BLOCKING FACTORIAL DESIGNS IN GREENHOUSE EXPERIMENTS

Steve Ferris

Steven G. Gilmour

BLOCKING FACTORIAL DESIGNS IN GREENHOUSE EXPERIMENTS

Steve Ferris
Biomathematics and Statistics Scotland
United Kingdom

Steven G. Gilmour†
The University of Reading
United Kingdom

† Address for correspondence: Department of Applied Statistics, The University of Reading, PO Box 240, Earley Gate, Reading RG6 6FN, UK. E-mail: S.G.Gilmour@reading.ac.uk

Abstract

Experiments in greenhouses usually have to be conducted with very limited resources. This makes it particularly important to control the between plot variation by appropriate use of blocking. Many greenhouse experiments are naturally laid out in a pattern that makes a class of designs known as semi-Latin squares useful. Their properties have been studied recently by a number of authors and this work is reviewed. Often, the experimental treatments will have a factorial structure. An example of a 2^3 structure is used to show how factorial treatments can be assigned to treatment labels to ensure that the appropriate information is obtained from the experiment.

Keywords: Glasshouse experiments; Semi-Latin square design; Tomatoes; Trojan square design.

1 Introduction

A class of designs known as Trojan squares or, more generally, semi-Latin squares can be extremely useful for experiments in many agricultural applications. These designs deserve to be more widely known and used. In this paper we describe the use and properties of Trojan square designs in the context of experiments in greenhouses on crops such as tomatoes, an application recently discussed by Edmondson (1998). In addition, we present some new work, on the use of these designs when the treatments have a factorial structure. The background is discussed in Section 2 and Trojan squares are described in Section 3. In Section 4 we present the results for an example

of a factorial treatment structure and some general comments are made in Section 5.

2 Experiments in Greenhouses

Protected crops such as tomatoes, cucumbers and peppers are an important part of horticulture in southern England and other parts of western Europe. Because of the cost of running commercial-scale greenhouses, experiments in the research station usually have to be done with fairly limited resources. Thus, for example, the facilities available may be a single compartment in a greenhouse. Within the greenhouse compartment, plots are often arranged in columns, which may come in pairs to facilitate access to the plots. A plan of a typical experimental layout of plots is given in Figure 1.

When considering blocking for experiments in such a layout, pairs of columns naturally form one blocking factor and rows naturally form another. That is to say that we can expect considerable variation between pairs of columns and between rows, so that in order to obtain efficient estimates of treatment comparisons we should use both of these as blocking factors. To simplify the terminology, we will refer to a pair of columns simply as "a column". There are two ways we can define blocks:

- Use two-way blocking, i.e. a row and column design, assuming the row and column effects are additive. Unlike most row and column designs, like Latin squares, there are two plots within each row$\times$column combination. With this blocking system, we should ensure that each treatment appears (as nearly as possible) equally often in each row and each treatment appears (as nearly as possible) equally often in each column.

- Use the row$\times$column combinations as blocks of size 2. After the initial definition of blocks, this ignores their spatial layout. With this blocking system, we should ensure that we have an incomplete block design which is as efficient as possible, i.e. a balanced, or as nearly balanced as possible, incomplete block design. This means that each pair of treatments should appear together in blocks as nearly as possible equally often.

The idea of Trojan square designs is to meet both of the above requirements simultaneously. Trojan squares can be useful in other areas of application. For example, The University of Reading's poultry house used to be arranged so that cages were stacked in a number of different brooders, with each tier in each brooder having two cages. Thus the layout looks just like Figure 1, except that this represents a side view, rather than a view from above. Bailey (1992) discussed a number of other applications, some in agriculture, some in other areas.

3 Trojan Square Designs

Consider running an experiment to compare 8 varieties of tomato in the facility shown in Figure 1. Label the varieties A, B, C, D, α, β, γ, δ, where the difference between Latin and Greek letters is meaningless as far as the varieties are concerned, i.e. all eight letters are interchangeable. If we use rows and columns as blocking factors, any design with each variety once in each row and once in each column will be fully efficient. For example, the design shown in Figure 2, which is known as an inflated Latin square, will be as good as any other.

