ELSEVIER

Journal of Statistical Planning and Inference 72 (1998) 99–107

journal of statistical planning and inference

On the identifiability of a supersaturated design

Jiahua Chen [a,*], Dennis K.J. Lin [b]

[a] *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

[b] *Department of Management Science and Information Systems Penn State University, University Park, PA 16802, USA*

Abstract

A supersaturated design investigates k factors in only n ($<k+1$) experimental runs. The goal for such a design is to identify, presumably only a few, relatively dominant effects with a cost as low as possible. While the construction of supersaturated designs has been widely explored, the data analysis aspect of such designs remains primitive. We study the following problem: How many dominant effects are allowed to make a meaningful data analysis possible for a supersaturated design with the maximum correlation ρ? The correlation here is defined as the cosine of the angle between two column vectors. The obtained results support the fundamental concern of the $E(s^2)$ criterion introduced by Booth and Cox (1962). Furthermore, under the normality assumption, we obtained a lower bound of the probability that the factor with the largest estimated effect has, indeed, the largest true effect. This bound depends on the relative size of the largest effect and the maximum correlation of the underlying design. Under some mild assumptions, we show that this probability is satisfactorily large. Consequently, by carefully constructing supersaturated designs, we not only save the cost of the experiment, but also make reliable inferences.

Keywords: Alias; Bias; Correlation; Estimability; Projection

1. Introduction

Many preliminary industrial screening experiments typically contain a large number of potentially relevant factors. Among them, only a few are believed to be active. The goal in such situations is to identify those (relatively) few dominant active factors with the least possible number of experimental runs. A (two-level) supersaturated design is a matrix of ± 1's with n rows and a large number of columns k (larger than n in general). Hence, it studies a large number of factors (k) with only a few runs (n).

First constructed systematically by Booth and Cox (1962), supersaturated designs have received a great deal of attention in the recent literature (see, Lin, 1991, 1993,

* Corresponding author.

0378-3758/98/$ – see front matter
PII: S0378-3758(98)00025-1

1995; Wu, 1993; Tang and Wu, 1993; Seewald, 1994; Nguyen, 1996; Deng et al., 1996; and Westfall et al., 1998). While the construction of supersaturated designs has been widely explored, the inferential aspect of such designs needs more investigation. This is the major interest of this paper.

Naturally, such a design will not allow us to estimate main effects of all k factors. This is, however, not always needed if we believe that only a small number (p) of them are active. The underlying requirement is that if there are only a few active factors, we should be able to identify them. Once these active factors are correctly identified, the design can then be projected onto a lower dimensional space. The resulting design is then an unsaturated design and ordinary data analysis can then be applied.

To assure that all active factors can be properly estimated after the projection, we must carefully select the supersaturated design. Suppose, for example, it is known that there are at most four active effects among all the factors. To assure identifiability, we need to select a supersaturated design such that any four columns of the design are linearly independent. We shall show that the correlations between the columns of the design are tools to guarantee linear independence. This is, in fact, the fundamental concern in Booth and Cox (1962)'s $E(s^2)$ criterion.

Furthermore, in order to project the design onto the set of active factors, it is necessary to identify those active factors correctly. Due to random noise and the partial aliasing structure of the design, we may not always select the correct set of active factors. However, we shall show that if the effects of the active factors are reasonably large as compared with inert factors and random noise, the probability of obtaining the correct set of active factors is satisfactorily high.

Our result here is different from the important work of *search design* in Srivastava (1975), who first showed that a necessary and sufficient condition for having *resolving power p* (i.e., identifying and estimating all p active factors) is that every $2p$ columns of the design have to be linearly independent. While his results guarantee the search of all active factors, we are interested in to what extent we can correctly identify *active* factors in the sense that when the factor with maximal effect is given, we calculate the probability that such a factor indeed has the largest estimated effect.

2. Estimability and correlation structure

Consider a supersaturated design in n experimental runs to investigate k ($\geqslant n$) factors. If $\boldsymbol{X}$ denotes the $n \times k$ design matrix (without intercept column), our model is

$$Y = \mu\mathbf{1} + \boldsymbol{X}\beta + \varepsilon,$$

where $\boldsymbol{Y}$ is the $(n \times 1)$ observable data vector; μ is the (scale) intercept term; $\mathbf{1}$ is an n-vector of 1's; β is a $(k \times 1)$ fixed parameter vector for the unknown factor effects; and ε is the noise vector assumed to be distributed as $\mathrm{N}(0, \sigma^2\boldsymbol{I}_n)$. Because k is larger than n, it is clear that the $\boldsymbol{X}$ matrix cannot be of full rank and orthogonality is only

possible for certain pairs of design columns. Here and hereafter, we say a matrix is of full rank when all of its column vectors are linearly independent.

