned under rule (vi).

5.3.3 Relative Efficiency of MSSP designs

Box and Jones (1992) compared the efficiency of two arrangements of split-plot design with that of a completely randomized design (CRD) and a strip-plot design. The estimates of error variance for each design were calculated and compared, assuming an uniformity trial. Kuehl (2000) explains the uniformity trial as an experiment in which the experimental units are measures, but have not been subjected to any treatments. In order to compare the efficiency of the MSSP designs to a CRD, the degrees of freedom for all the error terms need to be examined under the uniformity trial.

Example 5

Consider a replicated 3-MSSP design with factors *A*, *B*, and *C* corresponding to stages 1, 2, and 3, respectively. A fixed-effects linear model for this design can be represented by

$$
\begin{aligned}
y_{ijkl} = \mu + R_i + A_j + (St_{1_E})_{ij} + B_k + (AB)_{jk} + (St_{2_E})_{ijk} + \\
C_l + (AC)_{jl} + (BC)_{kl} + (ABC)_{jkl} + (St_{3_E})_{ijkl}
\end{aligned}
\qquad \textbf{(5.21)}
$$

where

μ is the overall mean,

a represents the levels of factor A; $j = 1, ..., a$,

b represents the levels of factor B; $k = 1, ..., b$,

c represents the levels of factor C; $l = 1, ..., c$,

r represents the number of replications; $i = 1, ..., r$,

(St_{1_E}) represents the error of stage 1 and $\sim N(0, \sigma_1^2)$,

(St_{2_E}) represents the error of stage 2 and $\sim N(0, \sigma_2^2)$, and

(St_{3_E}) represents the error of stage 3 and $\sim N(0, \sigma_3^2)$.

The effect $(RA)_{ij}$ represents the stage 1 error and effects $(RB)_{ik}$ and $(RAB)_{ijk}$ are pooled to form the error for stage 2. Similarly, effects $(RC)_{il}$, $(RAC)_{ijl}$, $(RBC)_{ikl}$, and $(RABC)_{ijkl}$ are pooled to form the error for stage 3. The degrees of freedom and the EMS of the error terms are shown in Table 5.6.

Table 5.6: Partial ANOVA Table for Example 5

Source	d.o.f.	EMS
R_i	(r-1)	
A_j	(a-1)	
St_{1e}	(r-1)(a-1)	E_{St1}
B_k	(b-1)	
AB_{jk}	(a-1)(b-1)	
$(St_{2e})_{ijk}$	a(r-1)(b-1)	E_{St2}
C_l	(c-1)	
AC_{jl}	(a-1)(c-1)	
BC_{kl}	(b-1)(c-1)	
ABC_{jkl}	(a-1)(b-1)(c-1)	
$(St_{3e})_{ijkl}$	ab(r-1)(c-1)	E_{St3}
Total	abcr-1	

Considering the uniformity trial, the degrees of freedom for stage 1 error is given by

$$dof(St_{1_E}) = dof(R) + dof(A) + dof(RA) = ar - 1. \quad \textbf{(5.22)}$$

The degrees of freedom for stage 2 error term is given by

$$dof(St_{2_E}) = dof(B) + dof(AB) + dof(RB) + dof(\mathrm{R}AB) = ar(b-1). \quad \textbf{(5.23)}$$

Finally, degrees of freedom for stage 3 error term is given by

$$\begin{aligned} dof(St_{3_E}) &= dof(C) + dof(AC) + dof(\mathrm{B}C) + dof(ABC) + dof(RC) + dof(\mathrm{R}AC) \\ &\quad + dof(RBC) + dof(\mathrm{R}ABC) \\ &= abr(c-1) \end{aligned} \quad . \textbf{(5.24)}$$

In addition, as a CRD has a single error term, the degrees of freedom for the error term, E_{CR}, equals the total degrees of freedom under uniformity trial. Hence,

$$dof(E_{CR}) = abcr - 1. \quad \textbf{(5.25)}$$

Therefore, if the uniformity trial were run as a CRD, an estimate of error variance for this design would be

$$(abcr-1)E_{CR} = (ar-1)E_{\mathrm{St}_1} + ar(b-1)E_{\mathrm{St}_2} + abr(c-1)E_{\mathrm{St}_3}. \quad \textbf{(5.26)}$$

For an unreplicated 2-level factorial design, $r = 1$, $a = b = c = 2$, and hence Equation 5.26 can be written as

$$7E_{CR} = E_{\mathrm{St}_1} + 2E_{\mathrm{St}_2} + 4E_{\mathrm{St}_3}\,. \tag{5.27}$$

As the left-hand side of Equation 5.27 is a weighted sum of the right-hand side, and since $E_{\mathrm{St}_1} > E_{\mathrm{St}_2} > E_{\mathrm{St}_3}$,

$$E_{\mathrm{St}_1} > E_{CR} > E_{\mathrm{St}_3}\,. \tag{5.28}$$

This implies that effects in stage 1 are less precisely estimated and effects in stage 3 are more precisely estimated as compared to effects in a CRD. Due to the nature of Equation 5.27, such a straightforward conclusion cannot be drawn on effects in stage 2. Depending on the ratios of the error terms of the stages, effects can be more or less precisely estimated. Let us consider an extreme case where the ratios of the errors are at least 7 (as 7 is the largest number in Equation 5.27).

$$\frac{E_{St_1}}{E_{St_2}} \geq 7 \Rightarrow E_{St_1} \geq 7E_{St_2}\,. \tag{5.29}$$

$$\frac{E_{St_2}}{E_{St_3}} \geq 7 \Rightarrow E_{St_2} \geq 7E_{St_3}\,. \tag{5.30}$$

Considering the lower bounds, Equations 5.27, 5.29, and 5.30 can be re-written as

$$E_{St_1} + 2E_{St_2} + 4E_{St_3} = 7E_{CR}. \quad \textbf{(5.31)}$$

$$E_{St_1} - 7E_{St_2} = 0. \quad \textbf{(5.32)}$$

$$E_{St_2} - 7E_{St_3} = 0. \quad \textbf{(5.33)}$$

Summing (7 × Equation 5.31) and (4 ×Equation 5.33) yields

$$7E_{St_1} + 18E_{St_2} = 49E_{CR}. \quad \textbf{(5.34)}$$

Summing Equation 5.34 and (7 × Equation 5.32) yields

$$67E_{St_2} = 49E_{CR} \Rightarrow \frac{67}{49}E_{St_2} = E_{CR}. \quad \textbf{(5.35)}$$

As 67/49 > 1,

$$E_{CR} > E_{\mathrm{St}_3}. \quad \textbf{(5.36)}$$

Equation 5.36 implies that effects in stage 2 are determined more precisely than in a CRD. Furthermore, the only definite extension to an *m*-MSSP would be

$$E_{\mathrm{St}_1} > E_{CR} > E_{\mathrm{St}_m}. \quad \textbf{(5.37)}$$

Depending on the ratios $\frac{E_{St_i}}{E_{St_{(i+1)}}}$ (i = 1,..., m-1), effects in the remaining stages are more or less precisely estimated than in a CRD. When the efficiency of m-MSSP design is compared to that of an m-stage split-lot design, it is easy to see that the effects in stage 1 of both designs would be determined with equal precision. However, the precision in estimating effects from other stages would highly depend on the ratios of adjacent stages in both the designs.

5.4 Analysis of MSFFSP Designs

Once the design has been finalized and the experiment executed, the next step is to calculate the effects of the factors and test them for significance. An ANOVA table gives a clear idea on the calculation of the effects in the model and the partitioning of the total variability in the model into its component parts. By observing the expected mean squares in the ANOVA table, various hypotheses that can be tested are identified. Let us look at the decomposition of the total sum of squares for an example.

Consider the model given in Example 5. Following the rules for obtaining the expected mean squares (EMS) from Montgomery (2005), Table 5.7 represents the EMS and degrees of freedom (dof) for the model in Example 5. It can be seen that the hypothesis of the means of the main effects and interactions between factors A, B, and C can be tested for significance against the appropriate error terms.

Table 5.7: EMS Table for Example 5

	r R i	a F j	b F k	c F l	dof	EMS
Source						
A_j	r	0	b	c	*(a-1)*	$\frac{rbc\sum A_j^2}{(a-1)} + bc\sigma_1^2 + c\sigma_2^2 + \sigma_3^2$
$(St1_E)_{ij}$	1	1	b	c	*(a-1)(r-1)*	$bc\sigma_1^2 + c\sigma_2^2 + \sigma_3^2$
B_k	r	a	0	c	*(b-1)*	$\frac{rac\sum B_k^2}{(b-1)} + c\sigma_2^2 + \sigma_3^2$
$(AB)_{jk}$	r	0	0	c	*(a-1)(b-1)*	$\frac{rc\sum AB_{jk}^2}{(a-1)(b-1)} + c\sigma_2^2 + \sigma_3^2$
$(St2_E)_{i(jk)}$	1	1	1	c	*a(r-1)(b-1)*	$c\sigma_2^2 + \sigma_3^2$
C_l	r	a	b	0	*(c-1)*	$\frac{rab\sum C_l^2}{(c-1)} + \sigma_3^2$
$(AC)_{jl}$	r	0	b	0	*(a-1)(c-1)*	$\frac{rb\sum AC_{jl}^2}{(a-1)(c-1)} + \sigma_3^2$
$(BC)_{kl}$	r	a	0	0	*(b-1)(c-1)*	$\frac{ra\sum BC_{kl}^2}{(b-1)(c-1)} + \sigma_3^2$
$(ABC)_{jkl}$	r	0	0	0	*(a-1)(b-1)(c-1)*	$\frac{r\sum ABC_{jkl}^2}{(a-1)(b-1)(c-1)} + \sigma_3^2$
$(St3_E)_{i(jkl)}$	1	1	1	1	*ab(r-1)(c-1)*	σ_3^2

Table 5.7 can be extended to the general case of m stages as follows. Let r be the number of replicates and K be an effect (main effect or a multi-way interaction effect) in stage k (k = 1,...., m). Let d^i be the degrees of freedom of the i^{th} effect and σ_i^2 be the variance of the error term for stage i. The expected mean square of the K^{th} fixed effect is given by

$$EMS(K) = \frac{r\prod_{i\neq k} d^i \sum K^2}{d^k - 1} + \sum_{j=k+1}^{m} (d^j d^{j+1} ... d^m)\sigma_{j-1}^2 + \sigma_m^2 . \quad \textbf{(5.38)}$$

The contrasts estimates for the effects can be calculated in the conventional way of subtracting the responses with "-" sign from the responses corresponding to a "+" sign. Once the contrasts for all effects have been obtained, they are tested for significance against appropriate error terms using the rules described earlier or from the EMS table. If the design in unreplicated, no error terms exist. One way to compensate for the loss of degrees of freedom for the error terms would be to pool higher-order interactions to create appropriate error terms. An alternate way would be to use normal probability plots for each error term to draw inferences on the significance on the effects. The number of normal probability plots equal the number of error terms in the model.

If there are random factors present in the model, the EMS table is modified according to the rules for EMS given by Montgomery (2005). The corresponding quadratic means effect would be replaced by the appropriate variance component for the random effect.

5.5 Catalog of MSFFSP Designs

Fractionating in a fractional factorial split-plot (FFSP) design can be classified as confounding within plots and split-plot confounding. In fractionating within plots, generators are chosen from the same plot. The highlight of this type of confounding is that the interactions between the whole plot and sub-plots are free of confounding if three or higher-order interactions are assumed negligible. However, the main effects of the

whole plot and sub-plots are confounded with two-way interactions. This type of design is useful in a robustness study, where the interaction effects between the whole plot and sub-plot are of main interest, e.g. a study of environmental effects on design factors. However, if the objective of the experiment is to estimate the main effects, this design will not be very helpful.

In split-plot confounding, the generators of the sub-plot involve the whole plot factors. In this case, the main effects of the sub-plot factors are free of confounding, whereas some of the two-way interactions terms are confounded with other two-way interactions. Due to the alias structure, this type of confounding is more often used in a setting where the sub-plot contains the design factors of interest. This design is not very useful, however, in robustness study experiments. Another striking difference between the two types of confounding is that even thought the overall resolution remains the same, the partial resolution of the sub-plot may increase by one. Bisgaard (2000) points out that, in general, split-plot confounding possibly increases the partial resolution of the sub-plot by at least one.

An interesting extension for an *m*-MSFFSP design is that a combination of the two types of confounding can be used to obtain a design having maximum number of clear (free of aliasing with higher-order interactions) effects. Depending on the process or experiment being investigated, these effects of interest can be either main effects of certain stages or interaction effects between stages. For instance in a three stage MSFFSP design, if the interaction effects between stages 1 and 2 along with the main effects of stage 3 are of primary interest, then split-plot confounding can be used to generate columns in stage 2, and confounding within plots can be used to generate columns in

stage 3. This advantage is valuable in constructing designs for processes that are conducted over multiple stages.

Given a set of experimental constraints, e.g., maximum number of runs, maximum allowable setting changes, number of hard-to-change factors, many MSFFSP can be designed. The final choice of which design to choose is to be made by the experimenter, in consultation with the designer. The objective of the experiment dictates the design to be chosen.

