Algorithmic Operations Research Vol.7 (2013) 71–82

Optimization in $2^m 3^n$ Factorial Experiments

G. S. R. Murthy [a]

[a]SQC & OR Unit, Indian Statistical Institute, Hyderabad 500007, India

D. K. Manna [b]

[b]SQC & OR Unit, Indian Statistical Institute, Kolkata 700108, India

Abstract

The need for adopting efficient designs in industrial experiments is well understood. Often situations arise where the existing designs such as orthogonal arrays are not suitable for designing required experiments. This paper deals with one such situation where there was a need for designing an asymmetrical factorial experiment involving interactions. Failing to get a satisfactory answer to this problem from the literature, the authors have developed an ad hoc method of constructing a design. It is transparent from the method of construction that the design provides efficient estimates for all the required main effects and interactions. The later part of this paper deals with the issues of how this method is extended to more general situations and how this ad hoc method is translated into a systematic approach. The method consists of formulating the construction problem as certain integer programming problems. It is believed that this method will be very useful in practical applications. The ideas are illustrated with a number of examples.

Key words: Efficient Design, Construction, Integer Programming.

1. Introduction

Fractional factorial experiments are widely used in industrial and other applications. In many instances, orthogonal arrays (OAs) provide a good basis for designing the experiments. Given the requirements in terms of main effects and interactions to be estimated, one has to first design the experiments, i.e., identify treatment combinations (TCs) which can lead to estimation of parameters of interest. Since there are several choices of designs, one is interested in using the most efficient design. Usually the efficiency of a design is measured in terms of run length, i.e., the number of TCs, and dispersion matrix of the estimators of parameters of interest.

A factorial experiment is said to be asymmetrical if there are at least two factors for which the number of levels considered for each of these factors is not the same. Designing asymmetrical fractional factorial experiments is relatively more difficult compared to the other category. The problem becomes more complex when the model involves interactions. Many researchers suggested methods for orthogonal plans (see Chakravarti [1956], Addleman [1962], Cheng [1989], Wang and Wu [1991,1992]). Orthogonal designs are efficient, but for a fixed run length such designs may not exist. In such cases one has to sacrifice orthogonality in favour of smaller run length. Anderson and Thomas [1979] proposed to derive resolution IV designs by collapsing the levels in foldover designs. Webb [1971] developed a number of catalogues for small incomplete experiments where each factor is tried at either two levels or three levels. Box and Draper [1971] studied the optimality of designs using $\mid X^t X \mid$ criterion (see Section 2). Mitchell [1974] proposed DETMAX algorithm for the construction of efficient designs with respect to $\mid X^t X \mid$ criterion. Wang and Wu [1992] introduced the concept of *near-orthogonality* and produced some construction methods for the main-effect plans. They also enlisted a number design layouts with varying run lengths for $2^m 3^n$ factorial experiments.

This article stems from the investigations of a problem we have encountered in an industry (see next section). It was required to conduct a $2^2 3^2$ factorial experiment in which certain interactions were to be estimated along with the main effects. Thus, an efficient design was to be constructed for the problem in question. Failing to get a satisfactory answer to this problem from the literature, we hit upon an ad hoc method for constructing an efficient design for this purpose. It is transpar-

Email: G. S. R. Murthy [murthygsr@gmail.com], D. K. Manna [dkmanna@hotmail.com].

ent from the method of construction, that certain main effects/interactions can be estimated orthogonally. We, then, developed a general methodology to convert our ad hoc method into a systematic one.

An interesting aspect of our methodology is that we formulate the problem as an Operations Research (OR) problem. It is shown that constructing efficient designs, under some restrictions, can be formulated as an integer programming problem (IPP) with either linear or quadratic objective function.

The present article is confined purely to 2^m3^n factorial experiments where it is required to estimate all the main-effects and some of the two-factor interactions.

In Section 2, we present the basic terminology and notation, and the specific problem we have encountered in the industry. In Section 3, our method of construction is described and is illustrated with some examples. In Section 4, we present the application of OR formulations in constructing the designs and illustrate the same with examples.

2. Background

In this Section we recall the relevant terminology and introduce the notation with the help of the following problem which we encountered in an aluminium alloy foundry.

2.1. *Problem*

In order to optimize the process parameters for reducing casting defects, it was planned to conduct an experiment with the following factors and levels : (A) Bath Temperature at 3 levels, (B) Phosphorous Content at 3 levels, (C) Charge Ratio at 2 levels, and (D) Filtering Method at 2 levels. Besides the main effects, interactions AB and AC were to be considered in the model. The problem was to construct an efficient design which allows the estimation of the 4 main effects and the above mentioned interactions.

The full-factorial experiment for the above problem involves 36 TCs. Any k of these 36 TCs will be referred to as a fraction of size k (k is called the run length). Say that a fraction of size k is regular if 36 is divisible by k. Say that a column of a fraction is *homogeneous* if all the levels in that column appear with same frequency; and say that a fraction is homogeneous if all the columns in the fraction are homogeneous. Throughout this paper, we confine our attention to regular homogeneous fractions only.

Let y_i denote the response due to ith TC. It is assumed that y_i's, $1 \leq i \leq k$, are uncorrelated, each with variance σ^2. If $A_3B_2C_1D_2$ is the ith TC, then under the model assumptions, the expected value of y_i is given by

$$\mathrm{E}(y_i) = \mu + a_3 + b_2 + c_1 + d_2 + (ab)_{32} + (ac)_{31}, \quad (1)$$

where μ is the general effect and $a, b, c, d, (ab)$ and (ac) denote respective factorial effects at their corresponding levels. Using reparametrization of these factorial effects, the above model can be remodeled as

$$\mathrm{E}(y) = X\beta \text{ and } D(y) = \sigma^2 I, \quad (2)$$

where $y = (y_1, y_2, \ldots, y_k)^t$ is the response vector and β is the vector whose elements correspond to μ, linear and quadratic components of the main effects and interactions of factorial effects considered in the above model; the ith row of the design matrix X, $X_{i.}$, corresponds to the i^{th} TC in the experiment and the jth column of X, $X_{.j}$, corresponds to β_j (the j^{th} element of β).

