COMMUN. STATIST.—THEORY METH., 27(6), 1289–1303 (1998)

A GRAPHICAL COMPARISON OF SUPERSATURATED DESIGNS

Sandy D. Balkin and Dennis K.J. Lin

The Pennsylvania State University
Department of Management Science and Information Systems
University Park, PA 16802

Key Words: boxplot, graphical comparison, Hadamard matrix, line plot

ABSTRACT

The search of construction methods for a "good" supersaturated design has received a great deal of attention in recent years. Many single-valued criteria have been proposed for comparison purposes. In this paper, we propose/review three criteria based on the information matrix of a projective design and provide two graphical methods for comparing different types of supersaturated designs: the line plot and the boxplot methods. Four different construction methods are compared using the proposed graphical techniques. These construction methods were proposed by Booth and Cox (1962), Lin (1993), Wu (1993) and Lin (1995). A thorough discussion on the proposed graphical techniques is given.

1. INTRODUCTION

A supersaturated design is a factorial design useful when the number of factors to be investigated are many and the experimental runs are expensive. Often in an experimentation, there are a large number of potentially relevant factors, but only a few are believed to have actual effects. Knowing the main effects of insignificant factors is usually not of interest and investigating these insignificant factors is a waste of resources. Supersaturated designs provide an approach for constructing an experimental design where the first order model is assumed and where the number of significant factors is small and the cost of an additional run is high.

In this paper we present a comparison of seven cases using four different supersaturated design construction methods proposed by Booth and Cox (1962), Lin (1993), Wu (1993) and Lin (1995), based on three criteria via tabular and graphical methods. Using these criteria, we can graphically identify and numerically support an optimal choice for design construction. In Section 2, we review the four different supersaturated designs. Section 3 presents the evaluation criteria which are compared via summary table and line plot in Section 4 and by boxplot in Section 5. The conclusions are then presented in Section 6.

2. FOUR SUPERSATURATED DESIGN CLASSES

Booth and Cox (1962) first suggest a systematic construction method for supersaturated design. Based on the $E(s^2)$ criterion, they provide seven supersaturated designs via computer search. These seven designs are for the cases:

$$(n,k) = (12,16), (12,20), (12,24), (18,24), (18,30), (18,36) \text{ and } (24,30),$$

where n is the number of experimental runs and k is the number of experimental factors.

Based on half-fraction Hadamard matrix, Lin (1993) proposes a simple construction method for a new class of supersaturated design. Besides the simplicity of construction, these designs are shown to exhibit nice mathematical properties. For this method, the value of n must be even. Wu (1993) makes use of the interaction columns of the Hadamard matrix in proposing another class of supersaturated design. Here, the value of n must be a multiple of four. To the best of our knowledge, these are the two most recent construction methods proposed in the literature since Booth and Cox's result. The first computer algorithm specifically for constructing supersaturated design is given by Lin (1995). This algorithm is based on the "exchange procedure," given the value of n and k (n can be even or odd).

Designs resulting from these four construction methods will be compared here. We next discuss the details of the construction methods, the criteria and the comparison methods.

A: **Booth and Cox (1962)** Booth and Cox first propose a systematic method for the construction of a supersaturated design based on computing the maximum cross-product between a trial set of vectors and new trial vectors with an equal number of +'s and -'s. The criterion is also known as $E(s^2) = \sum c_i c_j / \binom{k}{2}$ where c_i and c_j are design columns.

B: **Lin (1993)** Lin proposes a class of supersaturated designs constructed by using half fractions of Hadamard matrices, choosing a branching column from a given Hadamard matrix to split the whole matrix into two half fractions corresponding to those rows when the branching column has a value of + (or -). The resulting matrix has a dimension $(n, k) = (N/2, N-2)$ where N is the order of the Hadamard matrix.

C: **Wu (1993)** For a given Hadamard matrix of order N, one can generate all $\binom{N-1}{2}$ interaction columns. Combining the original matrix with all the interaction columns results in a supersaturated design of dimension $(n, k) = (N, N-1+\binom{N-1}{2})$.