On the other hand, if the blocks are used, ignoring the spatial arrangement, any good incomplete block design will be adequate. For 8 treatments in 16 blocks, each of 2 plots, no balanced incomplete block design exists, so a nearly balanced design is required. The design shown in Figure 3 is a partially balanced incomplete block design with two associate classes, i.e. there are two groups (Alphabets) of treatments, those in the same group appearing together in blocks the same number of times (zero times) and those in different groups appearing together in blocks the same number of times (once). This is as good as any other design if we ignore the spatial arrangement of blocks.

A Trojan square design achieves the advantages of both of the above designs simultaneously. It is constructed by using two (or more in general) mutually orthogonal Latin squares. A Trojan square design for the layout in our example is shown in Figure 4. In this design, each variety appears once in each row and once in each column and so is as good as the design in Figure 2 when using rows and columns as blocking factors. On the other hand, using blocks of size 2, it is identical to the design in Figure 3.

The appropriate randomization for a Trojan square design is to ran-

domly permute rows, randomly permute columns and randomly permute plots within blocks (row×column combinations). When our treatments are varieties, as in our example, and each comparison is of equal interest, then we will also randomly allocate the treatment labels to varieties. Note that the construction of the above design is mathematically the same as the construction of a Graeco-Latin square. However, here the symbols in each row×column combination apply to different plots, whereas in a Graeco-Latin square design, they apply to the same plot.

Trojan squares can be thought of as doubly resolvable (or Latinized) incomplete block designs. A *resolvable* incomplete block design is one in which blocks can be grouped, so that a group of blocks contains a single complete replicate of the treatments. Trojan squares are incomplete block designs with the blocks arranged so that those blocks in a row contain a complete replicate and those blocks in a column contain a complete replicate. Among doubly resolvable incomplete block designs, Cheng and Bailey (1991) showed that Trojan squares are:

- A-optimal, i.e. the average variance of all possible contrasts among treatments is minimized;
- D-optimal, i.e. the generalized variance of all possible contrasts among treatments is minimized;
- E-optimal, i.e. the maximum variance of any contrast among treatments is minimized.

Trojan squares have a long history, but their existence, construction and statistical properties have been studied in depth only recently - see Preece and Freeman (1983), Bailey (1988, 1990, 1992), Bailey and Royle (1997), Bailey and Chigbu (1997) and Edmondson (1998). They were introduced by Darby and Gilbert (1958) for greenhouse experiments and this application is discussed further by Bailey (1992) and Edmondson (1998).

Analysis of the data from a Trojan square design can be carried out directly using mixed models (SAS proc mixed, Genstat reml, Splus lme, etc.). To analyze data obtained from the design in Figure 4, an appropriate model would be

$$Y_{ijk(l)} = \mu + \rho_i + \gamma_j + (\rho\gamma)_{ij} + v_l + \epsilon_{ijk}, \qquad (1)$$

where $Y_{ijk(l)}$ is the response in plot k ($k = 1, 2$) in row i ($i = 1, \ldots, 4$) and column j ($j = 1, \ldots, 4$) with variety l ($l = 1, \ldots, 8$) applied, μ is the

overall mean, $\rho_i \overset{iid}{\sim} N(0, \sigma^2_\rho)$ is a random row effect, $\gamma_j \overset{iid}{\sim} N(0, \sigma^2_\gamma)$ is a random column effect, $(\rho\gamma)_{ij} \overset{iid}{\sim} N(0, \sigma^2_{\rho\gamma})$ is a random row$\times$column interaction effect, v_l ($\sum_{l=1}^{8} v_l = 0$) is the variety effect and $\epsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$ is the random plot effect. The row$\times$column interaction effect can also be described as a block effect nested within the row and column effects.

To analyze yield data from the above design in SAS, we could use the following commands:

```
proc mixed;
class ROW COLUMN VARIETY;
model YIELD = VARIETY;
random ROW|COLUMN;
```

or

```
proc mixed;
class ROW COLUMN BLOCK VARIETY;
model YIELD = VARIETY;
random ROW COLUMN BLOCK(ROW COLUMN);
```

