NTNU
Norwegian University of
Science and Technology

A Comparison of Factor Based Methods for Analysing Some Non-regular Designs

Ebenezer Afrifa-Yamoah

Master of Science in Mathematics (for international students)
Submission date: June 2016
Supervisor: John Sølve Tyssedal, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

To my lovely parents,

Howard and *Agnes*.

Summary

Non-regular designs have nice properties regarding run economy. However, standard methods of analysing regular designs are not applicable as a results of possible non-orthogonal contrast columns. We investigated three factor based methods of analysing non-regular factorial fractional designs and also performed follow-up runs in identifying the active factor subspace for an experiment. We studied the six factor 12-run PB and 16-run designs with some simulated models and also on a real data set from the metal cutting experiment by Garzon (2000). In our investigation, the 16-run design displayed a relatively significant performance in defining factor activities for models with four active factors over the 12-run PB design. The methods studied in this thesis were found to produce similar results in identifying one, two and three active factors. All the methods performed very well in identifying models with at most three active factors. However, for models with four active factors, the study revealed that the methods have shortcomings in identifying the correct active subspace. The Box-Meyer search estimated variance was lower compared to that of the other two methods. The projection based method is very simple to use, with much less intuition and was robust under various conditions of model's variability. It is not appropriate to use the method alone whenever the results indicate that three active factors are insufficient. However, follow-up experiments help to improve performance of the method. This study recommend the use of the factor based methods in defining factor activities for experiments.

Preface

In this master thesis, three factor based methods were investigated in identifying active factors in experiments using 12-run PB and 16-run designs. This is the result of the course MA3911 - Master Thesis in Mathematical Sciences, where the final result counts for 45.0 credits. The work was carried out during the fall semester and spring semester of 2015/16 academic year under the supervision of Prof. John S. Tyssedal. The completion of this master thesis concludes my Master of Science in Mathematics at the Norwegian University of Science and Technology (NTNU).

I thank John for his contributions and guidance during the supervision of this master thesis. His feedback and constructive discussions improved this thesis and my understanding. A greater source of my motivation with this master thesis has been his keen interest in the topic of study. I am grateful to the Norwegian Government for the Quota Scholarship Scheme, under which I studied this master degree. A special thanks to the lecturers and staff of the Department of Mathematical Sciences, NTNU for the cordial learning atmosphere during my studies. I also thank the staff of the International Office, NTNU for the support.

Table of Contents

List of Tables

List of Figures

Abbreviations

MEC	=	Maximum Estimation Capacity
2fi's	=	two-factor interactions
PB	=	Plackett-Burman
AF	=	Active Factors
JM	=	Jones-Montgomery
MD	=	Model Discrimination
DoE	=	Design of Experiment
ANOVA	=	Analysis of Variance

Chapter 1

Introduction

Experimentation has become the bedrock of understanding several phenomena of life and processes. Design and analysis of experiments have been incorporated into the scientific fields of enquiries, whereby processes of diverse complexities are simplified to promote efficiency and high production quality. Researchers are mostly interested in the influence of a set of factors, measured at some specified levels, on the response during experimentation. In the process industries, modeling and optimization (instead of the treatment comparison, which is the main objective in agricultural experiments), may be the main focus.

Scientific investigations mostly study the effects of many factors simultaneously. However, when constrained with limited time and resources, it is important to identify the factors that are highly influential. In many situations, several factors are presumed to influence both the location and dispersion of processes, but, normally a few of them are really vital; a condition referred to as factor sparsity. When an experiment is conducted, the factors with a significant amount of impact on the response are called active, otherwise inert. To identify these active factors requires efficient screening designs. A full factorial experiment allows all factorial effects to be estimated independently and is commonly used in practice. However, it is often too costly to perform this experiment when a large number of factors are involved. For example, an experiment involving 8 factors will require $2^8 = 256$ runs. As a results, interest in factorial designs with efficient properties of investigating several factors with fewer runs is on the rise.

A group of saturated experimental designs that are known to investigate several factors in fewer runs is the Plackett-Burman designs proposed by Plackett and Burman (1946), hereafter called PB designs. A design with s-levels and k factors is called saturated if the number of runs are equal to $k(s-1)+1$. For a two-level saturated design, this means that k factors can be investigated in $k+1$ runs. For instance, the 12-run PB design has the ability to investigate eleven factors in twelve runs and is one example of a non-regular factorial design employed in this thesis. The 12-run PB design is of projectivity, $P = 3$ and has the property that all $\binom{11}{3} = 165$ projections onto three factors are of just one type, that is, all the projections give a full 2^3 design plus the very best half fraction of a 2^3 design (Tyssedal and Niemi, 2014). Also, any of the $\binom{11}{4} = 330$ projections onto four factors can be obtained from an arbitrarily other selected projection onto four factors by interchanging rows, columns, or signs in columns, and all of them allow all main effects and two-factor interactions to be estimated. In particular, this means that for a model with k factors, $k \leq 4$, it is the same chance that these factors will be identified regardless of which factors they are. This very fair treatment of any set of k factors, $k \leq 4$, is a property that no other PB design has. Hence, the 12-run PB design has very attractive screening properties.

Statisticians are keenly interested in understanding the most efficient way of analysing this PB design. Standard methods of analysing regular designs are not applicable as a results of possible non-orthogonal contrast columns. This is because the problem of allocating individual effects to large contrast can easily occur without notice as a results of aliasing of effects. Several methods have been proposed in literature and they are broadly classified as factor-based or effect based search procedures. According to Tyssedal and Niemi (2014) the goal of a factor-based search is to identify the subspace of active factors of normally low dimension—typically 2, 3, or 4—within which most of the changes in the measured response occur. Then the functional relationship between the response and the factors may be investigated afterwards. The performance of such a procedure depends heavily on the projective properties of the design used. Factor based methods are less dependent on model assumptions. Examples of such methods can be found in Box and Meyer (1993), Tyssedal and Samset (1997), and Kulachi and Box (2003). On the other hand, the search for the most likely active main effects and two-factor interactions is the main focus of effect-based procedures. Examples of effect-based procedures can be found in the works of Hamada and Wu (1992) and Chipman et al. (1997). The assumptions of effect sparsity and heredity are strong guidance in the search for active effects. The heredity principle requires excluding an interaction to be in a model unless at least one (weak heredity) or both (strong heredity) of the parent main effects also are included in the model.

This thesis aims at investigating factor based methods of analysing non-regular factorial fractional designs and follow-up runs in identifying active factors in experiments. The 12-run PB and 16-run (by Jones and Montgomery (2010)) designs are the non-regular designs employed for this study. Three factor based methods of analysing designs are discussed and their analysis compared. The projective based approach by Tyssedal and Samset (1997), which exploits the projective properties of designs is compared with the Bayesian techniques proposed by Box and Meyer (1993) and the partial F approach introduced by Kulachi and Box (2003). Simulated models and a six factor metal cutting experiment conducted by Garzon (2000) are utilized in this thesis. The contribution of this thesis is two-fold: a) to study and compare the performance of these factor based techniques in analysing non-regular fractional factorial designs and b) to compare the performance of the 12-run PB and the 16-run designs in identifying active subspace.

This thesis is composed of five chapters, references and appendices. Chapter one introduces the research problem, research objectives and the significance of the research. Chapter two covers the theories relevant for the development of the methods used and reviews relevant literature. Chapter three presents the methods employed in this thesis. In chapter four the results of the analysis are presented and discussed. The summary and conclusions are presented in chapter five.

Chapter 2

Theory

2.1 Experimental Designs

Experimental design is a body of knowledge and techniques that enables an investigator to conduct better experiments, analyze data efficiently, and make the connections between the conclusions from the analysis and the original objectives of the investigation (Wu and Hamada, 2000). According to Telford (2007), it is a series of tests in which purposeful changes are made to the input variables of a system or process and the effects on response variables are measured. It dates back to the work of R. A. Fisher in the 1920s and 1930s at the Rothamsted Agricultural Experimental Station in the United Kingdom. Although the experimental design method was first used in an agricultural context, the method has been applied successfully in the military and in industry since the 1940s. To be mentioned is the work of Besse Day, at the U.S. Naval Experimentation Laboratory. Experimental design was employed there to establish the cause of bad welds at a naval shipyard during World

War II. George Box developed experimental design procedures for optimizing chemical processes. W. Edwards Deming taught statistical methods, including experimental design, to Japanese scientists and engineers in the early 1950s at a time when “Made in Japan” meant poor quality. Genichi Taguchi, the most well known of this group of Japanese scientists, is famous for his quality improvement methods. One of the companies where Taguchi first applied his methods was Toyota. Since the late 1970s, U.S. industry has again become interested in quality improvement initiatives, now known as “Total Quality” and “Six Sigma” programs. Design of experiments is considered an advanced method in the Six Sigma programs, which were pioneered at Motorola and General Electrics (GE).

Fisher (1935) demonstrated how valid conclusions could be drawn efficiently from experiments with natural fluctuations such as temperature, soil conditions and rainfall. Such variables are known as nuisance variables and could be known or unknown. The known nuisance variables usually cause systematic bias (e.g., batch-to-batch variation) whiles the unknown nuisance variables usually cause random variability in the results and are called inherent variability or noise. The problems posed by the two types of nuisance factors are addressed by the fundamental principles in design of experiments. The fundamental principles are randomization, replication, blocking, orthogonality and factorial experimentation. The randomization principle is a mean of protecting against unknown biases that distort results of experiments. By replication, the sample size increases as a means of improving the precision level of the experiment. Blocking is a method for increasing precision of an experiment by removing the effects of known nuisance factors. Orthogonal in design columns results in factor effects being uncorrelated and therefore can be more easily interpreted.

2.2 Factorial Experimental Designs at Two levels

Factorial designs are mostly used to investigate the effects of two or more factors simultaneously. They are either regular or non-regular. Full factorial two-level designs require $2 \times 2 \times \cdots \times 2 = 2^k$ runs, and thus are mostly referred to as 2^k *factorial designs*. The two factor levels are used to study first order and interaction effects of the response over a range of chosen factor levels. Factor effects can be estimated independently, however, many factors require that a large number of runs have to be performed. In some cases, it may be sufficient to perform only a fraction of the experimental runs because some factors may be assumed inert. The number of runs is then equal to 2^{k-p}, where p is the fraction of the design. When p equal to one, the design is referred to as a half fraction of the full factorial. These designs are known as fractional factorial designs. In these designs, effects are either estimated independently or fully aliased. However, the economy of run size makes 2^{k-p} designs often preferred compared to the full factorial designs. For illustration, consider a 2-level design with 5 factors. The full factorial design would require $2^5 = 32$ runs, whiles the half fraction, 2^{5-1}, would require 16 runs and are shown in Table 2.1 and Table 2.2 respectively.

The 2^{k-p} fractional design is constructed using p generators. A generator is the relation where the aliased effects are set equal to each other. This ensures that the signs in the design columns are equal. For instance, if the signs in the column for C are equal to the

Table 2.1: A full factorial 2^5 design

Run	A	B	C	D	E
1	-	-	-	-	-
2	+	-	-	-	-
3	-	+	-	-	-
4	+	+	-	-	-
5	-	-	+	-	-
6	+	-	+	-	-
7	-	+	+	-	-
8	+	+	+	-	-
9	-	-	-	+	-
10	+	-	-	+	-
11	-	+	-	+	-
12	+	+	-	+	-
13	-	-	+	+	-
14	+	-	+	+	-
15	-	+	+	+	-
16	+	+	+	+	-
17	-	-	-	-	+
18	+	-	-	-	+
19	-	+	-	-	+
20	+	+	-	-	+
21	-	-	+	-	+
22	+	-	+	-	+
23	-	+	+	-	+
24	+	+	+	-	+
25	-	-	-	+	+
26	+	-	-	+	+
27	-	+	-	+	+
28	+	+	-	+	+
29	-	-	+	+	+
30	+	-	+	+	+
31	-	+	+	+	+
32	+	+	+	+	+

signs for DE, then C and DE are aliased, and $C = DE$ is the generator for the design. Multiplying both sides by C yields the defining relation $I = CDE$. Another property of interest is the resolution, which is equal to R if all p-factor effects are aliased with effects comprising $R-p$ factors or more. The resolution is in simple terms equal to the number of letters in the shortest word in the defining relation. The design given in Table 2.2 has as its generator the relation $E = ABCD$. The defining relation thus becomes $I = ABCDE$. This design is a resolution V design.

Table 2.2: A 2^{5-1} fractional factorial design

Run	A	B	C	D	E
1	-	-	-	+	-
2	-	-	-	-	+
3	-	-	+	+	+
4	-	-	+	-	-
5	-	+	-	+	+
6	-	+	-	-	-
7	-	+	+	+	-
8	-	+	+	-	+
9	+	-	-	+	+
10	+	-	-	-	-
11	+	-	+	+	-
12	+	-	+	-	+
13	+	+	-	+	-
14	+	+	-	-	+
15	+	+	+	+	+
16	+	+	+	-	-

The analysis of such designs relies more heavily on regression modelling. In the regression model,

$$Y = X\beta + \epsilon \tag{2.1}$$

the design matrix **X**, plays a very important role in applications. For orthogonal columns of the design, the vector of estimators for the coefficients is given by;

$$\widehat{\beta} = (X'X)^{-1}X'Y = \begin{bmatrix} \frac{1}{n} & 0 & 0 & 0 \\ 0 & (x_1{}'x_1)^{-1} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & (x_k{}'x_k)^{-1} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ x_1{}'Y \\ \vdots \\ x_k{}'Y \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} \frac{Y_i}{n} \\ (x_1{}'x_1)^{-1}(x_1{}'Y) \\ \vdots \\ (x_k{}'x_k)^{-1}(x_k{}'Y) \end{bmatrix} \tag{2.2}$$

In an experiment, one chooses values for the explanatory variables (factors) $x_1, x_2, \cdots, x_k$ such that they are as favourable for the estimation as possible. In practice, factor levels are mostly recoded as -1 and 1 during the analysis. The process at high level is represented by 1 and the low level is represented by -1. This formulation yields orthogonal factor columns and computation of the coefficients for main effects and interaction effects are made easy. Main effects and interactions are defined according to Tyssedal (2011). For two-level designs the main effect of a factor is the expected average response when the factor is on the high level minus the expected average response when the factor is at the low level. The interaction between two factors is defined as half the main effect of a factor when the other is on the high level minus half the main effect of a factor when the other factor is at its low level. From the design presented in Table 2.2, we could formulate the design matrix, **X** as follows;

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \end{bmatrix}$$

2.2.1 Regular Designs

The class of 2^k designs and 2^{k-p} designs are called regular designs. A regular fractional factorial design can be specified in terms of a set of defining contrasts. A 2^{k-p} regular fractional factorial design is constructed from the full factorial design generated from the $k - p$ basic factors. Its aliasing structure is explicitly described by the defining contrast group (Wu and Hamada, 2000) generated by the p generators. Their alias structures are relatively simple in which a particular alias term (main effect or interaction) appears only once in association with a single data contrast, with a coefficient which is either 1 or -1. The regular design can be used to screen $N/2$ factors at projectivity $P = 3$ (Tyssedal et al., 2006). More discussion on projectivity in Section 3.3.

Suppose a two-level factorial design with 5 factors has to be performed in 8 runs. That is, the design of interest is a 2^{5-2} regular fractional factorial design. The 3 basic factors in a 2^{5-2} fractional factorial design are the three independent factors (A, B, C) of the base factorial design (a 2^3 full factorial design). The two added factors (D, E) are assigned to columns chosen from the remaining columns of the model matrix for the base factorial design. One possible assignment is $D = AC$ and $E = BC$. That is, the level settings of D and E are determined by the columns corresponding to AC and BC, respectively. Let I be the identity element (or, the column of 1's for the mean). Then, $I = ACD$ and $I = BCE$ are called the fractional generators. From every p independently chosen fractional generators, $2^p - p - 1$ more relations are derived. For example, $I = ABDE$ is derived from $I = ACD$ and $I = BCE$. The entire set of $2^p - 1$ relations,

$$I = ACD = BCE = ABDE,$$

forms the defining contrast subgroup, and the terms ACD, BCE and $ABDE$ are called words. The number of factors in a word is called the length of a word (or word-length).

Two distinct sets of fractional generators (or equivalently, defining contrast subgroups) generate distinct 2^{k-p} fractions of a 2^k full factorial design. That further introduces the notion of ranking among different 2^{k-p} fractions of a 2^k full factorial design. The ranking criteria are generally based on some operating assumptions that are common to many experiments:

- The effect sparsity principle: only a few effects in a factorial experiment are likely to be significant.
- The hierarchical ordering principle: lower order effects are more likely to be significant than higher order effects.

- The effect heredity principle: interactions involving significant main effects are more likely to be active than other interactions.

- Minimum aberration: For any two 2^{k-p} designs t_1 and t_2, let m be the smallest integer such that $B_m(t_1) \neq B_m(t_2)$, where $B_m(t_1)$ is the number of defining words of length m in the defining relation of design t_1. Then t_1 is said to have less aberration than t_2 if $B_m(t_1) < B_m(t_2)$. If there is no design with less aberration than t_1, then t_1 has minimum aberration.

The analysis of regular designs are straightforward and easy to perform. To identify active contrasts in regular designs without replicated rows, standard methods such as normal and half plot (Daniel, 1976) and Lenth's method (Lenth, 1989) are commonly used. The basis for the normal plot is that inert effects should normally be distributed with zero means and equal variances. The drawback of regular designs is that they only exist for the number of runs equal to a power of two. As a consequence, these designs are less economic and time efficient compared to some of the members of the non-regular designs. Also,effects are fully confounded, if the number of factors exceed p. However, their alias structure are relatively short compared to their non-regular counterparts.

2.2.2 Non-regular Design

Two-level designs that are not a 2^{k-p} design, are said to be non-regular. Li et al. (2003) defined a non-regular design as the one whose columns do not form an elementary Abelian group. An Abelian group, also called a commutative group, is a group in which the result of applying the group operation to two group elements does not depend on the order in which they are written. Non-regular designs such as PB designs and other orthogonal arrays are widely used in various screening experiments for their run size economy and flexibility (Wu and Hamada, 2000). Unlike regular designs, non-regular designs may exhibit a complex aliasing structure, that is, a large number of effects may neither be orthogonal nor fully aliased, which makes it difficult to interpret their significance. For instance, the 12-run PB design has every main effect potentially partially aliased with 45 two-factor interactions and a single two-factor interaction appears in the alias pattern of all main effect not involved with this two-factor interaction. For this reason, non-regular designs were traditionally used to estimate factor main effects only, but not their interactions. However, in many practical situations it is often questionable whether the interaction effects are negligible. Hamada and Wu (1992) demonstrated that some interactions could be entertained and estimated through their complex aliasing structure. They argued and justified that ignoring interactions can result in important effects being missed, spurious effects being detected, and estimated effects having reversed signs resulting in incorrectly recommended factor levels. One advantage of non-regular designs is their projective properties and that also they exist when the number of run, N, is a multiple of four.

2.3 Plackett-Burman Design

The evolution of non-regular designs came to light when Robin L. Plackett and J. P. Burman in 1946 while working in the British Ministry of Supply with a goal of finding experimental designs for investigating the dependence of some measured quantity on a number of independent variables (factors), each taking s levels, in such a way as to minimize the variance of the estimators of these dependencies using a limited number of experiments.

Plackett and Burman (1946) gave a large collection of two-level and three-level designs for multi-factorial experiments. These designs are often referred to as PB designs in the literature.

PB designs are saturated orthogonal design (fractional factorial) constructed on the basis of fractional replicates of a full factorial design (Montgomery, 2001). Further they are based on balanced incomplete blocks and can in N experiments (N number of runs) study $(k = N - 1)$ process variables, where N is a multiple of 4. For obtaining an orthogonal design matrix, the following conditions are necessary and sufficient:

1. The number of times each factor is adjusted to each of its levels must be the same;
2. The number of times every two factors, each at any one of its levels are encountered, must be the same;
3. The number of observations must be divisible by the square of the number of levels, defined as: $N = nl^2$

where n is an integer. When the above-stated conditions are available, the construction of an orthogonal matrix (experimental design) requires combinatorial operations only.

The PB design is probably the most well known non-regular designs. Plackett and Burman (1946) only included designs with $N \leq 100$, and they also omitted the design where $N = 92$. For PB designs where the number of runs is equal to a power of two the designs coincide with the regular ones, and the rest of the PB designs are non-regular. The 12-run PB design matrix is shown in Table 2.3.

Table 2.3: The 12-run Plackett and Burman design

Run	A	B	C	D	E	F	G	H	I	J	K	Observation
1	+	+	-	+	+	+	-	-	-	+	-	y_1
2	+	-	+	+	+	-	-	-	+	-	+	y_2
3	-	+	+	+	-	-	-	+	-	+	+	y_3
4	+	+	+	-	-	-	+	-	+	+	-	y_4
5	+	+	-	-	-	+	-	+	+	-	+	y_5
6	+	-	-	-	+	-	+	+	-	+	+	y_6
7	-	-	-	+	-	+	+	-	+	+	+	y_7
8	-	-	+	-	+	+	-	+	+	+	-	y_8
9	-	+	-	+	+	-	+	+	+	-	-	y_9
10	+	-	+	+	-	+	+	+	-	-	-	y_{10}
11	-	+	+	-	+	+	+	-	-	-	+	y_{11}
12	-	-	-	-	-	-	-	-	-	-	-	y_{12}

There are three basic methods for constructing these Plackett and Burman designs (Lin and Draper, 1992):

1. *Cyclic Generation*. Take a (specific) row of $N - 1$ plus and minus signs, provided by Plackett and Burman (1946). Construct $N - 2$ further rows by cyclicly permuting the signs in the first row. Add a row of all minus signs. This gives N rows (=runs) of ± 1 levels for $N - 1$ variables or factors (= columns). The $N = 12$ case, shown in 2.3, is developed in this manner. So are the designs for $N = 8, 16, 20, 24, 32, 36, 44, 48, 60, 68, 72, 80$ and 84.

2. *Doubling*. A block of plus and minus signs which we denote by **D** is given. The design is obtained by writing down;

$$\begin{bmatrix} u & D & D \\ -u & -D & D \end{bmatrix}$$

   where u denotes a unit column of all plus signs. Design for $N = 40, 56, 64, 88$ and 96 are obtained in this manner from those of 20, 28, 32, 44 and 48 runs, respectively.

Note that this method can be applied for any $N-$run design when N is a multiple of eight and an $(\frac{1}{2}N)$-run design is available.

3. *Block Permutation.* Several square blocks of plus and minus signs are given. Further rows are obtained by cyclic permutation of the blocks. A row of minus signs is then added. Designs for $N = 28, 52, 76$ and 100 are of this type.

2.3.1 Projection Properties of Plackett-Burman Designs

Box and Tyssedal (1996) defined a design to be of projectivity p if the projection onto every subset of p factors contains a full factorial design in p factors, possibly with some points replicated. It follows from these definitions that an orthogonal array of strength t is of projectivity t. The determination of the projectivity of an orthogonal two-level array were proven in three propositions by Box and Tyssedal (1996).

1. A saturated design obtained from a doubled $n \times n$ Hadamard matrix is always of projectivity $P = 2$ and only 2.

2. A saturated design obtained from cyclic orthogonal array is either a geometric factorial orthogonal array with $P = 2$ and only 2, or else has projectivity at least $P = 3$.

3. Any saturated two-level design obtained from an orthogonal array containing $n = 4m$ runs, with m odd, is of projectivity at least $P = 3$.

PB designs are saturated orthogonal arrays of strength two and all degrees of freedom are utilized to estimate main effects. An orthogonal array of N runs, m factors, s levels and strength t, denoted by $OA(N, s^m, t)$, is an $N \times m$ matrix in which each column has s symbols that appear equally often in the matrix. For example, the 12-run PB design in

Table 2.3 is an $OA(12, 2^{11}, 2)$. Orthogonal arrays of strength two allow all the main effects to be estimated independently and they are universally optimal for the main effects model (Cheng, 1980). A necessary condition for the existence of an $OA(N, s^m, 2)$ is that $N-1 \geq m(s-1)$. A design is called saturated if $N-1 = m(s-1)$ and supersaturated if $N-1 < m(s-1)$.