D: **Lin (1995)** Using the exchange procedure, Lin's algorithm examines all possible candidate columns with equal occurrence to minimize two criteria: the maximal pairwise correlation (the primary criterion) and the $E(s^2)$ (the secondary criterion).

3. CRITERIA USED

To compare the different supersaturated designs, Booth and Cox (1962) propose a criterion analogous to a commonly used determinant criterion. As an extension of the classical optimal design, Wu (1993) proposes the D, A, and E criteria for the situations where there are more than two active factors. For a review of all the related criteria on supersaturated design, see Deng and Lin (1994). Note that the projection property plays a crucial role here.

For any $n \times k$ supersaturated design $\mathbf{X}$, define a $n \times p$ matrix X_p to be a projective design with its p columns chosen from the $\mathbf{X}$ matrix. There are $\binom{k}{p}$ such submatrices. Denote $\lambda_1, \ldots, \lambda_p$ as the eigenvalues of the information matrix $\frac{1}{n} X_p' X_p$.

Three criteria will be considered here: the $a-$value, $d-$value and $r-$value.

1. The $a-$value is defined as the harmonic mean of the λ_i's

$$a = \left[\frac{1}{p} \sum_{i=1}^{p} \lambda_i^{-1} \right]^{-1},$$

and $a = 0$ if any $\lambda_i = 0$. The $a-$value of a projective design is always between 0 and 1, with 1 as the most desirable case and 0 the least desirable. Note that the classical $A-$optimality is defined as

$$\text{trace} \left(\frac{1}{n} X_p' X_p \right)^{-1} = \sum_i \lambda_i^{-1}$$

which is equivalent to the $a-$value defined above when $\lambda_i \neq 0$.

2. The $d-$value is defined as the geometric mean of the λ_i's,

$$d = \left(\prod_{i=1}^{p} \lambda_i \right)^{1/p}.$$

Similar to the $a-$value, the $d-$value is always between 0 and 1, with 1 as the most desirable case and 0 the least desirable. Note that $0 \leq a - \text{value} \leq d - \text{value} \leq 1$. The $d-$value is probably the most popular criterion in optimal design literature.

3. The $r-$value is defined as the absolute correlation coefficient between every pair of design columns c_i and c_j,

$$r - \text{value} = |\text{corr}(c_i, c_j)| = \left| \frac{c_i' c_j}{n} \right| \quad i \neq j.$$

The average of the $r_{ij}s$ is essentially the popular $E(s^2)$ criterion proposed by Booth and Cox (1962). Unlike the a and $d-$values, small $r-$values are more desirable in that $r - \text{value} = 0$ implies perfect orthogonality.

The geometric and harmonic means are used because they are insensitive to outliers in the sense that neither a large nor a small (though non-zero) eigenvalue will affect the measure of center as much as with the arithmetic mean. In the case of an orthogonal design, the resulting information matrix will have the number of columns on the diagonal and zero elsewhere.

4. SUMMARY TABLE AND LINE-PLOT

The goal of this investigation is to determine which method of supersaturated design construction is best in any particular situation. To evaluate each design method, we consider each of the seven cases given in Booth and Cox (1962) for

$$(n, k) = (12, 16), (12, 20), (12, 24), (18, 24), (18, 30), (18, 36) \text{ and } (24, 30)$$

and evaluate the criteria for the set of test vectors with columns $p = \{2, 3, 4, 5\}$. Note that Lin (1993) is not available for $(n, k) = (18, 36)$ and Wu (1993) is not available for $n = 18$ so no complete comparison studies can be made for these cases. Also note that for each situation, there are $\binom{k}{p}$ projective matrices to be evaluated. For example, when $p = 5$ and $(n, k) = (12, 20)$, there are $\binom{20}{5} = 15,504$ X_p matrices. Recently, Balkin and Lin (1998) have obtained the means and standard deviations of these seven cases. Their results for Case 1 are reproduced in Table I.