For the problem in question, β consists of 13 elements, namely, general effect (μ), linear/quadratic components of main effects (A_L, A_Q, B_L, B_Q, C, D), and the linear/quadratic components of interactions (A_LB_L, A_LB_Q, A_QB_L, A_QB_Q, A_LC, A_QC). Further, X is a $k \times 13$ matrix consisting of corresponding contrast vectors (CVs) along with a vector of 1's for μ .

In order that all the main effects and interactions (including μ) be estimable, a necessary condition is that X is of full column rank (see Rao [1974]). The least square estimator of β is given by $\hat{\beta} = (X^tX)^{-1}X^ty$ and the dispersion matrix of $\hat{\beta}$, denoted by D($\hat{\beta}$), is $\sigma^2(X^tX)^{-1}$. For the purpose of this paper, we shall assume, without loss of generality, that $\sigma^2 = 1$ throughout.

For the given requirements of estimating the factorial effects, there exist several choices of designs and one is interested in choosing an optimal design. A number of criteria (for optimality) have been developed (see Raghavarao [1971]) to compare and construct designs. We shall mention two such criteria here.

- Under the assumption that columns of X are normalized, $\mid X^tX \mid$ is a overall measure of efficiency (denoted as D-efficiency) of the design (see Rao [1974], Wang and Wu [1992]). Since the $\mid X^tX \mid$ is always less than or equal to product of the diagonal entries of X^tX, $\mid X^tX \mid$ attains its maximum when the columns of X are orthogonal. Furthermore, under orthogonality, the estimates of β_is attain the minimum variance. Therefore, the D-efficiency of the design is given by

$$D\text{-efficiency} = 100\times \mid X^tX \mid^{\frac{1}{p}} /k,$$

where p is the number of parameter to be estimated and k is the run length (see Mitchell [1974]).

- Another measure of efficiency is given by

$$I_F = p/\left[k\sum_{i=1}^{p} w_i \mathrm{V}(\hat{\beta}_i)\right], \qquad (3)$$

where k and p are as defined above ($p = 13$ in our problem), w_i's are some associated weights and $V(\hat{\beta}_i)$ is the variance of $\hat{\beta}_i$ which is the i^{th} diagonal entry of $(X^tX)^{-1}$ (see Webb [1971]).

For any 3-level factor F, we shall use the notation l_{F_L} and l_{F_Q} to denote CVs of the linear and quadratic components of the main effect of F. If G is another 3-level factor, then $F_LG_L,\ F_LG_Q,\ F_QG_L,\ F_QG_Q$ are the linear and quadratic components of FG interaction. Every component of a main effect or an interaction is represented by a contrast vector. For any two CVs u and v, we define the *nonorthogonality* between u and v as $\mid u^tv \mid$.

3. Construction Method

In this section, we will describe our method of constructing designs. It should be mentioned that this method of construction produces only regular fractions which are also homogeneous. In Subsections 3.1 and 3.2, we construct designs for our problem of Section 2.1. Subsection 3.3 provides an example to extend the methodology to more general situations.

It is clear that in order to estimate the main effects and interactions specified in our problem, we need at least 13 TCs. Any regular fraction of this 2^23^2 experiment must contain 2^i3^j TCs, $0 \le i,j \le 2$. Since the minimum number is 13, we go for the smallest regular fraction with a run length not less than 13.

3.1. *Construction in 18 runs*

To construct a design in 18 runs, we first look at 2^13^2 full-factorial. This layout has 3 columns (see columns (2), (3) and (4) of Table 1). Since it is a full-factorial layout, all the main effects and interactions (of A, B, C) are estimable and are orthogonal. We now augment these 3 columns with another 2-level column (column under D) to get a design layout for the required experiment. The main step of the method is that the new column is chosen so as to make the columns under D, B, A a 2^13^2 full-factorial experiment. The idea is to make the estimates of main effects and interactions involving factors $A,\ B$ and D uncorrelated.

Since columns under A, B and D form a full-factorial experiment, estimates of main effects and interactions of these factors will be uncorrelated. Above, we have already noted that the estimates of main effects and interactions of factors A, B and C are also uncorrelated.

The above method yields a design layout for the required experiment. However, it will be a *feasible design* provided it leads to estimability of all the required main effects and interactions. For instance, one of the candidates for the column under D is the column under C itself. But this choice results in complete confounding of the main effects of C and D.

There are several choices for the column under D. Following an ad hoc approach, this column is initially constructed, by trial and error method, so that the resulting design is a feasible one and that the correlation between the estimates of main effects of C and D is a minimum. A general and systematic approach to this problem is developed later. According to this approach, the column under D is constructed so that the nonorthogonality between the main effects of C and D (i.e., $\mid l_C^t l_D \mid$) is a minimum. While doing so, we also try to keep the nonorthogonalities of D with AC interaction as low as possible. Clearly we have an optimization problem at hand. We defer the details of this methodology to Section 4. The design thus constructed is given in Table 1.

Consider the design as constructed above in 18 runs. The nonsingularity of the resulting X^tX matrix indicates that the general effect μ and all the main effects and interactions of our interest are estimable. The dispersion matrix, $(X^tX)^{-1}$, is given in Table 2. The D-efficiency and the I_F-efficiency of this design are 115.70% and 98.11% respectively. It is evident from $(X^tX)^{-1}$ and I_F-efficiency, that our design is a reasonably good one for the given requirements.

In fact, the above design is D-optimal among the regular homogeneous ones with 18 runs. The proof is given in the Appendix. Furthermore, this design provides five degrees of freedom for estimating error.

It must be noted that we first tried to construct a suitable design using the existing literature on the subject.