Different criteria have been identified in literature based on which optimal designs have been proposed. Traditionally, the idea of resolution, first introduced by Box and Hunter (1961) has been used extensively in ranking deigns having the same number of factors and number of runs. Resolution can be described as the length of the shortest word in the defining relation. Typically, higher the resolution, better the design, as more information can be obtained. Fries and Hunter (1980) noted that designs having the same resolution could have a considerable difference in the ability to estimate effects and hence introduced the concept of minimum aberration. This criterion is an enhancement of the resolution criterion and seeks a design with the best possible word length pattern $W = (A_1, A_2, ...)$ where A_i is the number of words of length i in the defining relation. A design, D, is said to have minimum aberration if there is no other design, D_0, for which, for some integer r, $A_k(D_0) = A_k(D)$ for $k < r$ and $A_r(D_0) < A_r(D)$. Recently, Bingham and Sitter (1999, 2001), Huang et al. (1998), and Mukerjee and Fang (2000) provided a catalog of designs based on the minimum aberration criterion.

A typical nanomanufacturing investigation involves many factors and far less resources (time and money) to conduct a comprehensive examination of the factors of

interest. Moreover, in many situations, theoretical properties provide prior knowledge on the behavior of some factors. Chen et al. (1993) point out certain situations in which designs with minimum aberration do not provide the best design in terms of the ability to estimate effects of interest. For this reason, alternate criteria based on the common occurrences in nanomanufacturing experimentation are proposed. As reproducibility is a major concern at the nano-scale, most of the studies are directed towards identifying factors that cause the same experimental setup to yield different responses. These experiments can be categorized as classic robustness studies.

Keeping the nature of the experiments in mind, a catalog of *m*-MSFFSP designs is created and ranked according to the following criteria.

a) Robustness

This criterion refers to maximum number of clear two-way effects between stages. Only two-way interaction effects between stages are of interest here. The objective here is to test the effect of one factor on the response, given the levels of another factor. In this criterion, three- and higher-order interaction terms are assumed negligible.

b) Maximum number of "mixed" three-way interactions

In some cases, it is essential to consider few three-way interactions as well. For instance, if we are interested in examining the effect of one factor from stage 1 on the interaction effect of two other factors from stages 2 and 3, a three-way interaction term

needs to be considered. In this criterion, pure three-way interactions, and four- and higher-order interaction terms are assumed negligible.

These criteria, under no means, represent a complete set and should not be used if the objective of the study is different from the ones described above. However, similar criteria can be developed depending on the nature of the experiment.

As the motivation for these constructs of design originated from a robustness study, we provide a catalog of designs for similar experiments. Confounding within plots was used for all the designs in the catalog, as the effects of interest are second-order and certain third-order interactions. If a main effect is also an effect of interest to the researcher, then the "mixed" confounding described above needs to be considered.

All *m*-MSFFSP designs constructed using confounding within plots can be viewed as a series of fractional factorial designs. The catalog was hence created using the first part of the algorithm developed by Russell et al. (2004) for single stage designs. Designs are provided for MSFFSP experiments in three stages involving 10 and 11 factors as well as four stages involving 13 factors as they yield a manageable proportion of possible MSFFSP designs. As the designs are cataloged merely to illustrate the use of the optimality criteria, only a limited number of designs are considered. Designs having two factors in more than one stage are ignored, and the maximum number of factors in a stage is limited to seven. Furthermore, in any given stage, designs with maximum resolution are considered to allow the possibility of estimating main effects and pure interactions. Only designs having partial resolution of above II are considered. Finally,

designs having 32 and 64 runs are considered for three-stage designs and 64 and 128 runs for four-stage designs, as the more number of runs are generally not feasible.

Table 5.8 represents the fractional factorial designs, having different confounding patterns, used to create MSFFSP designs (under confounding within plots). The columns in each design are labeled 1 through 7. For every design, only defining relations yielding maximum resolutions are considered.

Table 5.8: FF Designs used for Catalog Creation

Number	Design	Defining relation (excluding generalized interactions)	Resolution
1	2^{2-1}	*I = 12*	II
2	2^{3-1}	*I = 123*	III
3	2^{3-2}	*I = 12 = 13*	II
4	2^{4-1}	*I = 1234*	III
5	2^{4-2}	*I = 12 = 13*	II
6		*I = 12 = 34*	II
7		*I = 12 = 134*	II
8	2^{5-1}	*I = 12345*	V
9	2^{5-2}	*I = 123 = 145*	III
10	2^{5-3}	*I = 12 = 13 = 14*	II
11		*I = 12 = 13 = 45*	II
12		*I = 12 = 13 = 145*	II
13		*I = 12 = 34 = 135*	II
14	2^{6-2}	*I = 1234 = 1256*	IV
15	2^{6-3}	*I = 123 = 145 = 246*	III
16	2^{6-4}	*I = 12 = 13 =14 = 15*	II
17		*I = 12 = 13 = 14 = 56*	II
18		*I = 12 = 13 = 14 = 156*	II
19		*I = 12 = 13 = 45 = 46*	II
20		*I = 12 = 13 = 45 = 146*	II
21		*I = 12 = 34 = 135 = 136*	II
22	2^{7-3}	*I = 1234 = 1256 = 1357*	IV
23	2^{7-4}	*I = 123 = 145 = 246 = 1247*	III

Using the information from Table 5.8, catalogs of three-stage MSFFSP designs having 10 and 11 factors are created and are listed in Table 5.9 and Table 5.10,

respectively. In these tables, n_1, n_2, and n_3 represent the number of factors in stages 1, 2, and 3, respectively, and k_1, k_2, and k_3 represent the number of generators for stages 1, 2, and 3, respectively. The tables also list the number of factors that are estimable under each of the two aforementioned criteria. For the designs listed in the catalogs, a maximum of 33 effects can be estimated under criterion (a) and 36 effects under criterion (b). In designs 9-18, a maximum of 39 effects can be estimated under criterion (a) and 45 effects under criterion (b). Finally, in designs 19-39, a maximum of 40 effects can be estimated under criterion (a) and 48 effects under criterion (b)

Table 5.9: Catalog of Three-Stage MSFFSP having 10 Factors

Design Number	Number of factors in			Degree of fractionation in			Number of runs	Overall resolution	Criterion a (max 33)	Criterion b (max 36)
	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3				
	n_1	n_2	n_3	k_1	k_2	k_3				
1	3	3	4	1	1	2	64	II	15	9
2	3	3	4	1	1	2	64	II	9	0
3	3	3	4	1	1	2	64	II	21	18
4	3	3	4	2	1	1	64	II	12	0
5	3	3	4	2	1	2	32	II	3	0
6	3	3	4	2	1	2	32	II	0	0
7	3	3	4	2	1	2	32	II	6	0

Table 5.10: Catalog of Three-Stage MSFFSP having 11 Factors

Design Number	Number of factors in			Degree of fractionation in			Number of runs	Overall resolution	Criterion a	Criterion b
	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3				
	n_1	n_2	n_3	k_1	k_2	k_3				
8	3	3	5	1	1	3	64	II	14	9
9	3	3	5	1	1	3	64	II	8	0
10	3	3	5	1	1	3	64	II	8	0
11	3	3	5	1	1	3	64	II	14	9
12	3	3	5	2	1	2	64	II	15	0
13	3	3	5	2	2	1	64	II	15	0
14	3	3	5	2	1	3	32	II	3	0
15	3	3	5	2	1	3	32	II	3	0
16	3	3	5	2	1	3	32	II	3	0
17	3	3	5	2	1	3	32	II	3	0
18	3	3	5	2	2	2	32	II	3	0
19	3	4	4	1	2	2	64	II	7	3
20	3	4	4	1	2	2	64	II	7	3
21	3	4	4	1	2	2	64	II	7	3
22	3	4	4	1	2	2	64	II	7	3
23	3	4	4	1	2	2	64	II	7	3
24	3	4	4	1	2	2	64	II	7	3
25	3	4	4	1	2	2	64	II	7	3
26	3	4	4	1	2	2	64	II	7	3
27	3	4	4	1	2	2	64	II	7	3
28	3	4	4	2	1	2	64	II	7	3
29	3	4	4	2	1	2	64	II	7	3
30	3	4	4	2	1	2	64	II	7	3
31	3	4	4	2	2	2	32	II	1	0
32	3	4	4	2	2	2	32	II	1	0
33	3	4	4	2	2	2	32	II	1	0
34	3	4	4	2	2	2	32	II	1	0
35	3	4	4	2	2	2	32	II	1	0
36	3	4	4	2	2	2	32	II	1	0
37	3	4	4	2	2	2	32	II	1	0
38	3	4	4	2	2	2	32	II	1	0
39	3	4	4	2	2	2	32	II	1	0

The defining relations for the designs listed in Table 5.9 and Table 5.10 are listed in Table 5.11. *A*, *B*, *C*, *D*, *E*, *F*, and *G* correspond to the factors in stage 1; *M*, *N*, *O*, *P*, *Q*, *R*, and *S* correspond to factors in stage 2; and *T*, *U*, *V*, *W*, *X*, *Y*, and Z correspond to factors in stage 3. The catalogs are aimed at assisting an experimenter is choosing the "best" design based on the optimality criteria.

Table 5.11: Defining Relations for the Three-Stage Deigns in the Catalog

Design Number	Defining Relation
1	(1) = ABC = MNO = ABCMNO = TU = ABCTU = MNOTU = ABCMNOTU = TV = ABCTV = MNOTV = ABCMNOTV = UV = ABCUV = MNOUV = ABCMNOUV
2	(1) = ABC = MNO = ABCMNO = TU = ABCTU = MNOTU = ABCMNOTU = VW = ABCVW = MNOVW = ABCMNOVW = TUVW = ABCTUVW = MNOTUVW = ABCMNOTUVW
3	(1) = ABC = MNO = ABCMNO = TU = ABCTU = MNOTU = ABCMNOTU = TVW = ABCTVW = MNOTVW = ABCMNOTVW = UVW = ABCUVW = MNOUVW = ABCMNOUVW
4	(1) = AB = AC = BC = MNO = ABMNO = ACMNO = BCMNO = TUVW = ABTUVW = ACTUVW = BCTUVW = MNOTUVW = ABMNOTUVW = ACMNOTUVW = BCMNOTUVW
5	(1) = AB = AC = BC = MNO = ABMNO = ACMNO = BCMNO = TU = ABTU = ACTU = BCTU = MNOTU = ABMNOTU = ACMNOTU = BCMNOTU = TV = ABTV = ACTV = BCTV = MNOTV = ABMNOTV = ACMNOTV = BCMNOTV = UV = ABUV = ACUV = BCUV = MNOUV = ABMNOUV = ACMNOUV = BCMNOUV
6	(1) = AB = AC = BC = MNO = ABMNO = ACMNO = BCMNO = TU = ABTU = ACTU = BCTU = MNOTU = ABMNOTU = ACMNOTU = BCMNOTU = VW = ABVW = ACVW = BCVW = MNOVW = ABMNOVW = ACMNOVW = BCMNOVW = TUVW = ABTUVW = ACTUVW = BCTUVW = MNOTUVW = ABMNOTUVW = ACMNOTUVW = BCMNOTUV
7	(1) = AB = AC = BC = MNO = ABMNO = ACMNO = BCMNO = TU = ABTU = ACTU = BCTU = MNOTU = ABMNOTU = ACMNOTU = BCMNOTU = TVW = ABTVW = ACTVW = BCTVW = MNOTVW = ABMNOTVW = ACMNOTVW = BCMNOTVW = UVW = ABUVW = ACUVW = BCUVW = MNOUVW = ABMNOUVW = ACMNOUVW = BCMNOUV
8	I = TU = TV = TW = UV = UW = VW = ABC = MNO = TUVW = ABCTU = ABCTV = ABCTW = ABCUV = ABCUW = ABCVW = MNOTU = MNOTV = NOTW = MNOUV = MNOUW = MNOVW = ABCMNO = ABCTUVW = MNOTUVW = ABCMNOTU = ABCMNOTV = ABCMNOTW = ABCMNOUV = ABCMNOUW = ABCMNOVW = ABCMNOTUVW
9	I = TU = TV = UV = WX = ABC = MNO = TUWX = TVWX = UVWX = ABCTU = ABCTV = ABCUV = ABCWX = MNOTU = MNOTV = MNOUV = MNOWX = ABCMNO = ABCTUWX = ABCTVWX = ABCUVWX = MNOTUWX = MNOTVWX = MNOUVWX = ABCMNOTU = ABCMNOTV = ABCMNOUV = ABCMNOWX = ABCMNOTUWX = ABCMNOTV
10	I = TU = TV = UV = ABC = MNO = TWX = UWX = VWX = ABCTU = ABCTV = ABCUV = MNOTU = MNOTV = MNOUV = TUVWX = ABCMNO = ABCTWX = ABCUWX = ABCVWX = MNOTWX = MNOUWX = MNOVWX = ABCMNOTU = ABCMNOTV = ABCMNOUV =
11	I = TU = VW = ABC = MNO = TVX = TWX = UVX = UWX = TUVW = ABCTU = ABCVW = MNOTU = MNOVW = ABCMNO = ABCTVX = ABCTWX = ABCUVX = ABCUWX = MNOTVX = MNOTWX = MNOUVX = MNOUWX = ABCTUVW = MNOTUVW = ABCMNOTU = ABCMNOVW = ABCMNOTVX = ABCMNOTWX = ABCMNOUVX = ABCMNOU
12	I = AB = AC = BC = MNO = TUV = TWX = UVWX = ABMNO = ABTUV = ABTWX = ACMNO = ACTUV = ACTWX = BCMNO = BCTUV = BCTWX = ABUVWX = ACUVWX = BCUVWX = MNOTUV = MNOTWX = MNOUVWX = ABMNOTUV = ABMNOTWX = ACMNOTUV = ACMNOTWX = BCMNOTUV = BCMNOTWX = ABMNOUVWX = ACMNO
13	I = AB = AC = BC = MN = MO = NO = ABMN = ABMO = ABNO = ACMN = ACMO = ACNO = BCMN = BCMO = BCNO = TUVWX = ABTUVWX = ACTUVWX = BCTUVWX = MNTUVWX = MOTUVWX = NOTUVWX = ABMNTUVWX = ABMOTUVWX = ABNOTUVWX = ACMNTUVWX = ACMOTUVWX = ACNOTUVWX = BCMNTUVWX = BCMOT