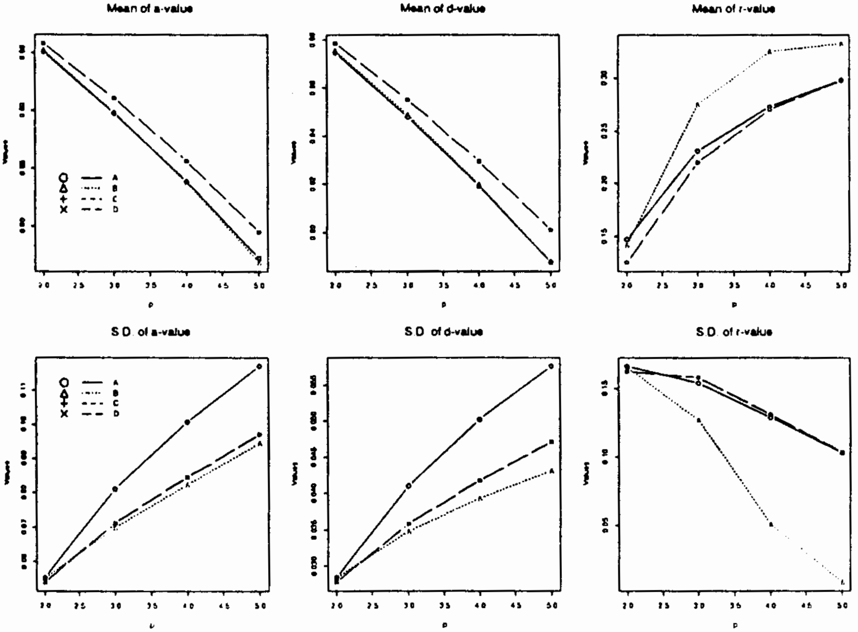
Due to the huge number of projective designs, only the summary statistics, such as the mean, of each criterion are reported in the literature. However, as we can see from the table, in many situations, the means can be very close to each other. Here, the standard deviation can be useful as a secondary criterion in separating one case from another. It is not easy to discern which of the methods is the best from such tables. To combat this difficulty, we create line plots of the means and

TABLE I

Means and Standard Deviations for CASE 1; n=12; k=16

STAT		a				d				r			
DESIGN		A	B	C	D	A	B	C	D	A	B	C	D
p = 2	$\binom{k}{p}=120$	0.9509	0.9528	0.9583	0.9583	0.9747	0.9757	0.9786	0.9786	0.1472	0.1417	0.125	0.125
		0.0554	0.0552	0.054	0.054	0.0285	0.0284	0.0278	0.0278	0.1662	0.1655	0.1621	0.1621
p = 3	$\binom{k}{p}=560$	0.8976	0.8986	0.9106	0.9106	0.9479	0.949	0.9551	0.9551	0.231	0.275	0.2202	0.2202
		0.0811	0.0699	0.0712	0.0712	0.041	0.0348	0.0358	0.0358	0.1539	0.1268	0.158	0.158
p = 4	$\binom{k}{p}=1820$	0.8385	0.8375	0.8565	0.8565	0.9191	0.9198	0.9295	0.9295	0.2729	0.3255	0.2701	0.2701
		0.1007	0.0824	0.0845	0.0845	0.0502	0.0393	0.0417	0.0417	0.1285	0.0506	0.1307	0.1307
p = 5	$\binom{k}{p}=4368$	0.7722	0.7683	0.7948	0.7948	0.8878	0.8875	0.9011	0.9011	0.2981	0.3331	0.2976	0.2976
		0.1171	0.0945	0.0971	0.0971	0.0576	0.0431	0.0471	0.0471	0.1025	0.0087	0.1031	0.1031