In this analysis the variety effect is estimated partly from within-block comparisons and partly from between block comparisons. More insight can be gained by defining *pseudo-factors*, ALPHABET and LETTER, where ALPHABET has two levels ($1 = A, B, C, D$; $2 = \alpha, \beta, \gamma, \delta$) and LETTER has four levels ($1 = A, \alpha$; $2 = B, \beta$; $3 = C, \gamma$; $4 = D, \delta$). The pseudo-factors are defined purely for convenience and have no physical meaning. In the model we can replace v_l with $v_l = a_m + l_n + (al)_{mn}$ using the obvious notation. This allows us to see how much information is obtained at the within block level and how much at the between block level. To analyze the data in Genstat, we can use the following commands:

```
blockstructure ROW*COLUMN
treatmentstructure VARIETY//(ALPHABET*LETTER)
anova YIELD
```

or

```
blockstructure (ROW+COLUMN)/BLOCK
treatmentstructure VARIETY//(ALPHABET*LETTER)
anova YIELD
```

This results in an analysis of variance with the structure shown in Table 1. If we used only rows and columns as blocking factors, the Row×Column and Plot strata in the analysis of variance would be combined. Thus all treatment comparisons would be made with 100% efficiency, but the residual variance would include the variation between blocks (having removed variation between rows and between columns). On the other hand, if we used only blocks in the analysis, the Row, Column and Row×Column strata would be combined into a single stratum for Blocks. Since the residual variance in the Blocks stratum would then include variation between rows and variation between columns, it will not provide much information if either of these two sources of variation is large. With the form of analysis in Table 1, if the variation between blocks is relatively small (having removed variation between rows and between columns) then we may get a considerable amount of inter-block information which can then be combined with the within-block information to get improved estimates of the treatment comparisons. If the variation between blocks is very large, we essentially end up with the same analysis as from the within-blocks analysis alone (i.e. treating Block as a fixed effect).

4 Factorial Treatment Structure

The design and analysis considered above are appropriate for experiments to compare unstructured treatments. What if the treatments have a factorial structure? For example, consider an experiment where the treatments have a 2^3 structure with factors Variety (P), Rate of aeration (Q) and Rate of nutrient feed (R). Edmondson (1998) described a similar experiment, but with a 4×2 structure. It is now no longer true that all treatment comparisons are of equal interest. Instead we are interested in the usual factorial contrasts, which themselves have the following order of priority: (i) main effects, (ii) two-factor interactions, (iii) three-factor interaction. We then have to consider which treatment should be allocated to which treatment label.

By complete enumeration of all the possible allocations, we discovered that there are two types of solution for this problem. Table 2 gives an example

of each type. The first type of design involves allocating the treatments so that a particular factorial contrast is obtained from the comparison between alphabets, thus ensuring that contrast is better estimated than the others, which are obtained from comparisons between letters. In the example in Table 2 (Design 1), the main effect of factor P is obtained from the main effect of the pseudo-factor Alphabet. The second type of design involves allocating the treatments so that four of the factorial contrasts are obtained partly from the comparison between alphabets and partly from comparisons between letters, thus ensuring that these contrasts are all slightly better estimated than the others, which are obtained entirely from comparisons between letters. In the example in Table 2 (Design 2), all three main effects and the three factor interaction are obtained partly from the main effect of the pseudo-factor Alphabet.

Another way to regard this allocation is to think of aliasing the contrasts of interest with the contrasts corresponding to the pseudo-factors, and in particular the contrast for the main effect of the pseudo-factor Alphabet. See Box, Hunter and Hunter (1978) for a fuller explanation of aliasing. Table 3 gives the coefficients of the contrasts for Design 1. This shows that, for example, the main effect of P is estimated from

$$(\hat{v}_\alpha + \hat{v}_\beta + \hat{v}_\gamma + \hat{v}_\delta) - (\hat{v}_A + \hat{v}_B + \hat{v}_C + \hat{v}_D),$$

where $\hat{v}_l$ is the estimate of the parameter v_l in the model given by equation (1). It is easily seen that the main effect of P is completely aliased (i.e. has correlation 1) with the Alphabet contrast, while the other effects are completely unaliased (i.e. have correlation 0) with the Alphabet contrast. Note that this aliasing is not harmful, as we will not actually estimate the Alphabet contrast. It is simply a device to illustrate how the estimates of the contrasts of interest relate to the analysis of variance in Table 1.

Table 4 gives the contrasts for Design 2. Here, the main effects and the three-factor interaction are all partially aliased with the Alphabet contrast. It is easily shown that they have correlation 0.5. The two-factor interactions are not aliased with the Alphabet contrast.