14	I = AB = AC = BC = TU = TV = TW = UV = UW = VW = MNO = ABTU = ABTV = ABTW = ABUV = ABUW = ABVW = ACTU = ACTV = ACTW = ACUV = ACUW = ACVW = BCTU = BCTV = BCTW = BCUV = BCUW = BCVW = TUVW = ABMNO = ACMNO = BCMNO = MNOTU = MNOTV = MNOTW = MNOUV = MNOUW = MNO
15	I = AB = AC = BC = TU = TV = UV = WX = MNO = ABTU = ABTV = ABUV = ABWX = ACTU = ACTV = ACUV = ACWX = BCTU = BCTV = BCUV = BCWX = TUWX = TVWX = UVWX = ABMNO = ACMNO = BCMNO = MNOTU = MNOTV = MNOUV = MNOWX = ABTUWX = ABTVWX = ABUVWX = ACTUWX = ACTVWX = ACUV
16	I = AB = AC = BC = TU = TV = UV = MNO = TWX = UWX = VWX = ABTU = ABTV = ABUV = ACTU = ACTV = ACUV = BCTU = BCTV = BCUV = ABMNO = ABTWX = ABUWX = ABVWX = ACMNO = ACTWX = ACUWX = ACVWX = BCMNO = BCTWX = BCUWX = BCVWX = MNOTU = MNOTV = MNOUV = TUVWX = MNOTWX
17	I = AB = AC = BC = TU = VW = MNO = TVX = TWX = UVX = UWX = ABTU = ABVW = ACTU = ACVW = BCTU = BCVW = TUVW = ABMNO = ABTVX = ABTWX = ABUVX = ABUWX = ACMNO = ACTVX = ACTWX = ACUVX = ACUWX = BCMNO = BCTVX = BCTWX = BCUVX = BCUWX = MNOTU = MNOVW = ABTUVW = AC
18	I = AB = AC = BC = MN = MO = NO = TUV = TWX = ABMN = ABMO = ABNO = ACMN = ACMO = ACNO = BCMN = BCMO = BCNO = UVWX = ABTUV = ABTWX = ACTUV = ACTWX = BCTUV = BCTWX = MNTUV = MNTWX = MOTUV = MOTWX = NOTUV = NOTWX = ABUVWX = ACUVWX = BCUVWX = MNUVWX = MOUVWX
19	I = MN = MO = NO = TU = TV = UV = ABC = MNTU = MNTV = MNUV = MOTU = MOTV = MOUV = NOTU = NOTV = NOUV = ABCMN = ABCMO = ABCNO = ABCTU = ABCTV = ABCUV = ABCMNTU = ABCMNTV = ABCMNUV = ABCMOTU = ABCMOTV = ABCMOUV = ABCNOTU = ABCNOTV = ABCNOUV
20	I = MN = MO = NO = TU = VW = ABC = MNTU = MNVW = MOTU = MOVW = NOTU = NOVW = TUVW = ABCMN = ABCMO = ABCNO = ABCTU = ABCVW = MNTUVW = MOTUVW = NOTUVW = ABCMNTU = ABCMNVW = ABCMOTU = ABCMOVW = ABCNOTU = ABCNOVW = ABCTUVW = ABCMNTUVW = ABCMOTUVW = ABCNOTUVW
21	I = MN = MO = NO = TU = ABC = TVW = UVW = MNTU = MOTU = NOTU = ABCMN = ABCMO = ABCNO = ABCTU = MNTVW = MNUVW = MOTVW = MOUVW = NOTVW = NOUVW = ABCTVW = ABCUVW = ABCMNTU = ABCMOTU = ABCNOTU = ABCMNTVW = ABCMNUVW = ABCMOTVW = ABCMOUVW = ABCNOTVW = ABCNOUVW
22	I = MN = OP = TU = TV = UV = ABC = MNOP = MNTU = MNTV = MNUV = OPTU = OPTV = OPUV = ABCMN = ABCOP = ABCTU = ABCTV = ABCUV = MNOPTU = MNOPTV = MNOPUV = ABCMNOP = ABCMNTU = ABCMNTV = ABCMNUV = ABCOPTU = ABCOPTV = ABCOPUV = ABCMNOPTU = ABCMNOPTV = ABCMNOPUV
23	I = MN = OP = TU = VW = ABC = MNOP = MNTU = MNVW = OPTU = OPVW = TUVW = ABCMN = ABCOP = ABCTU = ABCVW = MNOPTU = MNOPVW = MNTUVW = OPTUVW = ABCMNOP = ABCMNTU = ABCMNVW = ABCOPTU = ABCOPVW = ABCTUVW = MNOPTUVW = ABCMNOPTU = ABCMNOPVW = ABCMNTUVW = ABCOPTUV
24	I = MN = OP = TU = ABC = TVW = UVW = MNOP = MNTU = OPTU = ABCMN = ABCOP = ABCTU = MNTVW = MNUVW = OPTVW = OPUVW = ABCTVW = ABCUVW = MNOPTU = ABCMNOP = ABCMNTU = ABCOPTU = MNOPTVW = MNOPUVW = ABCMNTVW = ABCMNUVW = ABCOPTVW = ABCOPUVW = ABCMNOPTU = ABCMNOPT
25	I = MN = TU = TV = UV = ABC = MOP = NOP = MNTU = MNTV = MNUV = ABCMN = ABCTU = ABCTV = ABCUV = MOPTU = MOPTV = MOPUV = NOPTU = NOPTV = NOPUV = ABCMOP = ABCNOP = ABCMNTU = ABCMNTV = ABCMNUV = ABCMOPTU = ABCMOPTV = ABCMOPUV = ABCNOPTU = ABCNOPTV = ABCNOPUV
26	I = MN = TU = VW = ABC = MOP = NOP = MNTU = MNVW = TUVW = ABCMN = ABCTU = ABCVW = MOPTU = MOPVW = NOPTU = NOPVW = ABCMOP = ABCNOP = MNTUVW = ABCMNTU = ABCMNVW = ABCTUVW = MOPTUVW = NOPTUVW = ABCMOPTU = ABCMOPVW = ABCNOPTU = ABCNOPVW = ABCMNTUVW = ABCMOPTU

27	I = MN = TU = ABC = MOP = NOP = TVW = UVW = MNTU = ABCMN = ABCTU = MNTVW = MNUVW = MOPTU = NOPTU = ABCMOP = ABCNOP = ABCTVW = ABCUVW = MOPTVW = MOPUVW = NOPTVW = NOPUVW = ABCMNTU = ABCMNTVW = ABCMNUVW = ABCMOPTU = ABCNOPTU = ABCMOPTVW = ABCMOPUVW = ABCNOP
28	I = AB = AC = BC = TU = TV = UV = ABTU = ABTV = ABUV = ACTU = ACTV = ACUV = BCTU = BCTV = BCUV = MNOP = ABMNOP = ACMNOP = BCMNOP = MNOPTU = MNOPTV = MNOPUV = ABMNOPTU = ABMNOPTV = ABMNOPUV = ACMNOPTU = ACMNOPTV = ACMNOPUV = BCMNOPTU = BCMNOPTV = BCMNOPUV
29	I = AB = AC = BC = TU = VW = ABTU = ABVW = ACTU = ACVW = BCTU = BCVW = MNOP = TUVW = ABMNOP = ABTUVW = ACMNOP = ACTUVW = BCMNOP = BCTUVW = MNOPTU = MNOPVW = ABMNOPTU = ABMNOPVW = ACMNOPTU = ACMNOPVW = BCMNOPTU = BCMNOPVW = MNOPTUVW = ABMNOPTUVW = ACMNOPTU
30	I = AB = AC = BC = TU = TVW = UVW = ABTU = ACTU = BCTU = MNOP = ABTVW = ABUVW = ACTVW = ACUVW = BCTVW = BCUVW = ABMNOP = ACMNOP = BCMNOP = MNOPTU = MNOPTVW = MNOPUVW = ABMNOPTU = ACMNOPTU = BCMNOPTU = ABMNOPTVW = ABMNOPUVW = ACMNOPTVW = ACMNOPUVW = BCMNOP
31	I = AB = AC = BC = MN = MO = NO = TU = TV = UV = ABMN = ABMO = ABNO = ABTU = ABTV = ABUV = ACMN = ACMO = ACNO = ACTU = ACTV = ACUV = BCMN = BCMO = BCNO = BCTU = BCTV = BCUV = MNTU = MNTV = MNUV = MOTU = MOTV = MOUV = NOTU = NOTV = NOUV = ABMNTU = ABMNTV =
32	I = AB = AC = BC = MN = MO = NO = TU = VW = ABMN = ABMO = ABNO = ABTU = ABVW = ACMN = ACMO = ACNO = ACTU = ACVW = BCMN = BCMO = BCNO = BCTU = BCVW = MNTU = MNVW = MOTU = MOVW = NOTU = NOVW = TUVW = ABMNTU = ABMNVW = ABMOTU = ABMOVW = ABNOTU = ABNOVW = AB
33	I = AB = AC = BC = MN = MO = NO = TU = TVW = UVW = ABMN = ABMO = ABNO = ABTU = ACMN = ACMO = ACNO = ACTU = BCMN = BCMO = BCNO = BCTU = MNTU = MOTU = NOTU = ABTVW = ABUVW = ACTVW = ACUVW = BCTVW = BCUVW = MNTVW = MNUVW = MOTVW = MOUVW = NOTVW = NOUVW = ABM
34	I = AB = AC = BC = MN = OP = TU = TV = UV = ABMN = ABOP = ABTU = ABTV = ABUV = ACMN = ACOP = ACTU = ACTV = ACUV = BCMN = BCOP = BCTU = BCTV = BCUV = MNOP = MNTU = MNTV = MNUV = OPTU = OPTV = OPUV = ABMNOP = ABMNTU = ABMNTV = ABMNUV = ABOPTU = ABOPTV = ABO
35	I = AB = AC = BC = MN = OP = TU = VW = ABMN = ABOP = ABTU = ABVW = ACMN = ACOP = ACTU = ACVW = BCMN = BCOP = BCTU = BCVW = MNOP = MNTU = MNVW = OPTU = OPVW = TUVW = ABMNOP = ABMNTU = ABMNVW = ABOPTU = ABOPVW = ABTUVW = ACMNOP = ACMNTU = ACMNVW = ACOPTU =
36	I = AB = AC = BC = MN = OP = TU = TVW = UVW = ABMN = ABOP = ABTU = ACMN = ACOP = ACTU = BCMN = BCOP = BCTU = MNOP = MNTU = OPTU = ABTVW = ABUVW = ACTVW = ACUVW = BCTVW = BCUVW = MNTVW = MNUVW = OPTVW = OPUVW = ABMNOP = ABMNTU = ABOPTU = ACMNOP = ACMNTU =
37	I = AB = AC = BC = MN = TU = TV = UV = MOP = NOP = ABMN = ABTU = ABTV = ABUV = ACMN = ACTU = ACTV = ACUV = BCMN = BCTU = BCTV = BCUV = MNTU = MNTV = MNUV = ABMOP = ABNOP = ACMOP = ACNOP = BCMOP = BCNOP = MOPTU = MOPTV = MOPUV = NOPTU = NOPTV = NOPUV = AB
38	I = AB = AC = BC = MN = TU = VW = MOP = NOP = ABMN = ABTU = ABVW = ACMN = ACTU = ACVW = BCMN = BCTU = BCVW = MNTU = MNVW = TUVW = ABMOP = ABNOP = ACMOP = ACNOP = BCMOP = BCNOP = MOPTU = MOPVW = NOPTU = NOPVW = ABMNTU = ABMNVW = ABTUVW = ACMNTU = ACMNVW =
39	I = AB = AC = BC = MN = TU = MOP = NOP = TVW = UVW = ABMN = ABTU = ACMN = ACTU = BCMN = BCTU = MNTU = ABMOP = ABNOP = ABTVW = ABUVW = ACMOP = ACNOP = ACTVW = ACUVW = BCMOP = BCNOP = BCTVW = BCUVW = MNTVW = MNUVW = MOPTU = NOPTU = ABMNTU = ACMNTU = BCMNTU

Catalogs of 4-stage designs with 13 factors are presented in Table 5.12. Factors in stages 1, 2, and 3 are represented in a similar fashion as in the 3-stage designs. In addition, α, β, γ, δ, and ε represent the factors in stage 4. In these tables, n_1, n_2, n_3, and n_4 represent the number of factors in stages 1, 2, 3, and 4, respectively, and k_1, k_2, k_3, and k_4 represent the number of generators for stages 1, 2, 3, and 4, respectively. A maximum of 63 effects are estimable under criterion (a) and 135 effects under criterion (b). The corresponding defining relations for the designs in Table 5.12 are given in Table 5.13.