Table 1

Construction of a Fraction with 18 Runs

	2^13^2 full-factorial			Additional column	After rearranging rows and columns				
TC. No.	C	A	B	D	TC. No.	D	B	A	C
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	1	1	1	1	1	1	1	1	1
2	1	1	2	2	4	1	1	2	1
3	1	1	3	1	7	1	1	3	1
4	1	2	1	1	11	1	2	1	2
5	1	2	2	2	14	1	2	2	2
6	1	2	3	1	17	1	2	3	2
7	1	3	1	1	3	1	3	1	1
8	1	3	2	2	6	1	3	2	1
9	1	3	3	2	18	1	3	3	2
10	2	1	1	2	10	2	1	1	2
11	2	1	2	1	13	2	1	2	2
12	2	1	3	2	16	2	1	3	2
13	2	2	1	2	2	2	2	1	1
14	2	2	2	1	5	2	2	2	1
15	2	2	3	2	8	2	2	3	1
16	2	3	1	2	12	2	3	1	2
17	2	3	2	1	15	2	3	2	2
18	2	3	3	1	9	2	3	3	1

Table 2

Dispersion Matrix of Proposed Design in 18 runs $\times 10^{-2}$

Effect	μ	C	D	A_LC	A_QC	A_L	A_Q	B_L	B_Q	A_LB_L	A_LB_Q	A_QB_L	A_QB_Q
μ	5.56												
C	0.00	5.63											
D	0.00	-0.69	6.25										
A_LC	0.00	0.23	-2.08	9.03									
A_QC	0.00	-0.08	0.69	-0.23	2.85								
A_L	0.00	0.00	0.00	0.00	0.00	8.33							
A_Q	0.00	0.00	0.00	0.00	0.00	0.00	2.78						
B_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.33					
B_Q	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.78				
A_LB_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	12.50			
A_LB_Q	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.17		
A_QB_L	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.17	
A_QB_Q	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.39

Consequently, following Anderson and Thomas [1979], we constructed a resolution IV design (so as to estimate the interactions). This required 19 treatment combinations. The resulting layout and the corresponding dispersion matrix are given in Table $3(a)$ and Table $3(b)$ respectively. The D-efficiency and I_F-efficiency of this design are 103.89% and 80.88% respectively. Moreover, this design is irregular and nonhomogeneous. Compare the dispersion matrices (Table 2 and Table $3(b)$). In the latter design, many estimates are correlated.

3.2. *Construction in 12 runs.*

When there are restrictions on the run length, one often ignores certain higher order interactions so that the run length is reduced. There could be other reasons to ignore higher order interactions such as difficulty in giving practical interpretation to such interactions etc. For example, Webb [1971] argued in favor of considering only linear components of interaction involving 3-level factors. Supposing we can sacrifice a highest order in-

Table 3

(a)**: Resolution IV Design in 19 Runs**

TC No.	C	D	A	B	TC. No.	C	D	A	B
1	1	1	1	1	11	2	1	2	2
2	1	1	1	2	12	2	2	1	2
3	1	1	1	3	13	2	2	1	3
4	1	1	2	1	14	2	2	2	1
5	1	1	3	1	15	2	2	2	2
6	1	2	1	1	16	2	2	2	3
7	1	2	2	2	17	2	2	3	1
8	1	2	3	3	18	2	2	3	2
9	2	1	1	1	19	2	2	3	3
10	2	1	3	3					

(b)**: Dispersion Matrix of the Design of Table 3(a)** $\times 10^{-2}$

Effect	μ	C	D	A_LC	A_QC	A_L	A_Q	B_L	B_Q	A_LB_L	A_LB_Q	A_QB_L	A_QB_Q
μ	7.67												
C	1.51	6.76											
D	1.51	-1.58	6.76										
A_LC	-2.67	-0.46	-0.46	10.32									
A_QC	-0.55	-0.46	-0.46	-0.39	3.18								
A_L	-1.88	-1.39	-1.39	1.19	-0.40	9.74							
A_Q	-0.48	-0.60	-0.60	-1.31	1.34	-0.50	4.02						
B_L	-1.89	-1.54	-1.54	0.17	0.96	-0.24	1.56	9.87					
B_Q	-0.82	-0.45	-0.45	1.50	-0.61	1.42	-1.26	-0.35	3.64				
A_LB_L	-0.48	0.08	0.08	-1.72	0.07	-0.20	-0.25	-0.03	-0.71	10.70			
A_LB_Q	1.47	0.60	0.60	-0.97	0.02	-2.27	0.48	-0.79	-1.17	0.16	5.54		
A_QB_L	1.48	0.75	0.75	0.05	-1.34	-0.62	-1.58	-2.57	0.60	-0.01	-0.11	6.01	
A_QB_Q	-1.30	-0.39	-0.39	0.75	-0.10	0.76	0.25	0.66	0.50	-0.36	-0.62	-0.51	1.58

teraction in our problem, we can ask the question: Is there an efficient design in 12 runs?

Earlier it was observed that a minimum of 13 runs is necessary to estimate all the components of main effects and interactions. We assume that the highest order component of the interaction between A and B, viz., A_QB_Q, is absent. We proceed, as before, to construct the design by first writing down the $2^2 3^1$ (= 12) full-factorial layout (see columns (2), (3) and (4) of Table 4).

We then augment these 3 columns with another 3 - level column (under B) so that after rearranging the rows, the columns under B, D and C (see 2nd part of Table 4) form a $2^2 3^1$ full-factorial layout. Here we obtain the column under B by minimizing the nonorthogonality between A and B while maintaining the orthogonality of B with C, D and CD. This, again, is an optimization problem and the methodology is described in Section 4. The dispersion matrix of this design is given in Table 5. The D-efficiency and I_F-efficiency of the design are 84.92% and 54.55% respectively.

3.3. *Construction of Design for $2^3 3^3$ with Interactions in 24 runs.*

Consider an experiment with three 3-level factors, A, B, C and three 2-level factors D, E and F. Suppose it is required to estimate the interactions, AB, BC, AD, DE and EF. The minimum run length is 22. We construct a design in 24 runs. As before, we start with a $2^3 3^1$ full-factorial layout with the factors A, D, E and F. Next, we augment this layout with two more 3-level columns for B and C in two steps. First we augment the layout with the column for B so that the nonorthogonalities of B with A, AB and AD are minimized and that the columns of B, D, E and F form a full-factorial layout. From this, we get the layout for the factors A, B, D, E and F in 24 runs.

Finally, this layout is augmented with another 3-level column for C so that the nonorthogonalities of C with A, B, AB, BC and AD are minimized and that the columns of C, D, E and F form a full-factorial layout.

