ROW-COLUMN DESIGNS: NON-ADDITIVITY MAKES THEM HAZARDOUS TO USE

J.N. Srivastava and Y.C. Wang

Colorado State University
Fort Collins, CO 80523-1877

Abstract

Let there be \( v \) treatments, and two nuisance factors (rows and columns) at \( p \) and \( q \) levels respectively. The general model is: \( y_{ij} = \gamma_{ij} + \tau_k + \epsilon_{ij} \), where \( \tau_k \) is the effect of the \( k \)th treatment, \( \gamma_{ij} \) is the effect of the \((i,j)\) cell, and \( \epsilon_{ij} \)'s are independent random errors with zero mean and variance \( \sigma^2 \), where the \((i,j)\) cell has the \( k \)th treatment assigned to it. The nuisance factors are called “additive” if for all \((i,j)\), \( \gamma_{ij} \) is of the form \((\alpha_i + \beta_j)\), where \( \alpha_i \) and \( \beta_j \) are respectively the effects of the row \( i \) and column \( j \). There is no reason to assume that arbitrary nuisance factors arising in nature will be additive. This will be illustrated in the paper through examples of real data. In this paper, we shall systematically study the problem concerning non-additivity in row-column designs. A method for examining non-additivity and identifying the non-additive cells will be developed. It will show through our real examples that the effect of non-additivity will quite mislead the experimenters in the analysis of row-column designs in the past.

KEY WORDS: Latin square, Latin rectangle, Youden square, Additivity, Search linear model, Supertransversal.

1 INTRODUCTION

Row-column designs have been quite popular for a long time since Fisher introduced the Latin square (LS) design. There are a large number of papers in this field. The bibliography at the end lists a set of these papers. This list is not exhaustive but illustrative. The optimality of Latin square design under the assumption of additivity was proved by Wald (1943). Since then many similar results have been established. In particular, Kiefer (1975a) studied the A- and D-optimality of the generalized Youden square design.

In this investigation, we shall go into the fundamental question of studying the effect of non-additivity on the analysis of row-column designs. This work will be presented in several parts, since it is quite involved, and the questions are quite deep. In this paper, we shall systematically study the problem concerning the effect of non-additivity. We shall develop a technique to examine non-additivity and identify the non-additive cells in row-column designs. For different designs, the number of identifiable NA cells turns out to be
quite different. We shall also examine some real data, showing clearly that in the past, a closer analysis which does take into account the possibility of non-additivity of nuisance factors, would have led to different conclusions. Finally, we shall give some strategies to the experimenters for designing the experiments to eliminate the effect of non-additivity on the analysis of row-column designs.

The possibility of non-additivity of nuisance factors distorting the results from the analysis of a LS experiment (particularly for large sizes) was pointed out by Fisher and Yates (1948). Because of this, the “break up” of a large rectangle into several much smaller ones (for example, splitting a $(8 \times 8)$ square into 16 $(2 \times 2)$ squares) was suggested in Srivastava (1981, 1993), and in Srivastava (1977) and Srivastava and Beaver (1986), leading to Nested Multidimensional Block Designs. A study of the effect of non-additivity was started in Srivastava (1991); some alarming conclusions were reached there leading to the present full scale investigation.

In this paper, we first study the concepts of non-additivity and the number of non-additive (NA) cells. Next, we shall present some mathematical results which are needed in later work. Subsequently, the general (non-additive) linear model will be seen to be representable as a special case of the search linear model. Using the latter theory, the main results of this paper will be derived.

In the following sections, we shall develop the theory for the design $D_1$, which will be called a “Latin rectangle (LR) design”. The design $D_1$ is a $p \times q$ rectangle in which each of the $v$ treatments occurs $p/v$ times in each column. With respect to the rows, $D_1$ forms an $h$-BIBD, defined as follows. In an $h$-BIBD, for every fixed pair of a treatment and a block, either the treatment does not occur in the block at all, or occurs in exactly $h$ units in the block (where $h$ is a positive integer); furthermore if all the $h$ units in a block to which a treatment is assigned are pooled together to form a single “super-unit”, then with respect to these super-units, the design as a whole is a BIBD.

Thus, it is obvious that the ordinary BIBD’s are 1-BIBD’s. Also, if we take an ordinary BIBD, and split every unit into $h$ units (containing the same treatment) then we shall get a $h$-BIBD. This shows that there is a one-one correspondence between an $h$-BIBD and the 1-BIBD on which it is based.

It is clear from the above definition that Latin squares, F-squares, Youden rectangles, etc are all special cases of a $p \times q$ LR design that we are considering.

Now, the general model is

$$y_{ij} = \gamma_{ij} + \tau_k + \epsilon_{ij}, \quad (1.1)$$

where the symbols are defined as before. If the additive model holds, the $\gamma_{ij}$ is of the form

$$\gamma_{ij} = \alpha_i + \beta_j, \quad \text{for all } (i,j) \quad (1.2)$$
For all \((i, j)\), let
\[
\gamma_{..} = \frac{1}{pq} \sum_{i=1}^{p} \sum_{j=1}^{q} \gamma_{ij}, \quad \gamma_{i} = \frac{1}{q} \sum_{j=1}^{q} \gamma_{ij}, \quad \gamma_{..i} = \frac{1}{q} \sum_{j=1}^{q} \gamma_{ij},
\]
\[
\gamma_{j} = \frac{1}{p} \sum_{i=1}^{p} \gamma_{ij}, \quad \delta_{ij} = (\gamma_{ij} - \gamma_{i} - \gamma_{..} + \gamma_{..}), \quad \tau_{k} = \tau'_{k} - \tau.
\]
(1.3)

Thus, we have, for all \(i, j\),
\[
\gamma_{ij} = \mu' + \alpha_{i} + \beta_{j} + \delta_{ij}
\]
where
\[
\mu' = \gamma_{..}, \quad \alpha_{i} = \gamma_{i} - \gamma_{..}, \quad \beta_{j} = \gamma_{j} - \gamma_{..}
\]
(1.5)
The model (1.1) thus becomes
\[
y_{ij} = \mu + \alpha_{i} + \beta_{j} + \delta_{ij} + \tau_{k} + \varepsilon_{ij},
\]
(1.6a)
where
\[
\mu = \mu' + \tau.
\]
(1.6b)

Now, under the model (1.1) (or, equivalently, (1.6)), there is an intrinsic confounding between the \(\gamma\)’s and the \(\tau\)’s. To see this, let \(\hat{\tau}_{k} (k = 1, \cdots, v)\) be an estimate of \(\tau'_{k}\), where
\[
\hat{\tau}_{k} = \sum_{i=1}^{p} \sum_{j=1}^{q} a_{ij}^{(k)} y_{ij},
\]
(1.7)
where \(a_{ij}^{(k)}\) are real numbers. Now,
\[
E[\hat{\tau}_{k}] = \sum_{u=1}^{v} (\sum_{i=1}^{p} \sum_{j=1}^{q} a_{ij}^{(k)} \tau'_{u}) + \sum_{i=1}^{p} \sum_{j=1}^{q} a_{ij}^{(k)} \gamma_{ij},
\]
(1.8)
where, throughout this paper, \(\sum_{u} (u = 1, \cdots, v)\) will denote the sum over all cells \((i, j)\) in the \(p \times q\) LR, to which treatment \(u\) is assigned. Here, if \(k \neq l\), \((k, l = 1, \cdots, v)\), we have
\[
E[\hat{\tau}_{k} - \hat{\tau}_{l}] = \sum_{u=1}^{v} [\sum_{i=1}^{p} \sum_{j=1}^{q} (a_{ij}^{(k)} - a_{ij}^{(l)}) \tau'_{u}] + \sum_{i=1}^{p} \sum_{j=1}^{q} (a_{ij}^{(k)} - a_{ij}^{(l)}) \gamma_{ij}.
\]
(1.9)
Thus, for \((\hat{\tau}_{k} - \hat{\tau}_{l})\) to be an unbiased estimate of \((\tau'_{k} - \tau'_{l})\) for all values of the \(\gamma\)’s, it is necessary that
\[
a_{ij}^{(k)} - a_{ij}^{(l)} = 0, \quad \text{for all } (i, j) \text{ for which } \gamma_{ij} \neq 0.
\]
(1.10)

**Remark 1.1** This leads to the fact that if we have absolutely no knowledge about the \(\gamma\)’s, then there are unknown amount of biases present in the estimates that we may obtain using the data from our LR design.