(a) CASE 1: $(n, k) = (12, 16)$



(b) CASE 2: $(n, k) = (12, 20)$

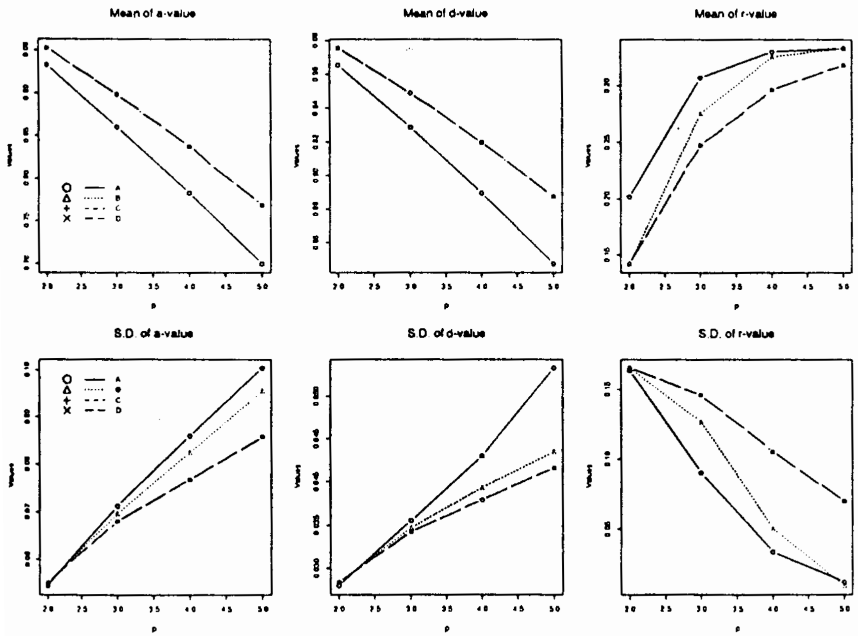


FIGURE I

Line plots for the means and the standard deviations of criteria values

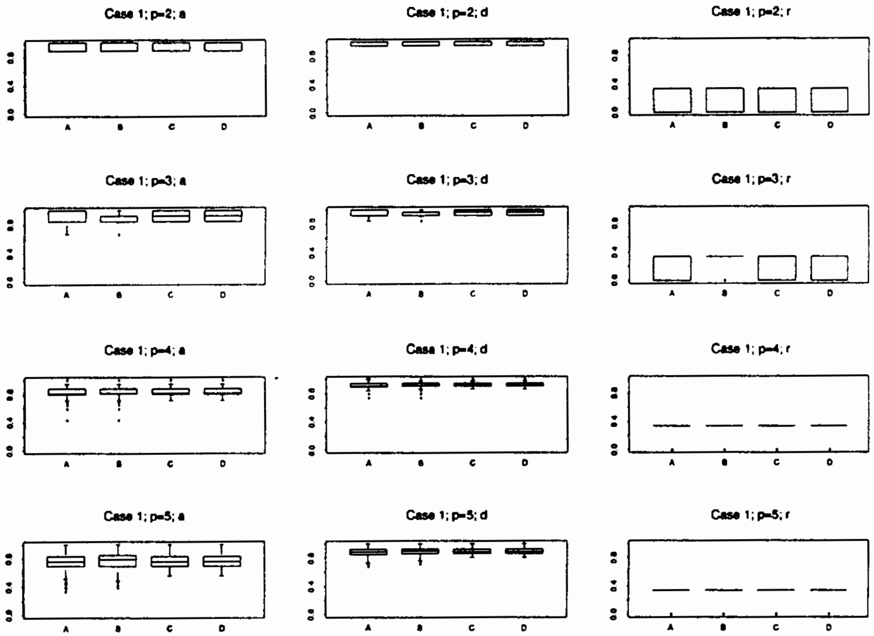


FIGURE II

Boxplots for CASE 1 of criteria values for each design and dimension

standard deviations of the criteria values for each of the four designs over the values of p. This allows us to quickly see how the design techniques behave relative to each other.