The efficiencies (in the Plot stratum) for estimating the factorial effects from the designs in Table 2 are shown in Table 5. As expected, Design 1 gives full efficiency for estimating the main effect of P, but only 50% efficiency for estimating all other effects. Similarly, Design 2 gives higher than 50% efficiency for the main effects and three-factor interaction and 50% efficiency for the two-factor interactions. The results for Design 2 may seem

disappointing. The relatively small increases in efficiency are due to the fact than the estimates of P, Q, R and PQR are correlated with each other. Design 1 allows orthogonal estimation of the effects, which slightly simplifies the interpretation of the results (although the correlations in Design 2 are only $-\frac{1}{7}$). On the other hand, Design 2 is factorially balanced (i.e. the factors can be interchanged without changing the efficiencies) which, as Bailey and Royle (1997) point out, makes interpretation simpler.

Presented with Table 5, most experimenters would probably prefer Design 1, especially if one of the factors was of particular interest. If the factors were all of equal interest and main effects were of particular interest (e.g. if interactions were expected to be small) then Design 2 might be preferred.

We can also consider versions of the optimality criteria, but with attention restricted to the contrasts of interest – the factorial contrasts. We consider:

- A_A- (usually called L-) optimality, i.e. minimizing the average variance of the contrasts of interest;
- D_A-optimality, i.e. minimizing the generalized variance of the contrasts of interest;
- E_A-optimality, i.e. minimizing the maximum variance of any contrast of interest.

Table 6 summarizes the optimality properties of Designs 1 and 2, showing L-, D_A- and E_A- optimality assuming the contrasts of interest are those corresponding to the factorial models of different orders. However, this is probably less informative than Table 5.

5 Summary and Conclusions

We can summarize our conclusions from this work as follows.

- Careful thought should be given to the appropriate form of blocking. Nested blocking structures should be considered if they might allow recovery of more inter-block information.
- Latinized block designs, such as Trojan squares, allow flexible blocking in these circumstances.

- Careful thought should be given to which treatment contrasts are of interest. If all contrasts are not of equal interest then the allocation of treatments to treatment labels is important.

- Allocation of treatments to labels should be done to meet the requirements of the experiment. Thus the contrasts of most interest can be estimated with higher precision than the others.

We believe that Latinized block designs, especially with factorial treatment structures, deserve much wider application than they have received up to now. The corresponding author would be glad to hear from anyone who has used, or is contemplating using, such a design.

Acknowledgements

This work was carried out while the first author was at The University of Reading, as part of his dissertation on the MSc in Biometry. He gratefully acknowledges financial support from the EPSRC. The second author presented the paper at the conference while he was on leave at the University of Wisconsin-Madison. He gratefully acknowledges financial support for attendance at the conference from the Research Board of The University of Reading and from the Statistical Services Centre at The University of Reading. We thank Rodney Edmondson of Horticulture Research International, Wellesbourne, UK, for sending a pre-print of his paper. Finally, we thank the referee for encouraging us to revise the paper to make it more accessable to a general readership.

References

Bailey, R.A. (1988) Semi-Latin squares. *Journal of Statistical Planning and Inference*, **18**, 299-312.

Bailey, R. A. (1990) An efficient semi-Latin square for twelve treatments in blocks of size two. *Journal of Statistical Planning and Inference*, **26**, 263-266.

Bailey, R.A. (1992) Efficient semi-Latin squares. *Statistica Sinica*, **2**, 413-437.

Bailey, R. A. and Chigbu, P. E. (1997) Enumeration of semi-Latin squares. *Discrete Mathematics*, **167/168**, 73-84.

Bailey, R. A. and Royle, G. (1997) Optimal semi-Latin squares with side six and block size two. *Proceedings of the Royal Society of London, Series A*,

453, 1903-1914.

Box, G. E. P., Hunter, W. G. and Hunter, J. S. (1978) *Statistics for Experimenters*. New York: Wiley.

Cheng, C.-S. and Bailey, R. A. (1991) Optimality of some two-associate-class partially balanced incomplete-block designs. *Annals of Statistics*, **19**, 1667-1671.

Darby, L. A. and Gilbert, N. (1958) The Trojan square. *Euphytica*, **7**, 183-188.

Edmondson, R.N. (1998) Trojan square and incomplete Trojan square designs for crop research. *Journal of Agricultural Science, Cambridge*, **131**, 135-142.