Note that once the active factors are identified, the whole design $\boldsymbol{X}$ is then projected onto a much lower dimension. Hence, the estimability of the effects of these factors depends on whether or not the projected design has full rank. We will show that the largest number of active factors which can be identified from a supersaturated design depends on the correlations between columns of $\boldsymbol{X}$. Let $\boldsymbol{A}=\{i_1,i_2,\ldots,i_p\}$ and $\boldsymbol{N}=\{i_{p+1},i_{p+2},\ldots,i_k\}$ denote indexes of active and inert factors, respectively, so that $\boldsymbol{A}\cup\boldsymbol{N}=\{1,2,\ldots,k\}$. Also, denote the projective design matrix as $\boldsymbol{X}_p$ and the columns of $\boldsymbol{X}$ as ξ_i, $i=1,2,\ldots,k$. For a two-level equal occurrence design, each design column contains an equal number of high- and low-level experimental values (denoted by + and −, respectively). Consequently, we can define $\mathrm{Corr}(\xi_i,\xi_j)=\xi_i'\xi_j/n$ for any $1\leqslant i,j\leqslant k$ (because $\sum_k \xi_{ik}=0$ and $\sum_k \xi_{ik}^2=n$); thus, we have:

Theorem 1. *If* $|\mathrm{Corr}(\xi_i,\xi_j)|<\rho=1/(p-1)$ *for all* $i\neq j$, *then* $\boldsymbol{X}_p$ *is of full rank.*

Proof. Write $\boldsymbol{X}_p'\boldsymbol{X}_p=(x_{ij})$ with $x_{ij}=n\mathrm{Corr}(\xi_i,\xi_j)$. Hence, (i) $x_{ii}=n$ and (ii) $|x_{ij}|<\rho n$, for all $i\neq j$. Therefore $x_{ii}>\sum_{j\neq i}|x_{ij}|$. This implies that $\boldsymbol{X}_p'\boldsymbol{X}_p$ is positive definite. Hence, X_p has full column rank. □

Note that if $X=(\xi_1,\ldots,\xi_k)$ is a supersaturated design, then $X=(\pm\xi_{\tau(1)},\ldots,\pm\xi_{\tau(k)})$ for any choice of ± signs and permutation function τ is an equivalent design. We will not distinguish equivalent designs in this paper.

When $\max|\mathrm{Corr}(\xi_i,\xi_j)|=\rho 1/(p-1)$ for $i\neq j$, there is no definite answer to this problem. A simple counterexample is the supersaturated design

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

Let $p=2$, we have $|\mathrm{Corr}(\xi_1,\xi_2)|=\rho=1/(p-1)$. However, $X_2=X$ is singular.

Despite this counterexample, X_p of most useful supersaturated designs has full rank when $|\mathrm{Corr}(\xi_i,\xi_j)|\leqslant 1/(p-1)$. For the 12-run supersaturated designs given by Lin (1991, 1993, 1995) and Wu (1993), it can be verified that any submatrices consisting of four columns have full rank.

Lemma 1. *Let* $X=(\xi_1,\ldots,\xi_k)$ *of size* $n\times k$ *be a supersaturated design with entries* ± 1. *If* $|\mathrm{Corr}(\xi_i,\xi_j)|\leqslant\rho=1/(p-1)$ *for all* $i\neq j$ *and a submatrix* $X_p=(\xi_1,\ldots,\xi_p)$ *is singular, then there is an equivalent submatrix of* X_p *such that* $X_p^{\mathrm{t}}X_p/n=(1+\rho)I-\rho\mathbf{1}\mathbf{1}^{\mathrm{t}}$.