Consider now the classical “additive” model:
\[
E[y_{ij}] = \mu + \alpha_{i} + \beta_{j} + \tau_{k}.
\]
(1.11)
A sufficient condition that (1.11) may hold is that in (1.6), we have
\[ \delta_{ij} = 0, \quad \text{for all} \ (i, j) \]  
(1.12)
Contrary to popular belief, (1.12) is not necessary, and (1.11) may hold even when \( \delta_{ij} \neq 0 \); this is illustrated by the following result which is easy to check.

**Theorem 1.1** Suppose that for all \( k \) and all \( (i, j) \), there exist constants \( \delta^*_{ij} \) such that
\[ \delta_{ij} = \delta^*_{ij} \]  
(1.13)
if the cell \((i, j)\) has treatment \( k \) assigned to it. Then, the “additive” model (1.11) holds even though the nuisance factors are not additive.

It is clear that under (1.13), estimates of \( (\tau'_k - \tau'_1) \) which are otherwise unbiased, will be biased by an amount \( (\delta^*_k - \delta^*_1) \).

It is true that (1.13) implies a “pattern” for values of \( \delta_{ij} \), which an arbitrary pair of nuisance factors may not give rise to; in general, the \( \delta \)'s will have no pattern. However, the distortion in the estimate of \( (\tau'_k - \tau'_1) \) may still be there. Indeed, in any “symmetric” design (such as a \( p \times q \) LR), the bias in \( (\hat{\tau}_k - \hat{\tau}_1) \), which denotes the BLUE (best linear unbiased estimate) of \( (\tau'_k - \tau'_1) \) will still be \( (\delta^*_k - \delta^*_1) \), if (1.13) is replaced by
\[ \delta_{ij} = \delta^*_k + \delta^*_0 \]  
(1.14)
where \( \delta^*_0 \) are such that
\[ \sum_u \delta^*_0 = 0, \quad \text{for} \ u = 1, \cdots, v \]  
(1.15)

The model (1.1) is representable as
\[ y = X\zeta + \delta + \epsilon \]  
(1.16a)
where
\[ y = (y_{11}, \cdots, y_{1q}; \cdots; y_{p1}, \cdots, y_{pq})', \]
\[ \zeta = (\mu; \alpha_1, \cdots, \alpha_p; \beta_1, \cdots, \beta_q; \tau_1, \cdots, \tau_v)', \]
\[ \delta = (\delta_{11}, \cdots, \delta_{1q}; \cdots; \delta_{p1}, \cdots, \delta_{pq})', \]
\[ \epsilon = (\epsilon_{11}, \cdots, \epsilon_{1q}; \cdots; \epsilon_{p1}, \cdots, \epsilon_{pq})' \]  
(1.16b)
and where the matrix \( X \ (n \times (1 + p + q + v)) \) is the incidence matrix of parameters with respect to the observations. The above model will be called “additive” if it is reducible to the form
\[ y = X\zeta + \epsilon, \]  
(1.17)
notice that this will happen when \( \delta = 0 \), or when \( \delta \) is not zero, but as of the form given by (1.13).

**Remark 1.2.** From the Bayesian view point, if the \((p-1)(q-1)\) independent elements of \( \delta \) have a continuous distribution, then the chance that (1.13) will hold is zero, and (1.17) will imply that \( \delta = 0 \)
As an example, the data from a $4 \times 4$ LS (Bliss (1967)) and the corresponding $X$ matrix are shown below; here the letters in the square stand for treatments, and the figures in parenthesis denote yields. Notice that the rows of $X$ are indexed by the cells of the $p \times q$ rectangle, and the columns by the elements of $\zeta$. We have

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
& 1 & 2 & 3 & 4 & \text{Mean} \\
\hline
\text{April} & B(24) & C(46) & D(34) & A(48) & 38 \\
23 & D(33) & A(58) & B(57) & C(60) & 52 \\
25 & A(57) & D(26) & C(60) & B(45) & 47 \\
26 & C(46) & B(34) & A(61) & D(47) & 47 \\
27 & & & & & \\
\hline
\text{Mean} & 40 & 41 & 53 & 50 & 46 \\
\hline
\end{array}
\]

\[
X = \begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{pmatrix}
\]

We shall use this example to illustrate many of the results of our papers. (This example is part of a factorial assay of insulin from the blood sugar response of rabbits with minor adjustments to avoid fractions. Each rabbit was injected with four doses, two of the standard at 0.6 and 1.2 units ($S_1 = A$ and $S_2 = B$) and two of the test preparation at unknown or presumably the same levels ($U_1 = C$ and $U_2 = D$) (i.e., $S_1$ and $S_2$ are two standard doses and $U_1$ and $U_2$ are two new doses which are being tested). Four rabbits (columns) were injected on each of the 4 days (rows) with the doses $A$ to $D$ of the LS. Each variate $y$ is the milligram percentage of blood sugar in a 1-ml sample taken from an ear vein 50 min after the injection of insulin, i.e., $y = \text{mg} \%$ blood sugar 50 min after injection.)
We shall use the following notations through this paper.

\[ I_n = \text{the } (n \times n) \text{ identity matrix}, \]
\[ J_{ext} = \text{the } (s \times t) \text{ matrix with 1 everywhere}, \]
\[ Y = (y_{ij}) = \text{the } (p \times q) \text{ matrix with } y_{ij} \text{ in the cell } (i,j), \text{ for all } i, j, \]
\[ \epsilon = (\epsilon_{ij}) = \text{the } (p \times q) \text{ matrix with } \epsilon_{ij} \text{ in the cell } (i,j), \text{ for all } i, j, \]
\[ \hat{\zeta} = (X'X)^{-1}X'y, \]
\[ \hat{y} = X\hat{\zeta} = X(X'X)^{-1}X'y, \]
\[ r = y - \hat{y} = (I_n - X(X'X)^{-1}X')y, \]
\[ X = (x'_{11}, \ldots, x'_{i_j}, \ldots; x'_{p_1}, \ldots, x'_{pq})', \]
\[ \Theta_{ri} = \text{the } (p \times q) \text{ matrix with 1 in the } i \text{th row and 0 elsewhere}, \]
\[ \Theta_{cj} = \text{the } (p \times q) \text{ matrix with 1 in the } j \text{th column and 0 elsewhere}, \]
\[ \Theta_{tk} = \text{the } (p \times q) \text{ matrix with 1 in those cells which have the } k \text{th treatment and 0 elsewhere}, \]
\[ \delta = (\delta_{11}, \ldots, \delta_{i_j}; \ldots; \delta_{p_1}, \ldots, \delta_{pq})', \]
\[ \Delta = ((\delta_{ij}) \text{ real matrix, whose } (i,j) \text{ element is } \delta_{ij}, \]
\[ r^{(\Delta)} = (r^{(\Delta)}_{11}, \ldots, r^{(\Delta)}_{i_j}; \ldots; r^{(\Delta)}_{p_1}, \ldots, r^{(\Delta)}_{pq})' = (I_n - X(X'X)^{-1}X')\delta, \]
\[ R_{\Delta} = ((r^{(\Delta)}_{ij}) \text{ matrix, whose } (i,j) \text{ element is } r^{(\Delta)}_{ij}, \]
\[ \mathcal{L} = \{ \Theta \mid \Theta = \sum_{i=1}^{p} a_{ri}\Theta_{ri} + \sum_{j=1}^{q} a_{cj}\Theta_{cj} + \sum_{k=1}^{r} a_{tk}\Theta_{tk}, \text{ where } a's \text{ are real numbers}, \}
\[ \mathcal{R} = \{ R_{\Delta} : \Delta \text{ is any } (p \times q) \text{ real matrix}, \]
\[ m = \min_{\Theta \in \mathcal{L}} \{ \text{ number of non-zero cells in } (\Delta + \Theta), \}
\]

where \( \hat{\zeta}, \hat{y} \) and \( r \) are the least square estimate of \( \zeta \), the fitted value of \( y \), and the residual vector, respectively, when we fit an additive model to the data \( y \), and where \( x'_{ij} \) is the row vector of \( X \) corresponding to the cell \( (i,j) \).