Preece, D.A. and Freeman, G.H. (1983) Semi-Latin squares and related designs. *Journal of the Royal Statistical Society, Series B*, **45**, 267-277.

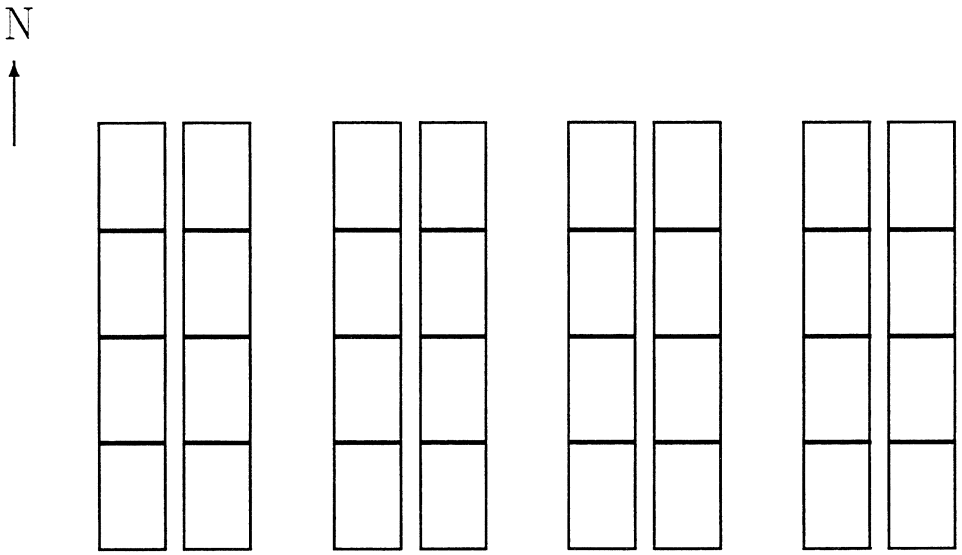

Figure 1: Typical greenhouse layout in a research station.

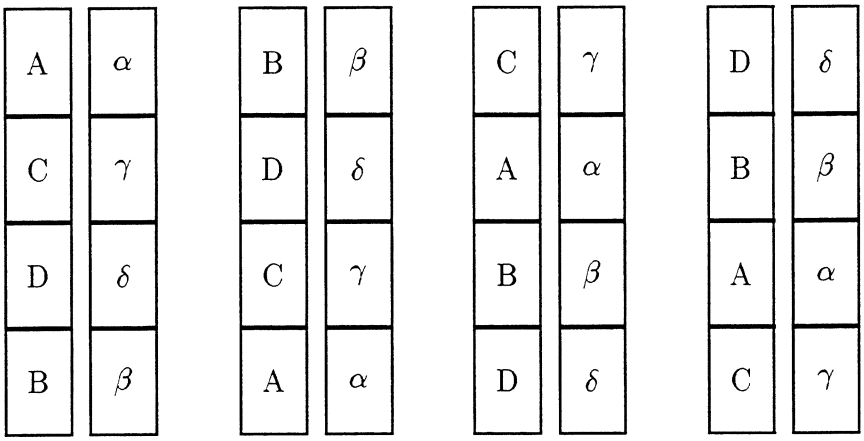

Figure 2: Inflated Latin square design to compare 8 varieties.

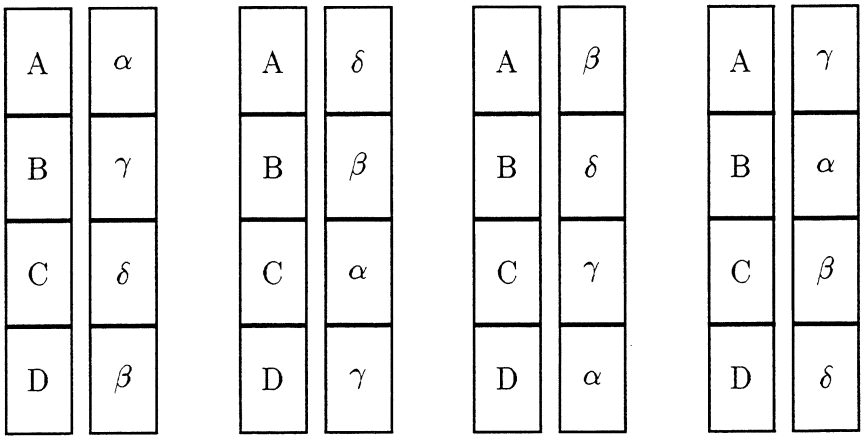

Figure 3: Partially balanced incomplete block design to compare 8 varieties.