For example, we can see from Figure I where the means and standard deviations for Cases 1 and 2 are plotted, that Design A performs worse with regards to the a and d-values. When the means of two designs are similar, one may use their standard deviation as a secondary criterion for comparison. For example, Designs B and C have a close mean in Case 2, but Design C performs better in the sense that it has a smaller standard deviation. The standard deviations indicate the "stability" of the projective designs. A small standard deviation implies that the projection results are more robust to the final columns chosen.

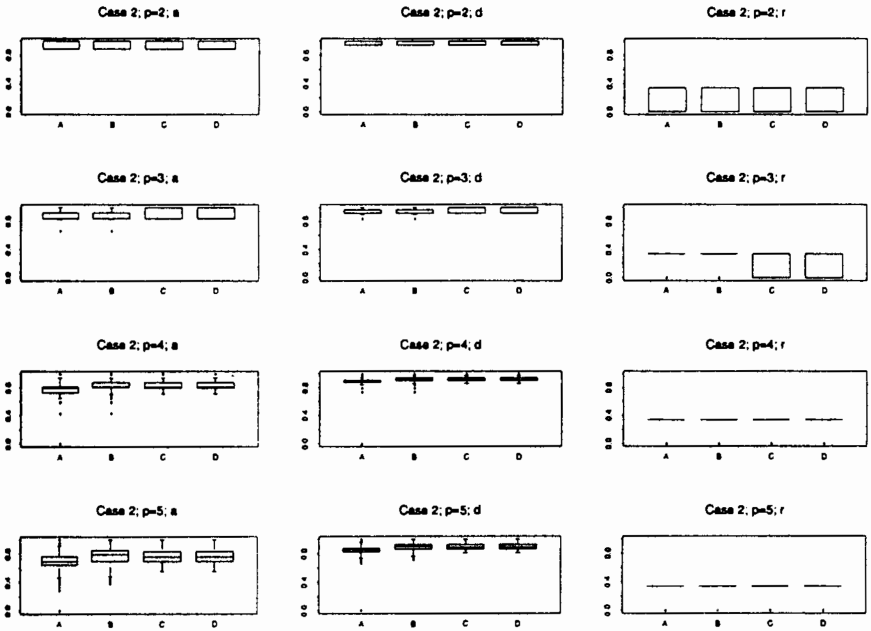


FIGURE III

Boxplots for CASE 2 of criteria values for each design and dimension

Based on the standard deviations plots (see Balkin and Lin (1998) for complete set), Design B out performs the others overall (i.e., Design B is most stable). Design A has the largest standard deviation in most situations.

However, we still lack information concerning the distribution of the criteria values over all the test vector combinations. In addition, we know that the mean and standard deviation are point estimate summary statistics that are not robust to outliers. This may cause problems if some of the test vector combinations' crossproducts are singular, resulting in eigenvalue(s) equal to 0. To learn more about the distribution of the criteria, a better approach is to use boxplots.