Orthogonal arrays include both regular and non-regular designs. For regular designs, the concepts of strength and resolution is that a design of resolution R is an orthogonal array of strength $t = R-1$. Design resolution measures the interdependence in effects in fractional factorial design. In other words, it describes how much the effects in a fractional factorial design are aliased with other effects. For fractional factorial design, one or more of the effects are confounded, meaning they cannot be estimated separately from each other. Resolution III, IV, and V designs are most common. The usual practice has been to use a fractional factorial design with the highest possible resolution. This is because higher resolution implies greater design strength. For example, it is usually better to choose a design where main effects are confounded with 3-way interactions (Resolution IV) instead of a design where main effects are confounded with 2-way interactions (Resolution III). For a regular design of resolution R, the projection onto any R factors must be either a full factorial or copies of a half-replicate of a full factorial. The projections for non-regular designs are more complicated.

PB designs are completely classified with respect to $R = 3$ and $R = 4$. Designs $N = 68, 72, 80$ and 84 are resolution IV. PB designs are of strength two, so the projection onto any two factors is a replicated full factorial. Lin and Draper (1992) studied the geometrical projection properties of the PB designs onto three or more factors. Their computer searches found all the projections of the $12-, 16-, 20-, 24-, 28-, 32-$ and $36-$run PB designs onto three factors. They found that these projections must have at least a copy of the full 2^3 factorial or at least a copy of a 2^{3-1} replicated or both. In particular, any projection onto three factors must contain a copy of a full factorial except for the $16-$ and $32-$run PB designs, which are regular designs. The important statistical implication of this finding is that if only at most three factors are truly important, then after identifying the active factors, all factorial effects among these active factors are estimable, regardless which three factors are important.

The 12-run PB design projects onto six replicates of a 2^1 design in one dimension, three replicates of a 2^2 design in every two dimensions and one-and-half replicate of a 2^3 design in every three dimensions. The 12-run PB design is of projectivity three. Wang and Wu (1995) found that its projection onto any four factors has the property that unbiased estimates are available for all the main effects and two-factor interactions if the higher-order interactions are negligible. If only a small subset of the factors are active it is of importance to know how well a design projects onto such a small subset. Also, if a subset of factors contains the only active factors, the difference between expected values in replicated runs is equal to zero. Hence, estimates of the variance within each group of replicated runs should be model independent.

Wang and Wu (1995) defined a design as having a hidden projection property if it allows some or all interactions to be estimated even when the projected design does not have the right resolution or other geometrical design property for the same interactions to be estimated. For the PB designs their hidden projection property is a result of complex aliasing between interactions and main effects. For instance, the 12-run PB design has any two-factor interaction, say CD to be orthogonal to main effects C and D, and partially aliased with all other main effects with correlation $\frac{1}{3}$ and $-\frac{1}{3}$. Thus, it is possible to estimate four main effects and all six two-factor interaction among them together.

2.4 16-runs Designs

Johnson and Jones (2011) discussed a classical-type construction of the 16-run design with 6, 7 and 8 factors for both regular and non-regular design with a 2^4 or a replicated 2^3 starting point. Additional factor columns were defined using familiar one term column generators or generators using weighted sums of effects. The construction was built around a design scheme proposed by Jones and Montgomery (2010) and hereafter called JM designs. For six factors, Jones and Montgomery (2010) compared the performance of the regular and non-regular designs in 16-runs for the "Photoresist" experiment in Montgomery (2001). This example established the superiority of the non-regular design in Table 2.5 to the 2^{6-2} resolution IV design (a regular design) in Table 2.4. Using the 2^{6-2} fractional factorial design, Montgomery (2001) identified factors A, B, C and E as active and one active two-factor interaction AB, CE or a combination of both, with additional runs required in order to break this alias. The same situation in the context of a specific non-regular design was considered by Jones and Montgomery (2010) through simulation.

CE was considered as active two-factor interaction in their setup. Their analysis identified the four active factors and the one and only active two-factor interaction CE.

The 2^{6-2} design was constructed having A,B,C and D as base factors. Factors E and F were set to $E = ABC$ and $F = BCD$ and thus the defining relation become $I = ABCE = BCDF = ADEF$. The design is as presented in Table 2.4.

Table 2.4: Photoresist Design

Run	A	B	C	D	E	F	Thickness
1	-1	-1	-1	-1	-1	-1	4524
2	1	-1	-1	-1	1	-1	4657
3	-1	1	-1	-1	1	1	4293
4	1	1	-1	-1	-1	1	4516
5	-1	-1	1	-1	1	1	4508
6	1	-1	1	-1	-1	1	4432
7	-1	1	1	-1	-1	-1	4197
8	1	1	1	-1	1	-1	4515
9	-1	-1	-1	1	-1	1	4521
10	1	-1	-1	1	1	1	4610
11	-1	1	-1	1	1	-1	4295
12	1	1	-1	1	-1	-1	4560
13	-1	-1	1	1	1	-1	4487
14	1	-1	1	1	-1	-1	4585
15	-1	1	1	1	-1	1	4195
16	1	1	1	1	1	1	4518

The non-regular design by Jones and Montgomery (2010) in Table 2.5 was constructed by first considering all 16 possible combinations of A, B, C and D. Hence, their design is a 2^4 full factorial design in A, B, C and D. Factors E and F were generated using the weighted sums of effects.

$$E = 1/2(AC + BC + AD - BD); F = 1/2(-AC + BC + AD + BD)$$

Johnson and Jones (2011) thought of this construction as a "principal" quarter fraction and gave the other three quarters fractions as;

$$E = 1/2(AC + BC + AD - BD); F = -1/2(-AC + BC + AD + BD)$$

$$E = -1/2(AC + BC + AD - BD); F = 1/2(-AC + BC + AD + BD)$$

$$E = -1/2(AC + BC + AD - BD); F = -1/2(-AC + BC + AD + BD)$$

Table 2.5: Non-regular alternative

Run	A	B	C	D	E	F	Thickness
1	1	1	1	1	1	1	4494
2	1	1	-1	-1	-1	-1	4592
3	-1	-1	1	1	-1	-1	4357
4	-1	-1	-1	-1	1	1	4489
5	1	1	1	-1	1	-1	4513
6	1	1	-1	1	-1	1	4483
7	-1	-1	1	-1	-1	1	4288
8	-1	-1	-1	1	1	-1	4448
9	1	-1	1	1	1	-1	4691
10	1	-1	-1	-1	-1	1	4671
11	-1	1	1	1	-1	1	4219
12	-1	1	-1	-1	1	-1	4271
13	1	-1	1	-1	-1	-1	4530
14	1	-1	-1	1	1	1	4632
15	-1	1	1	-1	1	1	4337
16	-1	1	-1	1	-1	-1	4319

2.4.1 Six-factor non-regular designs in 16 runs

Johnson and Jones (2011) from the JM design presented each of the 27 non-isomorphic 16-run, six-factor designs. Two or more designs are said to be isomorphic if one can be obtained from the other by relabeling the factors having the same number of levels, reordering the factor combinations and/or relabeling the levels of one or more factors.

Otherwise they are non-isomorphic. With a 2^4 full factorial design in A,B,C and D as the starting point, construction of 21 designs are possible. The construction schemes are categorized into three.

1. Classical: Under this scheme the factors E and F are confounded with two-, three-, or four-way interaction of A,B,C and D. Designs constructed are presented in Table 2.6.

Table 2.6: Classical construction scheme

Design	E	F
1	BC	ABC
2	AB	CD
3	AB	ACD
4	ABC	ABD

2. Hybrid: This scheme considers the situation where factor E is confounded and F correlated with two-, three-, or four-way interaction of A,B,C and D. Designs constructed are presented in Table 2.7.

Table 2.7: Hybrid construction scheme

Design	E	F
5	AB	$1/2[CD + ACD + BCD - ABCD]$
6	AC	$1/2[CD + ACD + BCD - ABCD]$
7	ABC	$1/2[CD + ACD + BCD - ABCD]$
8	AB	$1/2[AD + BD + CD - ABCD]$
9	AB	$1/2[AC + BC + AD - BD]$
10	BCD	$1/2[BD + ABD + CD - ACD]$
11	$ABCD$	$1/2[BD + ABD + CD - ACD]$
12	ABC	$1/2[AD + BD + ABCD - CD]$
13	ABD	$1/2[AD + BD + CD - ABCD]$
14	$1/2[AC + BC + BCD + AD - BD]$	$1/2[AC + BC - AD + BD]$
15	$1/2[AC + BC + BCD + AD - BD]$	$1/2[-AC + BC + AD + BD]$
16	$1/2[BD + ABD + BCD - ABCD]$	$1/2[BD - ABD + CD + ACD]$
17	$1/2[AD + BD + ACD - BCD]$	$1/2[AD - BD + CD + ABCD]$
18	$1/2[AD + ABD - CD + BCD]$	$1/2[AD + BD + ACD - BCD]$
19	$1/2[AC + ABC + AD - ABD]$	$1/2[AC + BC - AD + BD]$
20	$1/2[AC + ABD - CD + BCD]$	$1/2[AD + BD + ACD - BCD]$
21	$1/2[AB + AC - BD + CD]$	$1/2[AC + BC - AD + BD]$

3. Correlated: Under this scheme factor E and F are correlated with two-, three-, or four-way interaction of A,B,C and D. Designs constructed are presented in Table 2.7, elements 14 to 21.

The rest of the 6 cases of the construction start with a 2^3 full factorial design in A, B and C with various replication options for each of the designs. Designs constructed are presented in Table 2.8.

Table 2.8: Construction starting with a 2^3 full factorial design in A, B and C

Design	*D*	*E*	*Replication 1*	*Replication 2*
22	AC	ABC	$F = AB$ across both replicates	
23	AB	AC	$F = BC$	$F = ABC$
24	ABC		$E = AC; F = BC$	$E = BC; F = AC$
25	AB		$E = AB; F = BC$	$E = ABC; F = AC$
26	ABC		$E = AB; F = BC$	$E = BC; F = AC$
27			$D = AB; E = AC; F = ABC$	$D = AC; E = ABC; F = BC$

In this thesis, the 16-run design used is as presented in Table 2.9. The hybrid construction scheme is used in its construction with generators $E = 1/2(AD - BD + AC + BC)$ and $F = ABCDE$. This implies that $I = 1/2(ADE - BDE + ACE + BCE) = ABCDEF = 1/2(ADF + BDF - ACF + BCF)$. From the design construction, AB, CD and EF are free of aliasing with main effects. The alias structure for the main effect is as presented in Table 2.10. The main effects are orthogonal and no aliasing between two-factor interactions. As a results, three orthogonal subspaces can be investigated independently. The following linear relationships are important to be aware of; $A - B = DE - CF$, $A + B = CE + DF$, $C - D = BE - AF$, $C + D = BF + AE$, $E - F = AC - BD$ and $E + F = AD + BC$. The classical construction alternative to this design is a design with generators, $E = ABC$ and $F = ABD$ which gives

$I = ABCE = ABDF = CDEF.$

Table 2.9: The 16-run Design

Run	A	B	C	D	E	F
1	-1	-1	-1	-1	1	1
2	1	-1	-1	-1	-1	1
3	-1	1	-1	-1	1	-1
4	1	1	-1	-1	-1	-1
5	-1	-1	1	-1	-1	1
6	1	-1	1	-1	-1	-1
7	-1	1	1	-1	1	1
8	1	1	1	-1	1	-1
9	-1	-1	-1	1	1	-1
10	1	-1	-1	1	1	1
11	-1	1	-1	1	-1	-1
12	1	1	-1	1	-1	1
13	-1	-1	1	1	-1	-1
14	1	-1	1	1	1	-1
15	-1	1	1	1	-1	1
16	1	1	1	1	1	1

Table 2.10: Alias structure of the main effects of the 16-run design

Main effect	Aliased interaction
A, B	CE, DE, CF, DF
C, D	AE, BE, AF, BF
E, F	AC, AD, BC, BD

2.5 Data Analysis Methods for Factorial Designs

The methods of analysis of factorial designs can be classified as effect based or factor based. Effect based methods aim at identifying significant effects. The principle of effect heredity is often a precept. However, Tyssedal and Kulachi (2005) proposed an effect based method that does not depend on the heredity principle. Factor-based methods aim at identifying active factors and they are less dependent on model assumptions. Methods proposed by Box and Meyer (1993), Kulachi and Box (2003) and Tyssedal and Samset (1997) are the subjects of interest of this thesis. Both the effect-based and factor-based methods of analysis are viewed and applied from different perspectives. A sharp contrast can be made for the frequentist and the bayesian view. The typical approach to analysis is to calculate a set of contrasts, one for each column of the full design. Then a variety of different methods may be applied: normal probability plots, Daniel (1976); Box et al. (1978); Bayes plots, Box and Meyer (1986a); pseudo-standard error, Lenth (1989) to identify contrasts which are too large to attribute to noise alone. The strategy in analyzing factorial design is to identify as active those factors whose main effects can be associated with large contrasts, discounting the possibility of interactions.

2.5.1 Frequentist Approach

The probably best known frequentist approach is the one by Hamada and Wu (1992). It consists of three steps.

1. Entertain all the main effects and interactions that are orthogonal to the main effects. Use standard analysis methods such as ANOVA and half-normal plots to select significant effects.

2. Entertain the significant effects identified in the previous step and the two-factor interactions that consist of at least one significant effect. Identify significant effects using a forward selection regression procedure.

3. Entertain the significant effects identified in the previous step and all the main effects. Identify significant effects using a forward selection regression procedure.

Iterate between Steps 2 and 3 until the selected model stops changing.

This analysis strategy is based on two assumptions. The first assumption is the validity of the effect sparsity principle (Box and Meyer, 1986a). The second assumption is the validity of the weak effect heredity principle (Hamada and Wu, 1992). A motivational precept for the weak heredity is that it is often difficult to provide a good physical interpretation for a significant interaction AB without either A or B being significant.

2.5.2 Bayesian Approach

The Bayesian approach suggested by Box and Meyer (1993) considers all the possible explanations (models including interactions) of the data from a screening experiment and identifies those that fit the data well. The prior assumptions are as follows:

1. Effects calculated for inactive factors may be represented approximately as items from a normal distribution with mean zero and standard deviation σ.

2. For a proportion π of active factors the resulting effects are represented as item from a normal distribution with mean zero and a larger standard deviation $\gamma\sigma$.

The prior information is represented in two parameters: γ, the ratio of the standard deviation of the active to the inactive effects, and π, the percentage of active factors. Box et al. (2005) suggested to choose γ between 2 and 3 and $\pi = 0.25$, based on a survey of a number of published analyses of factorial designs. Recent study has confirmed that the results are not very sensitive to moderate changes in γ and π when active factors are present. A Bayesian framework is used to assign posterior probabilities to all models considered. Then these posterior probabilities are accumulated to marginal posterior probabilities for each factor. The technical details of the Bayesian analysis are complicated and given in

(Box and Meyer, 1993) or (Box et al., 2005). In practice, one can use the BsProb function in the R library BsMD, free downloadable from the R project homepage (http://www.R-project.org/), for the calculation. Chipman et al. (1997) proposed an effect based Bayesian approach that employs a Gibbs sampler to perform an efficient stochastic search of the model space.

2.5.3 Other Techniques Developed

There are further sophisticated analysis strategies proposed for experiments with complex aliasing. A projective based approach by Tyssedal and Samset (1997), which exploits the projective properties of the 12-run PB design has been proposed. Many other recent variable selection methods can also be used for similar purposes. For example, Yuan et al. (2007) suggested an extension of the general-purpose LARS (least angle regression), first proposed by Efron et al. (2004). Phoa et al. (2009) suggested the Dantzig Selector method, first proposed by Candes and Tao (2007), for factor screening.

2.6 Estimation and Confounding

Estimation of effects (main and interactions) is a major component in the analysis of experiments. When all interactions are assumed 'negligible', the estimation of main-effects is done as follows. For the $i-$th factor, attach to the entries in the response column the sign in the $i-$th column and divide the total by the divisor $N/2$, which is the number of plus signs in column i. For example, for the 12-run case of Table 2.3, the main effect of factor A is $l_A = (y_1 - y_2 + y_3 - y_4 - y_5 - y_6 + y_7 + y_8 + y_9 - y_{10} + y_{11} - y_{12})/6$.

The overall mean effect is obtained in a similar manner using the column u of plusses and the divisor 12 (N, in general). The estimated effects can also be seen from fitting the model $y_j = \beta_0 + \sum_{i}^{11} \beta_i + \epsilon_j$, $j = 1, 2, ..., 12$ by the standard least squares calculation $\widehat{\beta} = (X'X)^{-1}X'y = \frac{1}{12}X'y$, where X is a 12×12 matrix formed by a column of 1's adjoined to the block of ± 1 defined by Table 2.3, and y is a 12×1 column of y_i's. The main effect of factor A, l_A, is double the value of $\widehat{\beta}_A$, the $(i+1)th$ element of $\widehat{\beta}$. This is because $\widehat{\beta}_A$ measures a one unit effect, while l_A measures the two-unit effect from $x_A = -1$ to $x_A = 1$.

If we suppose that some interactions are not zero, the quantity $\widehat{\beta} = (X'X)^{-1}X'y$ will estimate the main-effect of factor A plus a linear combination of certain two-factor interactions. The alias structure can be obtained using calculations suggested by Box and Wilson (1951). Suppose we wish to fit the regression model,

$$E(y) = X\beta \tag{2.3}$$

by least squares. If the model is correct, then the estimator $\widehat{\beta} = (X'X)^{-1}X'y$ is unbiased. However, if the model is not correct, the estimator is biased. If the correct model takes the form;

$$E(y) = X\beta + X_A\beta_A \tag{2.4}$$

then the estimator of the regression coefficient vector is given by the estimator $\widehat{\beta} = (X'X)^{-1}X'y$, and it follows that;

$$E(\widehat{\beta}) = \beta + C\beta_A \tag{2.5}$$

where $C = (X'X)^{-1}X'X_A$. C is called the alias matrix or bias matrix. X_A is dependent on the true model which in general is unknown. The β consists of all the main effects, β_A consists of all two-factor interactions, and the alias matrix C is an $k \times k(k-1)/2$ matrix, where k is the number of factors. The size of C expands rapidly as k increases.

PB designs with power of two number of runs are 2^{k-p}_{III} fractional factorial designs and their alias relationships are easily obtained otherwise complex. For the 12-run PB design presented in Table 2.3, we construct the C for the block of only factor A; (see Lin and Draper (1992) for the complete alias structure).

Table 2.11: The first block of the alias matrix for the 12-run Plackett and Burman design

	A	B	C	D	E	F	G	H	I	J	K
AB	0	0	-1/3	1/3	1/3	-1/3	-1/3	1/3	-1/3	-1/3	-1/3
AC	0	-1/3	0	-1/3	-1/3	1/3	-1/3	1/3	-1/3	-1/3	1/3
AD	0	1/3	-1/3	0	-1/3	-1/3	-1/3	-1/3	-1/3	1/3	1/3
AE	0	1/3	-1/3	-1/3	0	1/3	1/3	-1/3	-1/3	-1/3	-1/3
AF	0	-1/3	1/3	-1/3	1/3	0	-1/3	-1/3	-1/3	1/3	-1/3
AG	0	-1/3	-1/3	-1/3	1/3	-1/3	0	-1/3	1/3	-1/3	1/3
AH	0	1/3	1/3	-1/3	-1/3	-1/3	-1/3	0	1/3	-1/3	-1/3
AI	0	-1/3	-1/3	-1/3	-1/3	-1/3	1/3	1/3	0	1/3	-1/3
AJ	0	-1/3	-1/3	1/3	-1/3	1/3	-1/3	-1/3	1/3	0	-1/3
AK	0	-1/3	1/3	1/3	-1/3	-1/3	1/3	-1/3	-1/3	-1/3	0

2.7 Follow-up Experiments

Design and analysis of experiments aim to arrive at valid conclusions. However, there are instances where results obtained are inconclusive, as to which factors are active. To resolve this ambiguity, additional runs are required. Also, some designs have complex aliasing structure and some active effects especially interaction are not being identified as such during analysis. Follow-up experiment is a technique for breaking the alias chain in a design with complex aliasing structure. Some follow-up experiment plans ensure that resolution IV designs are obtained from resolution III designs. Multiple techniques exist for augmenting an experimental design. They include full foldover, semifolding, D-optimal designs and Bayesian (MD-optimal). The choice of technique may be dependent on the analysis of the initial experiment, experimental objectives, availability of resources among others.

2.7.1 Full Foldover

According to Edwards et al. (2013), foldover is the most popular and classic approach for design augmentation. This design is such that levels of all of the factors have been reversed to form runs that are the mirror image of those in the original design in addition to a column where the first n entries are $+1$ and the last n entries are -1. Let **1** be $n \times 1$ vector of ones. The foldover, $\tilde{\mathbf{X}}$, of a two-level design **X** is given by

$$\begin{bmatrix} X & 1 \\ -X & -1 \end{bmatrix}$$

In general folding over a resolution III design produces a resolution IV design. To untangle the main effects from the interactions in the initial resolution III design, one can run a full foldover. Combining both blocks of runs produces a resolution IV design and all the main effects will be free and clear of two-factor interactions (2fi's). However, all the 2fi's maintain their aliasing. Although, the complex aliasing among estimated effects can be problematic, iterative procedures that exploit effect sparsity, hierarchy and heredity of such designs have been developed to allow identification of one or two active two-factor interactions (Wu and Hamada (2000), Miller and Sitter (2005)).

Consider the 2^{5-2} design (a resolution III design). The full factorial is presented in Table 2.1. The fractional factorial can be achieved in $2^3 = 8$ runs.

Table 2.12: A 2^{5-2} fractional factorial design

Run	A	B	C	D	E
1	-	-	-	+	+
2	+	-	-	-	-
3	-	+	-	-	+
4	+	+	-	+	-
5	-	-	+	+	-
6	+	-	+	-	+
7	-	+	+	-	-
8	+	+	+	+	+

We can increase the resolution of this design to IV if we augment the 8 original runs by adding 8 reversed runs of the original runs as shown in Table 2.13. This is known as a full foldover. The basic factors in a 2^{5-2} fractional factorial design again are the three independent factors (A, B, C) of the base factorial design (a 2^3 full factorial design). The two added factors (D, E) are assigned to columns chosen from the remaining columns of the model matrix for the base factorial design. The level settings of D and E are here determined by the columns corresponding to AB and AC, respectively.

Table 2.13: Reversed runs of a 2^{5-2} fractional factorial design

Run	A	B	C	D	E
9	+	+	+	-	-
10	-	+	+	+	+
11	+	-	+	+	-
12	-	-	+	-	+
13	+	+	-	-	+
14	-	+	-	+	-
15	+	-	-	+	+
16	-	-	-	-	-

2.7.2 Semifolding

Another foldover technique is the semifolding (John (2000), Mee and Peralta (2000)), that adds only half of a full foldover fraction. This is done by subsetting on a factorial effect. A common practice is to subset on a desirable level of an important active factor revealed in the initial experiment. The problem of multi-collinearity among main effects and two-factor interaction effects is avoided by subsetting on a main effect. Consider the 2^{5-2} design in Table 2.12. Suppose factor A is identified as an important active effect and the experimenter wishes to subset on the factor. When the fraction is semifolded on A (with desirable level as high), the points that were at the low level of A in the original fraction are repeated at the high level of A; the points that were at the high level are not repeated. Then the follow-up runs will be as presented in Table 2.14. According to John (2000), when the original fraction has resolution IV, then semifolding on A will result in the estimation of all two-factor interactions involved with A.

Table 2.14: Semifolding on A of 2^{5-2} fractional factorial design

Run	A	B	C	D	E
9	+	-	-	+	+
10	+	+	-	-	+
11	+	-	+	+	-
12	+	+	+	-	-

2.7.3 D-Optimal Designs

D-optimal designs are model specific designs. When used for augmentation it is driven by the best model(s) identified in the initial experiment. It is an iterative search algorithm that minimizes the covariance of the parameter estimates for a specified model. This is equivalent to maximizing the determinant $D = |X^T X|$, where X is the design matrix of model terms (the columns) evaluated at specific treatments in the design space (the rows). D-optimal designs do not require orthogonal design matrices, and as a result, parameter estimates may be correlated. Also, D-optimal suffers the setback of being subjective, in that follow-up runs are chosen solely on the account of improving estimation for a single model.