Table 5.12: Catalog of Four-Stage MSFFSP having 13 Factors

Design Number	Number of factors in				Degree of fractionation in				Number of runs	Overall resolution	Criterion a (max 63)	Criterion b (max 135)
	Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4				
	n_1	n_2	n_3	n_4	k_1	k_2	k_3	k_4				
1	3	3	3	4	2	1	1	2	128	II	15	9
2	3	3	3	4	2	1	1	2	128	II	9	0
3	3	3	3	4	2	1	1	2	128	II	21	18
4	3	3	3	4	2	2	1	1	128	II	12	0
5	3	3	3	4	2	2	2	0	128	II	0	0
6	3	3	3	4	2	2	2	1	64	II	0	0
7	3	3	3	4	1	2	2	2	64	II	3	0
8	3	3	3	4	1	2	2	2	64	II	0	0
9	3	3	3	4	1	2	2	2	64	II	0	0

Table 5.13: Defining Relations for the Four-Stage Deigns in the Catalog

Design Number	Defining Relation
1	*I = A B = A C = M N O = T U V = α β = α γ = BC = ABMNO = ABTUV = ABαβ = ABαγ = ACMNO = ACTUV = AC*αβ *= ACαγ = MNOTUV = MNO*αβ *= MNOαγ = TUVαβ = TUVαγ =* βγ
2	*I = AB = AC = MNO = TUV = αβ = γδ = BC = ABMNO = ABTUV = ABαβ = ABγδ = ACMNO = ACTUV = ACαβ = ACγδ = MNOTUV = MNO*αβ *= MNO*γδ *= TUV*αβ *= TUV*γδ = αβγδ
3	*I = AB = AC = MNO = TUV = αβ = αγδ = BC = ABMNO = ABTUV = ABαβ = ABαγδ = ACMNO = ACTUV = AC*αβ *= ACαγδ = MNOTUV = MNO*αβ *= MNO*αγδ *= TUV*αβ *= TUV*αγδ *= βγδ*
4	*I = AB = AC = MN = MO = NO = TUV =* αβγδ *= BC = ABMN = ABMO - ABTUV = AB αβγδ = ACMN = ACMO = ACTUV = ACαβγδ = NO = MNTUV = MNαβγδ = MOTUV = MOαβγδ = TUVαβγδ*
5	*I = AB = AC = MN = MO = TU = TV = BC = ABMN = ABMO = ABTU = ABTV = ACMN = ACMO = ACTU = ACTV = NO = MNTU = MNTV = MOTU = MOTV = UV*
6	*I = AB = AC = MN = MO = TU = TV =*αβγδ *= BC = ABMN = ABMO = ABTU = ABTV = ABαβγδ = ACMN = ACMO = ACTU = ACTV = ACαβγδ = NO = MNTU = MNTV = MNαβγδ = MOTU = MOTV = MOαβγδ = UV = TUαβγδ = TVαβγδ*
7	*I = ABC = MN = MO = TU = TV =* αβ *=* αγ *= ABCMN = ABCMO = ABCTU = ABCTV = ABCαβ = ABCαγ = NO = MNTU = MNTV = MNαβ = MNαγ = MOTU = MOTV = MOαβ = MOαγ = UV = TUαβ = TUαγ = TVαβ = TVαγ =* βγ
8	*I = ABC = MN = MO = TU = TV =* αβ *=* γδ *= ABCMN = ABCMO = ABCTU = ABCTV = ABCαβ = ABCγδ = NO = MNTU = MNTV = MNαβ = MNγδ = MOTU = MOTV = MOαβ = MOγδ = UV = TUαβ = TUγδ = TVαβ = TVγδ =* αβγδ
9	*I = ABC = MN = MO = TU = TV = αβ =* αγδ *= ABCMN = ABCMO = ABCTU = ABCTV = ABCαβ = ABCαγδ = NO = MNTU = MNTV = MNαβ = MNαγδ = MOTU = MOTV = MOαβ = MOαγδ = UV = TUαβ = TUαγδ = TVαβ = TVαγδ = βγδ*

5.6 Case Study: Applying the MSFFSP Design to a Nano-Scale Polymerization Process

We now return to the three-stage polymerization process explained in Section 5.1. The first stage involves preparation of the self-assembled monolayer (SAM), the second stage involves anchoring the catalyst to the SAM, and finally in the third stage, the polymerization takes place through synthesis. The three-stage procedure represented in Figure 5.1 is carried out to form films on the substrate that have thickness measured at

the nano-scale. Once the process is well understood, the properties of the film are to be studied as part of the long-term goals of the research. Hence, it is imperative to understand the relation of the factors involved in the polymerization process. A statistical model would help immensely in predicting the thickness of the films, given the factor settings. In addition, to obtain a film of a certain thickness, factors can be set at appropriate levels. Keeping the objectives in mind, a design is constructed for the polymerization process, and is explained below.

5.6.1 Design for the polymerization process

The process of preparing the self-assembled monolayers (SAMs) in stage I is studied extensively in past literature. The factors in Table 5.14 are known to affect the preparation of the SAM. Furthermore, current scientific knowledge and past literature indicate that the current operating levels of the factors listed in Table 5.14 are indeed the optimal settings for the stage. By operating at these levels, the SAMs prepared are robust to any further reactions. Hence, for all practical purposes, these factors are ignored in the current investigation.

Table 5.14: Factors in Stage I

Factors of interest	Current operation level
Amount of gold evaporated	200nm
Thickness of silicon wafer	2 inches
Amount of Cr	10nm
Temperature	room temp.
Time	24 hrs

In the stage II of the process, the SAM is immersed in a nitromethane solution so that catalyst can be anchored to the substrate for further reactions. The amount of catalyst in the solution dictates the anchoring process. The catalyst anchors itself through ligand exchange. The time for which the SAM is immersed in the solution is also known to affect the quality of catalyst anchorage. Once the catalyst is anchored, the substrate needs to be rinsed with a suitable choice of a solvent. Two solvents that are known to yield reasonable results are dichloromethane (CH_2Cl_2) and nitromethane (CH_3NO_2). Finally, the drying time for the process is also suspected to cause significant changes. It is important to note that any number of substrates can be processed (catalyst can be anchored) for a given setup of stage II. Table 5.15 summarizes the factors of interest in stage II and lists the levels at which they are to be investigated.

Table 5.15: Factors in Stage II

Factors of interest	Levels to be tested	
	"low" level	"high" level
Amount of catalyst (*M*)	1 mM	10 mM
Type of rinsing solvent (*N*)	CH_2Cl_2	CH_3NO_2
Drying time (*O*)	30 sec	60 sec
Reaction time (time substrate is immersed) (*A*)	4 hrs	12 hrs

In stage III, the catalyst-anchored SAM is exposed to Argon gas saturated with the monomer. The substrate can be exposed to the saturated gas in a couple of ways. One approach is to maintain a continuous flow of the gas from a chamber filled with excess of monomer to another chamber containing the substrate. On the other hand, to conserve monomer (which could be expensive if used in bulk), the second approach involves a single chamber filled with Argon and containing both the monomer and the substrate.

Polymers are formed on the substrate by the continuous reaction of the monomer with the initiator. Scientific knowledge of the process suggests that the type of gas flow used plays a vital role in the thickness of the polymer film. In addition, as in any chemical reaction, temperature at which the process is carried out and the reaction time determine the extent of the reaction. Due to the physical constraints of the setup, a maximum of four substrates can be treated (polymerized) in stage III when continuous flow of Argon is chosen. These factors, along with the levels, are summarized in Table 5.16.

Table 5.16: Factors in Stage III

Factors of interest	Levels to be tested	
	"low" level	"high" level
Reaction temperature (α)	10 ℃	25 ℃
Type of Argon flow (β)	continuous	static
Reaction time (T)	12 hrs	48 hrs

Due to limited resources, a maximum of one week (168 hours) can be devoted to the experiment. For this reason, factors A (stage II) and T (stage III) are said to be hard-to-change as minimum setting changes of their levels are preferred. A completely randomized design involves many setting changes of factors A and T and hence, is not considered. For the current investigation, if we ignore the process in stage I, then a conventional split-plot design could be employed to accommodate the hard-to-change factors. Considering stage II as the whole plot and stage III as the sub-plot, a $2^{4-1} \times 2^{3-1}$ split-plot design can be used. This design would require 32 runs and would involve 8 setting changes of factor A and 32 setting changes of factor T. The total time required to execute this design would be 1024 hours, which is unacceptable.

The factor *T* in stage III is harder-to-change than factor *A* in stage II. Hence, a split-plot design having stage II as the sub-plot and stage III as the whole plot would utilize fewer resources than the previous split-plot arrangement. Due to the incapability of the experimental setup to polymerize, more than four substrates at a time while maintaining continuous flow of Argon, this arrangement requires approximately half the time (436) as the previous arrangement but still exceeds the maximum.

The classic split-plot design is a special case of the *m*-MSFFSP design where *m* equals two. In order to accommodate the additional hard-to-change factor in each stage, a variation of the *m*-MSFFSP design can be considered. The variation demands that the hard-to-change factors be treated as separate stages, making the overall design a 4-MSFFSP. Following the variation, stages II and III are further separated into stages 1, 2, 3, and 4. To further minimize the setting changes, factor *T* corresponds to stage 1, factor *A* to stage 2, factors *M*, *N*, and *O* to stage 3, and finally, factors α and β to stage 4 of the design. Thus, the $2^1 \times 2^1 \times 2^{3-1} \times 2^{2-1}$ MSFFSP design would require only two setting changes of factor *T* and four setting changes of factor *A*. Moreover, there needs to be eight different setups at a given time, corresponding to the setup for the continuous flow of Argon. As only four such setups are possible, they need to be repeated, thus further increasing the time. Hence, if only two setting changes are required, the time consumed due to factor *A* is 1.5 times the number of setting changes. The experiment would then require 152 hours, which is well within the acceptable limit.

If factor *A* corresponded to stage 1 and factor *T* to stage 2, the entire experiment can be executed in only 136 hours. However, the execution of the experiment demands that 68 hours be spent continuously with the process to make the transition between

stages II and III uniform. Hence, the previous arrangement is chosen as it provides the experimenter more flexibility in executing the design, while taking only 24 more hours.

Even though the variation utilizes fewer resources, it has the disadvantage that the main effects of those factors being treated as separate stages can no longer be tested for significance. Thus, the above variation can only be applied to instances where the main effects of the hard-to-change factors are not of primary interest.

Once the design layout has been selected, an appropriate confounding pattern needs to be chosen so that the design is capable of estimating maximum number of effects of interest. In the polymerization experiment, the following effects are of interest:

- Main effects of factors M, α, and β
- Two-way interactions involving factor T: effects TA, TM, TN, TO, $T\alpha$, and $T\beta$.
- Two-way interactions involving factor A: effects AM, AN, AO, $A\alpha$, and $A\beta$.
- Two-way interactions in stage 3 involving factors N and O: effects NO, MN, and MO.
- Two-way interactions involving factors from different stages: effects $M\alpha$, $M\beta$, $N\alpha$, Nβ, $O\alpha$, $O\beta$.
- Two-way interactions between factors in stage 4: $\alpha\beta$
- Three-way interactions involving factors N and O and having at least two factors from different stages: effects $TN\alpha$, $TN\beta$, $TO\alpha$, $TO\beta$, AMN, AMO, $N\alpha\beta$, and $O\alpha\beta$.

In summary, there are 32 effects of interest for the experiment. An effect is said to be estimable if it is not aliased with main effects or two-way interactions not listed above. The catalogs created in Section 5.5 are mainly targeted towards robustness studies, where

only interaction effects are of interest and all effects are assumed to be of equal interest. In this case study, due to the experimenter's prior knowledge, some effects are suspected to be of greater importance than others. In addition, as these suspicious effects constitute main and interaction effects, a specific design not listed in the catalogs has to be constructed for the case study. As many interaction effects are of interest, confounding within plots can be used to generate columns. For the $2^1 \times 2^1 \times 2^{3-1} \times 2^{2-1}$ MSFFSP design being considered, the best design (in terms of number of effects estimable and maximum resolution) has the defining relation $I = MNO = \alpha\beta = MNO\alpha\beta$. The overall resolution of the design is II, and only the following seven effects are estimable: *TA*, *TM*, *TN*, *TO*, *AM*, *AN*, and *AO*.