Table 4

Construction of a Fraction with 12 Runs

	$2^2.3$ full-factorial			Additional column	After rearranging rows and columns				
TC. No.	C	D	A	B	TC. No.	B	D	C	A
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	1	1	1	1	1	1	1	1	1
2	1	1	2	3	8	1	1	2	2
3	1	1	3	2	5	1	2	1	2
4	1	2	1	2	12	1	2	2	3
5	1	2	2	1	3	2	1	1	3
6	1	2	3	3	7	2	1	2	1
7	2	1	1	2	4	2	2	1	1
8	2	1	2	1	11	2	2	2	2
9	2	1	3	3	2	3	1	1	2
10	2	2	1	3	9	3	1	2	3
11	2	2	2	2	6	3	2	1	3
12	2	2	3	1	10	3	2	2	1

Table 5

Dispersion Matrix for the Design with 12 Runs

Effect	μ	A_L	A_Q	B_L	B_Q	A_LB_L	A_LB_Q	A_QB_L	C	A_LC	A_QC	D
μ	0.093											
A_L	0.000	0.139										
A_Q	0.000	0.000	0.046									
B_L	0.000	-0.014	0.014	0.222								
B_Q	0.000	0.014	0.005	0.000	0.074							
A_LB_L	0.000	-0.042	0.014	-0.167	0.056	0.667						
A_LB_Q	0.019	-0.014	-0.005	0.000	0.037	0.111	0.111					
A_QB_L	0.019	-0.000	0.009	-0.056	-0.019	0.111	0.000	0.111				
C	0.009	-0.014	0.005	-0.056	0.019	0.222	0.056	0.056	0.167			
A_LC	0.000	0.000	0.000	0.125	0.042	-0.125	0.042	-0.083	-0.042	0.250		
A_QC	0.000	0.000	0.000	0.042	-0.042	-0.125	-0.042	0.000	-0.042	0.000	0.083	
D	0.009	-0.014	0.005	-0.056	0.019	0.222	0.056	0.056	0.083	-0.042	-0.042	0.167

4. OR Formulations

In Section 3, we have encountered the problem of augmenting a given set of columns with 2- or 3-level column so as to form a design layout. In this section, we formulate this problem as an integer programming problem (with linear or quadratic objective function). We first describe the procedure and the formulation, and then illustrate the same with some examples.

4.1. *Augmenting with 2-level columns*

Consider the problem of constructing the fourth column discussed in Subsection 3.1. It is required to construct a 2-level column which along with columns of A and B forms a full-factorial design. In addition to this, the resulting layout (with the columns) should provide estimability of all the main effects and interactions of interest.

The problem of choosing a 2-level column is equivalent to finding a 18×1 CV x so that it is orthogonal to l_{A_L}, l_{A_Q}, l_{B_L}, l_{B_Q}, $l_{A_LB_L}$, $l_{A_LB_Q}$, $l_{A_QB_L}$, $l_{A_QB_Q}$, and its nonorthogonality with l_C, l_{CA_L}, l_{CA_Q} is as close to zero as possible. In fact, if the nonorthogonalities ($\mid x^t l_C \mid$, $\mid x^t l_{A_LC} \mid$, $\mid x^t l_{A_QC} \mid$) are equal to zero, then the resulting design is an orthogonal design with which we can estimate all the main effects and interactions orthogonally. Therefore, *we expect that minimizing the nonorthogonality, in some sense, should lead us to estimability of the required main effects and interactions.*

Thus the problem of constructing the fourth column (which in turn gives us a design) is an optimization problem. Since we are interested in minimizing $\mid x^t l_C \mid$

, $\mid x^t l_{A_L C} \mid$ and $\mid x^t l_{A_Q C} \mid$ we may tackle the problem in several ways. We shall consider two types of objective functions, $f(x)$, here.

(i) Take $f(x) = \lambda_1 \mid x^t l_C \mid + \lambda_2 \mid x^t l_{A_L C} \mid + \lambda_3 \mid x^t l_{A_Q C} \mid$, where λ_is are some positive numbers (weights). Setting any of the λ_i's to be a very large positive number is equivalent to treating the corresponding component as a constraint. For example, if we set λ_1 to be a large positive number, then we are looking for a solution that satisfies $x^t l_C = 0$. This, in other words, means we are looking for a design which can provide orthogonal estimates for all the main effects while minimizing the nonorthogonality between D and AC.

(ii) Take $f(x) = (x^t l_C)^2 + (x^t l_{A_L C})^2 + (x^t l_{A_Q C})^2$. In this case, the resulting problem is a convex quadratic integer programming problem. Note that $f(x) = x^t P P^t x$ where $P = [l_C,\ l_{A_L C},\ l_{A_Q C}]$.

The complete formulation of the problem with objective function as discussed in (i) is given below. We use the standard notation a^+, a^-. That is, for any real a, $a^+ = \max(a, 0)$ and $a^- = \max(-a, 0)$; and for any vector x, x^+ and x^- are defined by $(x^+)_i = (x_i)^+$ and $(x^-)_i = (x_i)^-$. Further, we use e for the vector of 1's. The order of e will be clear from the context.

Formulation (F1):

$$\begin{array}{ll} \text{Minimize} & \lambda_1 u + \lambda_2 v + \lambda_3 w \\ \text{subject to} & \\ & \begin{array}{ll} Mx^+ - Mx^- & = 0, \\ e^t x^+ & = 9, \\ x^+ + x^- & = e, \\ -u \le l_C^t (x^+ - x^-) & \le u \\ -v \le l_{A_L C}^t (x^+ - x^-) & \le v \\ -w \le l_{A_Q C}^t (x^+ - x^-) & \le w \end{array} \end{array}$$

$u,\ v,\ w,\ x_i^+$ and x_i^- are nonnegative integers, $i = 1, 2, \ldots, 18$,
where $M = [l_{A_L}, l_{A_Q}, l_{B_L}, l_{B_Q}, l_{A_L B_L}, l_{A_L B_Q}, l_{A_Q B_L}, l_{A_Q B_Q}]^t$, and $\lambda_1,\ \lambda_2,\ \lambda_3$ are predetermined positive numbers.

For any solution $u,\ v,\ w,\ x^+, x^-$ of the above problem, define $x = x^+ - x^-$. Then x is the required CV and the 2-level column is given by $(e + x^+)$. The constraints $x^+ + x^- = e$ and $e^t x^+ = 9$ force the vector x to be a CV.

It may be noted that by taking the objective function $f(x) = (x^t l_C)^2 + (x^t l_{A_L C})^2 + (x^t l_{A_Q C})^2$ in this case, we actually find a design which optimizes the D-efficiency. This is because $\mid X^t X \mid = \alpha_1 - \alpha_2 f(x)$, where α_1 and α_2 are some positive constants.