2.7.4 MD-Optimal Design

A Bayesian approach for follow-up design in the context of model discrimination based on predictive densities for competing models was proposed by Meyer et al. (1996). Let M_i represent a model with $0 \leq m_i \leq k$ active factors and all interactions involving these factors up to a desired order. For some prior probability, π, that a factor is active, the prior probability of models, M_i, is given by $P(M_i) = \pi^{m_i}(1-\pi)^{k-m_i}$. The main effects and interactions in each model are assigned $N(0, \gamma^2\sigma^2)$ priors where γ is a scale factor. The intercept and error variance are assigned noninformative priors. After observing the data Y, one can obtain the posterior probability $P(M_i|Y)$ for each M_i. Thus, for l competing models, Meyer et al. (1996) define their model discrimination critirion as;

$$MD = \sum_{0 \leq i \neq j \leq l} P(M_i|Y)P(M_j|Y)I(f_i, f_j)$$

, where f_i denotes the predictive density conditional on Y and M_i, and $I(f_i, f_j) = \int_{-\infty}^{\infty} f_i \text{In}(f_i/f_j)$ is the Kullback-Leibler information. Follow-up experimental runs are selected that maximize MD. The MD criterion is implemented in the $BsMD$ package in the R software.

Chapter 3

Methods for Analysis

3.1 The Box-Meyer Method

When determining active factors and interactions among a set of factors in an experiment, various hypotheses might be considered. For instance, with five factors, V, W, X, Y and Z, one hypothesis is that a single factor is responsible for most of what is happening, in which case one need only to consider the five main effects. Assuming that two factors are responsible, 10 possible subsets can be considered; the subset of main effects X and Y with interaction XY, the subset of X and Z with interaction XZ, etc., are possibilities. Also, on the hypothesis that three factors, say X, Y and Z, may be active, the subset of main effects X, Y and Z with interactions XY, XZ and YZ and XYZ are considered together and so on. Box and Meyer (1993) considered an approach to consider all the possible explanations (including interactions) of the data from a screening experiment and identify the factors which fit the data well. The Bayesian framework is used to give an appropriate measure of fit to each model considered (posterior probability) that can be accumulated in various ways (marginal posterior probability).

Consider a set of $m + 1$ models, $M_0, ..., M_m$. Each model M_i has an associated vector of parameters θ_i so that the sampling distribution of data y, given the model M_i, is described by the probability density $p(y|M_i, \theta_i)$. The prior probability of the model M_i is $p(M_i)$, and the prior probability density of θ_i is $p(\theta_i|M_i)$. The predictive density of y, given model M_i, is written $p(y|M_i)$, and is given by the expression;

$$p(y|M_i) = \int_{\Theta_i} p(y|M_i, \theta_i)p(\theta_i|M_i)d\theta_i \tag{3.1}$$

Here Θ_i is the set on which $p(\theta_i|M_i)$ is defined. The posterior probability of the model M_i given the data y is then,

$$p(M_i|y) = \frac{p(M_i)p(y|M_i)}{\sum_{h=0}^{m} p(M_h)p(y|M_h)} \tag{3.2}$$

The posterior probabilities $p(M_i|y)$ provide a basis for model identification, tentatively plausible models are identified by their large posterior probability. For each model, M_i, we calculate $p(M_i)p(y|M_i)$ and then scale these quantities to sum to unity.

For the screening design situation with k factors, let M_i denote the model that a particular combination of f_i factors is active $0 \leq f_i \leq k$. There are 2^k models M_i starting from $i = 0$ (no active factors) to $i = 2^k - 1$ (k active factors). To model the condition of factor sparsity, let π be the prior probability that any one factor is active. For a screening experiment in which we typically expect to identify just a few (less than half) of the factors as important, appropriate values for π would be in the range from 0 and $\frac{1}{2}$. A nominal value of $\pi = 0.25$ has given sensible results in practice, and the individual experimenter can specify a different value based on the circumstances of a particular experiment (Box and Meyer, 1993). The prior probability $p(M_i)$ of the model M_i is then $\pi^{f_i}(1-\pi)^{k-f_i}$.

Let X_i be the matrix with columns for each effect under the model M_i using the convention of coded values -1 and $+1$ for two-level factors. X_i includes a columns of 1's for the mean and interaction columns up to any order desired. Let t_i be the number of such effects, excluding the mean. The dimensions of X_i are $n \times (1+t_i)$. Likewise, let β_i be the $(1+t_i) \times 1$ vector of true (regression) effects under M_i, and let y denote the $n \times 1$ vector of responses. The predictive density $p(y|M_i)$ is obtained by integrating;

$$p(y|M_i) = \int_0^\infty \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty p(y|M_i, \beta_i, \sigma) p(\beta_i|M_i, \sigma) p(\sigma|M_i) d\beta_i d\sigma \tag{3.3}$$

where the probability density of y given M_i is assumed to be the usual normal linear model;

$$p(y|M_i, \sigma, \beta_i) \propto \sigma^{-n} exp\left(\frac{-(y - X_i\beta_i)'(y - X_i\beta_i)}{2\sigma^2} \right) \tag{3.4}$$

$$p(\beta_i|M_i, \sigma) \propto \gamma^{-t_i} \sigma^{-t_i} exp\left(\frac{-\beta_i{}'\Gamma_i\beta_i}{2\sigma^2} \right) \tag{3.5}$$

$$p(\sigma|M_i) = p(\sigma) \propto \frac{1}{\sigma} \tag{3.6}$$

$$\Gamma_i = \frac{1}{\gamma^2}\begin{pmatrix} 0 & 0 \\ 0 & I_{t_i} \end{pmatrix} \tag{3.7}$$

The elements of β_i are assigned independent prior normal distribution with mean 0 and variance $\gamma^2\sigma^2$. A non-informative prior distribution is employed for the overall mean β_0 and $log(\sigma)$ so that $p(\beta_0, \sigma) \propto 1/\sigma$ where the likelihood is appreciable and negligible elsewhere.

Performing the integration with respect to β_i yields;

$$p(y|M_i,\sigma) \propto \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \gamma^{-t_i}\sigma^{-(n+t_i)} exp\left\{\frac{-((y - X_i\beta_i)'(y - X_i\beta_i) + \beta_i{}'\Gamma_i\beta_i)}{2\sigma^2}\right\} d\beta_i$$

$$= \gamma^{-t_i}\sigma^{-(n+t_i}|\Gamma_i + X_i{}'X_i|^{-1/2} exp\{-(y'(I - X_i(\Gamma_i + X_i{}'X_i)^{-1}X_i{}')y\}/2\sigma^2$$

$$\times \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \sigma^{-(t_i+1)}|\Gamma_i + X_i{}'X_i|^{1/2} exp\{-(\beta_i - \widehat{\beta}_i)'(\Gamma_i + X_i{}'X_i)(\beta_i - \widehat{\beta}_i)/2\sigma^2\}$$

where

$$\widehat{\beta}_i = (\Gamma_i + X_i{}'X_i)^{-1}X_i{}'y$$

The expression inside the integral is proportional to the density of a multivariate normal distribution with mean $\widehat{\beta}_i$ and covariance matrix $\sigma^2(\Gamma_i + X_i{}'X_i)^{-1}$ and so integrate to a constant leaving;

$$p(y|M_i,\sigma) \propto \gamma^{-t_i}\sigma^{-(n+t_i)}|\Gamma_i + X_i{}'X_i|^{-1/2} exp\{-(y'(I - X_i(\Gamma_i + X_i{}'X_i)^{-1}X_i{}')y\}/2\sigma^2 \tag{3.8}$$

Integrating out σ yields,

$$p(y|M_i) \propto \int_0^\infty \gamma^{-t_i}\sigma^{-(n-1)-1}|\Gamma_i + X_i{}'X_i|^{-1/2} \times exp\left\{ - \frac{(y'(I - X_i(\Gamma_i + X_i{}'X_i)^{-1}X_i{}')y}{2\sigma^2} \right\} d\sigma$$

and making a change of variables,

$$u = \frac{(y'(I - X_i(\Gamma_i + X_i{}'X_i)^{-1}X_i{}')y}{2\sigma^2}$$

results;

$$p(y|M_i) \propto \gamma^{-t_i}|\Gamma_i + X_i{}'X_i|^{-1/2}(y'(I - X_i(\Gamma_i + X_i{}'X_i)^{-1}X_i{}')y)^{-(n-1)/2} \times \int_0^\infty u^{(n-1)/2-1} exp - \{u\} du$$

The latter integral is a gamma function $(\Gamma((n-1)/2))$ and thus constant, leaving;

$$p(y|M_i) \propto \gamma^{-t_i}|\Gamma_i + X_i{}'X_i|^{-1/2}(y'(I - X_i(\Gamma_i + X_i{}'X_i)^{-1}X_i{}')y)^{-(n-1)/2} \tag{3.9}$$

Having observed the data vector y, the posterior probability of the model M_i can then be written, see [(Box and Meyer, 1986a)]

$$p(M_i|y) \approx C\left(\frac{\pi}{1-\pi}\right)^{f_i} \gamma^{-t_i} \frac{|X_0{}'X_0|^{1/2}}{|\Gamma_i + X_i{}'X_i|^{1/2}} \times \left(\frac{S(\widehat{\beta}_i) + \widehat{\beta}_i'\Gamma_i\widehat{\beta}_i}{S(\widehat{\beta}_0)}\right)^{-(n-1)/2} \tag{3.10}$$

where

$$S(\widehat{\beta}_i) = (y - X_i\widehat{\beta}_i)'(y - X_i\widehat{\beta}_i)$$

X_0 is design matrix of the active subspace under consideration and C is the normalization constant which forces all probabilities to sum to one and $\mathbf{I}_{t_i}$ is the $t_i \times t_i$ identity matrix.

The probabilities $p(M_i|y)$ can be accumulated to compute the marginal posterior probability P_j that factor j is active;

$$P_j = \sum_{M_i:factorjactive} p(M_i|y). \tag{3.11}$$

The probability P_j is just the sum of the posterior probabilities of all the distinct models in which the factor j is active. The probabilities P_j are thus calculated by direct enumeration over the 2^k possible models M_i. A large value for P_j would indicate that the factor j was active, and similarly, a value of P_j close to zero would indicate that the factor j was inert. After examining the P_j, the individual probabilities $p(M_i|y)$ may further identify specific combinations of factors that are most likely active.

3.2 Partial F Method

3.2.1 Motivation

According to Kulachi and Box (2003), experimentation in industry and engineering differ in three ways from the paradigm implied by statistics courses:

1. Data are frequently available in days, hours or sometimes even in minutes. Therefore, appropriate methods of statistical design and analysis must be used to exploit the advantage of such immediacy by using sequential experimentation to point the

trial to development and discovery.

2. Even for complicated methods of analysis, the computer is available to help to do this by facilitating fast analyses, hence, the experimenter's ability to perform iterative analysis of data and to see what it might imply in the light of alternative models and assumptions is greatly enhanced.

3. In many applications, it is true that not one but a number of responses are simultaneously available and that at any given stage of an investigation, the question whose answers will help the investigator the most is, *"which* factors affect which responses and in *what ways?"*

Most statistical methods has been based on one-shot assumption. A process is said to be a one-shot process, if all the important variables are supposed known at the beginning and after appropriate analysis, specific conclusions are drawn. However, in many engineering and industrial settings, this one-shot assumption does not hold and serve as a premise for the development of this method of analysis. A process of iterative investigation, which follows a trial often not predictable in advance. Engineering experimentation is described as an iterative with alternatives for a subsequent set of runs dependent on the results from a previous set. Figure 3.1 describes Kulachi and Box (2003) notion of iterative experimentation.

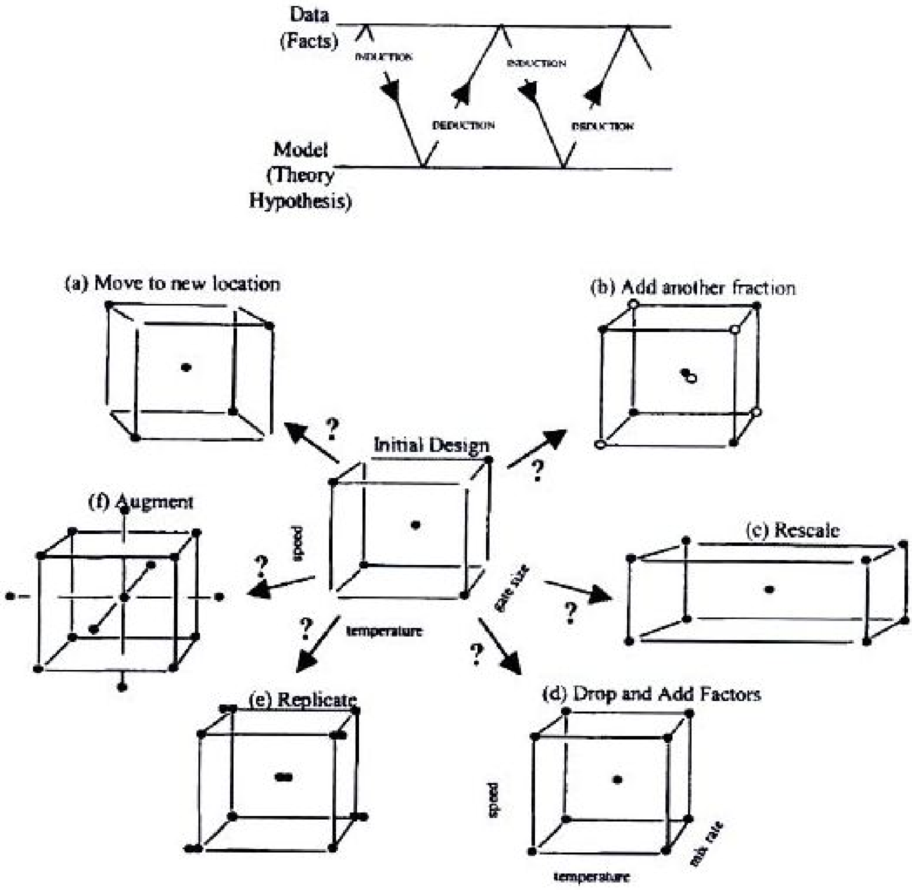

Figure 3.1: Iterative experimentation with alternatives for a subsequent set of runs depending on results from a previous set.

3.2.2 The Partial Analysis

This method aims at partitioning total variability in the data into its component parts. Data are generated through experimentation, and thus the concept of ANOVA is closely connected to design of experiments. ANOVA estimates three(3) sample variances: a total variance based on all the observation deviations from the grand mean, that is, the total sum of square, SS_{Total}, an error variance based on all the observation deviations from their corresponding treatment means, that is, the error sum of square, SS_{Error} and a treatment variance based on the deviations of treatment means from the grand mean, the result being multiplied by the number of observations in each treatment, that is, treatment sum of square, $SS_{Treatment}$. Therefore, we have $SS_{Total} = SS_{Error} + SS_{Treatment}$. The overall variability in the data is obtained by:

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2$$

and can be evaluated as;

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2 \tag{3.12}$$

where y_{ij} is the response variable; i, the number of factors, and j, the number of replications (runs) for the factors. The total variability in the data, as measured by the total sum of squares, can be partitioned into a sum of squares of the difference between and within treatment averages. The statistical significance of the experiment is determined by ratio of two variances which are independent of constant bias and scaling errors, as well as, the units used in measuring responses. There are $an = N$ total observations, thus SS_{Total} has $N - 1$ degrees of freedom. There are a factors, so $SS_{Treatment}$ has $a - 1$ degrees of freedom. Finally, for each factor, there are n replicates providing $n - 1$ degrees of freedom. In performing statistical analysis, the assumption of normal errors implies that the SS_{Total} is a normally distributed random variable; consequently SS_{Total}/σ^2 is distributed as chi-square with $N - 1$ degrees of freedom; SS_{Error}/σ^2 is chi-square with $N - a$ degrees of freedom and that $SS_{Treatment}/\sigma^2$ is also chi-square with $a - 1$ degrees of freedom. The $SS_{Treatment}/\sigma^2$ and SS_{Error}/σ^2 are independently distributed chi-square random variables. This follows from Cochran's theorem, since the degrees of freedom for $SS_{Treatment}$ and SS_{Error} add to $N-1$, the total number of freedom. Therefore, the ratio,

$$F_0 = \frac{(SS_{Treatment}/\sigma^2)/(a-1)}{(SS_{Error}/\sigma^2)/(N-a)} = \frac{MS_{Treatment}}{MS_{Error}} \tag{3.13}$$

is distributed as F with $a-1$ and $N-a$ degrees of freedom.

Kulachi and Box (2003) proposed a simplified and more intuitive version of the method of analysis of Box and Meyer (1993) (discussed in section 3.1), an approach they refered to as the partial analysis. Their approach of analysis is a mixture of the frequentist and the bayesian way of thinking. They argue that the posterior probabilities obtained using Box and Meyer method (1993) may be divided into two parts:

1. a penalty factor associated with the order of the model

2. a function depending only on the sums of squared errors.

Therefore, for a model of the same complexity, the evaluation of the sums of the squared error alone should produce similar results as the full analysis. Kulachi and Box (2003) in comparing models involving equal subset of factors among a set of factors, considered the use of the following two modified simple criteria.

$$SS_{RES} = \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 \tag{3.14}$$

$$F = \frac{\sum_{i=1}^{n} (\widehat{y}_i - \bar{y})^2 / a}{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 / n - a - 1} \tag{3.15}$$

where y_i is the observed response for the $i - th$ run, $\widehat{y}_i$ is the predicted response, $\bar{y}$ is the overall response mean.

3.3 Projection-based Method of Analysis

The properties of a factorial design when restricted to a subset of experimental factors is known as projective properties. Projective properties of a design enable the identification of active subspaces of factors without necessarily imposing restrictive assumptions on the underlying model. Further investigation of the functional relationship between design factors and responses can be undertaken on the active subspace identified. Non-regular designs have very favourable projective properties. For instance, the 12-run PB design projects onto six replicates of a 2^1 design in one dimension, three replicates of a 2^2 design in every two dimensions and one-and-half replicate of a 2^3 design in every three dimensions.

Tyssedal and Samset (1997) proposed a projective-based method of analysing factorial designs. They established that if the candidate set of factors under investigation contains the true active factors space, then replicated runs will have the same expected value. Thus, a model independent estimate of the error variance may be obtained regardless of any functional relationship. The ability of the subset of factors to explain the variation in the response may be evaluated with rather weak assumptions on the underlying model.

Consider the 12-run PB design with one active factor. Then for each of the eleven factors, there are six replicated runs on the high level and six likewise runs on the low level. From the two groups of points, two estimates of σ^2 with 5 degrees of freedom may be calculated and pooled to get an estimate of σ^2 with ten degrees of freedom. Altogether, 11 estimates of σ^2 will be obtained, one for each factor, and the factor associated with the lowest $\widehat{\sigma}^2$ would be judged most capable of explaining the variation in the data. Similarly, for two active factors we have for each of the $\binom{11}{5} = 55$ possibilities, 3 replicates for each of the four level combinations of the factors. Four estimates of σ^2 may be pooled to give a $\widehat{\sigma}^2$ with 8 degrees of freedom. The two factors associated with the smallest $\widehat{\sigma}^2$ would be considered to be the most capable of explaining the variation in the data. Similar explanation can be made for other active factors.

Implementation of an algorithm to calculate these estimates of σ^2 is easier for the 12-run PB design because of its cyclic nature of construction. Written codes for the projection based method in Wiik (2014) were sligthly modified in dimensions and was implemented in the R software for the analysis. For instance, if the design is as given in Table 2.3, the six pluses in column 1 is in run number $1, 3, 7, 8, 9$ and 11. Increasing each of these numbers by m, the new numbers modulus 11 gives us the position of pluses in column number $m + 1$(modulus 11). Hence only the vector of response values and the position of pluses in the first column is needed in order to create the 11 estimates of σ^2, assuming only one factor is active. The same idea may be exploited in investigating more than one active factor. In particular run s for any set of k factors will be identical to run $s + m$(modulus 11) in a different set of k factors if this set is obtained from the first one by shifting each factor to the right an equal amount m.

For three active factors only eight of the level combinations are needed. For the three first columns the four repeated runs are as given in Table 3.1.

Table 3.1: Run numbers for repeated runs in the three first columns in Table 2.3

(+ + -)	(+ - +)	(- + +)	(- - -)
1	3	2	6
8	11	10	12

Once the four replicated runs are detected for any set of three columns, we also get the position of the replicated runs for any three columns shifted to the right with the same amount. The following must be true;

$$\begin{aligned}\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2 &= \sum_{i:unrepeated}(y_i-\widehat{y}_i)^2+\sum_{i:repeated}(y_i-\widehat{y}_i)^2\\ &= \sum_{i:unrepeated}(y_i-\widehat{y}_i)^2+\sum_{i=1}^{m}\sum_{j=1}^{m_i}(y_{ij}-\widehat{y}_i)^2\end{aligned} \tag{3.16}$$

where $\widehat{y}_i$ is the fitted value for y_i, m is the number of repeated runs and m_i is the number of observations at each repeated run. From this it follows that $\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2$ is minimized when $\widehat{y}_i$ is chosen equal to y_i for all runs that are unrepeated and equal to $\bar{y}_i$, the average over the m_i observations at each repeated run, otherwise.

For the 12-run PB design, it is equivalent to successively assume that the design matrix in the regression model;

$$y = X\beta + \epsilon,$$

contains in addition to the intercept,

i) only one main effect column

ii) two main effect columns and their interaction column

iii) three main effects columns, their three two-factor interaction columns and their three-factor interaction column.

This method forces all interactions into the model. For orthogonal columns are satisfied (i) and (iii). Box and Tyssedal (1996) discussed the non-orthogonality of columns for (iii). For three main effect columns, say A, B and C, having defining generator $I = \pm ABC$ for eight experiments and $\mp ABC$ for the four others. This means that three-factor interaction columns is correlated with the column of only +1's and also have that each main effect column is correlated with the two-factor interaction column for which it is not involved. It follows that each column in **X** is correlated with exactly one other column with an equal amount of either $\frac{1}{3}$ or $-\frac{1}{3}$.

Chapter 4

Analysis and Results

The first design we use is the six-factor PB design matrix given in Table 4.1. According to Edwards et al. (2013), this six-factor PB design is ranked best based on the generalized minimum aberration criterion of Deng and Tang (1999). Its alias structure, assuming no three or higher order interaction, is as given in Table 4.2.