Let \( s \) be a non-negative integer. For \( s > 0 \), let
\[ U = \{(i_1,j_1),(i_2,j_2),\ldots,(i_s,j_s)\}. \]
be any arbitrary set of cells in the \( p \times q \) LR under study. For any set \( U \), the symbol \( \bar{U} \) denotes the complement of \( U \). If \( A \) is a \( (n \times t) \) matrix, then let \( A_U \) denote the \( (s \times t) \) submatrix of \( A \) whose rows are indexed by \( U \) and \( A_{\bar{U}} \) denote the \((n - s) \times t) \) submatrix of \( A \) with the \( s \) rows indexed by \( U \) deleted.

Let
\[ S^2_c(U) = \text{sum of squares due to error (SSE) when the cells in } U \text{ are ignored} \]
\[ = \begin{cases} S^2_c(i,j), & \text{if } U = (i,j), \\ S^2_c = \text{SSE based on the full data } y, & \text{if } U = \emptyset, \end{cases} \]
\[ \text{where } \emptyset \text{ is the null set, and where } (i,j)^{-} \text{ denotes the set of all cells excluding the cell } (i,j). \]
From the definition, it is easy to see that

$$E[r] = r^{(\Delta)}.$$  

(1.26)

Also, it is clear that given $\Delta$, the process of obtaining $\Theta$ so that $(\Delta + \Theta)$ has the smallest possible number of non-zero cells, corresponds to fitting an additive model to a given situation. In particular, if a $\Theta$ can be found such that $(\Delta + \Theta)$ has no non-zero cell, then the additive model fits perfectly.

For a given set of units $U$, let $y_U$ denote the portion of $y$ which comes from the set $U$. Similarly, let $\hat{y}_U$ be derived from $\hat{y}$. Define

$$r_U = y_U - \hat{y}_U$$  

(1.27)

Also, let $X_U$ be the portion of $X$ which corresponds to $U$. Then, the following result can be easily checked from the theory of linear models. (For simplicity, we have taken $X'X$ to be nonsingular.)

**Theorem 1.2** Under the notations introduced above, we have

$$S^2_e(U) = S^2_e - r^2_U[I_u - X_U(X'X)^{-1}X'_U]r_U,$$  

(1.28)

where $U$ contains a total of $s$ units. For $s = 1$, this reduces to

$$S^2_e(i,j) = S^2_e - \frac{r^2_{ij}}{[1 - \bar{z}(X'X)^{-1}z]}$$  

(1.29)

The situation where $m = 0$ is characterized by the following result, which is easy to check. The case $m > 0$ is discussed in the next section.

**Theorem 1.3** (a) $R$ is closed under linear combinations. (b) The following four statements are equivalent: (i) There exits $\zeta$ such that $X\zeta = \delta$, (ii) $\Delta \in \mathcal{L}$, (iii) $R_\Delta = 0$, and (iv) $m = 0$.

Consider now the overall situations. Suppose an experiment has been done using a $p \times q$ LR design. The model (1.1) shows that the $\gamma_{ij}$ are confounded with the $\epsilon_{ij}$. Two cases arise:

(i) The noiseless case, where $\sigma^2 = 0$, and hence the $\epsilon_{ij}$ are zero with probability one, and

(ii) The noisy case, where $\sigma^2 > 0$. The noiseless case is "unrealistic" since in practical life, $\sigma^2$ may usually be large. However, loosely speaking, it is clear that if we can not estimate the treatment contrasts too well (i.e., with a small mean square error) when $\sigma^2 = 0$, then the same would hold even more strongly when $\sigma^2 > 0$. Indeed, it will be seen in later, that many problems which are solvable when $\sigma^2 = 0$, are solvable only with a relatively small probability when $\sigma^2$ is large.

Thus, for simplicity, consider the noiseless case. For $Y$ as at (1.25), we recall that two different solutions (for $\zeta$ and $\delta$) exist, the estimate of $(\tau_1' - \tau_2')$ being $4$ in the first case,
and (−3) in the other! In this situation, we have an intrinsic confounding between certain possibly NA cells, with the result that for certain values of Y (like the one at (1.25)), it is impossible to identify (even when $\sigma^2 = 0$) the NA cells, and obtain the true value of certain treatment contrasts. Since we are quite helpless in such a situation, a very pertinent question is this. Under what condition, can we be certain that such situation will not arise? One of the objectives of this paper is to answer questions of this kind.

In real life, of course, we have $\sigma^2$ relatively large. Furthermore, we should expect most $\delta$'s to be non-zero. Our final objective is to answer questions for this general complex situation. To be able to achieve this goal, a systematic development is needed, so as to gain insight into this class of problems. Herein, we undertake the same.

We first show how the model (1.1) can be represented in terms of a Search Linear Model, introduced by Srivastava (1975). Using techniques from the latter theory, we consider the problem of (unique) identification of the NA cells. Necessary condition (in the form of upper bound on m) are obtained, so that such identification is possible. For $\sigma^2 > 0$, the probability of correct identification of the NA cells is studied. Finally, the general problem of design and analysis of experiments with more than one nuisance factors is discussed, and illustrations are given using real data.

2 SEARCH LINEAR MODEL

Let $y$ ($n \times 1$) be a vector of observations such that

$$\operatorname{Exp}[y] = A_1 \xi_1 + A_2 \xi_2,$$  \hspace{1cm} (2.1)

where $A_1$ ($n \times \nu_1$) and $A_2$ ($n \times \nu_2$) are known matrices, $\xi_1$ ($\nu_1 \times 1$) is an unknown vector of parameters, and $\xi_2$ ($\nu_2 \times 1$) is partially known in the following sense. It is known that there is an integer $m$ such that at most $m$ elements of $\xi_2$ are non-negligible, the remaining elements of $\xi_2$ being negligible, but it is not known which set of elements of $\xi_2$ is actually non-negligible and what the value of these non-negligible elements are. In actual applications, the integer $m$ would be considered small compared to $\nu_2$.

The model (2.1), called the "Search Linear Model", was introduced by Srivastava (1975), wherein the following fundamental result was established.

**Theorem 2.1** Suppose $\sigma^2 = 0$. Under the model (2.1), the elements of $\xi_1$ can be estimated (with variance zero), and $m$ non-negligible elements of $\xi_2$ can be identified (and estimated with variance zero) irrespective of which set of elements in $\xi_2$ is nonnegligible, if and only if

$$\operatorname{Rank} \left[ \begin{array}{c} A_1 \\ A_2 \\ \end{array} \right] = \nu_1 + 2m  \hspace{1cm} (2.2)$$

for any submatrix $A_2$ ($n \times 2m$) of $A_2$, where $2m \leq \min\{n - \nu_1, \nu_2\}$. If $\sigma^2 > 0$, the condition (2.2) is necessary, but is not sufficient in the sense that the correct identification can be done...
only with a probability which is less than one, and furthermore the variance of the estimate of the parameters will be positive.

**Remark 2.1** As emphasized in Srivastava (1996), it should be noted that identification of the true set of nonnegligible parameters in $\xi_2$ is possible for some cases (indeed, maybe a lot of cases, or even most cases), even though (2.2) may not hold. But, if we wish that such identification be possible in all cases, then (2.2) must hold. Thus, the moral is this. After the experiment is done, we should go ahead and project $y$ on the columns of $[A_1 : A_{21}]$, for any $A_{21}$ for which we believe that the corresponding elements of $\xi_2$ may be nonnegligible. To elaborate, consider first the case when $\sigma^2 = 0$. Suppose, we succeed in finding an $A_{21}$ (inside $A_2$) such that $y = [A_1 : A_{21}]\theta_1^T : \theta_{21}^T]^T$, for some $\theta$'s. If, we can find a competing set of parameters within $\xi_2$, such that for the corresponding submatrix $A_{22}$ of $A_2$, we have $y = [A_1 : A_{22}]\theta_1^{\ast T} : \theta_{22}^{\ast T}]^T$, for some $\theta^{\ast}$'s, then we have two solutions for $y$, and only further experiments (directed at discriminating between these two solutions) can help in establishing the correct conclusion. If, on the other hand, no such $A_{22}$ is found, then we have the unique solution already. If $\theta^2 \geq 0$, a perfect projection of $y$ on columns of $A_1$ and $A_{21}$ (or $A_{22}$) may not be possible, and we go by the minimum sum of squares due to error (as explained in various places elsewhere). The procedure outlined for $\sigma^2 = 0$ can still be used, (though the success obtained will depend much upon the value of $\theta$'s (or $\theta^{\ast}$'s) divided by $\sigma$.