Proof. If X_p is singular and $|\mathrm{Corr}(\xi_i,\xi_j)|\leqslant\rho=1/(p-1)$ for all $i\neq j$, we must be able to find a ξ_i, say ξ_1, such that $|\mathrm{Corr}(\xi_1,\xi_j)|=\rho$ for all $j\neq 1$, or else X_p is not singular by using the same proof of Theorem 1. Clearly, there is an equivalent design such that

$\mathrm{Corr}(\xi_1, \xi_j) = -\rho$ for all $j \neq 1$. So

$$X_p^t X_p/n = [\rho_{ij}] = \begin{bmatrix} 1 & -\rho \mathbf{1}^t \\ -\rho \mathbf{1} & W \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\rho \mathbf{1} & W - \rho^2 \mathbf{1}\mathbf{1}^t \end{bmatrix} \begin{bmatrix} 1 & -\rho \mathbf{1}^t \\ 0 & I \end{bmatrix}.$$

Now, we make an induction assumption that X_p is singular only if $\mathrm{Corr}(\xi_i, \xi_j) = \pm 1/(p-1)$ for $1 \leqslant i \neq j \leqslant p$. This assumption is obviously true for $p = 2$. Let us assume that it is also true for $p-1$. Then, the above X_p is singular if and only if $W - \rho^2 \mathbf{1}\mathbf{1}^t$ is singular. Note the order of $W - \rho^2 \mathbf{1}\mathbf{1}^t$ is $(p-1) \times (p-1)$. By induction assumption, we have, for $i \neq j$,

$$(\rho_{ij} - \rho^2)/(1 - \rho^2) = \pm 1/(p-2).$$

Solving this equation, we get $\rho_{ij} = -1/(p-1)$. □

Thus, for a 12-run supersaturated design with the maximum absolute correlation being $\frac{1}{3}$, if the submatrix $X_4 = (\xi_1, \ldots, \xi_4)$ is singular, then $\sum_{i=1}^4 \xi_i = 0$ for an equivalent X_4. Clearly, designs in Lin (1991, 1993, 1995) and Wu (1993) do not contain such structures. Therefore, the maximum estimable number of factors is at least four.

For a 20-run supersaturated design with maximum correlation $\frac{1}{4}$, if the submatrix $X_5 = (\xi_1, \ldots, \xi_5)$ is singular, then $\sum_{i=1}^5 \xi_i = 0$ after choosing a proper equivalent design. However, this is impossible since the summation of five ± 1's cannot be zero. We can summarize this example into the following theorem.

Theorem 2. *If* $|\mathrm{Corr}(\xi_i, \xi_j)| \leqslant \rho = 1/(p-1)$ *for all* $i \neq j$, *then* X_p *has full rank when* p *is odd.*

Theorems 1 and 2 ensure that if there are at most p active factors, when using a supersaturated design with maximum correlation less than or equal to $1/(p-1)$ for odd p, it is always possible to estimate all of them once their corresponding columns are identified. An important question that needs to be addressed is thus: assuming there are only p active factors, what is the probability these p factors have the largest estimated effects? The answer to this question depends on the relative size of their effects and also partially on the estimation method. We will discuss this problem in the next section.

3. Identifiability

The conventional point estimation for the β_i's is (see, Box, 1959, p. 177)

$$\hat{\beta}_i = (\bar{y}_{(i)}^+ - \bar{y}_{(i)}^-)/2 = x_i^t Y/n, \tag{1}$$

where $\bar{y}_{(i)}^+$ and $\bar{y}_{(i)}^-$ are the averages of the responses when factor x_i is used at its high-(+) and low-(−) levels, respectively. Consequently, we have $E(\hat{\beta}_i) = \beta_i + \sum_{j \neq i} \rho_{ij} \beta_j$, where ρ_{ij} is the correlation between columns ξ_i and ξ_j.

Suppose β_m is the maximal effect among all β_i's. We are interested in the probability that the estimate $\hat{\beta}_m$, of β_m, will be the largest among all estimates. To compute $\phi_m = \text{Prob}[\hat{\beta}_m > \max_{i \neq m} \hat{\beta}_i]$, the following two lemmas are needed.

Lemma 2. *Assume $Y_1, \ldots, Y_n$ are independent and normally distributed with the same variance and the design satisfies the condition that $\rho \leqslant 1/3$. Let $\hat{\beta}_i$, $i = 1, \ldots, k$ be defined as in Eq. (1); $\phi_{ma} = \text{Prob}[\hat{\beta}_m > \hat{\beta}_a]$; and $\phi_{mb} = \text{Prob}[\hat{\beta}_m > \hat{\beta}_b]$. Then*

$$\text{Prob}[\hat{\beta}_m > \max(\hat{\beta}_a, \hat{\beta}_b)] \geqslant \phi_{ma} \cdot \phi_{mb}.$$

Proof. Note that

$$\text{Prob}[\hat{\beta}_m > \max(\hat{\beta}_a, \hat{\beta}_b)] = \text{Prob}[\hat{\beta}_m > \hat{\beta}_a | \hat{\beta}_m > \hat{\beta}_b] \text{Prob}[\hat{\beta}_m > \hat{\beta}_b].$$