Table 4.1: Six-factor PB design

Run	A	B	C	D	E	F
1	-1	-1	-1	1	-1	1
2	-1	-1	1	-1	-1	1
3	-1	-1	1	-1	1	-1
4	-1	1	-1	-1	1	-1
5	-1	1	-1	1	1	1
6	-1	1	1	1	-1	-1
7	1	-1	-1	-1	1	1
8	1	-1	-1	1	-1	-1
9	1	-1	1	1	1	-1
10	1	1	-1	-1	-1	-1
11	1	1	1	-1	-1	1
12	1	1	1	1	1	1

Table 4.2: Alias structure of the six-factor PB design

$E(\widehat{A}) = A + \frac{1}{3}BC - \frac{1}{3}BD - \frac{1}{3}BE + \frac{1}{3}BF + \frac{1}{3}CD + \frac{1}{3}CE + \frac{1}{3}CF + \frac{1}{3}DE - \frac{1}{3}DF + \frac{1}{3}EF$
$\mathrm{E}(\widehat{B}) = B + \frac{1}{3}AC - \frac{1}{3}AD - \frac{1}{3}AE + \frac{1}{3}AF + \frac{1}{3}CD - \frac{1}{3}CE + \frac{1}{3}CF + \frac{1}{3}DE + \frac{1}{3}DF + \frac{1}{3}EF$
$\mathrm{E}(\widehat{C}) = C + \frac{1}{3}AB + \frac{1}{3}AD + \frac{1}{3}AE + \frac{1}{3}AF + \frac{1}{3}BD - \frac{1}{3}BE + \frac{1}{3}BF + \frac{1}{3}DE - \frac{1}{3}DF - \frac{1}{3}EF$
$\mathrm{E}(\widehat{D}) = D - \frac{1}{3}AB + \frac{1}{3}AC + \frac{1}{3}AE - \frac{1}{3}AF + \frac{1}{3}BC + \frac{1}{3}BE + \frac{1}{3}BF + \frac{1}{3}CE - \frac{1}{3}CF + \frac{1}{3}EF$
$\mathrm{E}(\widehat{E}) = E - \frac{1}{3}AB + \frac{1}{3}AC + \frac{1}{3}AD + \frac{1}{3}AF - \frac{1}{3}BC + \frac{1}{3}BD + \frac{1}{3}BF + \frac{1}{3}CD - \frac{1}{3}CF + \frac{1}{3}DF$
$\mathrm{E}(\widehat{F}) = F + \frac{1}{3}AB + \frac{1}{3}AC - \frac{1}{3}AD + \frac{1}{3}AE + \frac{1}{3}BC + \frac{1}{3}BD + \frac{1}{3}BE - \frac{1}{3}CD - \frac{1}{3}CE + \frac{1}{3}DE$

4.1 Simulated Models

In investigating the robustness of the methods described in Chapter 3, data were generated from model 4.1.

$$y = 2A + 0.8B + \frac{2}{0.4B + ABC + 2} + \epsilon \tag{4.1}$$

The data was generated from this non-regular model with noise that was assumed to be normally distributed with zero mean and equal variance. Tyssedal (2008b) used this functional relationship to study a 12 run PB design using the projective based method, with noise from a normally distributed error term with mean 0 and standard deviation, $\sigma = 0.3$. The active subspace of the design in Table 4.1 was investigated using simulated responses from model 4.1 with σ at $0.6, 0.8$ and 1.0. The response values are given in Appendix. The active factors were identified using the factor based methods considered. For one active factor subspace, six models were compared, for two active factor subspace, fifteen models were compared and twenty models for three active factor subspace. Table 4.3 display in chronological order the best five active factors. The Box-Meyer search was conducted with $p = 0.6$ and variance inflation factor, $g = 2.49$. The choice of p was motivated by the quest to be able to discriminate among models with up to four active factors and g was set to the default in the BsMD package. In the tables, models with zero probabilities are omitted. The residual standard error ($\widehat{\sigma}$) of the partial F-test for one, two and three active factors were evaluated at 10, 8 and 4 degrees of freedom respectively. The variance estimate from the projective based search were computed from the replicated runs. For one active factor, there are six replicated runs, for two active factors there are four runs that are replicated two times and for three active factors there are four runs replicated twice.

In Table 4.3, the results from the analysis of the 12-run design for model 4.1 with $\sigma = \{0.6, 0.8, 1.0\}$ are given. The three methods revealed similarities in identifying the active subspaces. The methods at various values of σ in most cases clearly discriminated the top ranked active subspace from the others. It was observed the Box-Meyer method in searching for one and two active factors performed extremely well in discriminating among models. However, in searching for three active factors, the Box-Meyer method displayed weak discriminatory power between the two topmost models, although in all cases it ranked the correct active subspace as the best. The low posterior probabilities assigned to the best active subspace for three factors indicate that the factor activity for factor C is very weak. This is because, the posterior probabilities for factors A and B (as active subspace) is very high at various σ, indicating very strong factor activity for these two factors. Tyssedal (2008b) reporting on this model at $\sigma = 0.3$ using the projection based search concluded on factors A, B and C as the active factors. The performance of the projective based and partial F methods from this study does not suggest the need for follow-up studies.

Table 4.3: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 12-run PB design with responses from model 4.1 with σ equal to 0.6, 0.8 and 1.0.

AF	Projective based search		Box-Meyer			Partial F		
	Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Post. Prob.	Factor	$\widehat{\sigma}$	F
				$\sigma = 0.6$				
1	B	5.19	B	5.18	1.0	B	2.279	73.66
	A	39.62				A	6.294	0.970
	E	42.87				E	6.548	0.137
	F	43.38				F	6.586	0.020
	D	43.44				D	6.591	0.005
2	A, B	1.63	A, B	1.69	0.886	A, B	1.277	86.18
	B, F	4.53	B, F	3.77	0.011	B, F	2.127	29.34
	B, E	5.65	B, E	4.58	0.004	B, E	2.376	22.99
	B, C	6.11	B, C	4.91	0.003	B, C	2.472	21.04
	B, D	6.45	B, D	5.16	0.002	B, D	2.541	19.78
3	A, B, C	0.45	A, B, C	0.682	0.041	A, B, C	0.667	138.8
	A, B, E	1.05	A, B, E	0.888	0.01	A, B, E	1.023	58.7
	A, B, F	1.24	A, B, F	0.938	0.007	A, B, F	1.112	49.6
	A, B, D	1.59	A, B, D	1.131	0.003	A, B, D	1.261	38.44
	B, E, F	3.45				B, E, F	1.858	17.42
				$\sigma = 0.8$				
1	B	4.86	B	4.9	1.0	B	2.204	83.19
	A	41.38				A	6.433	0.936
	D	45.21				D	6.724	0.107
	E	45.24				F	6.726	0.005
	F	45.24				E	6.726	0.004
2	A, B	1.18	A, B	1.39	0.945	A, B	1.087	1.25
	B, C	4.77	B, C	3.97	0.003	B, C	2.184	28.95
	B, F	4.96	B, F	4.11	0.002	B, F	2.227	27.75
	B, D	5.67	B, D	4.61	0.001	B, D	2.381	23.95
	B, E	5.79	B, E	4.70	0.001	B, E	2.407	23.37
3	A, B, C	0.32	A, B, C	0.659	0.019	A, B, C	0.570	198.6
	A, B, F	1.21	A, B, F	0.947	0.003	A, B, F	1.101	52.78
	A, B, E	1.64	A, B, E	1.116	0.001	A, B, E	1.279	38.92
	A, B, D	1.92	A, B, D	1.248	0.001	A, B, D	1.386	33.1
	B, E, F	2.70				B, E, F	1.643	23.38
				$\sigma = 1.0$				
1	B	4.67	B	4.72	1.0	B	2.161	83.81
	A	40.01				A	6.326	0.952
	D	43.73				D	6.613	0.022
	E	43.79				E	6.617	0.009
	C	43.80				C	6.618	0.005
2	A, B	0.98	A, B	1.23	0.966	A, B	0.999	146.5
	B, C	4.80	B, C	3.98	0.002	B, C	2.191	27.75
	B, E	5.09	B, E	4.18	0.001	B, E	2.255	26.06
	B, F	5.39	B, F	4.40	0.001	B, F	2.322	24.43
	B, D	5.55	B, D	4.51	0.001	B, D	2.355	23.67
3	A, B, C	0.54	A, B, C	0.743	0.005	A, B, C	0.738	114.4
	A, B, D	1.17	A, B, D	0.952	0.001	A, B, D	1.080	53.09
	A, B, F	1.51	A, B, F	1.051	0.001	A, B, F	1.227	40.98
	A, B, E	1.63	A, B, E	1.108	0.001	A, B, E	1.278	37.76
	B, E, F	3.18				B, E, F	1.782	19.15

Further studies of the three factor-based methods using the panel of models suggested by Tyssedal and Hussain (2016) was carried out. The search for the active subspace of the design in Table 4.1 using simulated responses from panel of models at various variability levels, that is, $\sigma = \{0.6, 0.8, 1.0\}$ was undertaken. The panel of models accounts for some possible model situations, such as, the number of active factors, number of terms, number of interaction, type of interactions, size of the terms and whether model obeys the heredity principles.

Model 1: y_1=$A + 2AB + 2AC+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 2: y_2=$A + 1.5B + 2C + AB + 1.5AC+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 3: y_3=$A + 1.5B + 2C + 1.5ABC+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 4: y_4=$2A + BC+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 5: y_5=$A + C + BC + CD+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 6: y_6=$2A + 3B + 2C + D + 3CD+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 7: y_7=$4A + B + C + D + 2AD+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

Model 8: y_8=$2A + 4C + 2BC + 2CD+ \epsilon$, $\epsilon \sim$ N(0,σ^2)

One, two and three active factor(s) subspaces respectively were identified using the factor based methods considered. Tables 4.4, 4.5, and 4.6 display in chronological order the best five subspaces with one, two and three active factors when $\sigma = \{0.6, 0.8, 1.0\}$ respectively. From the analysis, the methods displayed similar results. For the models in the panel with three active factors, the methods in most models identified the correct active subspace (A, B, C), and ranked it as the best among the five topmost subspaces. A surprising results for model 4 simulated at $\sigma = 0.8$ is that all the three methods ranked A, C, E instead of A, B, C as the most important active subspace, meanwhile factor E

was inert. Amidst that, the methods in most cases discriminated well among the models at the levels of σ studied.

Models 5 to 8 involved four factors, the methods' performance in identifying four active factor subspace for a six-factor 12-run PB design has been minimally exploited in literature. Table 4.7 presents the best three four active factors models identified by the methods studied. Three factor interactions were assumed negligible and thus for the partial F search, the residual standard error ($\widehat{\sigma}$) was evaluated with one degree of freedom for the 12-run PB design. Thus, the model has an intercept, four main effects and six two-factor interaction. For each four active subspace there exist one replicated run and thus the projective based search was implemented by comparing the variance for the replicated runs for all fifteen possible models. From the analysis, the methods with $\sigma = 0.6$, except for model 5, identified and ranked the correct active subspace (A, B, C, D) as the most important. However, with increased variability, the methods in most cases failed to identify the correct active subspace and also some slight disparity in model preference was observed among the methods, for instance the results for model 8 at $\sigma = 1.0$. The methods' did not display enough discriminatory powers among models and the results suggest a need for follow-up studies.

Table 4.4: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 12-run PB design with responses from the panel of models with $\epsilon \sim N(0, (0.6)^2)$

Mod	AF	Projective based search		Box-Meyer search			Partial F search			Mod	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F			Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
1	1	F	8.01	F	7.33	0.393	F	2.831	4.706	5	1	A	2.46	A	2.27	0.823	A	1.5688	11.24
		C	10.59	C	9.64	0.087	C	3.255	1.124			C	4.56	C	4.15	0.030	C	2.134	1.463
		A	10.97	A	9.98	0.072	A	3.312	0.743			B	4.95	B	4.51	0.019	B	2.226	0.541
		B	11.25	B	10.23	0.063	B	3.353	0.477			D	5.08	D	4.62	0.016	D	2.254	0.282
		D	11.54	D	10.49	0.055	D	3.397	0.211			F	5.16	F	4.69	0.015	F	2.272	0.119
	2	A, B	5.25	A, B	3.91	0.171	A, B	2.292	4.809		2	A, F	1.60	A, F	1.21	0.295	A, F	1.263	8.241
		C, F	7.08	C, F	5.22	0.035	C, F	2.66	2.884			A, C	2.22	A, C	1.65	0.052	A, C	1.489	5.182
		A, C	7.12	A, C	5.25	0.034	A, C	2.6691	2.848			A, E	2.36	A, E	1.76	0.038	A, E	1.536	4.71
		E, F	7.97				E, F	2,824	2.259			A, B	2.71	A, B	2.01	0.018	A, B	1.647	3.747
		B, F	8.71				B, F	2.952	1.84			C, D	2.86	C, D	2.12	0.014	C, D	1.691	3.418
	3	A, B, C	1.27	A, B, C	0.59	0.643	A, B, C	1.127	12.67		3	A, E, F	0.63	A, E, F	0.464	0.121	A, E, F	0.794	11.27
		B, C, F	3.17				B, C, F	1.781	4.734			A, D, F	0.81	A, C, D	0.582	0.035	A, D, F	0.901	8.624
		A, B, E	4.48				A, B, E	2.116	3.187			A, B, E	0.93	A, D, F	0.606	0.028	A, B, E	0.967	7.408
		A, B, F	5.16				A, B, F	2.272	2.69			A, C, D	0.99	B, C, D	0.674	0.016	A, C, D	0.993	6.992
		A, C, D	6.37				A, C, D	2.523	2.073			B, C, D	1.32				B, C, D	1.148	5.089
2	1	C	10.12	C	9.29	0.555	C	3.18	7.605	6	1	B	21.34	B	19.6	0.677	B	4.62	8.838
		B	12.71	B	11.61	0.163	B	3.565	4.015			A	32.07	A	29.2	0.076	A	5.663	2.536
		A	16.82	A	15.08	0.039	A	4.071	0.745			C	35.56	C	32.4	0.043	C	5.963	1.306
		F	16.82	F	15.08	0.036	F	4.101	0.586			E	39.02	E	35.5	0.026	E	6.247	0.302
		D	17.43	D	15.85	0.029	D	4.175	0.217			F	39.11	F	35.6	0.026	F	6.254	0.278
	2	A, C	4.64	A, C	3.55	0.533	A, C	2.155	10.12		2	C, D	14.36	C, D	10.8	0.209	C, D	3.789	6.668
		B, C	5.66	B, C	4.28	0.190	B, C	2.38	7.817			A, B	15.99	A, B	12.0	0.119	A, B	3.998	5.716
		C, F	9.76	C, F	7.22	0.011	C, F	3.124	3.415			B, C	20.05	B, C	14.9	0.036	B, C	4.478	4.016
		A, B	10.41	A, B	7.68	0.008	A, B	3.226	3.036			B, E	20.95	B, E	15.5	0.028	B, E	4.577	3.729
		B, F	11.06				B, F	3.325	2.702			B, F	24.99	B, D	19.3	0.013	B, F	4.999	2.695
	3	A, B, C	0.60	A, B, C	0.385	0.971	A,B, C	0.774	41.86		3	B, C, D	6.06	B, C, D	2.73	0.5	B, C, D	2.461	8.908
		B, C, F	3.35	B, C, F	1.391	0.001	B, C, F	1.83	7.027			A, B, E	8.33	A, B, E	4.90	0.02	A, B, E	2.886	6.325
		A, C, F	3.64	A, C, F	1.498	0.001	A, C, F	1.907	6.427			A, B, C	13.15				A, B, C	3.626	3.796
		B, C, E	6.39				B, C, E	2.528	3.408			C, D, F	16.01				C, D, F	4.001	3.016
		A, C, D	7.47				A, C, D	2.733	2.834			A, C, D	16.65				A, C, D	4.081	2.877
3	1	C	8.42	C	7.72	0.496	C	2.901	6.487	7	1	A	8.88	A	8.28	0.959	A	2.98	18.96
		B	10.80	B	9.85	0.130	B	3.286	2.853			C	21.86	C	19.92	0.008	C	4.675	1.767
		A	12.63	A	11.49	0.056	A	3.553	0.993			D	24.53	D	22.32	0.004	D	4.953	0.485
		F	13.29	F	12.09	0.042	F	3.646	0.442			E	24.87	E	22.62	0.004	E	4.987	0.343
		E	13.30	E	12.10	0.042	E	3.647	0.437			F	25.36	F	23.06	0.003	F	5.036	0.141
	2	B, C	6.54	B, C	4.86	0.092	B, C	2.558	4.405		2	A, D	3.01	A, D	2.47	0.885	A, D	1.736	25.77
		C, F	7.29	C, F	5.40	0.052	C, F	2.699	3.684			A, C	5.40	A, C	4.18	0.049	A, C	2.324	13.21
		B, F	8.02	B, F	5.92	0.031	B, F	2.832	3.103			A, E	9.66	A, E	7.24	0.002	A, E	3.108	6.21
		C, E	8.53				C, E	2.920	2.758			A, F	10.47	A, F	7.82	0.002	A, F	3.235	5.524
		A, C	8.61				A, C	2.934	2.708			A, B	11.06	A, B	8.25	0.001	A, B	3.326	5.084
	3	A, B, C	0.41	A, B, C	0.322	0.989	A, B, C	0.640	47.82		3	A, C, D	0.88	A, C, D	0.632	0.769	A, C, D	0.939	41.07
		B, D, F	3.84				B, D, F	1.96	4.591			A, B, D	2.33	A, B, D	1.278	0.017	A, B, D	1.525	15.23
		C, E, F	5.11				C, E, F	2.26	3.31			A, D, E	2.77	A, D, E	1.758	0.003	A, D, E	1.664	12.7
		A, C, E	5.46				A, C, E	2.337	3.058			A, C, E	2.99	A, D, F	1.877	0.002	A, C, E	1.729	11.72
		B, C, F	6.03				B, C, F	2.456	2.716			A, D, F	3.26				A, D, F	1.806	10.69
4	1	A	1.15	A	1.12	1.0	A	1.074	50.43	8	1	C	21.67	C	19.9	0.517	C	4.655	8.126
		E	6.70				E	2.588	0.404			A	24.35	A	22.3	0.276	A	4.935	6.129
		F	6.74				F	2.596	0.341			B	38.77	B	35.7	0.121	B	6.226	0.131
		D	6.87				D	2.622	0.138			D	38.97	D	35.4	0.022	D	6.242	0.08
		C	6.96				C	2.638	0.011			E	39.28	F	35.7	0.021	E	6.267	0.000
	2	A, E	0.90	A, E	0.73	0.146	A, E	0.949	23.15		2	A, C	8.43	A, C	6.25	0.763	A, C	2.903	12.87
		A, D	1.07	A, D	0.850	0.063	A, D	1.033	19.08			B, C	15.11	B, C	11.31	0.037	B, C	3.887	6.00
		A, F	1.09	A, F	0.864	0.058	A, F	1.043	18.70			C, D	17.19	C, D	12.81	0.019	C, D	4.146	4.948
		A, C	1.43	A, C	1.109	0.015	A, C	1.195	13.59			A, F	19.28	A, F	14.31	0.010	A, F	4.391	4.123
		A, B	1.44	A, B	1.118	0.014	A, B	1.2	13.45			A, E	22.15	A, E	16.37	0.005	A, E	4.706	3.245
	3	A, B, C	0.19	A, B, C	0.134	0.355	A, B, C	0.434	52.38		3	A, B, C	5.04	A, B, C	2.27	0.075	A, B, C	2.246	10.55
		A, E, F	0.81	A, E, F	0.372	0.001	A, E, F	0.902	11.66			A, C, D	6.31	A, C, D	2.72	0.027	A, C, D	2.512	8.323
		A, D, E	0.95	A, D, E	0.432	0.001	A, D, E	0.974	9.917			B, C, D	7.94	B, C, D	3.37	0.008	B, C, D	2.817	6.499
		A, C, E	1.14				A, C, E	1.067	8.175			A, C, F	9.54				A, C, F	3.089	5.31
		A, D, F	1.14				A, D, F	1.067	8.176			A, D, F	10.87				A, E, F	3.339	4.461

Table 4.5: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 12-run PB design with responses from the panel of models with $\epsilon \sim N(0,(0.8)^2)$

Mod	AF	Projective based search		Box-Meyer search			Partial F search			Mod	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F			Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
1	1	F	9.37	F	8.54	0.208	F	3.061	2.479	5	1	A	4.55	A	4.18	0.645	A	2.133	7.876
		A	10.12	A	9.22	0.137	A	3.182	1.546			C	6.76	C	6.17	0.076	C	2.601	2.019
		B	10.82	B	9.85	0.095	B	3.29	0.72			D	8.02	D	7.29	0.003	D	2.831	0.142
		C	11.52	C	10.48	0.068	C	3.395	0.145			B	8.06	B	7.33	0.029	B	2.839	0.085
		D	11.59	D	10.53	0.066	D	3.404	0.091			F	8.06	F	7.33	0.029	F	2.839	0.084
	2	A, B	4.15	A, B	3.08	0.144	C, F	2.576	1.969		2	A, E	2.57	A, E	1.94	0.402	A, E	1.603	7.881
		A, C	4.50	A, C	3.33	0.095	B, F	2.629	1.708			A, F	3.86	A, F	2.87	0.047	A, F	1.964	4.356
		B, F	6.41	B, F	4.70	0.014	A, F	2.642	1.647			A, C	3.96	A, C	2.94	0.041	A, C	1.99	4.176
		C, F	6.88				D, F	2.747	1.186			C, D	4.51	C, D	3.34	0.02	C, D	2.125	3.337
		E, F	7.26				E, F	2.747	1.185			B, C	5.30				B, C	2.302	2.449
	3	A, B, C	0.29	A, B, C	0.999	0.204	B, C, F	2.597	1.576		3	A, B, E	0.89	A, B, E	0.426	0.306	A, B, E	0.942	12.53
		A, C, E	4.20				A, C, F	2.612	1.528			A, E, F	1.09	A, E, F	0.492	0.139	A, E, F	1.043	10.10
		A, B, F	4.26				A, B, F	2.671	1.344			A, D, F	2.29				A, D, F	1.513	4.501
		A, C, F	4.56				C, D, F	2.731	1.17			B, C, E	2.62				B, C, E	1.617	3.869
		B, C, F	4.71				C, E, F	2.731	1.169			B, C, D	2.74				B, C, D	1.655	3.67
2	1	C	9.35	C	8.58	0.01	C	3.058	6.652	6	1	B	27.61	B	25.4	0.596	B	5.254	7.742
		B	11.62	B	10.61	0.156	B	3.409	3.407			A	38.62	A	35.2	0.097	A	6.215	2.681
		A	14.12	A	12.86	0.054	A	3.758	1.031			C	42.16	C	38.4	0.061	C	6.493	1.617
		F	14.84	F	13.50	0.041	F	3.853	0.493			F	47.48	F	43.2	0.032	F	6.891	0.315
		D	15.50	D	14.10	0.033	D	3.937	0.047			D	48.03	D	43.7	0.03	D	6.931	0.197
	2	A, C	3.92	A,C	3.00	0.684	A, C	1.98	10.58		2	C, D	16.78	C, D	12.6	0.283	C, D	4.097	7.061
		B, C	6.12	B, C	4.58	0.067	B, C	2.475	5.813			A, B	20.75	A, B	15.5	0.093	A, B	4.555	5.202
		C, F	8.40	C, F	6.22	0.012	C, F	2.898	3.515			B, C	23.71	B, C	17.6	0.046	B, C	4.87	4.219
		A, B	8.54	A, B	6.32	0.011	A, B	2.922	3.935			B, E	25.97	B, E	19.2	0.028	B, E	5.096	3.62
		B, F	10.71				B, F	3.272	2.183			B, F	32.29				B, F	5.685	2.39
	3	A, B, C	0.31	A, B, C	0.259	0.993	A, B, C	0.56	70.28		3	B, C, D	3.93	B, C, D	1.91	0.736	B, C, D	1.983	17.23
		A, C, F	2.22	A, C, F	0.966	0.001	A, C, F	1.491	9.443			A, B, E	9.34	A, B, E	4.05	0.012	A, B, E	3.057	6.918
		B, C, F	2.95				B, C, F	1.718	6.966			A, B, C	10.95				A, B, C	3.309	5.818
		A, B, F	4.94				A, B, F	2.222	3.935			A, C, D	11.97				A, C, D	3.459	5.276
		A, C, E	5.75				A, C, E	2.397	3.30			C, D, F	21.66				C, D, F	4.654	2.658
3	1	C	7.51	C	6.89	0.569	C	2.744	7.534	7	1	A	10.74	A	9.98	0.939	A	3.276	16.73
		B	10.16	B	9.28	0.111	B	3.188	2.957			C	24.44	C	22.27	0.011	C	4.943	1.741
		A	11.30	A	10.30	0.063	A	3.362	1.648			D	26.86	D	24.44	0.007	D	5.183	0.6812
		F	12.05	F	10.97	0.044	F	3.472	0.925			F	27.53	F	25.04	0.006	F	5.247	0.422
		D	13.16	E	11.96	0.027	E	3.628	0.005			E	28.42	E	25.84	0.005	E	5.331	0.096
	2	B, C	3.70	B, C	2.81	0.525	B, C	1.923	9.208		2	A, D	3.96	A, D	3.19	0.864	A, D	1.991	21.46
		C, F	5.28	C, F	3.95	0.081	C, F	2.298	5.648			A, C	7.27	A, C	5.56	0.041	A, C	2.696	10.49
		A, C	6.41	A, C	4.760	0.029	A, C	2.532	4.18			A, F	11.88	A, F	8.87	0.003	A, F	3.446	5.386
		B, F	8.19				B, F	2.862	2.69			A, E	12.95	A, E	9.64	0.002	A, E	3.598	4.721
		A, F	8.63				A, F	2.938	2.418			A, B	13.25	A, B	9.86	0.002	A, B	3.64	4.551
	3	A, B, C	0.71	A, B, C	0.406	0.871	A, B, C	0.8413	26.00		3	A, C, D	1.77	A, C, D	0.945	0.177	A, C, D	1.331	22.55
		B, C, F	2.24	B, C, F	0.930	0.009	B, C, F	1.496	7.835			A, B, D	2.48	A, B, D	1.274	0.034	A, B, D	1.574	15.97
		A, C, F	1.491				A, C, F	2.002	4.12			A, D, E	3.16	A, D, E	1.683	0.007	A, D, E	1.778	12.4
		C, E, F	4.09				C, E, F	2.022	4.03			A, C, E	3.80	A, D, F	2.01	0.003	A, C, E	1.949	10.22
		B, C, E	4.77				B, C, E	2.185	3.37			A, D, F	4.63				A, D, F	2.152	8.282
4	1	A	1.03	A	1.03	1.0	A	1.017	71.74	8	1	C	19.02	C	17.5	0.678	C	4.362	9.182
		E	8.16				E	2.856	0.364			A	26.31	A	24.00	0.118	A	5.129	3.872
		F	8.41				F	2.903	0.048			D	35.42	D	32.2	0.024	D	5.951	0.303
		C	8.43				C	2.903	0.029			B	35.65	B	32.4	0.023	B	5.971	0.235
		B	8.44				B	2.903	0.018			F	36.47	F	33.2	0.02	F	6.039	0.005
	2	A, E	0.92	A, E	0.763	0.084	A, E	0.960	27.91		2	A, C	11.04	A, C	8.36	0.28	A, C	3.322	8.353
		A, D	1.05	A, D	0.855	0.045	A, D	1.025	24.18			C, D	11.73	C, D	8.86	0.204	C, D	3.425	7.705
		A, F	1.17	A, F	0.945	0.026	A, F	1.084	21.32			B, C	12.81	B, C	9.63	0.129	B, C	3.579	6.829
		A, C	1.26	A, C	1.008	0.018	A, C	1.123	19.66			A, E	20.79	A, E	15.36	0.01	A, E	4.559	3.185
		A, B	1.27	A, B	1.015	0.017	A, B	1.128	19.48			C, F	23.15				C, F	4.811	2.587
	3	A, C, E	0.56	A, C, E	0.312	0.004	A, C, E	0.751	20.84		3	B, C, D	4.48	B, C, D	2.06	0.152	B, C, D	2.116	11.08
		A, D, E	0.93	A, D, E	0.447	0.001	A, D, E	0.963	12.45			A, C, D	5.98	A, C, D	2.58	0.045	A, C, D	2.445	8.147
		A, C, D	0.98				A, C, D	0.991	11.72			A, B, C	6.57	A, B, C	2.80	0.028	A, B, C	2.563	7.365
		A, B, C	1.09				A, B, C	1.043	10.52			C, D, F	6.93	C, D, F	2.92	0.022	C, D, F	2.633	6.948
		A, E, F	1.13				A, E, F	1.062	10.14			A, D, F	10.44				A, D, F	3.231	4.423