From (2.2), we can see that

$$2m \leq \min\{n - \nu_1, \nu_2\}. \quad (2.3)$$

If $2m > n - \nu_1$ or $\nu_2$, then (because of intrinsic confounding) the $m$ non-negligible elements in the model (2.1) can not be identified even if $\sigma^2 = 0$. We also recall from Srivastava (1996) that the model (1.1), in some situations, could be expressed in more than one form. In the context of this paper, we shall consider two forms, called models I and II. In Srivastava (1993), where the concept of nonadditivity of row-column designs was initially studied, model I was used. In Wang (1995), model II was studied more; the details of this study are presented in Wang (1996). In this paper, for the sake of completeness, results on model II will be summarized, and many results on model I will be derived. Under model I, we have

$$\text{Exp}[y] = A\xi + I_n\gamma, \quad (2.4)$$

where $\xi = (\mu; \alpha_1, \ldots, \alpha_{p-1}; \beta_1, \ldots, \beta_{q-1}; \tau_1, \ldots, \tau_{v-1})^T$, $A$ is the $(n \times \nu_1)$ design matrix for $\xi$, $n = pq$ and $\nu_1 = 1 + (p - 1) + (q - 1) + (v - 1)$, and where $\xi$ and $\gamma$ are unknown, except that the nature of $\gamma$ will be elaborated later. Note that $\gamma$ is $(pq \times 1)$ and its elements are the $\gamma_{ij} (i = 1, \ldots, p; j = 1, \ldots, q)$.

Model II is not unique; there are $pq$ of them, denoted by Model II $(i, j)$ (for $i = 1, \ldots, p; j = 1, \ldots, q$). In model II $(p, q)$, we have

$$\text{Exp}[y] = A\xi + B\delta^*, \quad (2.5)$$
where $A$ and $\xi$ are the same as in model I, and where $\xi^* = (\delta_{11}, \ldots, \delta_{1,q-1}; \ldots; \delta_{p-1,1}, \ldots, \delta_{p-1,q-1})$ and $B$ is the $(n \times \nu_2)$ design matrix for $\xi^*$, and where $\nu_2 = (p-1)(q-1)'$ and

$$\sum_{i=1}^p \delta_{ij} = \sum_{j=1}^q \delta_{ij} = 0 \quad \text{for any } i, j$$

(2.6)

Further more, we have

$$A = \left[ \begin{array}{cccc} J_{n \times 1} & A_\alpha & A_\beta & A_\tau \end{array} \right]$$

(2.7)

and

$$B = \begin{pmatrix} I_{p-1} \\ -J_{1 \times (p-1)} \end{pmatrix} \otimes \begin{pmatrix} I_{q-1} \\ -J_{1 \times (q-1)} \end{pmatrix}$$

(2.8)

where

$$A_\alpha = (n \times (p-1)) \text{ design matrix for } \alpha = \begin{pmatrix} I_{p-1} \\ -J_{1 \times (p-1)} \end{pmatrix} \otimes J_{q \times 1},$$

$$A_\beta = (n \times (q-1)) \text{ design matrix for } \beta = J_{p \times 1} \otimes \begin{pmatrix} I_{q-1} \\ -J_{1 \times (q-1)} \end{pmatrix},$$

$$A_\tau = (n \times (v-1)) \text{ design matrix for } \tau,$$

and where for any two matrices $C = ((c_{ij}))$ and $D$, the Kronecker product $\otimes$ is defined as

$$C \otimes D = ((c_{ij}D)).$$

(2.10)

The model (1.1) is equivalent to the model I or II $(p, q)$ because of relations (1.3) and (1.5), which say that $\sum_{i=1}^p \alpha_i = \sum_{j=1}^q \beta_j = \sum_{k=1}^v \gamma_k = \sum_{i=1}^p \delta_{ij} = \sum_{j=1}^q \delta_{ij} = 0$. Also, from (1.16a) and (1.16b), it follows that $A_\alpha$ is obtainable from $X_\alpha$ by ignoring the last column, and also subtracting it from the other columns. Similarly, for $A_\beta$ and $A_\tau$. The matrix $B$ in the model II $(p, q)$ depends only on the order of design. For the example (1.18), we have

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

(2.11)
We now explain the models \( (i, j) \). Recall the matrix \( \Delta \) of size \( (p \times q) \) whose elements are the \( \delta_{ij} \), and whose rows (and also whose columns) sum to zero. Thus, the matrix \( \Delta \) contains only \( (p-1) \times (q-1) \) independent elements. A set of such \( (p-1)(q-1) \) independent elements can be obtained by suppressing any row and column of \( \Delta \), since the suppressed elements can be easily obtained from the non-suppressed ones. In model \( (i, j) \), the \( i \)th row and \( j \)th column of \( \Delta \) are ignored, and the corresponding \( \delta^* \) contains the remaining elements of \( \Delta \); this is why the \( \delta^* \) in (2.5) corresponds to the model \( (p, q) \). (The \( \delta^* \) in (2.5) should, more precisely, be denoted by a symbol such as \( \delta^{*pq} \) so as to stress that we are using the model \( (p, q) \). However, in (2.5), we drop the superscripts in \( \delta^{*pq} \) in (2.5), for simplicity.

As mentioned above, the models I and II \( (i, j) \) (for all \( i, j \)) are equivalent as linear models in the sense that any one of them can be obtained from any other through reparametrisation. However, they are different considered as search linear models. This is seen by relating each such model to the basic structure of the search linear model given at (2.1), and noting that the \( \xi_2 \) and the \( A_2 \) are different in each case, and that in each case, the elements of \( \xi_2 \) are unrestricted real numbers (except that the number of nonnegligible elements of \( \xi_2 \) is bounded above by a number, which may be known or may be unknown).

We now consider the quantity \( m \) defined at (1.23). Notice that as \( \Theta \) varies over \( L \), the computation of \( \Delta + \Theta \) is like fitting different possible models of the form (1.16). So, when in (1.23), we take a minimum over \( \Theta \in L \), we are fitting the "best possible" model in the sense that the number of nonzero elements that we shall obtain in \( \delta \) (in (1.23)) is a minimum; and we will be able to explain \( E(y) \) by \( \zeta \) with a minimal use of the elements in \( \delta \). For this reason, the number \( m \) is called the number of NA cells in the \( (p \times q) \) LR-design under consideration, when the parameter values are given by the \( \delta_{ij} \) (or, equivalently, by the \( \gamma_{ij} \)).

In view of the above, the situation for this LR-design is representable by the (search linear) model I at (2.4), where it is given that \( \gamma_{p(p-1)} \) contains at most \( m \) nonzero elements. (Note that the value of \( m \) may or may not be known.)

For any particular case, a set \( S \) of \( m \) cells constitute "a set \( m \) NA cells", if there exist values of \( \gamma_{ij}((i,j) \in S) \) (with \( \gamma_{ij} = 0 \), when \( (i,j) \notin S \)), such that with this value of \( \gamma \), there exist a value of \( \zeta \) satisfying (2.4).

Notice that we may not have a unique "set of \( m \) NA cells". This will happen if there exist two values of \( \zeta \) (say, \( \xi_1 \) and \( \xi_2 \)), and two \( (n \times m) \) distinct submatrices (say \( I^{(1)} \) and \( I^{(2)} \)) of \( I_n \), along with two \( (m \times 1) \) vectors \( \gamma^{(1)} \) and \( \gamma^{(2)} \), such that
\[
A\xi_1 + I^{(1)}\gamma^{(1)} = A\xi_2 + I^{(2)}\gamma^{(2)}.
\] (2.12)

If (2.12) holds, then the rank condition (2.2) (for the model at (2.4)) will not be satisfied. Conversely, if the rank condition is not satisfied, then the set of \( m \) NA cells will not be unique, and more than one solutions (such as at (2.12) will arise.