We shall illustrate the above method with another example where it is required to construct a design with 12 TCs in Subsection 4.2.

4.2. *Example with a* $2^3 3^1$ *Experiment*

A design is to be constructed with four factors A, B, C and D with A at 3 levels, and B, C and D at two levels each. Furthermore, the interaction AB and BC are to be estimated. Here we have $2^3 3^1$ factorial experiment with 24 TCs in all. In order to estimate the main effects and interactions of interest, we need a minimum of 9 TCs. We shall construct a design with 12 TCs using our method. Table 6 presents a full-factorial experiment for A, B and C. It also presents the corresponding CVs for these factors.

In order to construct the required design, we need to augment the layout of Table 6 with another 2-level column as follows: construct a 2-level column for D so that the columns under $A,\ B,\ D$ form a $2^2 3^1$ full-factorial experiment (by doing this we ensure that AB interaction can be estimated orthogonally) and the nonorthogonalities of D with C and BC are minimized.

The mathematical formulation of this problem is given by:

$$\begin{array}{ll} \text{Minimize} & 100u + v \\ \text{subject to} & \\ & \begin{array}{ll} Mx^+ - Mx^- & = 0, \\ e^t x^+ & = 6, \\ x^+ + x^- & = e, \\ -u \le x^t l_C & \le u, \\ -v \le x^t l_{BC} & \le v, \end{array} \end{array}$$

$u,\ v,\ x_i^+$ and x_i^- are nonnegative integers, $i = 1, 2, \ldots, 12$,
where

$$M = [l_{A_L}, l_{A_Q}, l_B, l_{A_L B}, l_{A_Q B}, l_C\ ]^t.$$

The coefficient of u in the objective function is chosen to be 100 so that if there is a feasible solution x $(= x^+ - x^-)$ with $x^t l_C = 0$, then such a solution will emerge as an optimal solution, which in turn, leads to estimability of all the main-effects orthogonally.

The above problem is solved using the LINGO package and the solution is given by

$$x = (-1,\ 1,\ 1,\ -1,\ 1,\ -1,\ -1,\ 1,\ 1,\ -1,\ -1,\ 1).$$

Hence the column for D is $(1,\ 2,\ 2,\ 1,\ 2,\ 1,\ 1,\ 2,\ 2,\ 1,\ 1,\ 2)^t$. The dispersion matrix of the design is given in

Table 6

Layout for $A,\ B,\ C$ and the CVs

TC.No.	A	B	C	l_{A_L}	l_{A_Q}	l_B	l_{A_LB}	l_{A_QB}	l_C	l_{BC}
1	1	1	1	-1	1	-1	1	-1	-1	1
2	1	1	2	-1	1	-1	1	-1	1	-1
3	1	2	1	-1	1	1	-1	1	-1	-1
4	1	2	2	-1	1	1	-1	1	1	1
5	2	1	1	0	-2	-1	0	2	-1	1
6	2	1	2	0	-2	-1	0	2	1	-1
7	2	2	1	0	-2	1	0	-2	-1	-1
8	2	2	2	0	-2	1	0	-2	1	1
9	3	1	1	1	1	-1	-1	-1	-1	1
10	3	1	2	1	1	-1	-1	-1	1	-1
11	3	2	1	1	1	1	1	1	-1	-1
12	3	2	2	1	1	1	1	1	1	1

Table 7

Dispersion Matrix

Effect	μ	A_L	A_Q	B	A_LB	A_QB	C	BC	D
μ	0.083								
A_L	0	0.125							
A_Q	0	0	0.042						
B	0	0	0	0.083					
A_LB	0	0	0	0	0.125				
A_QB	0	0	0	0	0	0.042			
C	0	0	0	0	0	0	0.083		
BC	0	0	0	0	0	0	0	0.094	
D	0	0	0	0	0	0	0	0.031	0.094

D-Efficiency = 105.22%

I_F-Efficiency = 97.30%

Table 7. It can be seen from the dispersion matrix that all the main effects can be estimated orthogonally.

4.3. *Augmenting with 3-level columns*

Consider the problem of constructing a 3-level column discussed in Subsection 3.2. Given the columns of $A,\ C$ and D, the problem is to construct a 3-level column for B so that its three levels appear with the same frequency and the CVs of B (linear and quadratic) are orthogonal to $C,\ D$ and CD. In addition, the column should be chosen in such a way that it minimizes the nonorthogonalities of B with $A,\ AB$ and AC.

Constructing a homogeneous column for B is equivalent to constructing the linear (and hence quadratic) CV with 0s, 1s and -1s so that they appear with the same frequency. So for the problem in question, we must construct a (linear) CV x with four 1s, four zeros and four -1s satisfying the orthogonality constraints with $l_C,\ l_D$ and l_{CD}. This means $x\ (= l_{B_L})$ and the resulting quadratic component of $B\ (l_{B_Q})$ must be orthogonal to $l_C,\ l_D$ and l_{CD}. Observe that since x is the linear CV of B (l_{B_L}), the quadratic CV of B is given by $l_{B_Q} = 3(x^+ + x^-) - 2e$. Thus, in order that the columns of $B,\ C,\ D$ form a full-factorial design, we must have

$$M^t[l_{B_L}, l_{B_Q}] = 0, \text{ where } M = [l_C, l_D, l_C \bullet l_D].$$

Note that the $\bullet$ product between l_C and l_D is nothing but l_{CD} (for any two vectors u and v of same order, $u \bullet v$ is the vector w whose ith coordinate is $w_i = u_i v_i$). Since $l_{B_L} = x^+ - x^-$ and $l_{B_Q} = 3(x^+ + x^-) - 2e$, we have

$$\begin{aligned} &M^t[x^+ - x^-, 3(x^+ + x^-) - 2e] \\ &= [M^t(x^+ - x^-), 3M^t(x^+ + x^-)] \end{aligned}$$

Hence, the constraints reduce to

$$\begin{bmatrix} M^t & -M^t \\ M^t & M^t \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \end{bmatrix} = 0.$$