Hence, we need only to show

$$\text{Prob}[\hat{\beta}_m > \hat{\beta}_a | \hat{\beta}_m > \hat{\beta}_b] \geqslant \text{Prob}[\hat{\beta}_m > \hat{\beta}_a].$$

Let $Z_1 = \hat{\beta}_m - \hat{\beta}_a$ and $Z_2 = \hat{\beta}_m - \hat{\beta}_b$. Thus,

$$\text{Cov}(Z_1, Z_2) = \text{Var}(\hat{\beta}_m) - \text{Cov}(\hat{\beta}_m, \hat{\beta}_a) - \text{Cov}(\hat{\beta}_m, \hat{\beta}_b) + \text{Cov}(\hat{\beta}_a, \hat{\beta}_b) \geqslant 0.$$

This is because the absolute correlation between any two estimates is no larger than $\frac{1}{3}$. Now, from the joint normality of Z_1 and Z_2, we can write $Z_1 = aZ_2 + Z_3$ such that $a > 0$ and Z_3 is independent of Z_2. Also,

$$I(aZ_2 + Z_3 > 0) \geqslant I(Z_3 > 0)$$

at any sample point such that $Z_2 > 0$, where $I(\cdot)$ is the indicator random variable. Thus, by taking expectation conditional on $Z_2 > 0$ on both sizes,

$$\text{Prob}(Z_1 > 0 \mid Z_2 > 0) \geqslant \text{Prob}(Z_3 > 0 \mid Z_2 > 0) = \text{Prob}(Z_3 > 0).$$

Similarly, we have

$$I(aZ_2 + Z_3 > 0) \leqslant I(Z_3 > 0)$$

at any sample point such that $Z_2 \leqslant 0$. Thus, by taking expectation conditional on $Z_2 \leqslant 0$ on both sizes,

$$\text{Prob}(Z_1 > 0 \mid Z_2 \leqslant 0) \leqslant \text{Prob}(Z_3 > 0 \mid Z_2 \leqslant 0) = \text{Prob}(Z_3 > 0).$$

Hence,

$$\text{Prob}(Z_1 > 0 \mid Z_2 \leqslant 0) \leqslant \text{Prob}(Z_3 > 0) \leqslant \text{Prob}(Z_1 > 0 \mid Z_2 > 0).$$

This implies

$$\begin{aligned} \text{Prob}(Z_1 > 0) &= \text{Prob}(Z_1 > 0 \mid Z_2 > 0)\text{Prob}(Z_2 > 0) \\ &\quad + \text{Prob}(Z_1 > 0 \mid Z_2 \leqslant 0)\text{Prob}(Z_2 \leqslant 0) \\ &\leqslant \text{Prob}(Z_1 > 0 \mid Z_2 > 0) \end{aligned}$$

as required. □

Following Lemma 2, by induction we have $\phi_m \geqslant \prod_{i\neq m, i=1}^{k} \phi_{mi}$ where $\phi_{mi} = \text{Prob}[\hat{\beta}_m \geqslant \hat{\beta}_i]$. The induction can be done by showing

$$\text{Prob}(Z_1 > 0, Z_2 > 0, \ldots, Z_n > 0 \mid Z_{n+1} > 0) \geqslant \text{Prob}(Z_1 > 0, Z_2 > 0, \ldots, Z_n > 0)$$

in exactly the same way as is done in Lemma 2 by using properly defined normal random variables $Z_1, \ldots, Z_n$ and where Z_{n+1} is jointly normal and positively related to the other random variables.

Lemma 3. *Given m and i,*

$$\phi_{mi} \geqslant \Phi\left(\sqrt{\frac{n}{2(1-\rho_{mi})}} \cdot \delta_{mi}\right),$$

where $\delta_{mi} = (E\hat{\beta}_m - E\hat{\beta}_i)/\sigma$ *and* Φ *is the standard normal cumulative distribution function.*

Proof. First note that

$$\hat{\beta}_m - \hat{\beta}_i \sim \text{N}\left(E(\hat{\beta}_m - \hat{\beta}_i), \frac{2(1-\rho_{mi})}{n}\sigma\right).$$