As an example, consider again the square at (1.18). Let \( E(Y) \) denote \( E(y) \) written in the
form of this square, and let the yield corresponding to the different cells in (1.18) be given by

\[
E(Y) = \begin{bmatrix}
3 & 4 & 5 & 4 \\
2 & 3 & 4 & 5 \\
4 & 5 & 6 & 5 \\
5 & 4 & 5 & 4
\end{bmatrix}
\]  \hspace{1cm} (2.13)

Then, it can be checked that these values of \( E(y) \) are obtained in the model (2.4), for each of the following two values of \( \zeta \) and \( \gamma \).

(i) \( \zeta' = (0; 1, 1, 2, 2; 1, 1, 2, 2; 1, 1, 2, 2); \gamma_{21} = \gamma_{44} = -2; \) other \( \gamma's = 0; \)

(ii) \( \zeta' = (0; 1, 0, 2, 1; 1, 2, 3, 2; 1, 1, 1, 1); \gamma_{24} = \gamma_{41} = 2; \) other \( \gamma's = 0. \) \hspace{1cm} (2.14)

Indeed, it can be shown that if (2.12) holds, then there are an infinite number of solutions.

Now, consider the model II \((p, q)\) at (2.5). In this case, for a given value of \( E(y) \), "the number of NA cells" is defined to be the smallest number \( m_o(\equiv m_o(p, q)) \), such that the equation (2.5) can be satisfied with (some value of \( \zeta \)) and only \( m_o \) nonzero elements in \( \delta \), the rest of the \((n - m_o)\) elements in \( \delta \) being zero. Note that in this case, the word "cells" in "the number of NA cells" refers only to cells \((i, j)\) such that \( i \neq p \) and \( j \neq q \).

For a given value of \( E(y) \), the values of \( m \) and \( m_o(i, j) \) (for various \((i, j)\)) may be quite different. Indeed, the case "\( m = 1 \)" would correspond to \( m_o(i, j) = (p - 1)(q - 1)! \). On the other hand, "\( m = 4 \)" with

\[
\delta_{i_1 j_1} = -\delta_{i_1 j_2} = -\delta_{i_2 j_1} = \delta_{i_2 j_2}
\] \hspace{1cm} (2.15)

corresponds to

\[
m_o(g, h) = 4, m_o(g', h') = 1, m_o(g'', h'') = 2,
\] \hspace{1cm} (2.16)

where

\[
g \neq i_1 \text{ or } i_2, \ h \neq j_1 \text{ or } j_2, g' = i_1 \text{ or } i_2, \text{ and } h' = j_1 \text{ or } j_2.
\] \hspace{1cm} (2.17)

\[
g'' = i_1 \text{ or } i_2, \text{ and } h'' \neq j_1 \text{ or } j_2, \text{ or}
\]

\[
g'' \neq i_1 \text{ or } i_2, \text{ and } h'' = j_1 \text{ or } j_2.
\]
The situation where more than one solution for $E(y)$ exists can arise under model I as well.

For example, in the $4 \times 4$ LS at (1.18), we could have a different set of data $E(Y)$, where

$$E(Y) = \begin{pmatrix}
-2.5 & 6.0 & 0.5 & 2.0 \\
-1.0 & 1.5 & -3.5 & 11.0 \\
-7.5 & -0.5 & -4.0 & -4.0 \\
-5.0 & 3.0 & -1.0 & 5.0
\end{pmatrix}. \tag{2.18}$$

Now, let

$$\zeta_{(1)} = (0; 1.5, 2.0, -4.0, 0.5; -4.0, 2.5, -2.0, 3.5; 0.5, -3.5, 2.0, 1.0)'$$
$$\zeta_{(2)} = (0; 1.5, 2.0, -4.0, 0.5; -4.0, 2.5, -2.0, 3.5; -3.0, 0.0, 2.0, 1.0)' \tag{2.19}$$

Thus, it can be verified that the above $E(Y)$ can be obtained using $\zeta_{(1)}$ as the value of the $\zeta$, and two NA cells (so that, $m_o(4, 4) = 2$) (1,1) and (2,2), with $\delta_{11} = -\delta_{22} = 3.5$. Similarly, $E(Y)$ is also obtainable using $\zeta_{(2)}$ as the value of $\zeta$, with the two NA cells (2,3) and (3,1), where $\delta_{23} = -\delta_{31} = -3.5$. (In this example, for simplicity, we have taken $\epsilon = 0$.)

**Remark 2.2** Consider a row-column design with $m$ NA cells. Then, by the definition of $m$, it is clear that there exists a set of cells (say $U^1$) such that if $U^1$ is ignored, then the additive model will hold. As the example at (2.14) shows, sometimes more than one sets $U^1, U^2, \ldots$ (each with $m$ cells) may be present, such that (for each given $j$) if $U^j$ is ignored, then the additive model will hold. In the notation (1.25), this implies that for each $j$, if $\sigma^2 = 0$, then we shall have

$$S^2_\epsilon(U^j) = 0. \tag{2.20}$$

We now recall four decision rules ($R_1, \ldots, R_4$) for identification of parameters for $\sigma^2 \geq 0$. The first three were introduced by Srivastava (1975, 1976) in Search Theory, and the last one in Srivastava (1987).

The decision rule $R_1$ says that for fixed $m$, $U_0$ should be taken to be the set of NA cells if $S^2_\epsilon(U_0) \leq S^2_\epsilon(U)$, for all $U \neq U_0$.

Decision rule $R_2$ is as follows. Do $R_1$ with $m = 1$ first. Suppose this results in a cell (say, $c_1$) being identified as possibly NA. Let $F$ denote the set of all the $pq$ cells. For any $w_1, \ldots, w_k$, let $F(w_1, \ldots, w_k)$ be the set of the $(pq - k)$ cells obtained by excluding $w_1, \ldots, w_k$ from $F$. Now, apply rule $R_1$ to the set $F(c_1)$. Suppose this results in the cell $c_2$ being identified. Next, apply $R_1$ to the set $F(c_1, c_2)$. Repeat this procedure $m$ times; if this results in $U_0 = (c_1, \ldots, c_m)$ being identified as the NA units, then $R_2$ says that $U_0$ should be taken to be the NA set.

For $m = 1$, $R_1$ and $R_2$ are same. For $m > 1$, $R_1$ requires $\binom{n}{m}$ ANOVA computations, while $R_2$ requires only $mn - \binom{m}{2}$ computations, which is far less. Of course, $R_2$ is not
expected to be as accurate as $R_1$; however, for medium values of $m$, it may be useful when $\delta_{ij}$ are relatively large.

Decision rule $R_3$ is as follows. Fit the model (2.1) to the data, and obtain the $(p \times q)$ matrix showing the residuals for each cell. Choose those $m$ cells to be NA for which the absolute value of the residual are the largest among all the $pq$ cells.

When $m = 1$, rules $R_1$ and $R_3$ turn out to be very close (in studies on row-column designs); they are identical when $\sigma^2 = 0$. The rule $R_2$ may be called $R'_2$ if we use $R_3$ instead of $R_1$ at each of the $m$ stages of selecting a possible NA cell. Such usage will curtail the amount of computation that needs to be done. The rules $R_2$ and $R_3$ may succeed in some cases if the values of the non-zero $\delta$'s are large.

Decision rule $R_4$ (Intersection sieve) is a tentative rule which says the following: If $\theta_1, \theta_2, \theta_3$ are linear combinations of parameters such that the set of parameters involved in (2.1) satisfies $\theta_1 + \theta_2 = \theta_3 + \theta_2 = \pi$ (some real number), then it is likely (though, not certain) that $\theta_2 = \pi$, and $\theta_1 = \theta_3 = 0$. By "tentative", we mean that the user allows himself to ignore it under some non-predetermined circumstances. (In other words, the user allows his personal judgment to sometimes prevail over a "tentative" rule). A tentative decision rule for separating the negligible parameters from the non-negligible ones is called a "Sieve".

Decision rule $R_4$ was used by Srivastava (1987) for model-identification in factorial designs, and was applied in Wang (1995) for identifying the NA cells in a row-column design.