Next, consider the objective function. We wish to choose x so that the nonorthogonalities of B with A (i.e., $\mid l^t_{A_L} l_{B_L} \mid,\ \mid l^t_{A_Q} l_{B_L} \mid,\ \mid l^t_{A_L} l_{B_Q} \mid$ etc.), the nonorthogonalities of B with AB interaction (i.e., $\mid l^t_{B_L}(l_{A_L} \bullet l_{B_L}) \mid$, $\mid l^t_{B_Q}(l_{A_L} \bullet l_{B_L}) \mid$, $\mid l^t_{B_L}(l_{A_L} \bullet l_{B_Q}) \mid$, etc.) and the nonorthogonalities of B with AC interaction (i.e., $\mid l^t_{B_L}(l_{A_L} \bullet l_C) \mid$ etc.) are minimized. An interesting observation is that all these nonorthogonalities are linear functions of x^+ and x^-. To see this, first observe that for any vectors $p,\ q,\ u,$ and v (of same order),

(i**)** $p^t(u \bullet v) = u^t(p \bullet v)$,

(ii**)** $(p+q) \bullet (u+v) = p \bullet q + q \bullet u + p \bullet v + q \bullet v$.

So,

$$\begin{aligned} l^t_{B_L}(l_{A_L} \bullet l_{B_L}) &= l^t_{A_L}(l_{B_L} \bullet l_{B_L}) \\ &= l^t_{A_L}((x^+ - x^-) \bullet (x^+ - x^-)) \\ &= l^t_{A_L}(x^+ \bullet x^+ - x^+ \bullet x^- - x^- \bullet x^+ \\ &\quad + x^- \bullet x^-) \\ &= l^t_{A_L}(x^+ + x^-) \end{aligned}$$

and

$$\begin{aligned} l^t_{B_Q}(l_{A_L} \bullet l_{B_L}) &= l^t_{A_L}(l_{B_Q} \bullet l_{B_L}) \\ &= l^t_{A_L}((3(x^+ + x^-) - 2e) \bullet \\ &\quad (x^+ - x^-)) \\ &= l^t_{A_L}(x^+ - x^-) \\ &= l^t_{A_L} l_{B_L}. \end{aligned}$$

Similarly, we can show that other nonorthogonalities are also linear functions of x^+ and x^-.

For simplicity, we take the objective function as $\mid l^t_{A_L} l_{B_L} \mid + \mid l^t_{A_Q} l_{B_L} \mid + \mid l^t_{A_L} l_{B_Q} \mid$. The complete formulation is given below.

Formulation (F2):

$$\begin{array}{ll} \text{Minimize} & u + v + w \\ \text{subject to} & \\ & \begin{bmatrix} M^t & -M^t \\ M^t & M^t \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \quad = 0, \\ & e^t x^+ \qquad\qquad\qquad = 4, \\ & e^t x^- \qquad\qquad\qquad = 4, \\ & x^+ + x^- \qquad\qquad \leq e, \\ & -u \ \leq \ l^t_{A_L} l_{B_L} \qquad \leq u, \\ & -v \ \leq \ l^t_{A_Q} l_{B_L} \qquad \leq v, \\ & -u \ \leq \ l^t_{A_L} l_{B_Q} \qquad \leq w, \end{array}$$

$u,\ v,\ w,\ x^+_i,\ x^-_i$ are nonnegative integers, $i = 1, 2, \ldots, 12$.

The problem is solved using the LINGO package. The resulting layout and the corresponding dispersion matrix are given in Subsection 3.2. It must be mentioned that when we include the $A_Q B_Q$ component in the objective function, the program execution terminates with the conclusion that the problem is nonoptimal/infeasible.

4.4. *Enlisting all solutions*

Consider the problem of finding all columns (of 2 or 3-levels) those are orthogonal to a given set of columns. It is possible to solve this problem iteratively using the integer programming formulations. Wang and Wu [1991,1992] constructed a number of orthogonal and nearly orthogonal main effect plans for $2^m 3^n$ factorial experiments. In Wang and Wu [1992], it is shown that there are sixteen 2-level columns orthogonal to 3 given columns (one 3-level and two 2-level) – all having the same level in the first coordinate, and enlist all the 16 2-level columns (see Section 2, page 411 of Wang and Wu [1992] for details).

We illustrate our methodology to enlist all columns orthogonal to a given set of columns. Consider problem **F1** formulated in Subsection 4.1. Suppose $(x^+,\ x^-)$ is a feasible solution of **F1** ($u,\ v,\ w$ are not mentioned here as they are determined by x^+ and x^-). Define $\alpha = \{i : x^+_i > 0\}$ and $S = \{(z^+,\ z^-) : (z^+,\ z^-)$ is a feasible solution to **F1**$\}$. Augment the constraints of **F1** with $\sum_{i \in \alpha} x^+_i \leq 8$. Let $S^\star = \{(z^+,\ z^-) : (z^+,\ z^-)$ is a feasible solution to the augmented problem $\}$. Note that $S^\star = S \setminus \{(x^+,\ x^-)\}$. Thus by finding a feasible solution to the augmented problem, we get a new feasible solution to **F1**. By repeating the above process (by adding a new constraint in every iteration) we can generate all feasible solutions to **F1**. Solving the Wang and Wu's problem mentioned in the previous paragraph (with the additional constraint $x^+_i = 1$ to ensure that all columns have one in the first row), we find that there are exactly sixteen distinct solutions (columns) and in the 17th iteration, the augmented problem becomes infeasible.

5. Summary

Factorial Experiments are extensively used in industrial experimentation and other disciplines. The experiments often use asymmetrical factorial experiments. Constructing efficient designs for asymmetrical factorial experiments is, in general, a complex problem. In this article, we have considered $2^m 3^n$ factorial experiments with interactions. We have presented a methodology for construction of nearly orthogonal designs which are regular and homogenous. By formulating the problem of construction as an integer programming problem, we have shown that the designs can be constructed using OR packages. This has been demonstrated with examples. We have remarked that by minimizing the nonorthogonality, we expect that the resulting design will lead to estimability of the main effects and interactions of interest. In this connection we conclude this article with the following question : Is there a bound on the nonorthogonal objective functions considered in the formulations which will ensure nonsingularity of X^tX (in other words, estimability of parameters of interest)?

References

[1] Addleman S., Orthogonal Main-Effect Plans for Asymmetric Factorial Experiments, Technometrics, 4 (1962) 21-46.