Thus,

$$\begin{aligned}\text{Prob}[\hat{\beta}_m - \hat{\beta}_i \geqslant 0] &= P\left[\frac{(\hat{\beta}_m - \hat{\beta}_i) - E(\hat{\beta}_m - \hat{\beta}_i)}{\sqrt{\frac{2}{n}(1-\rho_{mi})}\sigma} \geqslant \frac{-E(\hat{\beta}_m - \hat{\beta}_i)}{\sqrt{\frac{2}{n}(1-\rho_{mi})}\,\sigma}\right] \\ &= \Phi\left[\frac{\sqrt{n}E(\hat{\beta}_m - \hat{\beta}_i)}{\sqrt{2(1-\rho_{mi})}\sigma}\right] = \Phi\left(\sqrt{\frac{n}{2(1-\rho_{mi})}} \cdot \delta_{mi}\right). \qquad \square\end{aligned}$$

Note that whether $\hat{\beta}_m$ is larger than $\hat{\beta}_i$ depends on several factors. One important factor is the value of ρ_{mi}. A positive ρ_{mi} makes $n/[2(1-\rho_{mi})]$ larger and hence improves the chance to identify β_m. On the other hand, a positive ρ_{mi} also brings $E\hat{\beta}_i$ closer to $E\hat{\beta}_m$ which makes δ_{mi} smaller and hence reduces the chance to identify β_m.

It is straightforward then from Lemmas 2 and 3 that:

Theorem 3. *The probability that* $\hat{\beta}_m$ *is the largest estimated effect is*

$$\phi_m \geqslant \prod_{i\neq m, i=1}^{k} \Phi\left(\sqrt{\frac{n}{2(1-\rho_{mi})}} \cdot \delta_{mi}\right).$$

The lower bounds given in Theorem 3 can be calculated directly for any specific design.

4. Simulation results

Table 1 shows some simulation results based on the supersaturated designs constructed by Lin (1991, 1993 and 1995) and Wu (1993). For each case, given the design and the number of factors, the simulations were conducted in the following way:

(1) Randomly select a number m from 1 to k. Let all $\beta_j = 0$ when $j \neq m$, and $\beta_m =$ either 1 or 2.

(2) Generate n of ε_i's from N(0, 1) to construct the responses $y_i = \sum_{u=1}^{k} x_{iu}\beta_u + \varepsilon_i$, $i = 1, 2, \ldots, n$.

(3) Obtain $\hat{\beta}_j$ by Eq. (1) for all j and record whether $\hat{\beta}_m$ is indeed the maximum of the $\hat{\beta}_i$'s.

(4) Repeat (1)–(3) 5000 times.

Note that the supersaturated designs constructed by half-fraction Hadamard matrices (Lin, 1993) can only examine $k = N - 2$ factors in $n = N/2$ runs, while the 12-run supersaturated design using interaction columns (Lin, 1991 and 1995; Wu, 1993) can study as many as 66 factors. The case $k = 10$ is not supersaturated, but is a reference benchmark to be compared with the performance of supersaturated designs. It is clear that in all the cases the lower bounds are satisfactorily large.

Supersaturated designs with $|\rho_{ij}| \leqslant \frac{1}{3}$ were recommended by Lin (1995). In fact, all designs discussed in Table 1 have such a property. In this case, we can extend Theorem 3 to:

Corollary 1. *If* $|\rho_{ij}| \leqslant \frac{1}{3}$, *then* $\phi_m \geqslant \prod_{i \neq m, i=1}^{k} \Phi\left(\sqrt{\frac{3n}{8}} \cdot \delta_{mi}\right)$.

Table 2 shows lower bound probabilities as given in Corollary 1 for various combinations of (n, k, δ_{mi}). Note that these probabilities do not depend on the design. Also, the probabilities given here, as expected, are smaller than the probabilities given in Table 1. In general, if $\delta_{mi} \geqslant 2$, the largest effect can always be correctly identified, a similar observation made by Lin (1995).

It is clear from Theorem 3 that if $\delta_{mi} < 1$, the lower bound probability given in Corollary 1 will result in a small probability. This is the case where the overall variation is larger than the factor effects. Supersaturated designs do not perform well in this case. In fact, most (unsaturated) designs have difficulty in this situation in identifying active factors. The top priority in this case should be variation reduction rather than factor identification. The probability of a Type I error occurring in supersaturated designs is an interesting issue, and, in fact, has been extensively discussed in Westfall et al. (1998).