Decision Rule $R_5$ is useful for LS designs, and is based on the concept of a "transversal", which is a set of $p$ cells of the LS, such that each row, column, and treatment occurs once in the set. Thus, in the LS at (1.18), $T_1$ and $T_2$ are both transversals, where $T_1$ is the set of cells 11, 22, 33, 44, and $T_2$ is 11, 24, 32, 43. This LS has a total of 8 transversals. The "yield" of a transversal is the sum of the yields of all the $p$ cells in it. Under the additive model, each transversal is expected to have the same yield. Under model I, by comparing yields of transversals, and using Rule $R_4$, it is sometimes possible to locate the NA cells. This works particularly well if the nonadditivity parameters are small in number, and their values have the same sign. Sometimes, probability plotting methods, using order statistics obtained by taking the set of yields of the various transversals, may be useful. This, however, will be discussed elsewhere.

Which model should one use in practice? The answer depends upon the situation. Model I is more intuitively evident and easier to grasp and interpret. The number of NA (nonadditivity) parameters in model II $(i, j)$ depend upon $(i, j)$. After the experiment is done, the value of $y$ should be projected on the parameters under model I and each model II $(i, j)$, and the most successful projection (i.e. one which gives a "small" error sum of squares, for the least number of NA parameters) may be accepted. We shall further comment on this in section 4.
3 NUMBER OF IDENTIFIABLE NA PARAMETERS

In this section we shall discuss bounds on $m$ and $m_o(i, j)$. We begin with model II.

**Theorem 3.1** Consider a $(p \times p)$ LS design. Then, for all $(i, j)$, $1 \leq i, j \leq p$, the NA parameters are identifiable under the model II $(i, j)$, only if

$$m_o(i, j) \leq p - 3$$  \hspace{1cm} (3.1)

**Proof:** Without loss of generality, we take $i = j = p$. Proof for other cases is similar. Let $a_{1i}$, $a_{2j}$, and $a_{3k}$ respectively denote the column of $A$ corresponding to the $i$th row, the $j$th column, and the $k$th treatment (for $i, j, k = 1, \ldots, p - 1$); the element of $a_{3k}$ with respect to $y_{gh}$ (i.e., the observation in the cell $(g, h)$, $g, h = 1, \ldots, p$) will be denoted by $a_{3k}(g, h)$. Similarly, let $b_{ij}(1 \leq i \leq p - 1, 1 \leq j \leq p - 1)$ denote the column of $B$ corresponding to the cell $(i, j)$, and let $b_{ij}(g, h)$ denote the element of $b_{ij}$ in the row corresponding to $(g, h)$.

Now, consider the LS design, and suppose treatment $d$ is assigned to the cell $(p, p)$ where $1 \leq d \leq p$. Two cases arise according as $d \neq p$, and $d = p$.

Case I. $d \neq p$. Let $u, v(1 \leq u, v \leq p - 1, u \neq d \neq v)$ be two treatments. We consider the columns $a_{3u}, a_{3v}$, and $b_{ij}$ (for $(i, j) \in \Gamma^*$, where $\Gamma^*$ is the set of cells $(i', j')$, for $1 \leq i', j' \leq p - 1$, such that either treatment $u$ or treatment $v$ is assigned to $(i', j')$, and we show that these columns are linearly dependent.

Now, each of the treatments $u$ and $v$ occur $p$ times in the LS, once in each row and in each column. In particular, each of these two treatments occur once in the $p$th row and the $p$th column, so that $\Gamma^*$ contains $2(p - 2)$ cells, half of which correspond to treatment $u$, and the other half to treatment $v$. Let $\Gamma^*$ be the disjoint union of $\Gamma_u^*$ and $\Gamma_v^*$, where $\Gamma_w^*(w = u, v)$ is the set of $(p - 2)$ cells in $\Gamma^*$ to which treatment $w$ is assigned. We proceed to show that

$$a_{3u} - a_{3v} = \sum_1(b_{ij} - \sum_2(b_{ij}),$$  \hspace{1cm} (3.2)

where $\sum_1$ is the sum over all $(i, j) \in \Gamma_u^*$, and $\sum_2$ is the sum over all $(i, j) \in \Gamma_v^*$.

To prove (3.2), we consider the $(g, h)$ element in the column vectors on each side of (3.2). Clearly, (3.2) holds if and only if, for all $1 \leq g, h \leq p$, we have

$$a_{3u}(g, h) - a_{3v}(g, h) = \sum_1(b_{ij}(g, h)) - \sum_2(b_{ij}(g, h))$$  \hspace{1cm} (3.3)
Now, we have
\[
\begin{align*}
  a_{3u}(g, h) &= 1, \text{ if treatment } u \text{ is assigned to the cell } (g, h) \\
  &= -1, \text{ if treatment } p \text{ is in the cell } (g, h) \\
  &= 0, \text{ otherwise.}
\end{align*}
\]  
\quad \text{(3.4a)}
\]

Similar is the case for \( a_{3v}(g, h) \). Also, for \((i, j) \in \Gamma_u^*\), we have
\[
\begin{align*}
  b_{ij}(g, h) &= 1, \text{ if } (i, j) = (g, h) \\
  &= 1, \text{ if } (g, h) = (p, p) \\
  &= -1, \text{ if } (g = i < p, h = p), \text{ or } (g = p, h = j < p), \\
  &= 0, \text{ otherwise.}
\end{align*}
\]  
\quad \text{(3.4b)}

Similar is the case when \((i, j) \in \Gamma_v^*\). Now consider the cell \((g, h)\), for \(1 \leq g, h \leq p\); the following subcases arise:
\[
\begin{align*}
  Ia & : (g, h) = (p, p); \\
  Ib & : (g, h) \text{ has treatment } p \text{ assigned to it;}
\end{align*}
\]  
\quad \text{(3.5)}

\[
\begin{align*}
  Ic & : (g, h) \text{ has treatment } u \text{ in it;}
\end{align*}
\]

\[
\begin{align*}
  Id & : (g, h) \text{ has treatment } v \text{ in it}
\end{align*}
\]

\[
\begin{align*}
  Ie & : (g, h) \text{ has treatment } w \text{ in it, where } w \neq u, v, p; \text{ and } (g, h) \neq (p, p)
\end{align*}
\]

For simplicity, we write
\[
\begin{align*}
  c(u, v, g, h) &= a_{3u}(g, h) - a_{3v}(g, h), \\
  d(g, h) &= -\sum_1 (b_{ij}(g, h)) - \sum_2 (b_{ij}(g, h)).
\end{align*}
\]  
\quad \text{(3.6)}

For \( Ia \), we have, from (3.4),
\[
\begin{align*}
  a_{3u}(p, p) - a_{3v}(p, p) = 0 - 0 = 0,
\end{align*}
\]  
\quad \text{(3.7a)}

\[
\begin{align*}
  \sum_1 [b_{ij}(p, p)] - \sum_2 [b_{ij}(p, p)] = 1 \times (p - 2) - 1 \times (p - 2) = 0.
\end{align*}
\]  
\quad \text{(3.7b)}

For \( Ib \), we get
\[
\begin{align*}
  a_{3u}(g, h) - a_{3v}(g, h) = (-1) - (-1) = 0,
\end{align*}
\]  
\quad \text{(3.8)}

Several subsubcases arise: (i) \( g \neq p \neq h \); (ii) \( g = p \neq h \) or \( g \neq p = h \). In case (i), \( b_{ij}(g, h) = 0 \), irrespective of whether \((i, j)\) cell has in it treatment \( u \) or \( v \); thus (3.3) holds. In case (ii), if \( g = p \neq h \), there are two distinct cells in column \( h \) containing treatments \( u \) and \( v \); the contribution from each of these two cells to \( c(i, j, u, v) \) is \((-1) - (-1) = 0 \). For all other cells \((i, j)\) in \( \Gamma^* \), the contribution is zero. Hence, (3.3) holds.

For \( Ic \), we have \( a_{3u}(g, h) = 1, a_{3v}(g, h) = 0 \). Now, two cases arise: (i) \( 1 \leq g, h < p \), and (ii) \( g = p \neq h \), or \( g \neq p = h \). In case (i), we have \( b_{gh}(g, h) = 1 \), and \( b_{ij}(g, h) = 0 \), for all other
(i, j) in \( \Gamma^* \). Hence, in case (i), (3.3) holds. Now, let \( g = p \neq h \). In column \( h \) of the LS, there is a cell (say, \((g_o, h)\)), in which treatment \( v \) occurs; thus \( b_{g_o, h}(g, h) = -1 \), and \( b_{ij}(g, h) = 0 \), for all other \((i, j) \in \Gamma^* \). Hence, (3.3) holds here as well. Similar is the case when \( g < p = h \). For \( \text{Id} \), result follows as for \( \text{Ic} \).