Table 4.6: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 12-run PB design with responses from the panel of models with $\epsilon \sim N(0,(1.0)^2)$

Mod	AF	Projective based search		Box-Meyer search			Partial F search			Mod	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F			Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
1	1	F	11.41	F	10.4	0.354	F	3.378	4.219	5	1	A	4.14	A	3.80	0.415	A	2.037	4.911
		A	14.63	A	13.3	0.092	A	3.825	1.093			C	5.44	C	4.96	0.096	C	2.333	1.371
		C	14.68	C	13.4	0.091	C	3.831	1.058			E	5.89	E	5.36	0.062	E	2.427	0.508
		E	16.21	E	14.7	0.053	E	4.027	0.01			B	6.11	B	5.56	0.051	B	2.473	0.120
		B	16.22	D	14.7	0.053	B	4.028	0.004			F	6.12	F	5.57	0.051	F	2.474	0.109
	2	A, B	7.30	A, B	5.44	0.180	A, B	2.702	4.741		2	A, E	1.598	A, E	1.91	0.237	A, E	1.598	5.411
		A, C	10.25	A, C	7.55	0.03	A, C	3.202	2.609			B, C	3.61	B, C	2.67	0.038	B, C	1.9	3.045
		C, F	11.27	C, F	8.28	0.018	C, F	3.357	2.134			A, F	4.04	A, F	2.97	0.021	A, F	2.009	2.441
		E, F	11.37				E, F	3.372	2.09			A, C	4.22				A, C	2.053	2.226
		B, F	11.91				B, F	3.451	1.875			A, B	4.52				A, B	2.125	1.9
	3	A, B, C	0.34	A, B, C	0.319	0.997	A, B, C	0.586	68.33		3	A, E, F	1.28	A, E, F	0.545	0.069	A, E, F	1.132	6.329
		A, B, F	6.01				A, B, F	2.451	3.289			A, B, E	1.74				A, B, E	1.318	4.52
		C, E, F	6.41				C, E, F	2.531	3.047			B, C, D	1.96				B, C, D	1.401	3.931
		B, C, F	6.68				B, C, F	2.585	2.897			B, C, E	2.23				B, C, E	1.493	3.395
		A, C, E	8.92				A, C, E	2.986	2.078			A, D, E	2.39				A, D, E	1.546	3.126
2	1	C	9.68	C	8.91	0.707	C	3.112	9.107	6	1	B	29.06	B	26.7	0.588	B	5.391	7.697
		B	15.14	B	13.8	0.064	B	3.891	2.221			A	40.54	A	37.0	0.098	A	6.3671	2.686
		A	16.98	A	15.46	0.034	A	4.121	0.894			C	43.49	C	39.6	0.067	C	6.595	1.826
		D	17.99	D	16.36	0.025	D	4.241	0.285			E	49.67	E	45.2	0.032	E	7.048	0.353
		E	18.22	E	16.56	0.023	E	4.268	0.155			F	50.16	F	45.6	0.031	F	7.083	0.252
	2	A, C	3.97	A, C	3.08	0.794	A, C	1.994	12.85		2	C, D	18.30	C, D	13.8	0.242	C, D	4.278	6.702
		B, C	7.84	B, C	5.85	0.023	B, C	2.8	5.197			A, B	22.36	A, B	16.7	0.084	A, B	4.729	5.00
		C, D	10.32	C, D	7.63	0.005	C, D	3.212	3.311			B, E	23.39	B, E	17.4	0.066	B, E	4.837	4.662
		C, E	10.35	C, E	7.65	0.005	C, E	3.217	3.292			B, C	26.0	B, C	19.3	0.038	B, C	5.099	3.926
		C, F	10.60	C, F	7.83	0.005	C, F	3.255	3.153			B, F	33.70				B, F	5.805	2.42
	3	A, B, C	1.89	A, B, C	0.866	0.217	A, B, C	1.376	13.39		3	B, C, D	6.67	B, C, D	2.92	0.271	B, C, D	2.582	10.45
		A, C, F	4.20	A, C, F	1.73	0.005	A, C, F	2.05	5.715			A, B, E	9.82	A, B, E	4.24	0.035	A, B, E	3.134	6.91
		A, C, E	5.62				A, C, E	2.371	4.13			A, B, C	11.10				A, B, C	3.331	6.05
		B, C, F	5.88				B, C, F	2.425	3.923			A, C, D	17.46				A, C, D	4.178	3.637
		B, C, E	6.35				B, C, E	2.519	3.593			C, D, E	18.48				C, D, E	4.299	3.403
3	1	C	10.13	C	9.31	0.678	C	3.182	8.643	7	1	A	12.39	A	11.6	0.974	A	3.521	21.49
		A	15.65	A	17.16	0.136	A	3.956	2.061			C	33.70	C	30.7	0.005	C	5.805	1.581
		B	16.70	B	15.21	0.046	B	4.086	1.304			F	37.63	F	34.2	0.003	F	6.134	0.372
		E	18.37	E	16.70	0.027	E	4.286	0.277			E	37.73	E	34.3	0.002	E	6.142	0.345
		F	18.74	F	17.04	0.024	F	4.33	0.07			D	38.00	D	34.6	0.002	D	6.164	0.270
	2	A, C	8.54	A, C	6.36	0.082	A, C	2.922	4.701		2	A, D	6.27	A, D	4.97	0.580	A, D	2.504	18.08
		C, F	8.94	C, F	6.65	0.064	C, F	2.991	4.368			A, C	8.79	A, C	6.78	0.105	A, C	2.965	12.13
		B, C	9.20	B, C	6.83	0.055	B, C	3.033	4.175			A, F	12.97	A, F	9.78	0.014	A, F	3.601	7.365
		B, F	11.73				B, F	3.425	2.698			A, E	13.78	A, E	10.36	0.01	A, E	3.712	6.776
		C, E	12.01				C, E	3.466	2.571			A, B	15.20	A, B	11.38	0.006	A, B	3.899	5.89
	3	A, B, C	2.20	A, B, C	1.025	0.385	A, B, C	1.485	11.66		3	A, B, D	2.12	A, B, D	1.32	0.21	A, B, D	1.456	25.73
		B, C, F	6.00				B, C, F	2.45	3.921			A, C, E	4.55	A, C, E	2.10	0.016	A, C, E	2.134	11.67
		C, E, F	6.55				C, E, F	2.559	3.545			A, C, D	4.77	A, C, D	2.14	0.015	A, C, D	2.184	11.12
		A, C, F	6.66				A, C, F	2.58	3.478			A, D, F	5.86	A, D, F	2.56	0.006	A, D, F	2.421	8.942
		A, B, F	8.90				A, B, F	2.983	2.46			A, B, C	10.71				A, B, C	3.272	4.636
4	1	A	1.99	A	1.88	0.992	A	1.411	29.16	8	1	C	20.4	C	18.7	0.5	C	4.517	7.344
		D	7.31	D	6.65	0.001	D	2.704	0.661			A	23.34	A	21.4	0.243	A	4.831	5.163
		B	7.69	B	7.00	0.001	B	2.774	0.131			D	34.56	D	31.4	0.029	D	5.879	0.237
		C	7.71	C	7.01	0.001	C	2.776	0.113			B	34.61	B	31.5	0.029	B	5.883	0.223
		F	7.74	F	7.04	0.001	F	2.783	0.065			E	35.25	E	32.0	0.026	E	5.937	0.037
	2	A, F	1.79	A, F	1.38	0.083	A, F	1.339	11.83		2	A, C	10.44	A, C	7.92	0.397	A, C	3.231	8.629
		A, D	1.85	A, D	1.42	0.07	A, D	1.361	11.36			C, D	12.46	C, D	9.37	0.157	C, D	3.53	6.799
		A, C	2.12	A, C	1.62	0.035	A, C	1.456	9.584			B, C	15.43	B, C	11.50	0.051	B, C	3.928	4.976
		A, B	2.30	A, B	1.74	0.023	A, B	1.516	8.643			A, E	19.29	A, E	14.27	0.016	A, E	4.392	3.448
		A, E	2.45	A, E	1.85	0.016	A, E	1.566	7.922			A, F	19.49	A, F	14.41	0.015	A, F	4.415	3.385
	3	A, B, C	0.76	A, B, C	0.352	0.047	A, B, C	0.875	13.99		3	A, C, D	4.30	A, C, D	1.96	0.187	A, C, D	2.074	11.18
		A, D, F	0.89	A, D, F	0.420	0.018	A, D, F	0.943	11.96			B, C, D	5.00	B, C, D	2.24	0.09	B, C, D	2.236	9.536
		A, B, F	1.29	A, B, F	0.548	0.004	A, B, F	1.137	8.036			A, B, C	6.29	A, B, C	2.69	0.033	A, B, C	2.507	7.47
		A, C, D	2.10				A, C, D	1.449	4.732			A, E, F	7.58	A, E, F	3.19	0.013	A, E, F	2.753	6.10
		A, B, E	2.16				A, B, E	1.47	4.579			A, B, E	8.71				A, B, E	2.951	5.231

Table 4.7: A comparison between factor based methods in identifying four active factors in the six-factor 12-run PB design with responses from the panel of models (model 5 to 8).

Model	Projective based search		Box-Meyer search			Partial F search		
	Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
				$\sigma = 0.6$				
5	A, D, E, F	0.03	A, C, E, F	0.083	0.341	A, D, E, F	0.184	153.6
	A, C, E, F	0.11	A, D, E, F	0.069	0.253	A, C, E, F	0.327	48.63
	A, B, E, F	0.70	A, B, E, F	0.135	0.024	A, B, E, F	0.834	7.409
6	A, B, C, D	0.50	A, B, C, D	0.408	0.887	A, B, C, D	0.703	81.24
	A, B, D, E	0.66	B, C, D, F	0.767	0.027	A, B, D, E	1.564	16.34
	B, C, D, F	3.98	A, B, D, E	0.995	0.007	B, C, D, F	1.998	9.966
7	A, B, C, D	0.005	A, B, C, D	0.30	0.03	A, B, C, D	0.069	5411
	A, D, E, F	0.044	A, D, E, F	0.312	0.024	A, D, E, F	0.209	589.8
	A, C, D, E	1.110	A, C, D, F	0.432	0.004	A, C, D, E	1.054	23.05
8	A, B, C, D	0.024	A, B, C, D	0.398	0.617	A, B, C, D	0.155	1630
	A, C, E, F	0.704	A, C, E, F	0.484	0.212	A, C, E, F	0.839	55.72
	A, B, C, F	4.960	A, C, D, F	0.935	0.006	A, B, C, F	2.227	7.819
				$\sigma = 0.8$				
5	A, B, E, F	0.030	A, D, E, F	0.079	0.344	A, B, E, F	0.178	256.9
	A, D, E, F	0.041	A, B, E, F	0.087	0.207	A, D, E, F	0.202	198.7
	A, B, D, E	0.211	A, B, C, D	0.094	0.131	A, B, D, E	0.459	38.54
6	A, B, D, E	0.046	A, B, C, D	0.425	0.522	A, B, D, E	0.214	1069
	A, B, C, D	0.268	A, B, D, E	0.558	0.118	A, B, C, D	0.518	182.5
	A, B, E, F	4.060	B, C, D, F	0.880	0.010	A, B, E, F	2.015	11.97
7	A, D, E, F	0.022	A, B, C, D	0.307	0.039	A, D, E, F	0.147	1334
	A, B, C, D	0.120	A, D, E, F	0.311	0.036	A, B, C, D	0.34	248.1
	A, C, D, F	0.681	A, C, D, F	0.312	0.036	A, C, D, F	0.825	42.01
8	A, C, E, F	0.120	A, C, E, F	0.348	0.516	A, C, E, F	0.347	302.8
	A, D, E, F	1.949	A, D, E, F	0.505	0.067	A, D, E, F	1.396	18.63
	B, C, D, F	4.024	A, B, C, D	0.508	0.031	B, C, D, F	2.006	8.969
				$\sigma = 1.0$				
5	A, B, D, E	0.053	B, C, D, E	0.133	0.190	A, B, D, E	0.230	117.4
	A, D, E, F	0.075	A, B, E, F	0.138	0.160	A, D, E, F	0.273	83.2
	B, C, D, E	0.25	A, D, E, F	0.138	0.160	B, C, D, E	0.50	24.7
6	A, B, C, D	0.163	A, B, C, D	0.501	0.885	A, B, C, D	0.404	314.4
	A, B, D, E	1.782	B, C, D, E	0.879	0.040	A, B, D, E	1.335	28.74
	A, B, C, E	3.00	A, B, D, E	1.022	0.018	A, B, C, E	1.733	17.03
7	A, B, C, E	0.046	A, B, D, F	0.525	0.144	A, B, C, E	0.215	841.7
	A, B, D, F	0.573	A, C, D, E	0.688	0.032	A, B, D, F	0.757	68
	A, B, D, E	1.418	A, B, C, E	0.695	0.031	A, B, D, E	1.191	27.4
8	A, C, E, F	0.168	A, B, C, D	0.383	0.459	A, C, E, F	0.410	210.2
	A, B, C, D	0.235	A, C, E, F	0.425	0.259	A, B, C, D	0.485	150.4
	A, C, D, F	1.520	A, C, D, F	0.479	0.135	A, C, D, F	1.233	23.15

The 16-run design of Johnson and Jones (2011) for six-factors displayed in Table 2.9 was also investigated in order to compare the search performance of the three factor-based methods considered for this thesis. Using simulated responses from model 4.1 at $\sigma = \{0.6, 0.8, 1.0\}$ given in Appendix, one, two and three active factor(s) respectively were identified. The residual standard error ($\widehat{\sigma}$) of the partial F-test for one, two and three active factors were evaluated at 14, 12 and 8 degrees of freedom respectively. The variance estimate from the projective based search were computed from the replicated runs. For one active factor, there are eight replicated runs, for two active factors there are four runs that are replicated four times and for three active factors there are eight runs replicated twice.

Table 4.8 displays in chronological order the best five one active factor, two and three active factors with $\sigma = \{0.6, 0.8, 1.0\}$ respectively. The results from the analysis of the 16-run design also reveal that for model 4.1 the three methods performed similar in identifying active subspaces. The methods at various values of σ in most cases clearly discriminated the top ranked active subspace from the others. It was observed that all the methods in searching for one and two active factors performed extremely well in discriminating among models. However, in searching for three active factors, the methods displayed relatively weak discriminatory power between the two topmost models, although in all cases each ranked the correct active subspace as the best. The low posterior probabilities assigned by the Box-Meyer search to most of the best active subspaces for three factors indicate that the factor activity for factor C is quite weak. This is because, the posterior probabilities for factors A and B (as active subspace) is very high at various σ, indicating strong factor activity for these two factors.

Table 4.8: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 16-run design with responses from model 4.1 with σ equal to 0.6, 0.8 and 1.0.

AF	Projective based search		Box-Meyer			Partial F		
	Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
$\sigma = 0.6$								
1	B	6.66	B	6.558	1.0	B	2.58	75.36
	A	37.33				A	6.11	1.935
	D	42.45				D	6.515	0.017
	E	42.47				E	6.586	0.020
	F	42.48				F	6.518	0.002
2	A, B	1.60	A, B	1.67	0.998	A, B	1.266	119.6
	B, E	7.32	B, E	6.60	0.000	B, E	2.706	23.07
	B, F	7.69			4	B, F	2.773	21.79
	B, D	7.71				B, D	2.776	21.74
	B, C	7.75				B, C	2.784	21.58
3	A, B, C	0.31	A, B, C	0.55	0.358	A, B, C	0.55	279.8
	A, B, F	1.20	A, B, F	1.029	0.004	A, B, F	1.095	69.69
	A, B, E	1.24	A, B, E	1.051	0.003	A, B, E	1.114	67.31
	A, B, D	2.14				A, B, D	1.462	38.59
	B, C, E	7.24				B, C, E	2.69	10.6
$\sigma = 0.8$								
1	B	5.73	B	5.69	1.0	B	2.393	91.44
	A	39.46				A	6.282	1.297
	C	42.78				C	6.541	0.109
	E	42.97				E	6.555	0.050
	F	42.99				F	6.557	0.040
2	A, B	2.40	A, B	2.30	0.922	A, B	1.549	79.86
	B, E	5.69	B, E	4.91	0.003	B, E	2.386	31.34
	B, C	6.12	B, C	5.24	0.002	B, C	2.473	28.91
	B, F	6.13	B, F	5.26	0.002	B, F	2.476	28.83
	B, D	6.54	B, D	5.58	0.001	B, D	2.558	26.76
3	A, B, C	0.67	A, B, C	0.755	0.329	A, B, C	0.818	127.6
	A, B, E	0.89	A, B, E	0.874	0.110	A, B, E	0.945	95.31
	A, B, F	2.76				A, B, F	1.66	30.17
	A, B, D	2.98				A, B, D	1.726	27.81
	B, C, F	4.76				B, C, F	2.182	16.98
$\sigma = 1.0$								
1	B	5.68	B	5.63	1.0	B	2.383	88.18
	A	37.82				A	6.15	1.349
	F	41.18				F	6.417	0.098
	E	41.34				E	6.434	0.021
	D	41.42				D	6.436	0.014
2	A, B	2.25	A, B	2.17	0.947	A, B	1.5	81.95
	B, F	6.16	B, F	5.26	0.001	B, F	2.481	27.44
	B, E	6.39	B, E	5.44	0.001	B, E	2.527	26.3
	B, F	6.43	B, F	5.48	0.001	B, C	2.536	26.09
	B, D	6.47	B, D	5.51	0.001	B, D	2.544	25.89
3	A, B, C	1.17	A, B, C	1.0	0.044	A, B, C	1.081	69.82
	A, B, F	1.99	A, B, D	1.44	0.003	A, B, F	1.411	40.48
	A, B, E	2.30	A, B, E	1.60	0.001	A, B, E	1.515	35
	A, B, D	3.12				A, B, D	1.766	25.45
	B, C, F	4.32				B, C, F	2.078	18.05

Using the panel of models suggested by Tyssedal and Hussain (2016), further studies on the search performance of the three factor based methods in identifying the active subspace of the design in Table 2.9 using simulated responses at $\sigma = \{0.6, 0.8, 1.0\}$ were performed. Tables 4.9, 4.10 and 4.11 display in chronological order the best five subspaces with one, two and three active factors with $\sigma = \{0.6, 0.8, 1.0\}$ respectively. From the analysis, the methods displayed similar results. For the models in the panel with three active factors, the methods in most cases identified the correct active subspace (A, B, C), and ranked it as best among the five topmost subspaces. For model 3 with $\sigma = 1.0$, all the three methods ranked A, C, E instead of A, B, C as the most important active subspace, meanwhile factor E was inert, however, the methods weakly discriminated among the two topmost models, an exception could be made for the Box-Meyer search for the high posterior probability (0.854) assigned to subspace A, C, E compared to 0.146 for A, B, C. In general, the methods in most cases discriminated well among the models at the levels of σ studied.