Confounding within plots could be employed in stage 3 as more interaction effects involving stage 3 factors are of interest than "pure" stage 3 effects. In addition, as we are interested in the main effects and "pure" interactions of stage 4 (α and β), we can consider split-plot confounding for the generator in stage 4. Thus, the generator in stage 3 has confounding within plots and the generator in stage 4 has split-plot confounding. Therefore, the overall design is said to have mixed confounding. Choosing $O = MN$ and $\beta = TAM\alpha$, the overall resolution of the design increases to III, and the defining relation is $I = MNO = TAM\alpha\beta = TANO\alpha\beta$. Note that the partial resolution of stage 4 increased from II to V. Using this design, 28 of the 32 effects of interest are estimable. Main effect *M* is confounded with *NO* and hence not estimable. If split-plot confounding were chosen in stage 3, then many more interactions involving stage 3 factors would not be estimable. Moreover, effects *AMN* and *AMO* are confounded with two-way interactions that are not of interest. The complete alias structure for the experiment is provided in Appendix 5A.

The corresponding layout of the design is shown in Table 5.17. It is interesting to note that a split-lot design would require only 76 hours of experimentation, which is much lesser than the time required by the 4-MSFFSP design. However, the only type of confounding permitted in a split-lot design would be confounding within plots, and hence only eight of the effects of interest would be estimable.

Table 5.17: Design for the Polymerization Experiment

Run	T	A	M	N	O=MN	α	β=TAMα	Treatment
1	-	-	-	-	+	-	+	*oβ*
2						+	-	*oα*
3			+	-	-	-	-	*m*
4						+	+	*mαβ*
5			-	+	-	-	+	*nβ*
6						+	-	*nα*
7			+	+	+	-	-	*mno*
8						+	+	*mnoαβ*
9		+	-	-	+	-	-	*ao*
10						+	+	*aoαβ*
11			+	-	-	-	+	*amβ*
12						+	-	*amα*
13			-	+	-	-	-	*an*
14						+	+	*anαβ*
15			+	+	+	-	+	*amnoβ*
16						+	-	*amnoα*
17	+	-	-	-	+	-	-	*to*
18						+	+	*toαβ*
19			+	-	-	-	+	*tmβ*
20						+	-	*tmα*
21			-	+	-	-	-	*tn*
22						+	+	*tnαβ*
23			+	+	+	-	+	*tmnoβ*
24						+	-	*tmnoα*
25		+	-	-	+	-	+	*taoβ*
26						+	-	*taoα*
27			+	-	-	-	-	*tam*
28						+	+	*tamαβ*
29			-	+	-	-	+	*tanβ*
30						+	-	*tanα*
31			+	+	+	-	-	*tamno*
32						+	+	*tamnoαβ*

one of the 16 groups

one of the 32 sub-groups

one of the 4 super-groups

one of the 2 super-super groups

Following the design layout, the experiment is conducted as follows:

- 16 groups of 2 SAMs each are formed
- 4 groups are randomly selected to form a super-group. Thus, there are 4 super-groups.
- 2 super-groups are selected at random.
- Each group within a super-group is immersed in a setup so that catalyst could be anchored. The 4 groups are randomly immersed in the 4 setups corresponding to 4 combinations of *M*, *N*, and *O*, forming 8 immersed-experimental units within a super-group.
- The 2 selected super-groups are randomly assigned to the 2 levels of factor *A*. The immersed-experimental units within each super-group are treated for corresponding reaction time and catalysts are thus anchored. Thus, 8 anchored-experimental units are formed in each of the 2 super-groups.
- Each group within a super-group is then divided into 2 sub-groups of one anchored-experimental unit each. The sub-groups are randomly assigned to the 2 setups corresponding to the combinations of α and β, forming 16 anchored-immersed-experimental units in all.
- The 16 anchored-immersed-experimental units within the 2 super-groups being considered are then combined to form a super-super-group and are treated for reaction time corresponding to a level of factor *T* in order to form films on the substrate. Thus, 16 synthesized-experimental units are formed.
- Thickness of the films are then measures on each of the 16 synthesized-experimental units.

- The other 2 super-groups are then selected and the above process is repeated to record thickness measurements of the remaining 16 synthesized-experimental units.

The experiment was randomized at four levels and specific instructions were provided to the researcher to execute the design. These instructions are provided in Appendix 5A.

5.6.2 Analysis of the Multistage Design

The linear model for the polymerization experiment, up to three-way interaction effects, can be represented as

$$
\begin{aligned}
y_{ijklmno} = {} & \mu + T_i + A_j + (St_{2_E})_{ij} + M_k + (TM)_{ik} + (AM)_{jk} + (TAM)_{ijk} + N_l \\
& + (TN)_{il} + (AN)_{jl} + (TAN)_{ijl} + (MN)_{kl} + (TMN)_{ikl} + (AMN)_{jkl} + \\
& + O_m + (TO)_{im} + (AO)_{jm} + (TAO)_{ijm} + (MO)_{km} + (TMO)_{ikm} + (AMO)_{jkm} \\
& + (NO)_{lm} + (TNO)_{ilm} + (ANO)_{jlm} + (MNO)_{klm} + (St_{3_E})_{ijklm} + \alpha_n + (T\alpha)_{in} \\
& + (A\alpha)_{jn} + (TA\alpha)_{ijn} + (M\alpha)_{kn} + (TM\alpha)_{ikn} + (AM\alpha)_{jkn} + (N\alpha)_{\boldsymbol{ln}} \\
& + (TN\alpha)_{i\boldsymbol{ln}} + (AN\alpha)_{j\boldsymbol{ln}} + (MN\alpha)_{k\boldsymbol{ln}} + (O\alpha)_{mn} + (TO\alpha)_{imn} + (AO\alpha)_{jmn} \\
& + (MO\alpha)_{kmn} + (NO\alpha)_{lmn} + \beta_o + (T\beta)_{io} + (A\beta)_{jo} + (TA\beta)_{ijo} + (M\beta)_{ko} \\
& + (TM\beta)_{iko} + (AM\beta)_{jko} + (N\beta)_{lo} + (TN\beta)_{ilo} + (AN\beta)_{jlo} + (MN\beta)_{klo} \\
& + (O\beta)_{mo} + (TO\beta)_{imo} + (AO\beta)_{jmo} + (MO\beta)_{kmo} + (NO\beta)_{lmo} + (\alpha\beta)_{no} \\
& + (T\alpha\beta)_{ino} + (A\alpha\beta)_{jno} + (M\alpha\beta)_{kno} + (N\alpha\beta)_{\boldsymbol{ln}\,o} + (O\alpha\beta)_{mno} + (St_{4_E})_{ijklmno}
\end{aligned}
\qquad \textbf{(5.39)}
$$

where

μ is the overall mean, and

$i, j, k, l, m, n, o = 1, 2$.

The table of EMS for the above model identifies the various hypotheses that can be tested for significance. For the sake of brevity, we consider a sample of each type of effect. For instance, T_i, A_j, M_k, and α_n represents the main effects from stages 1, 2, 3, and 4, respectively. $(TA)_{ij}$, $(TM)_{ik}$, $(T\alpha)_{in}$, $(AM)_{jk}$, $(A\alpha)_{jn}$, and $(M\alpha)_{kn}$ represent two-way interaction effects between stages1 and 2, 1 and 3, 1 and 4, 2 and 3, 2 and 4, and 3 and 4, respectively. Finally, $(TAM)_{ijk}$, $(TA\alpha)_{ijn}$, and $(AM\alpha)_{jkn}$ represent the three-way interaction effects between stages 1, 2, and 3, 1, 2, and 4, and 2, 3, and 4, respectively. The EMS for each of the above token effect was obtained using the EMS rules (Montgomery, 2005), and the EMS table is shown in Table 5.18.

Table 5.18: EMS Table for Polymerization Experiment

Source	2 F *i*	2 F *j*	2 F *k*	2 F *l*	2 F *m*	2 F *n*	2 F *o*	EMS
T_i	0	2	2	2	2	2	2	$64\sum T^2 + 64\sigma_1^2 + 32\sigma_2^2 + 4\sigma_3^2 + \sigma_4^2$
$(St1E)_i$	0	2	2	2	2	2	2	$64\,\sigma_1^2 + 32\,\sigma_2^2 + 4\sigma_3^2 + \sigma_4^2$
A_j	2	0	2	2	2	2	2	$64\sum A^2 + 32\,\sigma_2^2 + 4\sigma_3^2 + \sigma_4^2$
$(St2E)(ij)$	1	1	2	2	2	2	2	$32\,\sigma_2^2 + 4\sigma_3^2 + \sigma_4^2$
M_k	2	2	0	2	2	2	2	$64\sum M^2 + 4\sigma_3^2 + \sigma_4^2$
$(TM)_{ik}$	0	2	0	2	2	2	2	$32\sum\sum TM^2 + 4\sigma_3^2 + \sigma_4^2$
$(AM)_{jk}$	2	0	0	2	2	2	2	$32\sum\sum AM^2 + 4\sigma_3^2 + \sigma_4^2$
$(TAM)_{ijk}$	0	0	0	2	2	2	2	$16\sum\sum\sum TAM^2 + 4\sigma_3^2 + \sigma_4^2$
$(St3E)_{(ijklm)}$	1	1	1	1	1	2	2	$4\,\sigma_3^2 + \sigma_4^2$
α_n	2	2	2	2	2	0	2	$64\sum\alpha^2 + \sigma_4^2$
$(T\alpha)_{in}$	0	2	2	2	2	0	2	$32\sum\sum T\alpha^2 + \sigma_4^2$
$(A\alpha)_{jn}$	2	0	2	2	2	0	2	$32\sum\sum A\alpha^2 + \sigma_4^2$
$(TA\alpha)_{ijn}$	0	0	2	2	2	0	2	$16\sum\sum\sum TA\alpha^2 + \sigma_4^2$
$(M\alpha)_{kn}$	2	2	0	2	2	0	2	$32\sum\sum M\alpha^2 + \sigma_4^2$
$(AM\alpha)_{jkn}$	2	0	0	2	2	0	2	$16\sum\sum\sum AM\alpha^2 + \sigma_4^2$
$(St4E)_{(ijklmno)}$	1	1	1	1	1	1	1	σ_4^2

The values in Table 5.18 can also be confirmed by expanding the general expression for EMS given in Equation 4.32 . For instance, using Equation 4.32, the EMS for effect *T* is given by

$$EMS(T) = \frac{2^6 \sum T^2}{2-1} + (d^2 d^3 d^4)\sigma_1^2 + (d^3 d^4)\sigma_2^2 + d^4\sigma_3^2 + \sigma_4^2 . \qquad (5.40)$$

As $d^2 = 2$, $d^3 = 2^3$, and $d^4 = 2^2$,

$$EMS(T) = 64\sum T^2 + 64\sigma_1^2 + 32\sigma_2^2 + 4\sigma_3^2 + \sigma_4^2 . \qquad (5.41)$$

The EMS for other effects can be obtained in a similar way. From Table 5.18, it can be seen that the hypotheses on the means of the effects can be tested against appropriate error terms. Once the experiment is conducted and data obtained, significant effects could be identified with the help of normal probability plots, Table 5.18, and the rules developed in Section 5.3.2.

The rules for assigning contrasts were applied to the complex alias structure for the polymerization process. Specifically, 31 effects are divided into 4 categories to be tested against 4 error terms - 1 contrast is assigned to stage 1 error, 2 contrasts are assigned to stage 2 error, 12 contrasts are assigned to stage 3 error, and 16 contrasts are assigned to stage 4 error. The properties for MSFFSP designs were successfully applied to the motivating polymerization example. As the experiment is unreplicated, no degrees of freedom for the error terms exist and hence alternate methods have to be used. First, a normal probability plot can be used to identify potentially significant effects. When the usual assumptions hold, ANOVA can be conducted by pooling insignificant terms identified from the normal probability plot. If counter-intuitive results are observed, design can be extended to study the significant effects and factors from Stage I (formation of SAM) through an additional stage. Suitable conditions for film formation having desired thickness can be identified from separate analysis of each stage.