For \( \text{Ie} \), \( a_{3u}(g, h) = a_{3v}(g, h) = 0 \). Three cases arise: (i) \( 1 \leq g, h \leq p - 1 \), (ii) \( g = p \leq h \), (iii) \( g \neq p = h \). In (i), we have \( b_{ij}(g, h) = 0 \), for all \((i, j) \in \Gamma^* \), so that (3.3) holds. In (ii), there are \((p - 4)\) values of \( h \) possible; let \( h_0 \) be one of these \((p - 4)\) values. Now, in the \( h_0 \)th column of the LS, there exist exactly two cells, say \((g_1, h_0)\) and \((g_2, h_0)\) such that treatments \( u \) and \( v \) respectively occur in these. Then \( b_{g_1, h_0}(g, h_0) = -1 = b_{g_2, h_0}(g, h_0) \), and \( b_{ij}(g, h_0) = 0 \), for the other values of \((i, j) \in \Gamma^* \). Hence (3.3) holds, Similarly, for case (iii). Thus, (3.2) holds for Case I.

We now consider Case II. This holds because, without loss of generality, we can rename the rows and columns of our LS, so that the square has the standard representation, i.e., the first row of the LS has treatments in the order \((1, 2, \ldots, p - 1, p)\) and similarly for the first column. Then, treatment \( p \) cannot occur in the cell \((p, p)\).

This completes the proof of the theorem.

The condition (3.1) in the above theorem in the LS design is also sufficient; in other words, if (3.1) holds, then the NA parameters can be identified without ambiguity in all cases. (From remark 2.1, recall that even if (3.1) does not hold, unique identification of NA parameters may be possible in some cases (depending upon the set of NA parameters.) However, for the proof of the sufficiency of (3.1) for the LS design, and of similar other results for the \( F \)-Square design, Youden Square design, and the Lattice Squares, the reader is referred to Wang (1997).

We now consider bounds under model I. In place of (2.4), we shall use the model I in the form

\[
\text{Exp}(y) = X\zeta + \mathbf{I}_n \delta
\]

where \( X \) and \( \zeta \) are as at (1.16). Notice that in \( \zeta \), because of (1.3) and (1.5), we have

\[
\sum_1^p \alpha_i = \sum_1^p \beta_i = \sum_1^p \tau_k = 0,
\]

so that (2.4) and (3.9) are equivalent. Obviously, in (2.4), \( \zeta \) is a subvector of \( \zeta_1 \) and \( A \) is derived from \( X \). However, while the model (2.4) is a special case of (2.1), the model (3.9) is not so, since in the former, the elements of \( \zeta \) are all unknown, and in the latter they are unknown but are subject to the known linear dependence (3.10). As for (2.4), we assume that \( \delta \) has \( m \) nonzero elements which are to be identified.

The following notation will be used. For any matrix \( Q \), the "weight of \( Q \)" denoted by \( \text{wt}(Q) \), equals the number of nonzero elements in \( Q \). Also, for any real number \( q \), the symbol \([q]\) will denote the largest integer less than or equal to \( q \).
Theorem 3.2 A necessary condition (which is also sufficient if $\sigma^2 = 0$) that the $m$ NA parameters in $\zeta$ can be identified under the model (3.9) is that

$$wt(X\zeta) > 2m$$ (3.11)

for every non-zero value of $\zeta$, subject to (3.10).

Proof Suppose $\zeta_0$ is a nonzero value of $\zeta$ satisfying (3.10), and is such that the $wt(X\zeta_0) \leq 2m$. Then, clearly, there exist vectors $\omega_1(m \times 1)$ and $\omega_2(m \times 1)$ and two disjoint submatrices $\Omega_1(n \times m)$ and $\Omega_2(n \times m)$ of $I_n$, such that

$$X\zeta_0 = \Omega_1\omega_1 + \Omega_2\omega_2,$$ (3.12)

so that for any $\zeta^*_0(m \times 1)$ (satisfying (3.10), we get

$$X\zeta^*_0 + \Omega_1(-\omega_1) = X(\zeta^*_0 - \zeta_0) + \Omega_2(\omega_2),$$ (3.13)

which implies that (3.9) has two different solutions which are confounded. Hence, (3.11) is necessary. To prove sufficiency, notice that if $y$ has two different projections on the columns of $X$ and $I_n$ (subject to the given conditions), then a situation analogous to (3.13) would arise, which contradicts (3.11). This completes the proof.

Corollary 3.1 Consider model (3.9) for a $(p \times p)$ LS. A necessary (sufficient when $\sigma^2 = 0$) condition that the $m$ NA cells (i.e., parameters in $\delta$) can be identified and estimated is that there does not exist a non-zero vector $\zeta$ (satisfying (3.10)) such that

$$wt(X\zeta) \leq 2m.$$ (3.14)

Corollary 3.2 If there exists a non-zero vector $\zeta$, (subject to (3.10)), and an integer $a$ such that

$$wt(X\zeta) = a,$$ (3.15)

then the $m$ NA cells can be identified only if $m < [a/2]$.

We now try to obtain the smallest value of $a$ (i.e., for $wt(X\zeta)$) for a $p \times p$ LS design under the model (3.9). For $p = 3$, it was shown in Srivastava (1993) that no NA cells can be identified. So, we take $p \geq 4$. Consider the cell $(i, j)$ in the LS design to which treatment $k$ is applied. The expected value (under additivity) for the yield of the cell is $(\mu + \alpha_i + \beta_j + \tau_k)$. We wish to make these zero. Without loss of generality, we can assume that the first row of the LS is $(1, 2, \cdots, p - 1, p)$. Take

$$-\beta_l = \tau_l, \quad \text{for } l = 1, 2, \cdots, p.$$ (3.16)
Then, the value in the cells of the first row is \((\mu + \alpha_1)\). Notice that though
\[
\sum_{i=1}^{p} \alpha_i = 0, \text{ and } \mu \text{ is general,} \tag{3.17}
\]
we can assume that
\[
\mu = 0, \text{ and } (\alpha_1, \alpha_2, \cdots, \alpha_p) \text{ are unrestricted.} \tag{3.18}
\]
For example, \(\mu = 6, \text{ and } \alpha' = (3, -7, -2, 8, -2)\) is equivalent to \(\mu = 0, \alpha' = (9, -1, 4, 14, 4)\).
The following lemma is obvious.

**Lemma 3.1** (1). (3.17) is equivalent to \(\mu^* = 0, \alpha^{*'} = (\mu + \alpha_1, \cdots, \mu + \alpha_p)\) (of the form (3.18)).

(2). (3.18) implies \(\mu^* = \frac{1}{p} \sum \alpha_i, \alpha^{*'} = (\alpha_1 - \mu^*, \cdots, \alpha_p - \mu^*)\) (of the form (3.17)).

So, we take
\[
\mu = 0, \text{ and } (\alpha_1, \cdots, \alpha_p) \text{ are general with } \alpha_1 = 0. \tag{3.19}
\]
Then the expected yields in the first row in the LS are zero. Consider an element \(\psi\) in the \(i\)th \((i > 1)\) row and \(j\)th column. Its value is \(\alpha_i + q_{jk}\), where
\[
q_{jk} = \beta_j + \tau_k, \quad j \neq k \tag{3.20}
\]
where we assume that the treatment in the cell \((i,j)\) is \(k\).