[2] Anderson D.A. and Thomas A.M., Resolution IV Fractional Factorial Design for the General Asymmetric Factorial, Commun. Statist. - Theor. Meth., 8 (1979) 931-943.

[3] Box M.J. and Draper N.R., Factorial Designs, the $\mid X^tX \mid$ Criterion, and Some Related Matters, Technometrics, 13 (1971) 731-742.

[4] Chakravarti I.M. , Fractional Replication in Asymmetric Factorial Designs and Partially Balanced Arrays, Sankhya, 17 (1956) 143-164.

[5] Cheng C.S. , Some Orthogonal Main-Effect Plans for Asymmetric Factorials, Technometrics, 31 (1989) 475-477.

[6] Mitchell T.J. , An Algorithm for Construction of D-Optimal Designs, Technometrics, 16 (1974) 203-210.

[7] Mitchell T.J. , Computer Construction of D-Optimal First Order Designs, Technometrics, 16 (1974) 211-220.

[8] Raghavarao D. , Constructions and Combinatorial Problems in Design of Experiments, John Wiley, New York, 1971.

[9] Rao C.R. , Linear Statistical Inference and its Applications (2ed), John wiley, New York, 1974.

[10] Wang J.C. and Wu C.F.J. , An Approach to the Construction of Asymmetrical Orthogonal Arrays, JASA, 86 (1991) 450-456.

[11] Wang J.C. and Wu C.F.J. , Nearly Orthogonal Arrays with Mixed Levels and Small Runs, Technometrics, 34 (1992) 409-422.

[12] Webb S.R., Small Incomplete Factorial Experiment Design for Two- and Three-Level Factors, Technometrics, 13 (1971) 243-256.

Appendix

A :D-optimality of the design with layout in Table 1

We consider the layout in the right panel of Table 1. Let $\theta = (\alpha, \beta)$, where $\alpha = (D, C, A_LC, A_QC, A_L, A_Q)$ and β = (μ, B_L, B_Q, A_LB_L, A_LB_Q, A_QB_L, A_QB_Q) with μ as the general effect. Then θ is the vector of general, main and interaction effects to be estimated. It can be checked that θ is estimable.

For convenience, we use the following notation in the proof of our theorem below:

(1) For any two vectors g and h (of same order), the dot product $f = g \bullet h$ is a vector f defined by the coordinate-wise product $f_i = g_ih_i$; $< g, h >$ stands for the inner product between g and h.

(2) For simplicity and ease of understanding, we denote the columns of the design matrix by the corresponding factorial effects.

(3) For any vector g, g^+ is the vector whose i^{th} coordinate is $\max(g_i,\ 0)$; and g^- is the vector whose i^{th} coordinate is $\max(-g_i,\ 0)$.

Theorem : The design matrix corresponding to Table 1 is D-optimal among all homogeneous designs with 18 runs for estimating θ.

Proof : Let $G = [X : Y]$ be the design matrix for any homogeneous design where columns of X correspond to α and those of Y correspond to β. Let $H = [X_0 : Y_0]$ be the resulting design matrix of layout in Table 1. We will show that

$$\mid H^tH \mid \ = \ \max_G \mid G^tG \mid .$$

If possible, assume that $\mid G^tG \mid \ > \ \mid H^tH \mid$ for some G.

Claim 1. It can be checked that $X_0^tY_0 = 0$ and hence $\mid H^tH \mid = \mid X_0^tX_0 \mid \cdot \mid Y_0^tY_0 \mid$. Furthermore, $Y_0^tY_0$ is diagonal and

$$X_0^tX_0 = \begin{array}{c} \\ D \\ C \\ A_LC \\ A_QC \\ A_L \\ A_Q \end{array} \begin{array}{c} \begin{array}{cccccc} D & C & A_LC & A_QC & A_L & A_Q \end{array} \\ \begin{bmatrix} 18 & -2 & -4 & 4 & 0 & 0 \\ -2 & 18 & 0 & 0 & 0 & 0 \\ -4 & 0 & 12 & 0 & 0 & 0 \\ 4 & 0 & 0 & 36 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 36 \end{bmatrix} \end{array}$$

$$= \Sigma_0, \quad \text{(say).}$$

Claim 2. $\mid G^tG \mid \ \leq \ \mid X^tX \mid \cdot \mid Y^tY \mid \ \leq \ \mid X^tX \mid \cdot \mid Y_0^tY_0 \mid$. The last inequality follows from the fact that the diagonal entries of G^tG are same as those of H^tH, since G is homogeneous.

Let $\Sigma = X^tX = ((\sigma_{ij}))$. Also, let

$$\begin{aligned} \sigma_{25} &= < A_L, C > \; = x \\ \text{and } \sigma_{35} &= < A_L \bullet C, A_L > \; = \; < A_L \bullet A_L, C > \\ &= < A_L^+ \bullet A_L^-, C > \; = \; y. \end{aligned}$$

Then,

$$\begin{aligned} \sigma_{26} &= < A_Q, C > \; = \; < 3(A_L^+ + A_L^-) - 2e, C > \\ &= 3y, \\ \sigma_{36} &= < A_Q, A_L \bullet C > \; = \; < A_L, C > \; = \; x, \\ \sigma_{45} &= < A_Q \bullet C, A_L > \; = \; x, \\ \sigma_{46} &= < A_Q \bullet C, A_Q > \; = \; < A_Q \bullet A_Q, C > \\ &= < 4e - 3(A_L^+ + A_L^-), C > \\ &= -3y. \end{aligned}$$

Further, $\sigma_{23} = < C, A_L \bullet C > = 0$, $\sigma_{24} = < C, A_Q \bullet C > = 0$, $\sigma_{34} = < A_L \bullet C, A_Q \bullet C > = 0$ and $\sigma_{56} = < A_L, A_Q > = 0$.

Let $\lambda = \sigma_{12} = < D, C >$, $\delta = \sigma_{13} = < D, A_L \bullet C >$, $\tau = \sigma_{14} = < D, A_Q \bullet C >$, $\gamma = \sigma_{15} = < D, A_L >$, $\eta = \sigma_{16} = < D, A_Q >$.