Appendix 5A Alias Structure for the Polymerization Experiment

Complete alias structure for the polymerization experiment

The complete alias structure for the polymerization experiment is as follows (effects of interest are bold faced):

Generalized defining contrast:

$(1) = M\times N\times O = T\times A\times M\times\alpha\times\beta = T\times A\times N\times O\times\alpha\times\beta$

Confounded effects:

$T = T\times M\times N\times O = A\times M\times\alpha\times\beta = A\times N\times O\times\alpha\times\beta$

$A = A\times M\times N\times O = T\times M\times\alpha\times\beta = T\times N\times O\times\alpha\times\beta$

$\boldsymbol{T\times A} = T\times A\times M\times N\times O = M\times\alpha\times\beta = N\times O\times\alpha\times\beta$

$\boldsymbol{M} = N\times O = T\times A\times\alpha\times\beta = T\times A\times M\times N\times O\times\alpha\times\beta$

$\boldsymbol{T\times M} = T\times N\times O = A\times\alpha\times\beta = A\times M\times N\times O\times\alpha\times\beta$

$\boldsymbol{A\times M} = A\times N\times O = T\times\alpha\times\beta = T\times M\times N\times O\times\alpha\times\beta$

$T\times A\times M = T\times A\times N\times O = \boldsymbol{\alpha\times\beta} = M\times N\times O\times\alpha\times\beta$

$N = \boldsymbol{M\times O} = T\times A\times M\times N\times\alpha\times\beta = T\times A\times O\times\alpha\times\beta$

$\boldsymbol{T\times N} = T\times M\times O = A\times M\times N\times\alpha\times\beta = A\times O\times\alpha\times\beta$

$\boldsymbol{A\times N} = \boldsymbol{A\times M\times O} = T\times M\times N\times\alpha\times\beta = T\times O\times\alpha\times\beta$

$T \times A \times N = T \times A \times M \times O = M \times N \times \alpha \times \beta = \boldsymbol{O \times \alpha \times \beta}$

$\boldsymbol{M \times N} = O = T \times A \times N \times \alpha \times \beta = T \times A \times M \times O \times \alpha \times \beta$

$T \times M \times N = \boldsymbol{T \times O} = A \times N \times \alpha \times \beta = A \times M \times O \times \alpha \times \beta$

$\boldsymbol{A \times M \times N} = \boldsymbol{A \times O} = T \times N \times \alpha \times \beta = T \times M \times O \times \alpha \times \beta$

$T \times A \times M \times N = T \times A \times O = \boldsymbol{N \times \alpha \times \beta} = M \times O \times \alpha \times \beta$

$\boldsymbol{\alpha} = M \times N \times O \times \alpha = T \times A \times M \times \beta = T \times A \times N \times O \times \beta$

$\boldsymbol{T \times \alpha} = T \times M \times N \times O \times \alpha = A \times M \times \beta = A \times N \times O \times \beta$

$\boldsymbol{A \times \alpha} = A \times M \times N \times O \times \alpha = T \times M \times \beta = T \times N \times O \times \beta$

$T \times A \times \alpha = T \times A \times M \times N \times O \times \alpha = \boldsymbol{M \times \beta} = N \times O \times \beta$

$\boldsymbol{M \times \alpha} = N \times O \times \alpha = T \times A \times \beta = T \times A \times M \times N \times O \times \beta$

$T \times M \times \alpha = T \times N \times O \times \alpha = \boldsymbol{A \times \beta} = A \times M \times N \times O \times \beta$

$A \times M \times \alpha = A \times N \times O \times \alpha = \boldsymbol{T \times \beta} = T \times M \times N \times O \times \beta$

$T \times A \times M \times \alpha = T \times A \times N \times O \times \alpha = \boldsymbol{\beta} = M \times N \times O \times \beta$

$\boldsymbol{N \times \alpha} = M \times O \times \alpha = T \times A \times M \times N \times \beta = T \times A \times O \times \beta$

$\boldsymbol{T \times N \times \alpha} = T \times M \times O \times \alpha = A \times M \times N \times \beta = A \times O \times \beta$

$A \times N \times \alpha = A \times M \times O \times \alpha = T \times M \times N \times \beta = \boldsymbol{T \times O \times \beta}$

$T \times A \times N \times \alpha = T \times A \times M \times O \times \alpha = M \times N \times \beta = \boldsymbol{O \times \beta}$

$M \times N \times \alpha = \boldsymbol{O \times \alpha} = T \times A \times N \times \beta = T \times A \times M \times O \times \beta$

$T \times M \times N \times \alpha = \boldsymbol{T \times O \times \alpha} = A \times N \times \beta = A \times M \times O \times \beta$

$A \times M \times N \times \alpha = A \times O \times \alpha = T \times N \times \beta = T \times M \times O \times \beta$

$T \times A \times M \times N \times \alpha = T \times A \times O \times \alpha = \boldsymbol{N \times \beta} = M \times O \times \beta$

Appendix 5B Instructions for Polymerization Experimentation

Detailed instructions on conducting the polymerization experiment

1. **Label** the SAMs with random numbers between 1 and 32.

2. **Immerse SAMs** labeled 31 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

3. **Immerse SAMs** labeled 3 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

4. **Immerse SAMs** labeled 25 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

5. **Immerse SAMs** labeled 16 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

6. **Immerse SAMs** labeled 4 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

7. **Immerse SAMs** labeled 23 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

8. **Immerse SAMs** labeled 20 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

9. **Immerse SAMs** labeled 2 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

10. **Remove SAMs** labeled 31, 3, 25, 16, 4, 23, 20, and 2 from their respective solutions after 4 hours [corresponding to (-) level of factor *A*].

11. **Rinse SAMs** labeled 4, 23, 20, and 2 with CH_3NO_2 [corresponding to (+) level of factor *N*].

12. **Rinse SAMs** labeled 31, 3, 25, and 16 with CH_2Cl_2 [corresponding to (-) level of factor *N*].

13. **Dry SAMs** labeled 31, 3, 20, and 2 for 60 seconds [corresponding to (+) level of factor *O*].

14. **Dry SAMs** labeled 25, 16, 4, and 23 for 30 seconds [corresponding to (-) level of factor *O*].

15. **Immerse SAM** labeled 3 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

16. **Immerse SAM** labeled 31 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

17. **Immerse SAM** labeled 4 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

18. **Immerse SAM** labeled 23 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

19. **Immerse SAM** labeled 25 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

20. **Immerse SAM** labeled 16 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

21. **Immerse SAM** labeled 20 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

22. **Immerse SAM** labeled 2 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

23. **Remove SAMs** labeled 31, 3, 25, 16, 4, 23, 20, and 2 after 48 hours [corresponding to (+) level of factor *T*].

24. **Measure the thickness** of the films on SAMs labeled 31, 3, 25, 16, 4, 23, 20, and 2 [corresponds to the response].

Processing complete for 8 substrates at this point

25. **Immerse SAMs** labeled 1 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

26. **Immerse SAMs** labeled 32 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

27. **Immerse SAMs** labeled 29 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

28. **Immerse SAMs** labeled 30 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

29. **Immerse SAMs** labeled 27 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

30. **Immerse SAMs** labeled 5 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

31. **Immerse SAMs** labeled 9 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

32. **Immerse SAMs** labeled 14 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

33. **Remove SAMs** labeled 1, 32, 27, 5, 29, 30, 9, and 14 from their respective solutions after 12 hours [corresponding to (+) level of factor *A*].

34. **Rinse SAMs** labeled 29, 30, 9, and 14 with CH_3NO_2 [corresponding to (+) level of factor *N*].

35. **Rinse SAMs** labeled 1, 32, 27, and 5 with CH_2Cl_2 [corresponding to (-) level of factor *N*].

36. **Dry SAMs** labeled 1, 32, 9, and 14 for 60 seconds [corresponding to (+) level of factor *O*].

37. **Dry SAMs** labeled 27, 5, 29, and 30 for 30 seconds [corresponding to (-) level of factor *O*].

38. **Immerse SAM** labeled 27 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

39. **Immerse SAM** labeled 5 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

40. **Immerse SAM** labeled 9 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

41. **Immerse SAM** labeled 14 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

42. **Immerse SAM** labeled 29 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

43. **Immerse SAM** labeled 30 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

44. **Immerse SAM** labeled 1 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

45. **Immerse SAM** labeled 32 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

46. **Remove SAMs** labeled 1, 32, 27, 5, 29, 30, 9, and 14 after 48 hours [corresponding to (+) level of factor *T*].

47. **Measure the thickness** of the films on SAMs labeled 1, 32, 27, 5, 29, 30, 9, and 14 [corresponds to the response].

Half the experiment (16 substrates) has been completed at this point

48. **Immerse SAMs** labeled 19 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

49. **Immerse SAMs** labeled 10 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

50. **Immerse SAMs** labeled 12 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

51. **Immerse SAMs** labeled 26 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

52. **Immerse SAMs** labeled 28 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

53. **Immerse SAMs** labeled 11 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

54. **Immerse SAMs** labeled 7 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

55. **Immerse SAMs** labeled 21 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

56. **Remove SAMs** labeled 19, 10, 28, 11, 12, 26, 7, and 21 from their respective solutions after 4 hours [corresponding to (-) level of factor *A*].

57. **Rinse SAMs** labeled 12, 26, 7, and 21 with CH_3NO_2 [corresponding to (+) level of factor *N*].

58. **Rinse SAMs** labeled 19, 10, 28, and 11 with CH_2Cl_2 [corresponding to (-) level of factor *N*].

59. **Dry SAMs** labeled 19, 10, 7, and 21 for 60 seconds [corresponding to (+) level of factor *O*].

60. **Dry SAMs** labeled 28, 11, 12, and 26 for 30 seconds [corresponding to (-) level of factor *O*].

61. **Immerse SAM** labeled 28 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

62. **Immerse SAM** labeled 11 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

63. **Immerse SAM** labeled 21 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

64. **Immerse SAM** labeled 7 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

65. **Immerse SAM** labeled 12 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

66. **Immerse SAM** labeled 26 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

67. **Immerse SAM** labeled 10 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

68. **Immerse SAM** labeled 19 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

69. **Remove SAMs** labeled 19, 10, 28, 11, 12, 26, 7, and 21 after 12 hours [corresponding to (-) level of factor *T*].

70. **Measure the thickness** of the films on SAMs labeled 19, 10, 28, 11, 12, 26, 7, and 21 [corresponds to the response].

Processing complete for 24 substrates at this point

71. **Immerse SAMs** labeled 13 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

72. **Immerse SAMs** labeled 18 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

73. **Immerse SAMs** labeled 8 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

74. **Immerse SAMs** labeled 24 in a solution containing 1mM catalyst [corresponding to (-) level of factor *M*].

75. **Immerse SAMs** labeled 15 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

76. **Immerse SAMs** labeled 6 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

77. **Immerse SAMs** labeled 22 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

78. **Immerse SAMs** labeled 17 in a solution containing 10mM catalyst [corresponding to (+) level of factor *M*].

79. **Remove SAMs** labeled 13, 18, 15, 6, 8, 24, 22, and 17 from their respective solutions after 12 hours [corresponding to (+) level of factor *A*].

80. **Rinse SAMs** labeled 8, 24, 22, and 17 with CH_3NO_2 [corresponding to (+) level of factor *N*].

81. **Rinse SAMs** labeled 13, 18, 15, and 6 with CH_2Cl_2 [corresponding to (-) level of factor *N*].

82. **Dry SAMs** labeled 13, 18, 22, and 17 for 60 seconds [corresponding to (+) level of factor *O*].

83. **Dry SAMs** labeled 15, 6, 8, and 24 for 30 seconds [corresponding to (-) level of factor *O*].

84. **Immerse SAM** labeled 18 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

85. **Immerse SAM** labeled 13 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

86. **Immerse SAM** labeled 8 in a chamber at 25°C with static Argon [corresponding to (+,+) levels of factors (α, β)].

87. **Immerse SAM** labeled 24 in a chamber at 10°C with continuous flow of Argon [corresponding to (-,-) levels of factors (α, β)].

88. **Immerse SAM** labeled 15 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

89. **Immerse SAM** labeled 6 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

90. **Immerse SAM** labeled 22 in a chamber at 25°C with continuous flow of Argon [corresponding to (+,-) levels of factors (α, β)].

91. **Immerse SAM** labeled 17 in a chamber at 10°C with static Argon [corresponding to (-,+) levels of factors (α, β)].

92. **Remove SAMs** labeled 13, 18, 15, 6, 8, 24, 22, and 17 after 24 hours [corresponding to (-) level of factor *T*].

93. **Measure the thickness** of the films on SAMs labeled 13, 18, 15, 6, 8, 24, 22, and 17 [corresponds to the response].

Processing complete for all 32 substrates at this point

6. CONCLUSIONS AND DISCUSSION

6.1 Summary

The ideas presented in this thesis address the areas of nanomanufacturing challenges, experimental designs with missing observations, and multistage split-plot experiments. The motivating factor has been the severe lack of experimental knowledge needed to effectively discover new products and processes at the nano-scale.

The first part of this research focused on the nano-scale lubrication experiments and the problem of missing observations stemming from it. Bayesian algorithms were developed that performed much better than the traditional non-Bayesian ones, even in the case of mismatches. A polymerization process conducted over multiple stages was studied in the second part of the research. Various properties and characteristics of split-plot designs applied to multiple stages were studied, and general expressions were obtained for a given number of stages.

Through these designs, we illustrated how DOE techniques can be used in nanomanufacturing design and engineering by helping researchers understand the product and process dynamics. There are undoubtedly countless other nanotechnology–related projects that can also benefit from these methods. For instance, missing observations could occur in an experiment similar to the one conducted by Sun et al. (2004) to investigate the conditions of the surface treatment aimed at improving the rheological behavior of the nanosilica composite no-flow underfill.

A brief summary of the research contributions is presented in the next section. Directions for the advancement of this research are proposed in the subsequent section.

Finally, a few potential applications of the proposed designs to fields other than nanotechnology are also mentioned.

6.2 Research Contributions

Although commendable research has taken place in the field of nanotechnology, there remains a wide gap in the translation of the technology to the commercial world. Through this research, we endeavor to solve a "piece of the puzzle" and motivate other researchers to further explore this area. There are many areas in nanomanufacturing that demand inter-disciplinary work in order to develop new, better and more robust products and processes.