Table 4.9: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 16-run design with responses from the panel of models with $\epsilon \sim N(0, (0.6)^2)$

Mod	AF	Projective based search		Box-Meyer search			Partial F search			Mod	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F			Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
1	1	A	9.26	A	8.66	0.188	A	3.043	2.534	5	1	C	2.62	C	2.46	0.481	C	1.62	6.151
		F	9.41	F	8.80	0.087	F	3.068	2.266			A	3.09	A	2.90	0.141	A	1.759	3.079
		E	9.87	E	9.22	0.117	E	3.141	0.743			E	3.56	E	3.32	0.050	E	1.886	0.859
		D	10.92	D	10.20	0.055	D	3.305	0.011			F	3.63	F	3.38	0.044	F	1.904	0.581
		B	10.94	B	10.20	0.055	B	3.307	0.001			D	3.75	D	3.50	0.034	D	1.936	0.102
	2	A, C	4.91	A, C	3.99	0.434	A, C	2.215	6.398		2	C, D	1.95	C, D	1.58	0.135	C, D	1.397	5.027
		A, B	6.14	A, B	4.96	0.084	A, B	2.478	4.311			B, C	2.07	B, C	1.67	0.088	B, C	1.438	4.517
		A, F	8.99	A, F	7.22	0.005	A, F	2.998	1.677			A, C	2.27	A, C	1.83	0.045	A, C	1.505	3.783
		A, E	9.55				A, E	3.091	1.339			C, E	2.68				C, E	1.636	2.585
		A, D	10.79				A, D	3.284	0.732			C, F	2.69				C, F	1.641	2.544
	3	A, B, C	0.23	A, B, C	0.222	1.0	A, B, C	0.477	94.69		3	B, C, D	1.43	B, C, D	0.795	0.033	B, C, D	1.194	4.154
		A, B, F	4.22				A, B, F	2.055	4.036			C, D, E	1.44				C, D, E	1.199	4.115
		A, C, E	5.01				A, C, E	2.238	3.225			B, C, E	1.69				B, C, E	1.3	3.323
		A, B, E	5.13				A, B, E	2.265	3.121			C, D, F	1.71				C, D, F	1.309	3.266
		A, C, F	5.31				A, C, F	2.305	2.974			A, C, D	1.72				A, C, D	1.313	3.239
2	1	C	9.14	C	8.58	0.567	C	3.024	7.774	6	1	B	20.94	B	19.6	0.525	B	4.576	6.81
		B	10.84	B	10.14	0.162	B	3.292	4.38			A	25.60	A	24.0	0.116	A	5.068	2.967
		A	12.91	A	12.06	0.044	A	3.593	1.427			C	27.16	C	25.4	0.076	C	5.212	2.043
		E	13.59	E	12.70	0.030	E	3.687	0.649			D	29.47	D	27.5	0.042	D	5.429	0.788
		F	13.61	F	12.71	0.030	F	3.689	0.634			E	31.13	E	29.0	0.028	E	5.579	0.005
	2	A, C	5.99	A, C	4.87	0.283	A, C	2.447	7.09		2	A, B	18.08	A, B	14.6	0.057	A, B	4.252	4.034
		B, C	6.70	B, C	5.44	0.125	B, C	2.588	5.913			C, D	18.44	C, D	14.9	0.050	C, D	4.294	3.879
		C, E	7.16	C, E	5.80	0.076	C, E	2.676	5.271			B, C	19.79	B, C	16.0	0.029	B, C	4.449	3.339
		A, B	9.29	A, B	7.46	0.011	A, B	3.048	3.146			C, E	22.21				C, E	4.713	2.54
		C, F	9.82				C, F	3.134	2.76			B, F	20.79				B, F	4.56	2.988
	3	A, B, C	0.21	A, B, C	0.245	1.0	A,B, C	0.463	131.8		3	B, C, D	9.72	B, C, D	5.42	0.124	B, C, D	3.117	5.266
		A, C, E	5.77				A, C, E	2.402	3.733			C, D, E	13.20				C, D, E	3.633	3.575
		A, C, F	5.84				A, C, F	2.416	3.733			C, D, F	14.58				C, D, F	3.818	3.128
		A, B, F	6.53				A, B, F	2.555	3.215			A, C, D	18.06				A, C, D	4.25	2.304
		B, C, F	6.80				B, C, F	2.608	3.038			B, C, E	18.88				B, C, E	4.345	2.155
3	1	C	6.73	C	6.63	0.731	C	2.595	9.429	7	1	A	6.94	A	6.64	0.999	A	2.634	36.47
		B	9.32	B	8.71	0.066	B	3.052	2.932			E	23.59				E	4.857	0.843
		A	9.78	A	9.41	0.046	A	3.127	2.136			D	23.87				D	4.886	0.670
		F	11.22	F	10.47	0.017	F	3.350	0.060			F	24.07				F	4.906	0.547
		D	11.26	D	10.51	0.042	D	3.356	0.011			B	24.09				B	4.908	0.535
	2	B, C	5.56	B, C	4.51	0.099	B, C	2.358	5.458		2	A, D	2.12	A, D	1.91	0.992	A, D	1.455	51.12
		C, E	5.85	C, E	4.74	0.068	C, E	2.418	4.991			A, E	5.33	A, E	4.46	0.002	A, E	2.309	17.88
		A, C	6.09	A, C	4.93	0.051	A, C	2.468	4.635			A, F	6.95				A, F	2.636	12.79
		B, E	6.59	B, E	5.32	0.028	B, E	2.567	3.98			A, B	7.02				A, B	2.649	12.63
		A, E	6.84	A, E	5.52	0.022	A, E	2.615	2.708			A, C	7.38				A, C	2.717	11.82
	3	A, B, C	0.12	A, B, C	0.166	0.964	A, B, C	0.341	192.7		3	A, B, D	1.55	A, B, D	1.05	0.106	A, B, D	1.363	25.78
		A, B, E	0.29	A, B, E	0.258	0.036	A, B, E	0.538	76.86			A, D, F	1.86	A, D, F	1.20	0.059	A, D, F	1.363	25.78
		B, C, F	5.57				B, C, F	2.36	2.904			A, C, D	2.00	A, C, D	1.29	0.023	A, C, D	1.414	23.88
		B, C, E	5.74				B, C, E	2.396	2.782			A, D, E	2.45	A, D, E	1.55	0.009	A, D, E	1.566	19.25
		B, D, E	6.56				B, D, E	2.562	2.29			A, E, F	5.98				A, E, F	2.446	7.219
4	1	A	2.11	A	2.02	0.998	A	1.453	32.42	8	1	C	13.18	C	12.5	0.970	C	3.63	18.44
		E	6.37				E	2.524	1.39			A	25.36	A	23.7	0.008	A	5.036	2.853
		F	6.74				F	2.597	0.535			E	29.39	E	27.4	0.003	E	5.421	0.545
		C	6.85				C	2.617	0.315			F	29.88	F	27.9	0.002	F	5.466	0.308
		D	7.00				D	2.646	0.006			D	30.50	D	28.5	0.002	D	5.523	0.014
	2	A, E	1.65	A, E	1.38	0.193	A, E	1.286	15.76		2	A, C	9.22	A, C	7.59	0.287	A, C	3.036	12.00
		A, F	2.12	A, F	1.750	0.032	A, F	1.456	11.41			B, C	10.82	B, C	8.85	0.090	B, C	3.289	9.171
		A, C	2.22	A, C	1.83	0.023	A, C	1.491	10.70			C, D	10.85	C, D	8.88	0.088	C, D	3.294	9.133
		A, B	2.46	A, B	2.01	0.011	A, B	1.568	9.29			C, E	12.33	C, E	10.05	0.035	C, E	3.511	7.56
		A, D	2.462	A, D	2.018	0.011	A, D	1.569	9.284			C, F	13.05	C, F	10.62	0.023	C, F	3.612	6.921
	3	A, B, C	0.23	A, B, C	0.189	0.988	A, B, C	0.484	58.69		3	A, B, C	6.91	A, B, C	3.93	0.052	A, B, C	2.629	7.695
		A, B, E	1.35				A, B, E	1.162	9.24			A, C, D	6.99	A, C, D	4.00	0.046	A, C, D	2.643	7.600
		A, E, F	1.85				A, E, F	1.361	6.414			C, D, E	7.88	C, D, E	5.09	0.007	B, C, D	2.808	6.602
		A, D, E	1.92				A, D, E	1.385	6.155			C, D, F	8.52	B, C, D	5.27	0.006	C, D, F	2.919	6.025
		A, C, E	2.17				A, C, E	1.474	5.303			B, C, D	9.42				B, C, D	3.069	5.340

Table 4.10: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 16-run design with responses from the panel of models with $\epsilon \sim N(0, (0.8)^2)$

Mod	AF	Projective based search		Box-Meyer search			Partial F search			Mod	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F			Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
1	1	F	9.52	F	8.90	0.135	F	3.086	1.444	5	1	C	3.68	C	3.45	0.440	C	1.919	5.945
		A	9.65	A	9.02	0.122	A	3.107	1.238			A	4.22	A	3.95	0.161	A	2.055	3.407
		E	9.91	E	9.26	0.100	E	3.148	0.837			E	4.80	E	4.48	0.062	E	2.191	1.31
		B	10.38	B	9.68	0.072	B	3.221	0.176			F	4.85	F	4.53	0.057	F	2.203	1.144
		C	10.43	C	9.73	0.069	C	3.229	0.108			B	5.21	B	4.86	0.034	B	2.282	0.114
	2	A, B	3.88	A, B	3.17	0.814	A, B	1.969	8.637		2	B, C	2.41	B, C	1.96	6.277	B, C	1.553	6.154
		A, C	7.68	A, C	6.18	0.005	A, C	2.771	2.382			B, C	3.10	A, C	2.50	0.044	A, C	1.76	3.911
		A, F	9.95	A, F	7.98	0.001	A, F	3.154	0.927			C, D	3.30	C, D	2.66	0.028	C, D	1.816	3.431
		E, F	10.38				E, F	3.222	0.721			C, F	3.51				C, F	1.873	2.980
		A, E	10.54				A, E	3.247	0.649			C, E	3.73				C, E	1.930	2.579
	3	A, B, C	0.44	A, B, C	0.328	1.0	A, B, C	0.660	47.12		3	A, B, C	1.73	A, B, C	0.994	0.061	A, B, C	1.316	4.916
		A, B, F	2.91				A, B, F	1.706	6.079			B, C, D	1.97	B, C, D	1.158	0.019	B, C, D	1.403	4.192
		A, B, E	3.13				A, B, E	1.768	5.576			C, D, E	2.45				C, D, E	1.564	3.151
		A, B, D	5.60				A, B, D	2.366	2.608			C, D, F	2.64				C, D, F	1.624	2.836
		A, C, F	6.70				A, C, F	2.645	1.861			B, C, E	2.61				B, C, E	1.616	2.877
2	1	C	7.73	C	7.28	0.825	C	2.781	11.38	6	1	B	19.62	B	18.4	0.565	B	4.429	7.118
		B	11.34	B	10.61	0.049	B	3.368	3.298			A	25.33	A	23.7	0.085	A	5.033	2.352
		A	12.44	A	11.62	0.025	A	3.527	1.773			C	25.62	C	24.0	0.078	C	5.062	2.165
		E	13.66	E	12.76	0.012	E	3.696	0.360			D	28.42	D	26.5	0.036	D	5.331	0.575
		F	13.76	F	12.84	0.012	F	3.709	0.263			E	29.33	E	27.4	0.029	E	5.416	0.122
	2	A, C	4.67	A, C	3.83	0.477	A, C	2.16	10.02		2	C, D	17.47	C, D	14.1	0.050	C, D	4.180	3.903
		C, E	5.70	C, E	4.64	0.112	C, E	2.387	7.484			A, B	17.89	A, B	14.4	0.042	A, B	4.230	3.718
		B, C	5.90	B, C	4.80	0.086	B, C	2.429	7.083			B, C	18.20	B, E	14.7	0.037	B, E	4.266	3.589
		C, F	8.65	C, F	6.98	0.005	C, F	2.941	3.564			B, F	20.47				B, C	4.524	2.748
		C, D	8.82	C, D	7.11	0.005	C, D	2.969	3.422			C, E	20.75				C, E	4.524	2.748
	3	A, B, C	0.81	A, B, C	0.559	0.992	A, B, C	0.901	33.3		3	B, C, D	8.09	B, C, D	4.61	0.242	B, C, D	2.844	6.176
		A, C, E	4.15				A, C, E	2.036	5.622			C, D, E	11.61	C, D, E	6.46	0.019	C, D, E	3.408	3.952
		A, C, F	4.96				A, C, F	2.228	4.504			C, D, F	14.88				C, D, F	3.858	2.832
		B, C, E	6.13				B, C, E	2.475	3.434			B, C, F	16.03				B, C, F	4.004	2.549
		B, C, F	6.19				B, C, F	2.488	3.385			B, C, E	18.16				B, C, E	4.261	2.117
3	1	C	7.05	C	6.64	0.896	C	2.655	13.39	7	1	A	7.78	A	7.44	0.998	A	2.790	34.17
		B	11.10	B	10.38	0.032	B	3.331	3.407			F	25.31				C	5.031	0.818
		A	12.67	A	11.33	0.012	A	3.559	1.248			C	25.58				C	5.058	0.664
		E	13.71	E	12.80	0.007	E	3.703	0.086			D	25.72				D	5.071	0.586
		D	13.76	D	12.84	0.006	D	3.709	0.037			B	25.89				B	5.088	0.488
	2	B, C	5.06	B, C	4.13	0.0269	B, C	2.249	8.729		2	A, D	3.29	A, D	2.86	0.937	A, D	1.815	33.96
		C, E	5.83	C, E	4.74	0.096	C, E	2.414	7.051			A, E	6.72	A, E	5.58	0.005	A, E	2.593	14.59
		A, C	6.87	A, C	5.58	0.029	A, C	2.622	5.361			A, C	7.24	A, C	5.98	0.004	A, C	2.690	13.27
		B, E	7.82	B, E	6.33	0.011	B, E	2.797	4.227			A, F	7.36	A, F	6.08	0.003	A, F	2.712	12.99
		C, F	8.00	C, F	6.47	0.009	C, F	2.829	4.046			A, B	8.00	A, B	6.598	0.002	A, B	2.829	11.62
	3	A, B, C	0.27	A, B, C	0.271	0.996	A, B, C	0.520	101		3	A, C, D	1.91	A, C, D	1.32	0.289	A, C, D	1.383	26.86
		A, B, E	0.84	A, B, E	0.571	0.004	A, B, E	0.916	31.75			A, D, E	2.73	A, D, F	1.850	0.036	A, D, E	1.652	18.48
		B, C, F	5.70				B, C, F	2.388	3.696			A, D, F	3.07	A, D, E	1.89	0.030	A, D, F	1.753	16.3
		B, C, E	5.21				B, C, E	2.282	4.157			A, C, E	7.05	A, B, D	1.99	0.014	A, C, E	2.656	6.452
		B, D, E	7.17				B, D, E	2.677	2.706			A, B, F	7.24				A, B, F	2.691	6.255
4	1	A	2.16	A	2.07	0.999	A	1.470	37.10	8	1	C	15.46	C	14.6	0.953	C	3.932	16.13
		E	7.32				E	2.705	1.107			A	29.17	A	27.3	0.009	A	5.401	1.968
		F	7.63				F	2.762	0.480			E	31.89	E	29.8	0.005	E	5.647	0.610
		B	7.78				C	2.7906	0.193			F	32.01	F	29.9	0.004	F	5.658	0.549
		C	7.89				C	2.809	0.008			B	33.12	B	30.9	0.003	B	5.755	0.065
	2	A, E	1.62	A, E	1.35	0.245	A, E	1.272	18.76		2	C, D	11.48	C, D	9.14	0.237	C, D	3.388	9.932
		A, F	2.21	A, F	1.82	0.027	A, F	1.485	12.70			B, C	11.98	B, C	9.80	0.140	B, C	3.461	8.965
		A, C	2.30	A, C	1.89	0.020	A, C	1.515	12.04			A, C	13.25	A, C	10.80	0.067	A, C	3.640	7.722
		A, B	2.36	A, B	1.94	0.016	A, B	1.537	11.60			C, F	14.46	C, F	11.76	0.036	C, F	3.803	6.739
		A, D	2.49	A, D	2.04	0.011	A, D	1.578	10.79			C, E	15.26	C, E	12.39	0.024	C, E	3.906	6.176
	3	A, B, C	0.65	A, B, C	0.416	0.196	A, B, C	0.805	23.21		3	B, C, D	7.46	B, C, D	4.34	0.082	B, C, D	2.731	7.782
		A, B, E	0.82	A, B, E	0.5090	0.043	A, B, E	0.908	18.00			C, D, F	8.42	A, C, D	5.34	0.017	C, D, F	2.901	6.766
		A, E, F	1.81				A, E, F	1.347	7.559			C, D, E	8.57	C, D, E	5.87	0.008	C, D, E	2.927	6.624
		A, D, E	1.93				A, D, E	1.390	7.027			A, C, D	9.52				A, C, D	3.086	5.843
		A, C, E	2.15				A, C, E	1.465	6.208			B, C, E	11.66				B, C, E	3.414	4.566

Table 4.11: A comparison between factor based methods for identifying one, two and three active factor(s) in the six-factor 16-run design with responses from the panel of models with $\epsilon \sim N(0, (1.0)^2)$

Mod	AF	Projective based search		Box-Meyer search			Partial F search			Mod	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F			Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
1	1	F	9.02	F	8.44	0.249	F	3.004	3.537	5	1	A	4.55	A	4.26	0.313	A	2.132	4.564
		A	9.19	A	8.60	0.216	A	3.032	3.210			C	4.81	C	4.50	0.207	C	2.193	3.558
		E	10.71	E	10.00	0.070	E	3.272	0.786			B	5.52	B	5.16	0.074	B	2.349	1.291
		D	11.02	D	10.28	0.057	D	3.319	0.368			F	5.69	F	5.32	0.059	F	2.386	0.821
		B	11.25	B	10.50	0.049	B	3.354	0.072			E	5.90	E	5.50	0.046	E	2.428	0.318
	2	A, C	5.57	A, C	4.52	0.244	A, C	2.360	5.472		2	A, C	3.87	A, C	3.12	0.040	A, C	1.967	3.271
		A, B	6.17	A, B	4.98	0.115	A, B	2.484	4.550			B, C	4.02	B, C	3.24	0.030	B, C	2.006	2.99
		A, F	8.07	A, F	6.49	0.016	A, F	2.84	2.541			C, D	4.16	C, D	3.35	0.023	C, D	2.04	2.758
		C, F	9.21				C, F	3.034	1.731			A, F	4.40				A, F	2.098	2.394
		D, F	9.85				D, F	3.139	1.353			A, E	4.70				A, E	2.167	1.989
	3	A, B, C	1.27	A, B, C	0.778	0.950	A, B, C	1.128	16.16		3	B, C, F	2.31				B, C, F	1.521	4.066
		A, B, F	3.60				A, B, F	1.898	5.133			C, D, F	2.51				C, D, F	1.584	3.661
		A, C, F	4.58				A, C, F	2.139	3.801			A, B, C	2.69				A, B, C	1.639	4.501
		A, C, E	5.57				A, C, E	2.361	2.913			A, C, D	3.48				A, C, D	1.866	2.320
		A, B, E	5.92				A, B, E	2.433	2.676			B, C, D	3.71				B, C, D	1.926	2.108
2	1	C	10.52	C	9.85	0.405	C	3.243	5.782	6	1	B	23.07	B	21.6	0.240	B	4.803	3.484
		B	11.59	B	10.84	0.198	B	3.404	3.955			C	24.44	C	22.9	0.157	C	4.944	2.506
		A	13.26	A	12.39	0.073	A	3.641	1.690			A	25.46	A	23.8	0.115	A	5.046	1.839
		E	14.30	E	13.25	0.041	E	3.782	0.544			D	26.50	D	24.8	0.086	D	5.148	1.219
		F	14.33	F	13.37	0.041	F	3.785	0.521			E	28.66	E	26.7	0.048	E	5.353	0.078
	2	A, C	7.43	A,C	6.02	0.157	A, C	2.725	5.339		2	C, D	12.12	C, D	9.87	0.536	C, D	3.481	7.096
		C, E	8.26	C, E	6.68	0.072	C, E	2.874	4.395			C, E	20.50	C, E	16.50	0.011	C, E	4.528	2.559
		B, C	8.45	B, C	6.83	0.061	B, C	2.906	4.211			B, C	21.58	B, C	17.37	0.008	B, C	4.646	2.229
		A, B	9.13	A, B	7.37	0.034	A, B	3.022	3.591			A, B	22.66				A, B	4.760	1.860
		B, E	10.84				B, E	3.292	2.398			B, F	24.03				B, F	4.902	1.596
	3	A, B, C	1.62	A, B, C	0.994	0.944	A, B, C	1.273	17.19		3	C, D, E	5.777	C, D, E	3.49	0.568	C, D, E	2.403	8.838
		A, C, E	5.24				A, C, E	2.289	4.527			B, C, D	7.28	B, C, D	4.24	0.132	B, C, D	2.698	6.775
		A, B, F	7.18				A, B, F	2.680	2.994			A, C, D	11.71				A, C, D	3.422	3.778
		B, C, F	7.29				B, C, F	2.700	2.933			C, D, F	13.21				C, D, F	3.635	3.218
		A, B, E	7.76				A, B, E	2.785	2.688			B, C, E	18.22				B, C, E	4.269	2.019
3	1	C	8.40	C	7.88	0.560	C	2.898	7.412	7	1	A	11.03	A	10.4	0.982	A	3.321	20.63
		B	10.13	B	9.48	0.140	B	3.182	3.756			B	24.97	B	23.3	0.002	B	4.997	1.291
		A	11.31	A	10.57	0.062	A	3.363	1.898			F	25.03	F	23.4	0.002	F	5.003	1.257
		F	12.67	F	11.83	0.027	F	3.559	0.192			C	25.27	C	23.6	0.002	C	5.027	1.111
		D	12.67	D	11.83	0.027	D	3.560	0.189			D	26.52	D	24.8	0.002	D	5.150	0.395
	2	B, E	5.73	B, E	4.66	0.243	B, E	2.394	6.458		2	A, D	5.62	A, D	4.71	0.818	A, D	2.371	18.63
		B, C	6.62	B, C	5.36	0.085	B, C	2.573	5.056			A, F	10.13	A, F	8.28	0.012	A, F	3.183	8.565
		C, E	6.92	C, E	5.60	0.061	C, E	2.631	4.658			A, B	10.14	A, B	8.29	0.012	A, B	3.185	8.544
		A, C	7.99	A, C	6.45	0.021	A, C	2.827	3.499			A, C	10.31	A, C	8.42	0.010	A, C	3.311	8.347
		C, F	9.54				C, F	3.088	2.287			A, E	11.20	A, E	9.12	0.006	A, E	3.346	7.369
	3	A, B, E	0.56	A, B, E	0.417	0.854	A, B, E	0.751	44.43		3	A, B, D	4.08	A, B, D	2.42	0.136	A, C, D	2.019	12.24
		A, B, C	0.77	A, B, C	0.528	0.146	A, B, C	0.880	32.04			A, C, D	4.51	A, C, D	2.65	0.070	A, C, D	2.123	10.96
		B, C, E	5.07				B, C, E	2.251	3.928			A, D, F	5.92	A, D, F	3.35	0.018	A, D, F	2.433	8.074
		B, C, F	5.91				B, C, F	2.432	3.199			A, D, E	5.84	A, D, E	3.40	0.016	A, D, E	2.416	8.201
		B, D, E	6.11				B, D, E	2.472	3.060			A, B, F	5.55				A, B, F	2.356	8.186
4	1	A	2.70	A	2.56	0.988	A	1.643	22.15	8	1	C	14.36	C	13.6	0.975	C	3.79	19.46
		F	6.68				F	2.584	0.612			A	27.93	A	36.1	0.007	A	5.285	3.207
		B	6.73				B	2.594	0.491			F	33.27	F	31.1	0.002	F	5.768	0.446
		E	6.82				E	2.612	0.296			E	33.66	E	31.4	0.002	E	5.802	0.277
		C	6.83				C	2.614	0.280			B	34.30	B	32.0	0.002	B	5.857	0.012
	2	A, F	2.72	A, F	2.22	0.038	A, F	1.648	7.973		2	A, C	9.30	A, C	7.68	0.435	A, C	3.049	13.24
		A, B	2.82	A, B	2.30	0.029	A, B	1.679	7.535			C, D	12.19	C, D	9.97	0.061	C, D	3.491	9.143
		A, D	2.84	A, D	2.31	0.027	A, D	1.684	7.469			B, C	12.43	B, C	10.17	0.053	B, C	3.526	8.888
		A, E	2.91	A, E	2.37	0.023	A, E	1.705	7.184			C, F	14.03	C, F	11.44	0.022	C, F	3.746	7.416
		A, C	2.98	A, C	2.43	0.019	A, C	1.727	6.907			C, E	14.77	C, E	12.01	0.015	C, E	3.843	6.851
	3	A, B, C	0.67	A, B, C	0.42	0.596	A, B, C	0.820	19.600		3	A, C, D	6.94	A, C, D	4.06	0.068	A, C, D	2.634	8.754
		A, B, E	1.92	A, B, E	1.08	0.001	A, B, E	1.384	6.128			C, D, F	7.25	A, B, C	4.24	0.049	C, D, F	2.693	8.326
		A, B, F	2.37				A, B, F	1.538	4.745			A, B, C	7.42	C, D, F	5.20	0.010	A, B, C	2.723	8.12
		A, C, F	3.35				A, C, F	1.830	3.016			A, C, E	9.00				A, C, E	3.000	9.143
		A, D, E	3.35				A, D, E	1.831	3.014			A, C, F	9.43				A, C, F	3.071	6.139

The models 5 to 8 involved four factors, the methods' performance in identifying four active factor subspace for a six-factor 16-run design were studied. The 16-run design has a special projection properties. Studying its aliasing structure, factors AB, CD, and EF are free of aliasing with main effects. The subspaces $ABCD$, $ABEF$ and $CDEF$ form a full 2^4 factorial design and can be studied using standard methods. Table 4.7 presents the best three four active factors models identified by the methods studied. Models with all four main effects and six two-factor interactions were investigated and thus the partial F search evaluated the residual standard error ($\widehat{\sigma}$) with five degrees of freedom. From the analysis, all the methods correctly identified the most important active subspace for the simulated models except for model 6 with $\sigma = 0.8$ and $\sigma = 1.0$. The methods appreciably discriminated between the models compared, however, for model 6, all the methods only weakly discriminated among the models. The results from model 6 suggest the need for follow-up studies.