If there is a set of factors which are active, a bound can be found as follows. Recall that for any two events E_1 and E_2,

$$\text{Prob}(E_1 E_2) \geqslant \text{Prob}(E_1) + \text{Prob}(E_2) - 1.$$

Let A be the set of active factors. For any $m \in A$, define $\psi_m = \text{Prob}(\hat{\beta}_m \geqslant \hat{\beta}_i, \text{for } i \notin A)$. Then the probability that $\hat{\beta}_{m'}$ and $\hat{\beta}_{m''}$ are larger than any of the estimated null effects

Table 1
Successful identification probabilities in 5000 simulations

Design	Run size	β	Number of factors (k)					
			10	20	30	40	50	60
HFHM	12	1	0.9462	0.9062	—	—	—	—
		2	1	1	—	—	—	—
	18	1	0.9888	0.9794	0.9694	—	—	—
		2	1	1	1	—	—	—
	24	1	—	0.9950	0.9934	0.9912	—	—
		2	—	1	1	1		
IntCol	12	1	0.9560	0.9026	0.8556	0.8092	0.7842	0.7530
		2	0.9999	1.0000	1.0000	1.0000	0.9999	0.9999

HFHM = Supersaturated designs using half Fraction Hadamard matrices (Lin, 1993).
IntCol = Supersaturated designs using interaction columns (Lin, 1991 and Wu, 1993).

Table 2
Lower bound probabilities given by Corollary 1

Run size	δ_{mi}	Number of factors (k)				
		10	20	30	40	50
12	1	0.8574	0.7237	0.6092	0.5134	0.4327
	2	0.9999	0.9998	0.9997	0.9996	0.9995
16	1	0.9374	0.8725	0.8121	0.7558	0.7034
	2	1.000	1.000	1.000	1.000	1.000
20	1	0.9726	0.9430	0.9143	0.8865	0.8595
	2	1.000	1.000	1.000	1.000	1.000
24	1	0.9879	0.9747	0.9616	0.9487	0.9360
	2	1.000	1.000	1.000	1.000	1.000

is at least $\psi_{m'} + \psi_{m''} - 1$. If both $\psi_{m'}$ and $\psi_{m''}$ are larger than 99%, this bound is 98%. Obviously, we have $\psi_m \geqslant \phi_m$ for any $m \in A$. Hence, a lower bound of ψ_m can be obtained from the last theorem.

More generally, we have

Theorem 4. *Let $A = \{1, 2, \ldots, p\}$ correspond to p active factors in the design. Assume the conditions in Theorem 1 are satisfied. Then the probability that $\hat{\beta}_i$, $i = 1, \ldots, p$, are the p largest factor estimates is no less than*

$$\psi_1 + \psi_2 + \cdots + \psi_p - (p - 1).$$

For example, if $p = 4$ and each of ψ_i, $i = 1, \ldots, p$, is larger than 99%, this bound becomes 96% – a very satisfactorily large lower-bound probability. However, the bound decreases rapidly when p increases or the ψ_i's become smaller. Of course, for such cases, a supersaturated design is not recommended.

Note that the results given above can be straightforwardly extended to the reverse case to find Prob$[\hat{\beta}_w \leqslant \hat{\beta}_i]$ where β_w is the minimal effect among all β_i's.

Acknowledgements

We thank the reviewers for valuable comments and suggestions that led to a substantial improvement of this article. Chen is partially supported by the Natural Sciences and Engineering Research Council of Canada. Lin is partially supported by the National Science Foundation via Grant DMS-9204007. We also thank participants of the Bose Memorial Conference (1995, Fort Collins, CO) for many constructive comments.

References

Booth, K.H.V., Cox, D.R., 1962. Some systematic supersaturated designs. Technometrics 4, 489–495.

Box, G.E.P., 1959. Discussion on "random balance experimentation" by F.E. Satterthwaite. Technometrics 1, 174–180.

Deng, L.Y., Lin, D.K.J., Wang, J.N., 1996. Supersaturated designs using hadama matrices. Revised for Biometrika.

Lin, D.K.J., 1991. Supersaturated designs. Working Papers, 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.

Seewald, W., 1994. Analysis of experimental data when only one factor is important. Technical Report, No. 9405, Ciba-Geigy Ltd, Mathematical Applications, Switzerland.

Srivastava, J.N., 1975. Designs for searching non-negligible effects. In: Srivastava, J.N., (Ed.), A Survey of Statistical Design and Linear Models. North-Holland, New York, pp. 507–519.

Tang, B., Wu, C.F.J., 1993. A method for construction of supersaturated designs and its $E(s^2)$ optimality. Preprint.

Wu, C.F.J., 1993. Construction of supersaturated design through partially aliased interactions. Biometrika 80, 661–669.