Traditional algorithms based on minimizing sum of squares of the residuals have shown to yield reasonable results for *n*-way factorial designs. However, their performance plummets when applied to a full (2^k) factorial design, more so in a fractional factorial (2^{n-k}) design. To overcome these limitations, Bayesian-based algorithms are proposed in this research, which make use of *a priori* knowledge about the significant effects in the underlying process. These algorithms also have the capability to accommodate any prior belief held by the researcher investigating the process. These proposed algorithms not only yield more accurate estimates of the missing observations, but also provide close approximations to the null model (i.e., the model obtained when no observations are missing).

In addition, to facilitate present-day experimentation in nanomanufacturing, multistage fractional factorial split-plot designs are proposed for processes carried out over multiple stages. These designs are particularly valuable in circumstances in which

some experimental units need to be treated over all stages before processing on other units can commence. These multistage designs overcome the severe limitation of the commonly used split-lot designs by providing greater flexibility in the choice of confounding patterns. In addition, multistage designs require fewer runs than split-lot designs, while yielding greater information about the factors of interest. Two design optimality criteria based on common occurrences in nanomanufacturing experimentation are also proposed. Catalogs of designs for certain operational set-ups (number of factors and runs) are provided and ranked according to optimality criteria. These catalogs of designs are intended to aid practitioners in choosing a design, given the operational constraints for the experiment considered.

In summary, using these designs and integrating them into the modeling and production for nanomanufacturing research will yield strategic advantages by speeding the research and development cycle, stretching the experimental budget, and helping to create more reliable, robust, and better performing products. It is also believed that as advanced nanotechnology applications are explored with experimental design, there will be new questions that call for modifications or perhaps completely new constructs of experimental designs that will simultaneously advance the field of DOE.

6.3 Future Work

In this research, Bayesian algorithms were proposed to estimate missing observations in a nano-scale lubrication experiment. Some of the properties and characteristics of these algorithms were studied, along with the performance over different conditions for data

sets that resemble the lubrication experiment. From this work, we identify some potential areas for future research. These include the following:

- Nano-scale lubrication experiments could be conducted in a laboratory setting. Responses could then be measured on those conditions for which molecules already exist. Once the incomplete data set is formed, one of the Bayesian algorithms can be applied to obtain the estimates of the missing observations. Using these estimates, the characteristics of the molecules can be studied and recommendations can be provided to achieve desired degree of lubrication.
- As the motivating example for the development of Bayesian algorithms was a 2^{4-1} fractional factorial design, the performance of the algorithms was tested for data sets of the same size (eight runs). A recommendation for future study would be to perform a similar investigation as the one in Section 4.8 but for larger data sets. For different data sets, the effects of incorrect specification of the active factors and positions of missing observations can be examined. Thus, a more informed recommendation on the best performing Bayesian algorithm can be offered.
- Effects of variance in a data set for which the estimates of the missing observations are obtained can be examined. This study can be facilitated by using data sets having different values of standard deviations in the above performance study of the Bayesian algorithms.
- In the present study of the consequence of incorrect specification of active factors by the screening algorithm on the performances of Bayesian algorithms, an assumption

is that all "true" active effects are statistically significant at the same level of significance. In other words, all true effects have the same *p*-value in the presence of a complete data set. A future direction of research could be the sensitivity of this assumption. For instance, an answer to the question "up to what level of significance, do the Bayesian algorithms yield satisfactory estimates of missing observations?" can have some research potential.

A multistage polymerization process was considered and a MSFFSP design was proposed to investigate the effects of the process, while using few of the limited resources. Moreover, based on the common occurrences in the field of nanomanufacturing, two optimality criteria were proposed in Section 5.5 to rank various MSFFSP designs. The following directions of future work can be considered with respect to multistage designs.

- The design proposed in Section 5.6 for the multistage polymerization process can be executed by following the systematic directions given in Appendix 5B. Once the experiment is conducted, the thickness of the films on the 32 substrates can be measured. Using the film thickness of the 32 substrates as the response, the MSFFSP design can be analyzed following the appropriate procedure given in Section 5.6.2.
- The catalogs of design for processes conducted over three and four stages are presented in Section 5.5. The designs in these catalogs were ranked based on two common criteria, namely, robustness and mixed three-way interactions. Depending

on the objective of a study, different criteria can be used to evaluate its efficiency and adequacy. A few examples of the criteria are maximum number of clear effects, maximum number of clear main effects and two-way interactions within a stage, and maximum number of clear effects that are estimated with greater precision. A recommendation for future research would be to examine some of these design optimality criteria and rank the MSFFSP designs accordingly.

- The type of fractionation considered in the catalogs is "confounding within plots" as robustness studies were of primary interest. In future, "mixed confounding" can be used by removing the constraint of choosing columns from the same plot for some stages. These designs could potentially maximize the estimation of the effects of interest.
- Finally, a recommendation for future study would be to combine the ideas from Chapters 4 and 5. Missing observations in fractional factorial designs were considered in this research. The principle behind estimating the missing observations was reduction in the sum of squares of the error term. Missing observations can potentially occur in designs having more than one error term. An important research consideration would be to investigate the extension of the Bayesian algorithms to yield estimates for more advanced designs such as the split-plot or MSFSSP designs.

6.4 Application to Other Fields

We also hope that current and future extensions will be applied to fields other than nanomanufacturing. In particular, experimental designs have recently found extensive

application in the field of pharmaceutical research and development of new drugs. In 2003, the U.S. Department of Health and Human Services Food and Drug Administration promoted implementation of a Process Analytical Technology (PAT) (see FDA Draft Guidance, 2003) entitled “Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance”. This framework advocates the use experimental designs along with other statistical tools (process monitoring and control, multivariate data analysis, etc) to facilitate process understanding, continuous improvement and develop risk-free mitigation strategies. It is further stated that efficient experimentation can serve as building blocks of knowledge that grow to accommodate a higher degree of complexity throughout the life of a product.

Pharmaceutical industries like Eli Lilly and Company, Merck and Co., Inc., etc. are involved in discovery of new drugs by creating or finding the right combination of molecules that meet specific needs. Once these medicines are discovered and proven safe for human consumption, they need to be manufactured in large quantities with strict quality standards, usually set by International Organization for Standardization (ISO). Thus, a typical sequence in the introduction of a medicine involves a phase of drug discovery in which effective research yields newer compounds targeted at specific diseases or patient needs. The second phase involves drug development, where the discovered drugs are tested and manufactured for commercial use with the help of an optimized process for scaling-up.

During drug discovery, the robustness of a process under investigation to small changes in external conditions (example purification or crystallization conditions) is studied through experimental designs. In addition, DOEs are also used to identify critical

process variables that lead to a desired range of output values. Due to the high uncertainty involved with compounds and molecules in the discovery stage, there exists a high chance of not being able to traverse the entire space of a designed experiment. In other words, the all the runs in the experiment cannot be performed, thus yielding missing observations.

Similar uncertainties may also exist in the field of biotechnology, specifically in the study of performance of particular cells. Often, a statistical approach is adopted to study cell performance and provide optimal settings for the same (for example, see DeLong et al., 2004). Some of the conditions to which these cells are subjected, as part of a factorial design approach, might not be conducive, and hence, no results are obtained corresponding to those conditions. Missing observations in the above circumstances can be handled using one of the Bayesian algorithms proposed.

Lastly, researchers in the field of polymer research have widely begun to apply statistical approaches for the optimization of polymer formulations. Gulmus and Moneke (2005) note that the development of new polymer formulations for commercial and industrial applications, such as plastics, requires significant time, materials and labor costs. Moreover, the development process is conducted over multiple steps that include extrusion of relatively large amounts of polymer, its granulation, moulding of parts and manual testing of product performance. To expedite the development of new formulations, multistage designs (MSSP or MSFFSP) can be used effectively while maintaining the low-cost of the process. By treating the reaction time or labor costs or the available resources as hard-to-change factors, plastic formulations can be investigated and optimized faster and at lower costs than with traditional factorial designs.

REFERENCES

Ahmed, W., Ahmed, E., Maryan, C., Jackson, M. J., Ogwu, A. A., Ali, N. Neto, V. F., and Gracio, J. (2006). "Time Modulated CVD Process Optimized Using the Taguchi Method," *Journal of Materials Engineering and Performance*, 15, 2, 236-241.

Allen, F. E. and Wishart, J. (1930). "A Method of Estimating the Yield of a Missing Plot in Field Experimental Work," *Journal of Agricultural Science*, 20, 399-406.

Anderson, A. B., Basilevsky, A., and Hum, D. P. J. (1983). "Missing Data: A Review of the Literature," *Handbook of Survey Research*, 415-494, P. H. Rossi, J. D. Wright, and A. B. Anderson (Editors).

Anderson, M. M. (2005). "Path Creation in the Making - The Case of Nanotechnology," *DRUID Conference*, Copenhagen, Denmark.

Ashurst, W. R., Yau, C., Carraro, C., Lee, C., Maboudian, R., and Dugger, M. T. (2001). "Dichlorodimethylsilane as an Anti-stiction Monolayer for MEMS: A Comparison to the Octadecylyrichlosilane Self-Assembled Monolayer," *Journal of MEMS*, 10, 1, 41-49.

Bain, C. D., Troughton, E. B., Tao, Y., Evall, J., Whitesides, G. M., and Nuzzo, R. G. (1989). "Formation of Monolayer Films by the Spontaneous Assembly of Organic Thiols from Solution onto Gold," *Journal of the American Chemical Society*, 111, 321-335.

Bartlett, M. S. (1937). "Some Examples of Statistical Methods of Research in Agriculture and Applied Biology," *Supplement to Journal of the Royal Statistical Society*, 4, 2, 137-183.

Basumallick, A., Das, G. C., and Mukherjee, S. (2003). "Design of Experiments for Synthesizing *in situ* Ni-SiO_2 and Co-SiO_2 Nanocomposites by Non-isothermal Reduction Treatment," *Nanotechnology*, 14, 903-906.

Bhushan, B., editor (2004). *Springer Handbook of Nanotechnology*, Springer-Verlag Berlin Heidelberg, Germany.

Bingham, D. and Sitter, R. (1999). "Minimum Aberration Two-Level Fractional Factorial Split-Plot Designs," *Technometrics*, 41, 62-70.

Bingham, D. and Sitter, R. (2001). "Design Issues in Fractional Factorial Split-Plot Experiments," *Journal of Quality Technology*, 33, 2-15.

Bingham, D. and Sitter, R. (2003). "Fractional Factorial Split-Plot Designs for Robust Parameter Experiments," *Technometrics*, 45, 80-89.

Bisgaard, S. (2000). "The Design and Analysis of 2^{k-p} x 2^{q-r} Split-Plot Experiments," *Journal of Quality Technology*, 32, 39-56.

Boal, A. K., Bachand, G. D., Rivera, S. B., and Bunker, B. C. (2006). "Interactions between Cargo-Carrying Biomolecular Shuttles," *Nanotechnology*, 17, 349-354.

Box, G. E. P. (1990). "A Simple Way to Deal with Missing Observations from Designed Experiments," *Report # 57*, Center for Quality and Productivity Improvement, University of Wisconsin-Madison, Madison, WI.

Box, G. E. P. and Hunter, J. S. (1961). "The 2^{k-p} Fractional Factorial Designs Part I," *Technometrics*, 3, 3, 311-351.

Box, G. E. P. and Jones, S. (1992). "Split-Plot Designs for Robust Product Experimentation," *Journal of Applied Statistics*, 19, 3-26.

Box, G. E. P. and Meyer, R. D. (1993). "Finding Active Factors in Fractionated Screening Experiments," *Journal of Quality Technology*, 25, 94-105.

Buck, S. F. (1960). "A Method of Estimation of Missing Values in Multivariate Data Suitable for Use with an Electronic Computer," *Journal of the Royal Statistical Society, Series B,* 22, 302- 307.

Bulmahn, B. J. and Chauhan, C. K. (1994). "Missing Data - What Do We Do Now?," *48th Annual Quality Congress Proceedings*, 571-576.

Busnaina, A., Barry, C., and Miller, G., organizers (2003). "*NSF Workshop on Three Dimensional Nanomanufacturing: Partnering with Industry; Conclusions and Report,"* last accessed at http://www.nano.neu.edu/nsf_workshop.html on August 21, 2005.

Butler, N. A. (2004a). "Construction of Two-Level Split-Lot Fractional Factorial Designs for Multistage Processes," *Technometrics*, 46, 4, 445-451.

Butler, N. A. (2005). "Classification of Efficient Two-Level Fractional Factorial Designs of Resolution IV or More," *Journal of Statistical Planning and Inference*, 131, 1, 145-159.

Chauhan, C. K. and Bulmahn, B. J. (1993). "Missing Observations in Fractional Factorial Designs," *Midwest Decision Sciences Institute Proceedings*, 147-149.

Chen, C., Wang, S., Ko, R., Kuo, Y., Uang, K., Chen, T., Liou, B., and Tsai, H. (2006). "The Influence of Oxygen Content in the Sputtering Gas on the Self-Synthesis of Tungsten Oxide Nanowires on Sputter-Deposited Tungsten Films," *Nanotechnology*, 17, 217-223.

Chen, J., Sun, D. X., and Wu, C. F. J. (1993). "A Catalogue of Two-Level and Three-Level Fractional Factorial Designs with Small Runs," *International Statistical Review – Special Issue on Statistics in Industry*, 61, 1, 131-145.