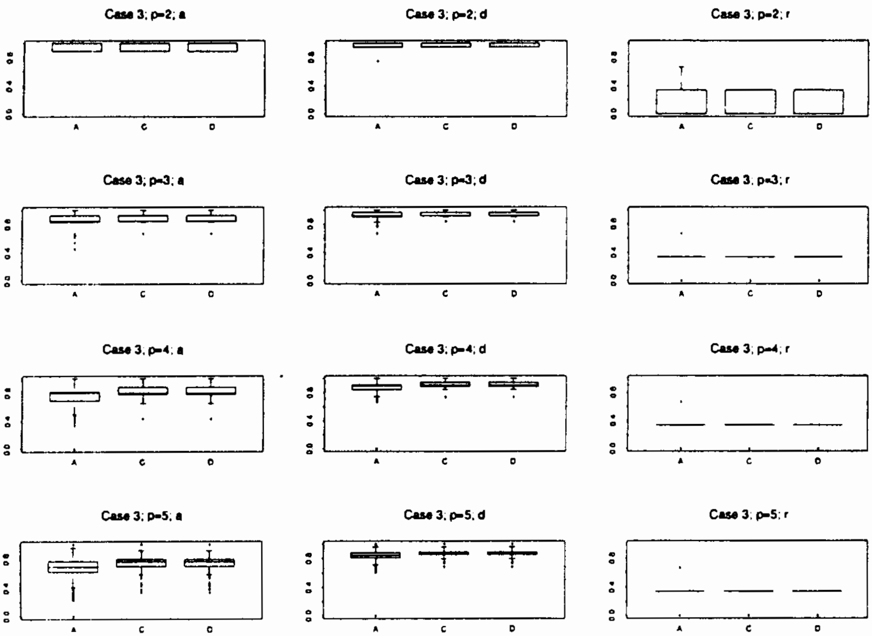


FIGURE IV

Boxplots for CASE 3 of criteria values for each design and dimension

5. BOXPLOTS AND FURTHER DISCUSSION

In order to get an idea of the distribution of the criteria values, we create modified boxplots of the criteria values for each case and each design. A modified boxplot is created by drawing lines at the first, second and third quartiles denoted by $Q1$, $Q2$ and $Q3$, completing the box by connecting the first and third quartiles, and extending whiskers to the smallest criterion value greater than $Q1 - 1.5 \times IQR$ and to the greatest criterion value less than $Q3 + 1.5 \times IQR$ where IQR is the inter-quartile range. Values lying outside of the whisker limits are deemed possible outliers and denoted by a dot. The statistics used in creating a modified boxplot are known to be robust to outliers.

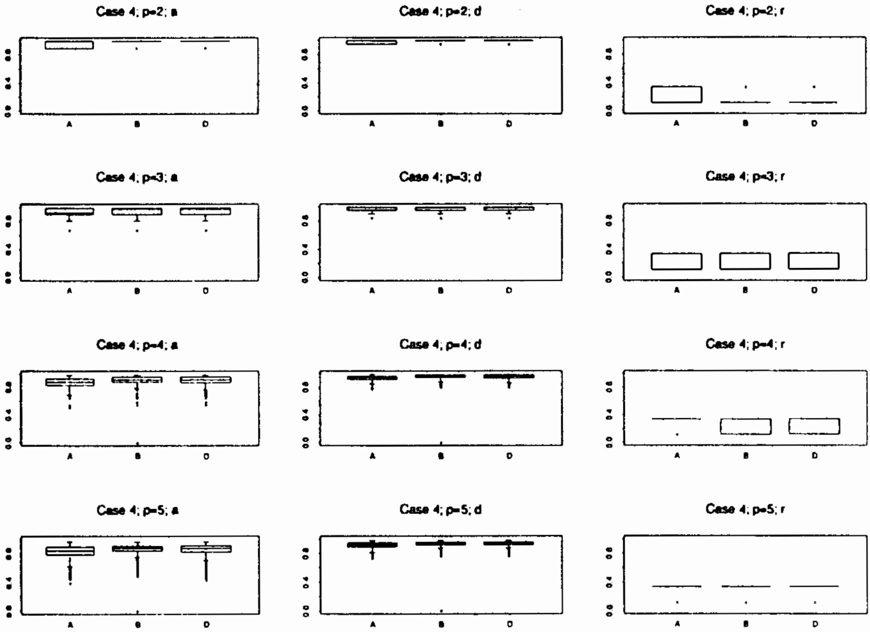


FIGURE V

Boxplots for CASE 4 of criteria values for each design and dimension

From these plots, we get an idea about the distribution of the criteria. For instance, we can see if criterion values are symmetric or skewed, concentrated about a single value or spread out, and most importantly, if there are any combinations of test vectors whose crossproduct is singular. This last fact is an important piece of information hidden by the summary table.