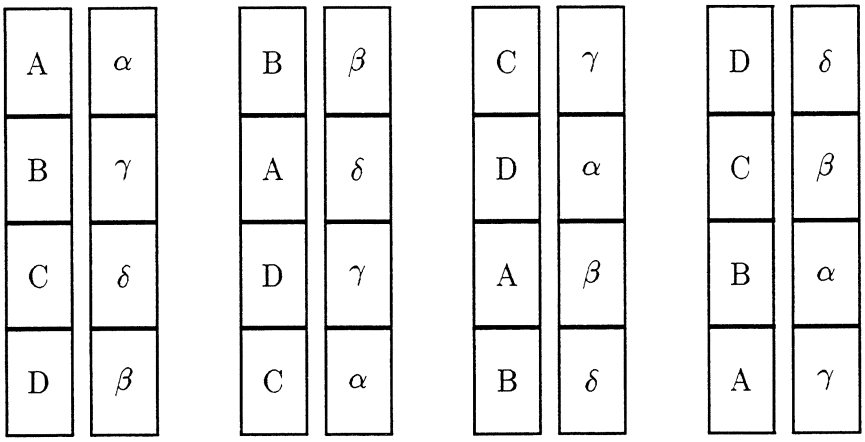

Figure 4: Trojan square design to compare 8 varieties.

Stratum	Source	df	Efficiency (%)
Row	Residual	3	
Column	Residual	3	
Row×Column	Letter	3	50
	Alphabet×Letter	3	50
	Residual	3	
Plot	Alphabet	1	100
	Letter	3	50
	Alphabet×Letter	3	50
	Residual	9	

Table 1: Outline analysis of variance for the Trojan square design in Figure 4.

	Design 1			Design 2		
Label	P	Q	R	P	Q	R
A	0	0	0	0	0	0
B	0	0	1	0	0	1
C	0	1	0	0	1	0
D	0	1	1	1	0	0
α	1	0	0	1	1	1
β	1	0	1	1	1	0
γ	1	1	0	1	0	1
δ	1	1	1	0	1	1

Table 2: Two allocations of 2^3 treatments to treatment labels.

	Alphabet	P	Q	R	PQ	PR	QR	PQR
A	-1	-1	-1	-1	1	1	1	-1
B	-1	-1	-1	1	1	-1	-1	1
C	-1	-1	1	-1	-1	1	-1	1
D	-1	-1	1	1	-1	-1	1	-1
α	1	1	-1	-1	-1	-1	1	1
β	1	1	-1	1	-1	1	-1	-1
γ	1	1	1	-1	1	-1	-1	-1
δ	1	1	1	1	1	1	1	1

Table 3: Coefficients of the contrasts of interest and the pseudo-factor contrast from Design 1

	Alphabet	P	Q	R	PQ	PR	QR	PQR
A	-1	-1	-1	-1	1	1	1	-1
B	-1	-1	-1	1	1	-1	-1	1
C	-1	-1	1	-1	-1	1	-1	1
D	-1	1	-1	-1	-1	-1	1	1
α	1	1	1	1	1	1	1	1
β	1	1	1	-1	1	-1	-1	-1
γ	1	1	-1	1	-1	1	-1	-1
δ	1	-1	1	1	-1	-1	1	-1

Table 4: Coefficients of the contrasts of interest and the pseudo-factor contrast from Design 2

	Efficiency (%)	
Effect	Design 1	Design 2
P	100	57.1
Q	50	57.1
R	50	57.1
PQ	50	50
PR	50	50
QR	50	50
PQR	50	57.1

Table 5: Efficiencies for the designs in Table 2.

Model	Design 1	Design 2
3rd order	L-, D_A-optimal	L-, D_A-, E_A-optimal
2nd order	L-, D_A-optimal	E_A-optimal
1st order	L-, D_A-optimal	E_A-optimal

Table 6: Optimality of the designs in Table 2.