Chipman, H. (1998). "Handling Uncertainty in Analysis of Robust Design Experiments," *Journal of Quality Technology*, 30, 11-17.

Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs*, John Wiley and Sons Inc., New York, NY.

Cool, A. (2000). "A Review of Methods Dealing with Missing Data," *Paper Presented at The Annual Meeting of the Southwest Educational Research Association*, Dallas, TX.

Daniel, C. (1959). "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments," *Technometrics*, 1, 4, 311-341.

DeLong, B., Lin, N., Ross, S., and Caple, M. (2004). "Medium Optimization Kit for CHO Cells: Using Factorial Matrix Design of Experiment to Increase Cell Performance," *Sigma-Aldrich Corporation*, St. Louis, MO.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm (with discussion)," *Journal of the Royal Statistical Society, Series B,* 39, 1-38.

Dewey, A., Ren, H., and Zhang, T. (2000). "Behavioral Modeling of Microelectromechanical Systems (MEMS) with Statistical Performance-Variability Reduction and Sensitivity Analysis," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* 47, 2, 105-113.

Doumandis, H. (2002). "The Nanomanufacturing Programme at the National Science Foundation," *Nanotechnology*, 13, 248-252.

Draper, N. R. and Stoneman, D. M. (1964). "Estimating Missing Values in Unreplicated Two-Level Factorial and Fractional Factorial Designs," *Biometrics*, 20, 3, 443-458.

Dronavajjala, K., Rajagopalan, R., Uppili, S., Allara, D., and Foley, H. (2006). "A Simple Technique to Grow Polymer Brushes Using in-situ Surface Ligation of Organometallic Initiator," submitted to *Journal of the American Chemical Society*.

FDA Draft Guidance. (2003). *Guidance for Industry PAT – A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance*, Washington D.C: Food and Drug Administration, Center for Drug Evaluation and Research (CDER).

Feynmann, R. P. (1960). "There's Plenty of Room at the Bottom," *Engineering Science*, 23, 22-36.

Finney, D. J. (1946). "Recent Developments in the Design of Field Experiments I. Split-Plot Confounding," *Journal of Agricultural Science*, 36, 56-62.

Fries, A. and Hunter, W. G. (1980). "Minimum Aberration 2^{k-p} Designs," *Technometrics*, 40, 314-326.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis,* Chapman and Hall, London.

Goh, T. N. (1997). "Some Strategies for Experimentation under Operational Constraints," *Quality and Reliability Engineering International*, 13, 279-283.

Grisolia, J., Shalchain, M., Assayag, G. B., Coffin, H., Bonafos, C., Schamm, S., Atarodi, S. M., and Claverie, A. (2005). "The Effects of Oxidation Conditions on Structural and Electric Properties of Silicon Nanoparticles Obtained by Ultra-Low-Energy Ion Implantation," *Nanotechnology*, 16, 2987-2992.

Gulmus, S. A. and Moneke, M. (2005). "High Speed Development of New Plastics," *Dutch Polymer Institute*, Eindhoven, Netherlands.

Healy, M. J. R. and Westmacott, M. (1956). "Missing Values in Experiments Analyzed on Automatic Computers," *Applied Statistics*, 5, 203-206.

Hering, F. and Wang, M. (1998). "Efficiency Comparison of Split-Block Design versus Split-Plot Design," *Listy Biometryczne/Biometrical Letters*, 35, 27-35.

Hinklemann, K. and Kempthorne, O. (1994). *Design and Analysis of Experiments Volume I: Introduction to Experimental Design*, John Wiley and Sons Inc., New York, NY.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). "Bayesian Model-Averaging: A Tutorial," *Statistical Science*, 14, 4, 382-417.

Huang, H., Tan, O. K., Lee, Y. C., Tse, M., S., Guo, J., and White, T. (2006). "Effects of Plasma Treatment on the Growth of SnO_2 Nanorods from SnO_2 Thin Films," *Nanotechnology*, 17, 743-746.

Huang, P., Chen, D., and Voelkel, J. O. (1998). "Minimum-Aberration Two-Level Split-Plot Designs," *Technometrics*, 40, 4, 314-326.

Jagarkal, S. G., Hossain, M. M., Agonafer, D., Lulu, M., and Reh, S. (2004). "Design Optimization and Reliability of PWB Level Electronic Package," *Proceedings of Inter Society Conference on Thermal Phenomena*, Las Vegas, NV, 368-376.

Jancar, B. and Suvorov, D. (2006). "The Influence of Hydrothermal-Reaction Parameters on the Formation of Chrysotile Nanotubes," *Nanotechnology*, 17, 25-29.

Ji, J., Tay, F. E. H., Miao, J., and Sun, J. (2006). "Characterization of Silicon Isotropic Etch by Inductively Coupled Plasma Etcher for Microneedle Array Fabrication," *Journal of Physics*, Conference Series 34, 1137-1142.

John, J. A. and Prescott, P. (1975). "Estimating Missing Values in Experiments," *Applied Statistics*, 24, 190-192.

Kaiser, J. (1990). "The Robustness of Regression and Substitution by Mean Methods in Handling Missing Values," *Paper Presented at the Second Islamic Countries Conference on Statistical Science*, Malaysia.

Kempthorne, O. (1952). *The Design and Analysis of Experiments*, John Wiley and Sons Inc., New York, NY.

Kharbas, H., Nelson, P., Yuan, M., Gong, S., and Turng, L. (2003). Effects of Nano-Fillers and Process Conditions on the Microstructure and Mechanical Properties of Microcellular Injection Molded Polyamide Nanocomposites," *Polymer Composites*, 24, 6, 655-671.

Kim, K. J. and Shahinpoor, M. (2003). "Ionic Polymer-Metal Composites: II. Manufacturing Techniques," *Smart Materials and Structures*, 12, 65-79.

Kuehl, R. O. (2000). *Design of Experiments - Statistical Principles of Research Design and Analysis*, 2nd edition, Duxbury, Belmont, Pacific Grov, CA.

Kutner, M. H, Nachtsheim, C. J., Neter, J. and Li, W. (2005). *Applied Statistical Linear Models*, 5th edition, McGraw-Hill Irwin, New York, NY.

Little, R. J., and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*, John Wiley and Sons Inc., New York, NY.

Liu, H., Weiss, R. E., Jennrich, R. I., and Wenger, N. S. (1999). "PRESS Model Selection in Repeated Measures Data," *Computational Statistics and Data Analysis*, 30, 169-184.

Mee, R. W. and Bates, R. L. (1998). "Split-Lot Designs: Experiments for Multistage Batch Processes," *Technometrics*, 40, 127-140.

Miller, A. (1997). "Strip-Plot Configurations of Fractional Factorials," *Technometrics*, 39, 153-161.

Milliken, G. A., Shi, X., Mendicino, M., and Vasudev, P. K. (1998). "Strip-Plot Design for Two-Step Processes," *Quality and Reliability Engineering International*, 14, 197-210.

Montgomery, D.C. (2005). *Design and Analysis of Experiments*, 6th Edition, John Wiley and Sons Inc., New York, NY.

Mukerjee, R. and Fang, K. (2000). "Fractional Factorial Split-Plot Designs with Minimum Aberration and Maximum Estimation Capacity," *Statistica Sinica*, 12, 885-903.

Nembhard, H. B. and Acharya, N. N. (2006). "Bayesian Algorithms for Missing Observations in Experimental Design for Nanomanufacturing," submitted to *Journal of Applied Statistics*.

Nembhard, H. B., Acharya, N. N., Aktan, M., and Kim, S. (2005). "Design Issues and Analysis of Experiments in Nanomanufacturing," Chapter 20, *CRC Press Handbook of Industrial and Systems Engineering*, Taylor and Francis Group, Boca Raton, FL.

Park, S. (2004). "Statistical Design of Experiments and Analysis on Gate Poly-Silicon Critical Dimension," *IEEE Transactions on Semiconductor Manufacturing*, 17, 3, 362-374.

Peng, C. Y. J., Harwell, M., Liou, S. M., and Ehman, L. H. (2002). "Advances in Missing Data Methods and Implications for Education Research," *Annual meeting of Chinese American Educational Research and Development Association*, New Orleans, LO.

Plackett, R. L. and Burman, J. P. (1946). "The Design of Optimum Multifactorial Experiments," *Biometrica*, 33, 305-325.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1997). "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association*, 92, 437, 179-191.

Rajagopal, R. and Del Castillo, E. (2005). "Model-Robust Process Optimization using Bayesian Model Averaging," *Technometrics*, 47, 2, 152-163.

Raymond, M. R. and Roberts, D. M. (1987). "A Comparison of Methods for Treating Incomplete Data in Selection Research," *Educational and Psychological Measurement*, 47, 13-26.

Ren, H., Jog, A, and Fair, R. B. (2001). "Statistical Optimal Design of Microelectromechanical Systems (MEMS)," *International Conference on Modeling and Simulation of Microsystems MSM*, San Diego, CA, March 19-21.

Rubin, D. B. (1972). "A Non-Iterative Algorithm for Least Squares Estimation of Missing Values in any Analysis of Variance Design," *Applied Statistics*, 21, 136-141.

Rubin, D. B. (1976). "Noniterative Least Squares Estimates, Standard Errors and F-Tests for Analyses of Variance with Missing Data," *Journal of the Royal Statistical Society*, *Series B*, 38, 270-274.

Russell, K. G., Lewis, S. M., and Dean, A. M. (2004). "Fractional Factorial Designs for the Detection of Interactions between Design and Noise Factors," *Journal of Applied Statistics*, 31, 5, 545-552.

Samset. O. and Tyssedal, J. (1998). "Study of the Box-Meyer Method for Finding Active Factors in Screening Experiments," *Technical Report*, Department of Mathematical Sciences, The Norwegian University of Science and Technology, Norway.

Schafer, J. L. and Graham, J. W. (2002). "Missing Data: Our View of State of the Art," *Psychological Methods*, 7, 2, 147-177.

Schafer, J. L. and Schenker, N. (2000). "Inference with Imputed Conditional Means," *Journal of the American Statistical Association,* 95, 144-154.

Shearer, P. (1973). "Missing Data in Quantitative Designs," *Applied Statistics*, 22, 135-140.

Strawhecker, K., Asay, D. B., and Kim, S. H. (2006). "Gas-Phase Lubrication of MEMS Devices: Using Alcohol Vapor Adsorption Isotherm for Lubrication of Silicon Oxides," *Dekker Encyclopedia of Chemical Processing*, 1143-1150.

Sullivan, J. P., Friedmann, T. A., and Hjort, K. (2001). "Diamond and Amorphous carbon MEMS," *MRS Bulletin*, 26, 4, 309-311.

Sun, Y., Zhang, Z., and Wong, C. P. (2004). "Fundamental Research on Surface Modification of Nano-size Silica for Underfill Applications," *9th International Symposium on Advanced Packaging Materials*, 132-138.

Tabachnick, B. G. and Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.), Needham Heights, MA: Allyn and Bacon.

Taguchi, G. (1987). *System of Experimental Design*, UNIPUB, New York, NY.

The Royal Society and The Royal Academy of Engineering (2003). "Nanotechnology: Views of Scientists and Engineers," *Report of a Workshop as Part of the Nanotechnology study*, last accessed on August 10, 2006 at www.nanotec.org.uk/.

Unalan, H. E. and Chhowalla, M. (2005). "Investigation of Single-Walled Carbon Nanotube Growth Parameters using Alcohol Catalytic Chemical Vapour Deposition," *Nanotechnology*, 16, 2153-2163.

Vivacqua, C. (2003). "*Using Strip-Block Designs as an Alternative to Reduce Costs of Experimentation in Robust Product Design and Multistage Processes*," PhD Thesis, IE Department, University of Wisconsin – Madison, Madison, WI.

Yates, F. (1933). "The Analysis of Replicated Experiments when the Field Results are Incomplete," *Empirical Journal of Experimental Agriculture,* 1, 129-142.

Yates, F. (1937). “The Design and Analysis of Factorial Experiments,” *Imperial Bureau of Social Science*, Technical Communication 35, Harpenden, England.

Yong, V. and Hahn, H. T. (2005). “Dispersant Optimization using Design of Experiments for SiC/Vinyl Ester Nanocomposites,” *Nanotechnology*, 16, 354-360.

Zhang, T., Xu, Z., Tao, D. L., Teng, F., Li, F. S., Zheng, M. J., and Xu, X. R. (2005). “The Influence of ZnO Nanorods on the Poly-(3, 4-Ethylenedioxythiophene): Poly-(Styrenesulphonic Acid) Buffer Layer in a Polymer Light-Emitting Diode,” *Nanotechnology*, 16, 2861-2864.

VITA

Navinchandra Acharya was born in Mumbai, India on November 19, 1977. He completed his Bachelor's degree in Biomedical Engineering in 2000 from Thadomal Shahani Engineering School, which is a part of Mumbai University, India. He received his Masters degree in Biomedical Engineering from Rensselaer Polytechnic Institute, Troy, NY. During his Master's degree, he worked as an intern at General Electric Medical Systems in Waukesha, WI. Following his Master's degree, he pursued his Ph.D. in Industrial Engineering at The Pennsylvania State University, State College, PA. Navinchandra Acharya has published a book chapter and has submitted four research papers based on this dissertation.