Table 4.12: A comparison between factor based methods in identifying four active factors in the six-factor 16-run design with responses from the panel of models (model 5 to 8).

Model	Projective based search		Box-Meyer search			Partial F search		
	Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	$\widehat{\sigma}$	F
				$\sigma = 0.6$				
5	A, B, C, D	0.30	A, B, C, D	0.134	0.755	A, B, C, D	0.548	17.120
	B, C, D, E	1.08				B, C, D, E	1.04	4.384
	B, C, D, F	2.31				B, C, D, F	1.40	3.286
6	A, B, C, D	0.25	A, B, C, D	0.373	0.57	A, B, C, D	0.502	172.3
	C, D, E, F	0.30	C, D, E, F	0.387	0.43	C, D, E, F	0.543	147.1
	B, C, D, E	6.90				B, C, D, E	2.626	5.85
7	A, B, C, D	0.52	A, C, D, F	0.410	0.047	A, B, C, D	0.720	67.0
	A, C, D, F	0.54	A, B, C, D	0.404	0.019	A, C, D, F	0.732	64.9
	A, B, D, E	0.99	A, B, D, E	0.542	0.006	A, B, D, E	0.994	34.9
8	A, B, C, D	0.17	A, B, C, D	0.339	0.999	A, B, C, D	0.407	258
	A, C, D, F	5.26				A, C, D, F	2.294	7.626
	A, C, D, E	6.07				A, C, D, E	2.464	6.540
				$\sigma = 0.8$				
5	A, B, C, D	0.43	A, B, C, D	0.191	0.681	A, B, C, D	0.656	16.570
	B, C, D, F	1.21				B, C, D, F	1.101	5.557
	A, C, D, E	1.99				A, C, D, E	1.406	3.218
6	C, D, E, F	0.49	C, D, E, F	0.436	0.64	C, D, E, F	0.698	84.49
	A, B, C, D	0.59	A, B, C, D	0.471	0.36	A, B, C, D	0.770	69.34
	B, C, D, F	0.59				B, C, D, F	2.561	5.817
7	A, B, C, D	0.60	A, B, C, D	0.448	0.124	A, B, C, D	0.775	61.99
	A, C, D, F	1.37				A, C, D, F	1.170	26.89
	A, C, D, E	2.01				A, C, D, E	1.418	18.15
8	A, B, C, D	0.57	A, B, C, D	0.498	0.993	A, B, C, D	0.754	81.40
	B, C, D, E	5.75				B, C, D, E	2.397	7.61
	B, C, D, F	6.33				B, C, D, F	2.515	6.86
				$\sigma = 1.0$				
5	A, B, C, D	0.81	A, B, C, D	0.322	0.180	A, B, C, D	0.897	9.985
	A, C, D, F	2.50				A, C, D, F	1.582	2.873
	C, D, E, F	2.62				C, D, E, F	1.62	2.715
6	C, D, E, F	1.76	C, D, E, F	0.848	0.591	C, D, E, F	1.326	22.46
	A, B, C, D	2.48	B, C, D, E	1.172	0.144	A, B, C, D	1.574	15.79
	B, C, D, E	2.75	A, B, C, D	1.086	0.093	B, C, D, E	1.659	14.15
7	A, B, C, D	0.26	A, B, C, D	0.338	0.981	A, B, C, D	0.505	149.5
	A, B, D, F	2.69				A, B, D, F	1.639	13.71
	A, C, D, E	2.93				A, C, D, E	1.711	12.55
8	A, B, C, D	0.80	A, B, C, D	0.583	0.954	A, B, C, D	0.893	59.8
	A, C, D, F	4.49				A, C, D, F	2.118	10.22
	C, D, E, F	5.65				C, D, E, F	2.378	8.002

4.2 Real data-Metal cutting experiment

One of the motivations for this thesis was to explore data which were originally collected as part of an investigation to compare the techniques of a full factorial design in six factors (Garzon, 2000). In this case study, the objective was the optimisation of the milling operation as a typical metal cutting process through DoE. The aim of the optimization was to define the best machining conditions in a manufacturing environment. The full factorial experiment involved six factors, each set at two levels. Table 4.13 displays the factors and their levels. Eight repetitions of each factor combination were performed. Note that repeated measurements and not replications were carried out. The mean and standard deviation of the repetitions were calculated. Mønness et al. (2007) determined that a reciprocal transformation of the original response (surface roughness) is necessary, because a high correlation was identified between the mean and standard deviation. This transformation is employed for this study.

Table 4.13: Metal Cutting factors and levels

Factor	Low (-)	High (+)
Tool Speed (A)	2700 rpm	3200 rpm
Workpiece speed (B)	203 mm/min	330 mm/min
Depth of cut (C)	0.5 mm	1.0 mm
Coolant (D)	off	On
Direction of cut (E)	Conventional	Climbing
Number of cut (F)	1	2

Edwards et al. (2013) used Lenth's method to analyze the unreplicated 64-run data, and obtained a pseudo standard error of 0.01338. Controlling the individual error rate of $\alpha = 0.05$, a critical value of 2.014 was found. Eleven effects: D, E, F, CD, CE, CF, DE, EF, BCD and DEF were found significant. The final model is presented in

Table 4.14.

Table 4.14: Final Model from full factorial ($R^2 = 95\%$)

Intercept	Estimate	t-ratio
Constant	0.825	76.33
D	0.216	20.04
E	0.137	12.68
F	-0.062	-5.71
CD	0.032	2.98
CE	0.037	3.4
CF	0.036	3.37
DE	-0.190	-17.62
DF	0.056	5.21
EF	0.029	2.67
BCD	0.028	2.66
DEF	-0.051	-4.75

The search performance of the factor based methods in identifying active factors with the metal cutting data was explored. The transformed data is as presented in Table 4.15 for the six-factor 12 run PB design. Edwards et al. (2013) study revealed factor activity for $C - F$, however the posterior probabilities for the two topmost models suggested some ambiguity as to which factors are active. Using a particular fraction and follow-up runs, they fitted a 2fi model in $C - F$, and found D and DE as significant. Mønness et al. (2007) employed two disjoint PB design in analyzing this experiment found that none of the two designs yielded a significant results, as measured by the F-test. However, one of the designs revealed factor D as significant, assuming only model with main effects. Table 4.16 presents the results from the analysis employing the factor based methods. From the analysis, all the methods performed well in identifying the important active subspaces for one, two, three and four active factors. All the three methods reported factor activity for factors $C - F$. The projection based and partial F could not discriminate well among the five top models in both cases. This results seems to suggest that a follow-up experiment appear to be necessary. However, the Box-Meyer methods performed relatively well in discriminating among the two topmost models.

Table 4.15: Six-factor 12-run PB design with response for the metal cutting experiment

Run	A	B	C	D	E	F	Response
11	-1	-1	-1	1	-1	1	1.1
60	-1	-1	1	-1	-1	1	0.08
17	-1	-1	1	-1	1	-1	1.02
53	-1	1	-1	-1	1	-1	0.96
19	-1	1	-1	1	1	1	0.82
5	-1	1	1	1	-1	-1	1.1
59	1	-1	-1	-1	1	1	0.76
54	1	-1	-1	1	-1	-1	1.16
29	1	-1	1	1	1	-1	0.93
36	1	1	-1	-1	-1	-1	0.75
34	1	1	1	-1	-1	1	0.05
28	1	1	1	1	1	1	1.06

Table 4.16: A comparison between factor based methods for identifying one, two, three and four active factor(s) in the six-factor 12-run design with responses from the metal cutting experiment

AF	Projective based search		Box-Meyer search			Partial F search		
	Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	σ	F
1	D	0.10	D	0.092	0.402	D	0.318	5.372
	F	0.12	F	0.110	0.157	F	0.3465	2.918
	C	0.14	E	0.128	0.066	E	0.3752	1.016
	E	0.14	C	0.128	0.066	C	0.3752	1.016
	A	0.15	A	0.140	0.041	A	0.3927	0.074
2	D, E	0.05	D, E	0.036	0.433	D, E	0.218	8.168
	D, F	0.05	D, F	0.041	0.213	D, F	0.234	6.795
	C, D	0.09	C, D	0.066	0.015	C, D	0.300	3.09
	C, F	0.10				C, F	0.321	2.346
	C, E	0.11				C, E	0.325	2.240
3	C, D, E	0.01	D, E, F	0.007	0.818	C, D, E	0.088	28.29
	D, E, F	0.01	C, D, E	0.010	0.095	D, E, F	0.091	26.4
	C, D, F	0.02				C, D, F	0.138	11.06
	C, E, F	0.02				C, E, F	0.150	9.235
	A, D, E	0.06				A, D, E	0.253	2.900
4	C, D, E, F	0.0004	C, D, E, F	0.001	0.732	C, D, E, F	0.021	344.5
	B, C, D, E	0.002	A, D, E, F	0.002	0.097	A, D, E, F	0.042	86.04
	A, D, E, F	0.002	B, C, D, E	0.002	0.038	B, C, D, E	0.042	86.04
	B, C, E, F	0.004	A, C, D, E	0.003	0.015	B, C, E, F	0.064	38.18
	A, C, D, E	0.009	B, C, E, F	0.003	0.005	A, C, D, F	0.092	18.25

The 16-run six factor design was also used in investigating the performance of the methods in active factors identification. The transformed data is as presented in Table 4.17. Table 4.18 presents the results from the analysis. From the analysis, the methods in identifying one, two, and three active factors mostly agreed in their choice especially for the top ranked model. For three active factors, B, D, E was ranked the most important by the 16-run design compared to D, E, F by the 12-run design. Model discrimination by the methods is not obvious. For four active factor subspace, the methods identified B, D, E, F as the most active subspace compared to C, D, E, F by the 12-run design.

Table 4.17: Six-factor 16-run design and response for the metal cutting experiment

Run	A	B	C	D	E	F	Response
7	-1	-1	-1	-1	1	1	0.94
40	1	-1	-1	-1	-1	1	0.08
53	-1	1	-1	-1	1	-1	0.96
36	1	1	-1	-1	-1	-1	0.75
60	-1	-1	1	-1	-1	1	0.08
38	1	-1	1	-1	-1	-1	0.35
27	-1	1	1	-1	1	1	0.96
37	1	1	1	-1	1	-1	0.96
23	-1	-1	-1	1	1	-1	1.02
47	1	-1	-1	1	1	1	1.03
48	-1	1	-1	1	-1	-1	0.90
54	1	1	-1	1	-1	1	0.98
63	-1	-1	1	1	-1	-1	0.95
29	1	-1	1	1	1	-1	0.93
22	-1	1	1	1	-1	1	1.15
28	1	1	1	1	1	1	1.06

Table 4.18: A comparison between factor based methods for identifying one, two, three and four active factor(s) in the six-factor 16-run design with responses from the metal cutting experiment

AF	Projective based search		box-Meyer search			Partial F search		
	Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	σ	F
1	D	0.08	D	0.078	0.414	D	0.2885	6.489
	E	0.09	E	0.085	0.211	E	0.3024	4.704
	B	0.10	B	0.091	0.130	B	0.3121	3.514
	A	0.11	A	0.111	0.030	A	0.3447	0.354
	F	0.12	F	0.113	0.027	F	0.3472	0.151
2	D, E	0.03	D, E	0.024	0.935	D, E	0.1698	15.73
	B, D	0.05	B, D	0.039	0.026	B, D	0.2176	8.006
	B, E	0.05	B, E	0.041	0.017	B, E	0.2243	7.298
	A, D	0.09				D, F	0.2927	2.635
	D, F	0.09				E, F	0.3189	1.590
3	B, D, E	0.01	B, D, E	0.011	0.450	B, D, E	0.1044	10.04
	D, E, F	0.01	D, E, F	0.012	0.089	D, E, F	0.1120	5.197
	B, E, F	0.03	A, D, E	0.021	0.003	B, E, F	0.1654	7.767
	A, D, E	0.03				A, D, E	0.1843	6.03
	C, D, E	0.04				C, D, E	0.1913	5.515
4	B, D, E, F	0.009	B, D, E, F	0.004	0.221	B, D, E, F	0.095	18.47
	B, C, E, D	0.02	B, C, D, E	0.006	0.005	B, C, E, D	0.128	9.933
	A, D, E, F	0.02	A, D, E, F	0.007	0.002	A, D, E, F	0.129	9.813

4.3 Follow-up runs

The gains in models' discriminatory powers of the methods through follow-up runs was explored. The full foldover and the semi-foldover designs for both 12- and 16-run designs for the metal cutting experiment were considered. The semi-foldover design is constructed by subsetting on the most active factor identified in the initial analysis using the original design. From the initial analysis of the metal cutting experiment for the 12- and 16-run designs in Tables 4.16 and 4.18 respectively, we conducted the follow-up schemes for both designs. We performed subsetting on the high level of factor D based on the initial analysis for the 12- and 16-run designs in the semi-foldover scheme. Tables 4.19 and 4.20 present the additional runs for the 12- and 16-run respectively.

Table 4.19: Additional runs for the follow-up schemes for the six-factor 12-run PB design with response for the metal cutting experiment

Full Foldover								Semi foldover							
Run	A	B	C	D	E	F	Response	Run	A	B	C	D	E	F	Response
37	1	1	1	-1	1	-1	0.96	9	-1	-1	1	1	-1	1	1.03
57	1	1	-1	1	1	-1	1.03	25	-1	-1	1	1	1	-1	1.02
52	1	1	-1	1	-1	1	0.98	55	-1	1	-1	1	1	-1	0.96
26	1	-1	1	1	-1	1	1.29	47	1	-1	-1	1	1	1	1.03
38	1	-1	1	-1	-1	-1	0.35	42	1	1	-1	1	-1	-1	1.19
59	1	-1	-1	-1	1	1	0.76	10	1	1	1	1	-1	1	1.09
5	-1	1	1	1	-1	-1	1.1								
27	-1	1	1	-1	1	1	0.96								
24	-1	1	-1	-1	-1	1	0.2								
3	-1	-1	1	1	1	1	1.06								
23	-1	-1	-1	1	1	-1	1.02								
14	-1	-1	-1	-1	-1	-1	0.64								

Table 4.20: Additional runs for the follow-up schemes for the six-factor 16-run design with response for the metal cutting experiment

Full Foldover								Semi foldover							
Run	A	B	C	D	E	F	Response	Run	A	B	C	D	E	F	Response
50	1	1	1	1	-1	-1	1.21	10	1	1	1	1	-1	1	1.09
58	-1	1	1	1	1	-1	1.15	5	-1	1	1	1	-1	-1	1.10
26	1	-1	1	1	-1	1	1.29	15	1	-1	1	1	1	1	0.98
3	-1	-1	1	1	1	1	1.06	25	-1	-1	1	1	1	-1	1.02
57	1	1	-1	1	1	-1	1.03	35	1	1	-1	1	1	1	0.90
19	-1	1	-1	1	1	1	0.82	13	-1	1	-1	1	-1	1	1.05
54	1	-1	-1	1	-1	-1	1.16	20	1	-1	-1	1	1	-1	0.99
11	-1	-1	-1	1	-1	1	1.10	12	-1	-1	-1	1	-1	-1	1.15
34	1	1	1	-1	-1	1	0.05								
64	-1	1	1	-1	-1	-1	0.25								
6	1	-1	1	-1	1	1	0.85								
17	-1	-1	1	-1	1	-1	1.02								
32	1	1	-1	-1	1	1	0.84								
24	-1	1	-1	-1	-1	1	0.20								
44	1	-1	-1	-1	1	-1	1.07								
14	-1	-1	-1	-1	-1	-1	0.64								

Tables 4.21 and 4.22 respectively display the results for the analysis from the follow-up experiment for the 12- and 16-run designs respectively. The follow-up schemes could not appreciably improve the discriminatory powers for the factor based methods. Similar to the initial analysis of the 12-run PB design, all the methods revealed for both the full foldover and the semi-foldover that C, D, E, F is the most important active subspace. Edwards et al. (2013) analysis of the full foldover indicated factor activity for $D - F$ and weakly for C. Also, they reported for the semi-foldover (subset on the high level of D) that a factor activity is shown in $D - F$.

The analysis of the follow-up plans for the 16-run design however, identified factor activity for B, D, E and F. Although, Mønness et al. (2007) and Edwards et al. (2013) both found main effects of factors B and C as insignificant in the full factorial results. The 12-run weakly identified a factor activity for factor C. On the contrary, the 16-run identified a weakly factor activity for factor B. Mønness et al. (2007) investigated a 2^{6-2}_{IV} design with generators $E = BCD$ and $F = ACD$, found D, E, F as significant.

Table 4.21: A comparison between factor based methods for identifying one, two, and three and four active factor(s) in the follow-up schemes for the six-factor 12-run design for the metal cutting experiment

Plan	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	σ	F
Full Foldover	1	D	0.07	D	0.069	0.963	D	0.268	15.45
		E	0.11	E	0.105	0.007	E	0.3317	2.444
		F	0.12	F	0.110	0.004	F	0.3397	1.304
		C	0.12	C	0.117	0.002	C	0.3495	0.0165
		B	0.12				B	0.3495	0.014
	2	D, E	0.03	D, E	0.026	0.998	D, E	0.1717	23.73
		D, F	0.06	D, F	0.056	0.001	D, F	0.2538	7.25
		C, D	0.08				C, D	0.2765	5.058
		B, D	0.08				B, D	0.2788	4.861
		A, D	0.08				A, D	0.2802	4.752
	3	D, E, F	0.01	D, E, F	0.012	0.843	D, E, F	0.1180	25.3
		C, D, F	0.02	C, D, E	0.02	0.003	C, D, E	0.1443	16.16
		A, D, E	0.03				A, D, E	0.1838	9.088
		B, D, E	0.03				B, D, E	0.1846	8.989
		B, D, F	0.07				B, D, F	0.2657	3.155
	4	C, D, E, F	0.01	C, D, E, F	0.001	0.682	C, D, E, F	0.0913	30.62
		A, D, E, F	0.02	A, D, E, F	0.002	0.007	A, D, E, F	0.1263	15.26
		B, D, E, F	0.08	B, C, D, E	0.002	0.0018	B, D, E, F	0.2828	6.543
Semi-foldover	1	D	0.05	D	0.062	0.876	D	0.2251	11.76
		F	0.10	F	0.092	0.027	F	0.3131	2.428
		C	0.11	C	0.10	0.013	C	0.3265	0.950
		A	0.11	E	0.103	0.011	E	0.3308	0.518
		B	0.11	B	0.106	0.008	B	0.336	0.011
	2	D, E	0.03	D, E	0.025	0.756	D, E	0.170	16.18
		D, F	0.03	D, F	0.029	0.213	D, F	0.1838	13.17
		A, D	0.05	C, D	0.045	0.005	C, D	0.2314	6.583
		B, D	0.07	A, D	0.058	0.001	A, D	0.2672	3.769
		B, D	0.07				B, D	0.2725	3.445
	3	D, E, F	0.005	D, E, F	0.006	0.993	D, E, F	0.0691	52.67
		C, D, E	0.006	C, D, E	0.011	0.004	C, D, E	0.0784	40.52
		C, D, F	0.01				C, D, F	0.1094	20.14
		A, D, E	0.03				A, D, E	0.1712	7.381
		B, D, E	0.03				B, D, E	0.1812	6.431
	4	C, D, E, F	0.003	C, D, E, F	0.001	0.612	C, D, E, F	0.075	31.03
		B, C, D, E	0.008	A, D, E, F	0.002	0.057	A, D, E, F	0.083	25.30
		A, D, E, F	0.018	B, C, D, E	0.002	0.018	B, D, E, F	0.094	19.81

Table 4.22: A comparison between factor based methods for identifying one, two, and three and four active factor(s) in the follow-up schemes for the six-factor 16-run design for the metal cutting experiment

Plan	AF	Projective based search		Box-Meyer search			Partial F search		
		Factor	$\widehat{\sigma}^2$	Factor	$\widehat{\sigma}^2$	Pos. Prob.	Factor	σ	F
Full Foldover	1	D	0.08	D	0.079	0.978	D	0.2857	17.91
		E	0.11	E	0.105	0.012	E	0.3297	5.978
		F	0.13	F	0.123	0.001	F	0.3560	0.853
		A	0.13				A	0.3608	0.046
		B	0.13				B	0.3609	0.022
	2	D, E	0.03	D, E	0.024	1.0	D, E	0.1618	40.47
		D, F	0.08				D, F	0.2801	7.277
		C, D	0.08				C, D	0.2900	6.166
		A, D	0.09				A, D	0.2946	5.689
		B, D	0.09				B, D	0.2955	5.596
	3	D, E, F	0.02	D, E, F	0.013	0.367	D, E, F	0.1265	31.46
		C, D, E	0.02	C, D, E	0.019	0.001	C, D, E	0.1561	19.50
		A, D, E	0.03				A, D, E	0.1699	15.92
		B, D, E	0.03				B, D, E	0.1732	15.19
		C, D, F	0.09				C, D, F	0.2927	3.094
	4	C, D, E, F	0.02	C, D, E, F	0.016	0.001	C, D, E, F	0.1528	16.14
		A, D, E, F	0.03				A, D, E, F	0.175	11.81
		B, D, E, F	1.39				B, D, E, F	0.178	11.39
Semi-foldover	1	D	0.06	D	0.053	0.943	D	O.2349	14.23
		B	0.08	B	0.077	0.012	B	0.2837	2.834
		F	0.08	E	0.079	0.009	E	0.2870	2.274
		E	0.09	A	0.084	0.004	A	0.2971	0.658
		C	0.09	F	0.086	0.004	F	0.2995	0.283
	2	D, E	0.02	D, E	0.017	0.995	D, E	0.1383	28.17
		B, D	0.03	B, D	0.027	0.005	B, D	0.1758	14.89
		D, F	0.05				D, F	0.2342	5.487
		A, D	0.06				A, D	0.2372	6.679
		B, E	0.06				B, E	0.2384	5.060
	3	B, D, E	0.008	B, D, E	0.006	0.902	B, D, E	0.090	32.95
		D, E, F	0.009	D, E, F	0.008	0.048	D, E, F	0.099	26.63
		A, D, E	0.02				A, D, E	0.1384	12.63
		C, D, E	0.02				C, D, E	0.1450	11.29
		B, E, F	0.03				B, E, F	0.1670	7.96
	4	B, D, E, F	0.008	B, D, E, F	0.005	0.111	B, D, E, F	0.089	24.17
		A, D, E, F	0.01	A, D, E, F	0.007	0.001	A, D, E, F	0.108	16.00
		B, C, D, E	0.01				B, C, D, E	0.114	13.97

Chapter 5

Discussion and Conclusion

5.1 Discussion

This thesis aimed at investigating three factor based methods of analysing some non-regular factorial fractional designs and follow-up runs in identifying active factors in experiments. The 12-run PB and 16-run (by Jones and Montgomery (2010)) designs for six factors were the non-regular designs employed for this study. Factor based methods of analysing designs are discussed and their analysis compared.