Figures II-VIII display the boxplots for Cases 1-7 respectively. The distributions of all three criteria are given. The conclusion from the line plots can be confirmed by the boxplots, such as the highest/lowest mean and standard deviation, etc. There are other important messages that can not be found from the line plot. The boxplots for Cases 6 and 7 when $p = 5$ do not appear because they are too computationally intensive at this point of time. For example, the data file for Case 6 when $p = 5$ is

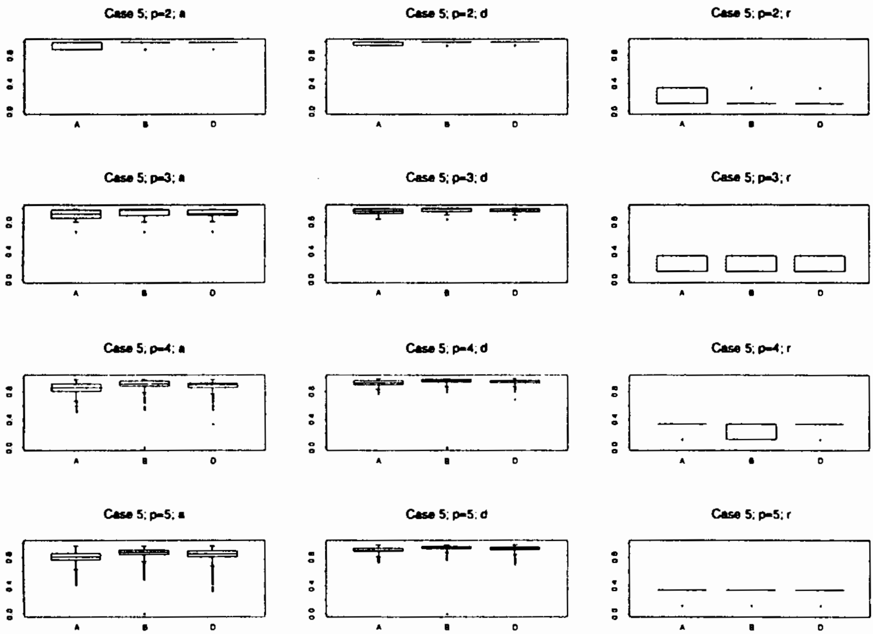


FIGURE VI

Boxplots for CASE 5 of criteria values for each design and dimension

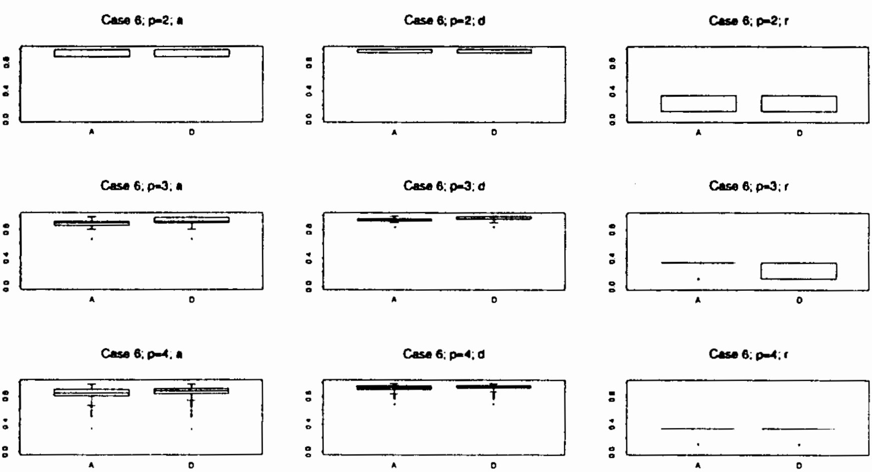


FIGURE VII

Boxplots for CASE 6 of criteria values for each design and dimension

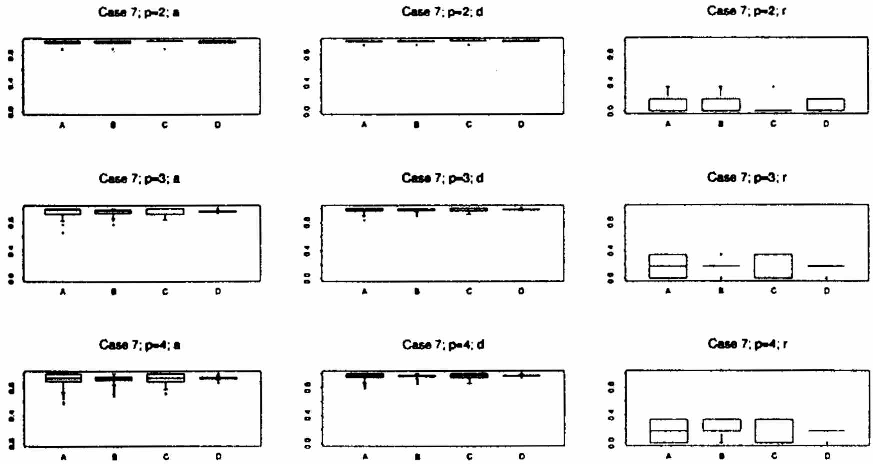


FIGURE VIII

Boxplots for CASE 7 of criteria values for each design and dimension

63 megabytes. To perform the necessary calculations to produce the boxplots would require more computer power than we currently have access to.

For an analysis example, from Figure III, Design A has some singular projections for p=4 and 5 (or its resolution rank is only 3, see Deng, Lin and Wang (1994). This is not desirable at all, but cannot be seen from the summary tables or the line plots. Similarly, Design A has a high r-value for Case 3 (see Figure IV), although its $E(s^2)$ (or equivalently, the average r-value) is reasonably low. It is also interesting to note that Design B has a higher median in most of cases (even though its means are not particularly high). Such boxplots provide a better comparison and can detect many important characteristics.

6. CONCLUSION

For the supersaturated designs tested in this study, no single-valued criterion is capable of characterizing their properties. If it is to be used, the line plot will provide a better visual comparison than the summary table. When the projection properties are important, as in supersaturated design context, the boxplot can provide more information concerning the distribution of the criteria over the set of test vectors, than the summary table.