With the above notation and calculations, we can write

$$\Sigma = \begin{array}{c} \\ D \\ C \\ A_LC \\ A_QC \\ A_L \\ A_Q \end{array} \begin{array}{c} \begin{array}{cccccc} D & C & A_LC & A_QC & A_L & A_Q \end{array} \\ \begin{bmatrix} 18 & \lambda & \delta & \tau & \gamma & \eta \\ \lambda & 18 & 0 & 0 & x & 3y \\ \delta & 0 & 12 & 0 & y & x \\ \tau & 0 & 0 & 36 & x & -3y \\ \gamma & x & y & x & 12 & 0 \\ \eta & 3y & x & -3y & 0 & 36 \end{bmatrix} \end{array}$$

Observation 1. Note that, with 18 runs, $\lambda = < D, C >$ is such that either $|\lambda| = 2$ or $|\lambda| \geq 6$.

Next,

$$|\Sigma| \leq \begin{vmatrix} 18 & \lambda \\ \lambda & 18 \end{vmatrix} \cdot \begin{vmatrix} 12 & 0 & y & x \\ 0 & 36 & x & -3y \\ y & x & 12 & 0 \\ x & -3y & 0 & 36 \end{vmatrix}$$

If $|\lambda| \geq 6$, then the right hand side of the above inequality is less than or equal to $|\Sigma_0|$. Hence, as $|G^tG| > |H^tH|$, $|\lambda| = 2$.

Claim 3. $|\lambda| = 2$, $x = 0$ and $y = 0$.

We have already shown that $|\lambda| = 2$. Now,

$$|\Sigma| \leq 18 \cdot \begin{vmatrix} 18 & 0 & 0 & x & 3y \\ 0 & 12 & 0 & y & x \\ 0 & 0 & 36 & x & -3y \\ x & y & x & 12 & 0 \\ 3y & x & -3y & 0 & 36 \end{vmatrix}$$

$$\begin{aligned} &= 18 \times 6 \left[\left(3x^2 + 9y^2 - 24 \times 36\right)^2 - 12^2 \times 36^2 \right] \\ &= \; \Delta, \quad \text{say}. \end{aligned}$$

Let $P = \{i : i\text{th coordinate of } A_L = 0\}$. Then

$$\begin{aligned} y &= < A_L \bullet C, A_L > \\ &= < A_L^+ + A_L^-, C > \\ &= 2r - 6, \end{aligned}$$

where r is the number of -1's in $\{C_i : i \in P\}$.

Let $Q = \{i : i^{th} \text{ coordinate of } A_L = -1\}$ and let s be the number of -1's in $\{C_i : i \in Q\}$. Then $x = 2r + 4s - 18$.

Clearly, $0 \leq r, s \leq 6$ and $r + s \leq 9$. With these restrictions, we can work out the values of x and y, and evaluate Δ. It turns out that

$$\Delta < |\Sigma_0| \qquad \text{for all} \;\; (x, y) \neq (0, 0).$$

Since $|\Sigma| > |\Sigma_0|$, it follows that $x = y = 0$. Thus,

$$\Sigma = \begin{bmatrix} 18 & \lambda & \delta & \tau & \gamma & \eta \\ \lambda & 18 & 0 & 0 & 0 & 0 \\ \delta & 0 & 12 & 0 & 0 & 0 \\ \tau & 0 & 0 & 36 & 0 & 0 \\ \gamma & 0 & 0 & 0 & 12 & 0 \\ \eta & 0 & 0 & 0 & 0 & 36 \end{bmatrix}$$

where λ is equal to -2 or 2.

Suppose $\lambda = 2$. Without loss of generality, we can write the contrast vectors for D and C in four blocks as shown in the Table A.1.

In the remaining part of the table, we have the frequencies of elements in the contrast vector for A_L corresponding to the blocks.

Using the above table, we compute the quantities :

$$\begin{aligned} \delta &= < C \bullet D, A_L > &&= 2(t - r), \\ \tau &= < C \bullet D, A_Q > &&= 2(t + r) - 4s, \\ \gamma &= < D, A_L > &&= 2(t_1 + t_2 - r_1 - r_2), \\ \eta &= < D, A_Q > &&= 6(t_1 + t_2 + r_1 + r_2 - 6), \\ x &= < C, A_L > &&= 2(t_1 - t_2 - r_1 + r_2 - t + r), \\ y &= < A_L \bullet C, A_L > &&= 6(t_1 - t_2 + r_1 - r_2 - t \\ & && \quad -r + 6). \end{aligned}$$

The restrictions on the frequencies are

Table A.1

Contrast Vectors of C and D

Block	D	C	Frequency
I	1	1	$r_1 = \# -1$
	1	1	$s_1 = \# 0$
	1	1	$t_1 = \# 1$
	1	1	
	1	1	
II	−1	−1	$r - r_1 = \# -1$
	−1	−1	$s - s_1 = \# 0$
	−1	−1	$t - t_1 = \# 1$
	−1	−1	
	−1	−1	
III	1	−1	$r_2 = \# -1$
	1	−1	$s_2 = \# 0$
	1	−1	$t_2 = \# 1$
	1	−1	
IV	−1	1	$6 - r - r_2 = \# -1$
	−1	1	$6 - s - s_2 = \# 0$
	−1	1	$6 - t - t_2 = \# 1$
	−1	1	

$$
\begin{gathered}
0 \le r, s, t \le 6, \\
r + s + t = 10, \\
0 \le r_1 \le \min\{r, 5\}, \\
0 \le s_1 \le \min\{s, 5\}, \\
0 \le t_1 \le \min\{t, 5\}, \\
r_1 + s_1 + t_1 = 5, \\
0 \le r_2 \le \min\{6 - r, 4\}, \\
0 \le s_2 \le \min\{6 - s, 4\}, \\
0 \le t_2 \le \min\{6 - t, 4\}, \\
r_2 + s_2 + t_2 = 4.
\end{gathered}
$$

By setting $x = 0$ and $y = 0$, it has been checked by enumeration, through a simple computer program, that under the above restrictions, $\mid \Sigma \mid \le \mid \Sigma_0 \mid$ for every feasible combination of $(r, s, t, r_1, s_1, t_1, r_2, s_2, t_2)$ (there are only 144 feasible combinations). This contradicts our assumption about the existence of G satisfying $|G^t G| > |H^t H|$.

Similar arguments hold for $\lambda = -2$ also. This completes the proof.

Received 8-1-2013; accepted 17-2-2013