The projection based search method by Tyssedal and Samset (1997) was investigated. According to Tyssedal and Samset (1997), taking advantage of the projection properties of a design may increase the efficiency of the all subset selections procedure and even further improvement gained if the design is cyclically constructed. This method has often been reported in literature (Tyssedal et al. (2006), Tyssedal (2008b), Wiik (2014), Tyssedal and

Hussain (2016)). The results from this study suggest the projection based search as a useful tool in the analysis of fractional factorial non-regular design. In our simulation studies, we chose $\sigma = \{0.6, 0.8, 1.0\}$ and the method in all cases studied correctly identified the active subspace for models with at most three active factors. For both the six-factor 12-run and 16-run designs, it was observed that the projection based method was able to identify the correct factors for both low and the high level of variability in model 4.1 and the panel of models. In the follow-up design, the projection based search improved significantly in discriminating among the two topmost model. The problem of ambiguity in model selection is solved when follow-up techniques are employed. The full foldover and semi-foldover for both the 12-run and 16-run designs performed well in distinguishing the most active subspace from the others for the metal cutting experiment. From our search, we observed that the projection based method is less affected by increasing the error variance. The findings in this thesis agrees with Wiik (2014) that a major setback of this method is that it may provide no sign of error whenever three active factors are not enough. As a results, whenever the method indicates that three active factors are insufficient, the results cannot be trusted without support from other methods. The fact that the projection based method does not assume any particular form of the model except for sparsity, it may point to models that were not identified by other methods, in particular, the Box-Meyer method.

The Box-Meyer search by Box and Meyer (1993) has problem of overfitting as the number of active factors become large due to the high p, however, it has the ability to appreciably reduce the models to be investigated under satisfactorily conditions. In our search allowing for the number of interactions in the simulated models and $p = 0.6$ for both the 12-

and 16-run designs, the method performed appreciably well in identifying the correct active space for models with three active factors, for instance, in the case of model 4.1 and models 1 to 4 in the panel of models even for data with high variability. However, very low posterior probabilities were assigned to the correct subspace for model 4.1 and model 4 in the panel of models. This could be attributed to the inclusion of factors with weak factor activity. For models with four active factors, the method in most case failed to rank the models' active subspaces as the most important, but in some cases where the correct active subspaces were identified, a high posterior probability was assigned. The method in the case of models with four active factors tends to favour parsimonious models more than the correct model. The additional parameter in the posterior probability were in some cases helpful in explaining the ambiguities in top ranked models.

In the metal cutting experiment, although there seems to be ambiguity in model discrimination using the variance estimate for both designs, however, the posterior probabilities assigned to the topmost model were significantly high compared to the top five models for one, two and three active factors. However, in the 16-run design the posterior probabilities for the topmost model for two and three active factors were distinct from the others and were appreciably large. But for four active factor the posterior probability for the top ranked model were very low.

The partial F search by Kulachi and Box (2003) uses a simple F criterion in defining active subspace. The method produced results very similar to that of the projection based search. From the analysis of the 12- and 16-run designs, the method in all cases studied correctly identified the active subspace. The method was able to identify the correct factors for both low and the high level of variability in model 4.1 and the panel of models. In the metal cutting experiment, the method performed very well agreeing with the results from the other methods. Although, models were not clearly discriminated in the initial design, the method identified the active subspace reported in literature as the most important. The method significantly improved in discriminating among models in the follow-up schemes.

5.2 Conclusion

The methods studied in this thesis were found to produce similar results in identifying one, two and three active factors in the 12- and 16-run experiments. In this investigation, the 16-run design, with runs advantage did not display any significant performance in defining factor activities for experiments over the 12-run PB design for models with three active factors. However, for models with four active factors, the 16-run design having five degrees of freedom for the error, relatively performed better compared to the 12-run with one degree of freedom for the error. All the methods performed very well in identifying models with at most three active factors. However, for models with four active factors, the study revealed that the methods have shortcomings in identifying the correct active subspace. It was found that the projection based search and the partial F search produced very similar results. There was no clear sign of which of the two methods to generally prefer for identifying three active subspace. The Box-Meyer search estimated variances were lower compared to that of the other two methods. The dilemma in using the Box-Meyer search will always be the choice of the p and g parameters. Based on the study conducted in this thesis, it is concluded that the factor based methods performed reasonably well in defining factor activities in experiments with at most three active factors. The projection based method is very simple to use, with much less intuition and was robust under various conditions of model's variability. However, it is not appropriate to use the method alone whenever the results indicate that three active factors are insufficient. This is due to the fact that it may provide wrong results with no sign of error in cases where more than three factors are active. However, follow-up experiments help to improve performance of the method. This study recommend the use of the factor based methods in defining factor activities for experiments.

Bibliography

Box, G., Meyer, R. D., 1993. Finding the active factors in fractionated screening experiments. Journal of Quality Technology 25, 94–105.

Box, G. E. P., Hunter, W. G., Hunter, J. S., 1978. Statistics for Experimenters. John Wiley and Sons; New York.

Box, G. E. P., Hunter, W. G., Hunter, J. S., 2005. Statistics for Experimenters: Design, Innovation, and Discovery, 2nd Edition. Wiley: New York.

Box, G. E. P., Meyer, R. D., 1986a. An analysis for unreplicated fractional factorials. Technometrics 28, 11–18.

Box, G. E. P., Tyssedal, J., 1996. Projective properties of certain orthogonal arrays. Biometrika 83 (4), 950–955.

Box, G. E. P., Wilson, W. G., 1951. On the experimental attainment of optimal conditions. Journal of Royal Statistical Society Series B 13, 1–38, discussion 38–45.

Candes, E., Tao, T., 2007. The dantzig selector: statistical estimation when p is much larger than n. Annals of Statistics 35, 2313–2351.

Cheng, C. S., 1980. Orthogonal arrays with variable numbers of symbols. Annals of Statistics 8, 447–453.

Chipman, H., Hamada, M., Wu, C. F. J., 1997. A bayesian variable-selection approach for analyzing designed experiments with complex aliasing. Technometrics 39, 372–381.

Daniel, C., 1976. Applications of Statistics to Industrial Experimentation. John Wiley and Sons; New York.

Deng, L. Y., Tang, B., 1999. Generalized resolution and minimum aberration criteria for plackett-burman and other nonregular factorial designs. Statistica Sinica 9, 1071–1082.

Edwards, D. J., Wesse, M. L., Palmer, G. A., 2013. Comparing methods for design follow-up: revisiting a metal-cutting case study. Appl. Stochastic Models Bus. Ind. 30, 464–478.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. Annals of Statistics 32, 407–499.

Fisher, R. A., 1935. The Design of Experiments. Ohio and Boyd, Edinburgh, Scotland.

Garzon, I. E., 2000. Optimisation for product and process improvement: investigation of Taguchi tools and genetic algorithm. Ph.D. thesis, University of Newcastle Upon Tyne, UK.

Hamada, M., Wu, C. F. J., 1992. Analysis of designed experiments with complex aliasing. Journal of Quality Technology 24, 130–137.

John, P., 2000. Breaking alias chains in fractional factorials. Communications in Statistics-Theory and Methods 29, 2143–2155.

Johnson, M. E., Jones, B., 2011. Classical design structure of orthogonal designs with six to eight factors and sixteen runs. Qual. Reliab. Engng. Int. 27, 61–70.

Jones, B., Montgomery, D. C., 2010. Alternatives to resolution iv screening designs in 16 runs. International Journal of Experimental Design and Process Optimisation 1 (4), 285–295.

Kulachi, M., Box, G. E. P., 2003. Catalysis of discovery and development in engineering and industry. Quality Engineering 15 (3), 513–517.

Lenth, R. V., 1989. Quick and easy analysis of unreplicated factorials. Technometrics 31, 469–473.

Li, W., Lin, D. J. K., Ye, K. Q., 2003. Optimal foldover plans for two-level non-regular orthogonal designs. Technometrics 45 (4), 347–351.

Lin, D. K., Draper, N. R., 1992. Projection properties of plackett and burman designs. Technometrics 34, 423–428.

Mee, R., Peralta, M., 2000. Semifolding 2^{k-p} designs. Technometrics 42 (2), 122–134.

Meyer, R. D., Steinberg, D. M., Box, G., 1996. Follow-up designs to resolve confounding in multifactor experiments. Technometrics 38 (4), 303–313.

Miller, A., Sitter, R. R., 2005. Using folded-over nonorthogonal designs. Technometrics 47 (4), 502–513.

Montgomery, D. C., 2001. Design and Analysis of Experiments, 5th Edition. John Wiley and Sons, Inc.

Mønness, E., Linsley, M. J., Garzon, I. E., 2007. Comparing different fractions of a factorial design: A metal cutting case study. Appl. Stochastic Models Bus. Ind. 23, 117–128.

Phoa, F., Pan, Y., Xu, H., 2009. Analysis of supersaturated designs via the dantzig selector. Journal of Statistical Planning and Inference 139, 2362–2372.

Plackett, R. L., Burman, J. P., 1946. The design of optimum multifactorial experiments. Biometrika 33, 305–325.

Telford, J. K., 2007. A brief introduction to design of experiments. John Hopkins Apl Technical Digest 27 (3), 224–232.

Tyssedal, J., 2008b. "Projectivity in Experimental Design" in Encyclopedia of Statistics in Quality and Reliability, eds. F. Ruggeri, R.S. Kennet and F. W. Faitin. New York: Wiley.

Tyssedal, J., 2011. Two-level factorial designs and how to block them.

Tyssedal, J., Grinde, H., Røstad, C. C., 2006. The use of a 12-run plackett-burman design in the injection moulding of a technical plastic component. Qual. Reliab. Engng. Int. 22, 651–657.

Tyssedal, J., Hussain, S., 2016. Factor screening in non-regular two-level designs based on projection-based variable selection. Journal of Applied Statistics 43 (3), 490–508.

Tyssedal, J., Kulachi, M., 2005. Analysis of split-plot designs with mirror image pairs as sub-plots. Quality and Reliability Engineering 15 (3), 509–513.

Tyssedal, J., Niemi, R., 2014. Graphical aids for the analysis of two-level nonregular designs. Journal of Computational and Graphical Statistics 23 (3), 678–699.

Tyssedal, J., Samset, O., 1997. Analysis of the 12-run plackett and burman design. Preprint Statistics (8/1997.

Wang, J. C., Wu, C. F. J., 1995. A hidden projection property of plackett-burman and related designs. Statistica Sinica 5, 235–250.

Wiik, H. E., 2014. Methods of Analyzing the 12-run Plackett-Burman Design. Ph.D. thesis, Norwegian University of Science and Technology (NTNU), Dept. of Mathematical Sciences.

Wu, C. F. J., Hamada, M., 2000. Experiments: Planning, Analysis, and Parameter Design Optimization. Wiley, New York.

Yuan, M., Joseph, V., Lin, Y., 2007. An efficient variable selection approach for analyzing designed experiments. Technometrics 49, 430–439.

Appendix

Data generation

Six factors in 12-runs

A = (-1, -1, -1, -1, -1, -1, 1, 1, 1, 1, 1, 1)

B = (-1, -1, -1, 1, 1, 1, -1, -1, -1, 1, 1, 1)

C = (-1, 1, 1, -1, -1, 1, -1, -1, 1, -1, 1, 1)

D = (1, -1, -1, -1, 1, 1, -1, 1, 1, -1, -1, 1)

Responses from Model 4.1

$y_{12(0.6)}$ = (-6.69, -5.09, -3.98, 6.83, 7.62, 4.88, -0.56, -0.93, -2.33, 8.25, 9.68, 10.94)

$y_{12(0.8)}$ = (-5.96, -5.41, -4.64, 7.29, 6.44, 6.12, -0.87, -0.33, -2.98, 8.35, 11.11, 10.12)

$y_{12(1.0)}$ = (-6.75, -4.42, -5.54, 6.22, 5.94, 6.94, -0.82, -0.96, -2.71, 8.66, 10.66, 8.93)

Simulation codes for responses from the panel of models.

```
for (i in 1:12){ y[i] <- A[i]+2*A[i]*B[i]+2*A[i]*C[i]+ rnorm(1,0,sd)}

for (i in 1:12){ y[i] <- A[i]+1.5*B[i]+2*C[i]+A[i]*B[i]+1.5*A[i]*C[i]+rnorm(1,0,sd)}

for (i in 1:12){ y[i] <- A[i]+1.5*B[i]+2*C[i]+1.5*A[i]*B[i]*C[i]+ rnorm(1,0,sd)}

for (i in 1:12){ y[i] <- 2*A[i]+B[i]*C[i]+ rnorm(1,0,sd)}

for (i in 1:12){ y[i] <- A[i]+C[i]+B[i]*C[i]+C[i]*D[i]+ rnorm(1,0,sd)}
```

```
for (i in 1:12){ y[i] <- 2*A[i]+3*B[i]+2*C[i]+D[i]+3*C[i]*D[i]+ rnorm(1,0,sd)}

for (i in 1:12){ y[i] <- 4*A[i]+B[i]+C[i]+D[i]+2*A[i]*D[i]+ rnorm(1,0,sd)}

for (i in 1:12){ y[i] <- 2*A[i]+4*C[i]+2*B[i]*C[i]+2*C[i]*D[i]+ rnorm(1,0,sd)}
```

Responses from the panel of models

$y_{12(1,0.6)}$ = (2.72, 0.39, -0.92, -1.21, -1.60, -4.66, -1.90, -4.75, 0.13, 0.45, 5.56, 5.14)

$y_{12(2,0.6)}$ = (-1.86, -1.38, -1.33, -1.46, -1.29, 0.84, -5.07, -5.15, 2.02, -0.76, 6.23, 8.42)

$y_{12(3,0.6)}$ = (-6.92, 1.09, 1.76, -0.30, 0.96, 0.80, -0.45, -1.19, -0.37, -1.37, 6.94, 6.11)

$y_{12(4,0.6)}$ = (-0.97, -2.47, -3.62, -3.04, -2.92, -1.07, 2.88, 2.79, 0.56, 0.86, 2.82 2.41)

$y_{12(5,0.6)}$ = (-1.98, -3.16, -1.71, -1.45, -3.48, 1.69, 1.43, -0.22, 1.81, 0.25, 1.59, 3.27)

$y_{12(6,0.6)}$ = (-9.60, -5.29, -6.28, 1.98, -2.70, 6.88, -0.95, -4.72, 3.66, 4.80, 2.59, 10.84)

$y_{12(7,0.6)}$ = (-6.62, -2.40, -2.50, -2.80, -4.68, -3.01, -0.83, 4.42, 7.42, 0.21, 2.81, 8.91)

$y_{12(8,0.6)}$ = (-6.77, -2.51, -2.29, -5.86, -10.32, 4.95, 2.00, -1.24, 5.26, -2.43, 6.39, 9.54)

$y_{12(1,0.8)}$ = (3.25, -1.37, -2.15, 0.22, -1.26, -5.01, -2.02, -3.29, 1.02, 1.40, 6.23, 4.06)

$y_{12(2,0.8)}$ = (-1.51, -1.46, -1.57, -0.98, -2.45, -0.20, -5.09, -4.54, 1.70, -0.66, 6.71, 6.92)

$y_{12(3,0.8)}$ = (-4.99, 0.93, -1.08, -1.54, -0.55, 2.16, -0.92, -0.75, -0.29, -1.88, 6.47, 7.25)

$y_{12(4,0.8)}$ = (-1.11, -2.99, -3.01, -2.48, -3.51, -1.96, 3.06, 3.15, 1.44, 1.61, 4.14, 1.38)

$y_{12(5,0.8)}$ = (-1.85, -1.60, -2.32, -2.61, -4.94, 1.57, 2.48, -0.95, 1.41, 0.08, 0.73, 5.22)

$y_{12(6,0.8)}$ = (-9.70, -6.85, -7.58, 1.20, -4.17, 8.57, -1.36, -5.51, 4.78, 4.34, 4.03, 10.44)

$y_{12(7,0.8)}$ = (-6.93, -3.45, -2.97, -3.18, -5.76, -2.62, -2.79, 5.15, 7.31, 0.52, 2.77, 8.57)

$y_{12(8,0.8)}$ = (-5.81, -1.05, -3.37, -5.35, -8.79, 6.91, 1.60, -2.02, 5.66, -2.51, 5.03, 9.74)

$y_{12(1,1.0)}$ = (3.81, -0.82, -0.49, -3.80, -2.77, -6.71, -2.89, -4.08, 0.16, -0.96, 5.59, 5.25)

$y_{12(2,1.0)}$ = (-1.88, -1.67, -0.13, -1.54, -2.30, 0.10, -6.07, -4.94, 3.98, -0.19, 5.01, 8.31)

$y_{12(3,1.0)}$ = (-7.41, 0.27, 3.17, -2.06, -1.06, 0.81, -1.47, 0.62, 0.29, -1.27, 6.63, 8.59)

$y_{12(4,1.0)}$ = (-2.25, -4.00, -2.65, -2.23, -2.49, -1.19, 1.45, 3.15, 0.94, -0.87, 2.87, 4.03)

$y_{12(5,1.0)}$ = (0.41, -0.51, -1.92, -1.54, -4.87, 2.14, 1.86, -0.30, 0.51, 1.23, 1.56, 4.49)

$y_{12(6,1.0)}$ = (-10.90, -7.11, -6.54, -0.22, -2.67, 7.17, -1.63, -7.70, 5.76, 5.49, 3.60, 10.35)

$y_{12(7,1.0)}$ = (-10.77, -3.53, -1.55, -4.96, -5.27, -2.26, 0.74, 6.56, 7.63, 1.00, 2.69, 9.60)

$y_{12(8,1.0)}$ = (-10.77, -3.53, -1.55, -4.96, -5.27, -2.26, 0.74, 6.56, 7.63, 1.00, 2.69, 9.60)

Six factors in 16-runs

A = c(-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1)

B = c(-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1)

C = c(-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1)

D = c(-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1)

Responses from Model 4.1

$y_{16(0.6)}$ = (-6.63, -0.58, 6.16, 8.30, -5.12, -2.66, 4.46, 11.12, -5.55, -1.04, 6.41, 9.39, -3.91, -2.55, 3.94, 11.78)

$y_{16(0.8)}$ = (-6.90, -1.63, 6.25, 9.50, -3.46, -1.92, 4.86, 11.74, -7.98, -2.03, 6.87, 7.52, -3.34, -2.59, 5.30, 9.64)

$y_{16(1.0)}$ = (-4.96, -2.62, 7.84, 8.26, -5.52, -2.50, 3.85, 11.49, -6.58, -1.32, 6.76, 9.45, -2.56, -1.35, 4.24, 10.24)

Responses from the panel of models

$y_{16(1,0.6)}$ = (2.32, -3.70, -1.27, 1.49, -1.29, 1.69, -5.67, 4.88, 3.17, -2.81, -1.09, 1.02, -2.12, 1.62, -4.60, 4.63)

$y_{16(2,0.6)}$ = (-1.81, -5.27, -1.66, 0.53, -1.55, 2.08, 0.38, 7.50, -2.47, -6.33, -1.67, -0.25, -1.72, 1.22, -0.15, 7.04)

$y_{16(3,0.6)}$ = (-6.10, 0.002, -0.39, -1.77, 1.06, 0.56, 1.22, 5.93, -6.33, 0.32, 0.33, -0.79, 1.15, 0.09, 1.13, 6.00)

$y_{16(4,0.6)}$ = (-0.06, 3.31, -3.06, 0.60, -4.33, 1.05, -0.78, 3.51, -0.13, 3.70, -2.36, 1.60, -3.57, 0.82, -1.59, 2.61)

$y_{16(5,0.6)}$ = (-0.09, 1.70, -1.67, -0.14, -1.45, 0.38, 0.12, 1.59, -1.55, -0.60, -4.02, -2.45, -0.58, 2.301, 2.28, 2.61)

$y_{16(6,0.6)}$ = (-5.59, -0.25, 0.27, 4.48, -7.55, -2.97, -1.61, 2.64, -8.32, -5.33, -2.83, 1.71, 0.29, 4.89, 6.93, 11.35)

$y_{16(7,0.6)}$ = (-5.03, -0.83, -2.46, 1.50, -3.58, 0.63, -2.20, 2.31, -6.25, 5.39, -5.47, 5.67, -5.48, 6.26, -3.04, 9.22)

$y_{16(8,0.6)}$ = (-2.02, 1.80, -6.49, -2.09, -2.20, 2.50, 1.31, 6.10, -5.13, -2.37, -9.13, -5.53, 2.15, 6.25, 4.73, 10.58)

$y_{16(1,0.8)}$ = (3.05, -2.60, -1.40, 1.65, -0.68, 0.84, -5.89, 3.79, 4.08, -2.22, -1.78, 1.12, 0.64, -0.63, -5.16, 4.75)

$y_{16(2,0.8)}$ = (-3.31, -4.88, -2.56, -1.27, -0.16, 1.91, -0.16, 7.02, -2.49, -5.14, -0.65, 0.69, -1.81, 2.80, 0.88, 7.42)

$y_{16(3,0.8)}$ = (-6.85, -1.46, -1.18, -1.30, 2.26, 0.09, 0.91, 5.81, -6.50, -1.68, 0.06, -0.83, 1.66,

-0.12, 1.39, 7.13)

$y_{16(4,0.8)}$ = (-0.55, 2.39, -3.30, 0.51, -3.26, 0.04, -2.40, 4.86, -1.84, 2.67, -2.21, 1.21, -4.20, 1.68, -0.79, 3.94)

$y_{16(5,0.8)}$ = (0.08, 2.12, -1.62, -0.27, -1.36, 1.17, 0.86, 2.19, -1.08, 0.62, -5.09, -2.22, -0.15, 2.05, 2.68, 3.82)

$y_{16(6,0.8)}$ = -4.84, -1.27, 1.24, 4.51, -8.56, -3.48, -0.56, 4.09, -8.40, -4.26, -2.01, -0.66, 1.29, 5.10, 5.62, 10.62)

$y_{16(7,0.8)}$ = (-6.37, -0.92, -3.73, 1.07, -2.01, 0.45, -1.57, 3.58, -7.20, 5.93, -6.28, 7.51, -4.60, 5.87, -2.60, 7.39)

$y_{16(8,0.8)}$ = (-0.93, 3.01, -6.17, -2.81, -1.77, 0.98, 1.30, 6.00, -6.06, -2.31, -10.33, -6.70, 3.13, 6.17, 4.97, 10.10)

$y_{16(1,1.0)}$ = 2.73, -1.35, -0.50, 1.97, -1.92, 2.91, -4.97, 5.42, 2.00, -3.55, -1.49, 0.82, 0.13, 1.12, -6.59, 3.80)

$y_{16(2,1.0)}$ = (-2.43, -3.94, -0.48, -0.25, -0.32, 1.13, -0.44, 8.23, -1.77, -7.25, -2.44, 1.83, -2.05, 1.97, -0.66, 6.62)

$y_{16(3,1.0)}$ = -5.74, 0.34, -0.08, -1.37, 2.02, 0.12, 1.39, 6.09, -7.30, -1.19, -0.01, -0.76, 0.12, -1.03, 0.002, 6.73)

$y_{16(4,1.0)}$ = (0.05, 5.70, -2.71, 0.25, -3.19, 1.83, -1.75, 3.16, -0.49, 3.40, -2.41, -0.05, -2.61, -0.09, -1.39, 2.24)

$y_{16(5,1.0)}$ = (-0.66, 0.52, -1.99, -0.15, -4.20, 0.60, 0.17, 1.30, -4.32, -0.19, -2.67, -1.40, -1.52, 1.85, 3.49, 4.00)

$y_{16(6,1.0)}$ = (-3.47, -0.37, -1.24, 4.81, -6.39, -5.47, -1.31, 2.92, -6.53, -4.36, -2.83, -0.82, 1.90, 7.60, 7.03, 10.22)

$y_{16(7,1.0)}$ = (-5.37, -0.84, -2.84, 1.00, -1.46, 1.43, 1.34, 2.76, -8.06, 3.97, -4.48, 8.50, -5.24, 6.70, -1.57, 9.13)

$y_{16(8,1.0)}$ = (-3.52, 1.34, -5.96, -2.01, -1.65, 3.02, 0.56, 6.55, -6.51, -2.00, -10.82, -4.95, 1.54, 5.48, 6.43, 10.50)