A new criterion based on the harmonic mean of eigenvalues, called the a-value, similar to the traditional $A-$criterion, is proposed here. Accompanied with the well known a-value and r-value, the projection properties of a supersaturated design can be revealed. Note that all of these three criteria have a value between 0 and 1 for a fair comparison. For the designs tested in this study, Design A seems to perform worse than others, in terms of these three criteria.

ACKNOWLEDGEMENTS

The authors wish to thank co-editor Pi-Erh Lin and two anonymous referees for their valuable comments which considerably added to the quality of this manuscript.

This research was partially supported by the National Science Foundation via Grant DMS-9704711.

BIBLIOGRAPHY

Balkin, S. D. and Lin, D. K. J. (1998), "A Graphical Comparison of Supersaturated Designs," The Pennsylvania State University, Department of Management Science and Information Systems Working Paper Series, No. 98-01.

Booth, K.H.V. and Cox, D.R. (1962), "Some Systematic Supersaturated Designs," *Technometrics*, **4**, 489-495.

Deng, L.Y. and Lin, Dennis K.J. (1994), "Criteria for Supersaturated Designs," *Proceedings of the Section on Physical and Engineering Sciences*, American Statistical Association, 124-128.

Deng, L.Y., Lin, Dennis K.J. and Wang, J.N. (1994), "Supersaturated Design Using Hadamard Matrix," *IBM Research Report*, **RC19470**, IBM Watson Research Center.

Lin, Dennis K.J. (1991), "Systematic Supersaturated Designs," Working Paper, **No. 264**, College of Business Administration, University of Tennessee.

Lin, D. K. J. (1993), "A New Class of Supersaturated Designs," *Technometrics*, **35** 28-31.

Lin, D. K. J. (1995), "Generating Systematic Supersaturated Designs," *Technometrics*, **37**, 213-225.

Plackett, R. L. and Burman, J. P. (1946), "The Design of Optimum Multifactorial Experiments," *Biometrika*, **33** 303-325.

Wu, C. F. J. (1993), "Construction of Supersaturated Designs Through Partially Aliased Interactions," *Biometrika*, **80**, 661-669.