As an example, we look at the following \(4 \times 4\) LS:
\[
\begin{bmatrix}
A & B & C & D \\
B & C & D & A \\
C & D & A & B \\
D & A & B & C
\end{bmatrix} \tag{3.21}
\]
Then, the matrix for \(q_{ij}\) is
\[
\begin{bmatrix}
q_{12} & q_{23} & q_{34} & q_{41} \\
q_{13} & q_{24} & q_{31} & q_{42} \\
q_{14} & q_{21} & q_{32} & q_{43}
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\theta & \theta & \theta & -3\theta \\
2\theta & 2\theta & -2\theta & -2\theta \\
3\theta & \theta & \theta & \theta
\end{bmatrix}, \tag{3.22}
\]
where \(\theta = q_{12} = q_{23} = q_{34}\). Now, by choosing \(\alpha_2 = -\theta, \alpha_3 = 2\theta, \alpha_4 = -\theta\), we obtain (for \(wt(X\xi)\) a = 4). Hence, by Corollary 3.2,

**Definition 3.1.** Two LS \(L_1\) and \(L_2\) are called isomorphic if \(L_2\) can be obtained from \(L_1\) by a permutation of rows and/or a permutation of columns, and/or a permutation of symbols of treatments.
Lemma 3.2. For the $4 \times 4$ LS, there are only four non-isomorphic LS's.

Proof. Without loss of generality, we assume that the first row and the first column of the LS are respectively $(A, B, C, D)$ and $(A, B, C, D)'$. Then, it is easy to check that there are only four non-isomorphic squares as below:

\[
L_{41} = \begin{bmatrix}
A & B & C & D \\
B & A & D & C \\
C & D & A & B \\
D & C & B & A \\
\end{bmatrix}, \quad L_{42} = \begin{bmatrix}
A & B & C & D \\
B & A & D & C \\
C & D & B & A \\
D & C & A & B \\
\end{bmatrix}, \quad (3.23)
\]

\[
L_{43} = \begin{bmatrix}
A & B & C & D \\
B & C & D & A \\
C & D & A & B \\
D & A & B & C \\
\end{bmatrix}, \quad L_{44} = \begin{bmatrix}
A & B & C & D \\
B & D & A & C \\
C & A & D & B \\
D & C & B & A \\
\end{bmatrix}.
\]

Lemma 3.3. Let $L_1$ and $L_2$ be any two isomorphic $p \times p$ LS's. Then,

\[
\text{MNIC in } L_1 = \text{MNIC in } L_2. \quad (3.24)
\]

The proof of the Lemma is directly obtained by the definition of MNIC and the concept of isomorphism. Therefore, we only consider non-isomorphic LS's for MNIC.

Theorem 3.3. For any $4 \times 4$ LS,

\[
\text{MNIC} = 1 \quad (3.25)
\]

Proof. From Lemma 3.2 and 3.3, we only need to prove the theorem for the four non-isomorphic LS's. The square $L_{43}$ has already been dealt with at the example (3.21). We use (3.16) and (3.20) in other cases as well. Consider $L_{41}$, and let $q_{12} = q_{34} = \theta$, and $q_{23} = \phi$. Then in place of (3.22), we obtain

\[
\begin{bmatrix}
0 & 0 & 0 & 0 \\
\theta & -\theta & \theta & -\theta \\
\theta + \phi & \theta + \phi & -\theta - \phi & -\theta - \phi \\
2\theta + \phi & \phi & -\phi & -2\theta - \phi \\
\end{bmatrix}.
\]

Now, choosing $\theta = 0, \alpha_1 = \alpha_2 = 0, \alpha_3 = \phi = \alpha_4$, we find that (3.25) holds. For $L_{42}$, we again take $q_{12} = q_{34} = \theta, q_{23} = \phi$, and obtain

\[
\begin{bmatrix}
0 & 0 & 0 & 0 \\
\theta & -\theta & \theta & -\theta \\
\theta + \phi & \theta + \phi & -\phi & -2\theta - \phi \\
2\theta + \phi & \phi & -\theta - \phi & -\theta - \phi \\
\end{bmatrix}.
\]

20
Letting $\theta = 0$, the situation is found to be the same as for $L_{41}$. For $L_{44}$, we take $q_{12} = q_{24} = q_{43} = \theta$, obtaining

$$
\begin{bmatrix}
0 & 0 & 0 & 0 \\
\theta & \theta & -3\theta & \theta \\
3\theta & -\theta & -\theta & -\theta \\
2\theta & 2\theta & -2\theta & -2\theta \\
\end{bmatrix}
$$

so that $\alpha_1 = 0, \alpha_2 = -\theta, \alpha_3 = \theta, \alpha_4 = 2\theta$, gives the result. This completes the proof.

We now prove that if there is only one NA cell in any one of the four non-isomorphic LS’s, then the NA cell can be identified by using the $R_1$ method. For simplicity, it is sufficient to illustrate the proof for the case of one Latin Square, and for one cell within it. The proof for the other cases follows on the same lines. (Indeed, all the cases could be combined into one. We do not opt for that, since although it would be seemingly general, the main point will be obscured.) Consider $L_{44}$, and suppose that the cell (1,2) is NA, and that $\delta$ is the nonadditivity parameter. Then, when we fit an additive model to $y$-values, we have the residual matrix.

$$
R_\Delta = ((r_{ij}^\Delta)) = -\frac{1}{8}\delta
\begin{bmatrix}
-1 & 3 & -1 & -1 \\
-1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 \\
\end{bmatrix}
$$

(3.26)

Thus, using (1.29), we get

$$
S_c^2 = \frac{2}{3}\delta^2,
S_c^2(i,j)^- = S_c^2 - \frac{8}{3} r_{ij}^2 = \begin{cases} 
0 & \text{if } (i,j) = (1,2) \\
\frac{1}{3}\delta^2 & \text{if } (i,j) \neq (1,2)
\end{cases}
$$

(3.27)

So the NA cell can be identified by $R_1$.

We now consider Latin Squares based on permutations.

Thus, let $p = 5$ and $\theta$ be the permutation given by

$$
\theta = \begin{pmatrix} 
1 & 2 & 3 & 4 & 5 \\
5 & 4 & 1 & 2 & 3 \\
\end{pmatrix}
$$

(3.28)

This means that $\theta$ changes 1 to 5, 2 to 4, etc. Then, $\theta(1) = 5, \theta(5) = 3 = \theta^2(1), \theta(3) = 1 = \theta^3(1)$. Thus, if $\alpha_0 = (1,2,3,4,5),$

$$
\theta(\alpha_0) = (5,4,1,2,3) \quad \theta^2(\alpha_0) = (3,2,5,4,1) \\
\theta^3(\alpha_0) = (1,4,3,2,5) \quad \theta^4(\alpha_0) = (5,2,1,4,3) \\
\theta^5(\alpha_0) = (3,4,5,2,1) \quad \theta^6(\alpha_0) = (1,2,3,4,5) = \alpha_0
$$

(3.29)
Definition 3.2. A permutation $\theta$ is called full cycle if

$$\theta^j(1) \begin{cases} 1, & \text{if } j = p \\ \neq 1, & \text{if } j < p \end{cases}$$ (3.30)

For example, if $p = 5$ and

$$\theta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \end{pmatrix}, \quad \theta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 2 & 1 \end{pmatrix},$$ (3.31)

then $\theta_2$ is full cycle, and $\theta_1$ is not.

Definition 3.3. A $p \times p$ LS is called a permutation LS (PLS) if it is of the form

$$[\alpha_0 : \theta(\alpha_0) : \theta^2(\alpha_0) : \cdots : \theta^{p-1}(\alpha_0)] = [\alpha_0 : \alpha_1 : \cdots : \alpha_{p-1}]$$ (3.32)

where $\alpha_0 = (1, 2, 3, \cdots, p)'$, and $\theta$ is some full cycle permutation of $\alpha_0$. Clearly, (3.32) is a LS if and only if $\theta$ is full cycle. Also, note that we have defined the PLS at (3.32) using $\alpha_0$ for the first column. We could equivalently define a PLS by using $\alpha_0'$ as the first row. Furthermore, a LS obtained by permuting rows and/or columns of a PLS is also defined to be a PLS.

As an example, for $\theta_2$ given by (3.31), we get the PLS:

$$\begin{bmatrix} 1 & 4 & 2 & 3 & 5 \\ 2 & 3 & 5 & 1 & 4 \\ 3 & 5 & 1 & 4 & 2 \\ 4 & 2 & 3 & 5 & 1 \\ 5 & 1 & 4 & 2 & 3 \end{bmatrix}$$ (3.33)

Theorem 3.4. For any PLS of size $(p \times p)$ we have (under Model I):

$$\text{MNIC} \leq (p^2 - 9)/8, \text{ for } p \text{ odd} ,$$ (3.34a)

$$\text{MNIC} \leq [p^2/8], \text{ for } p \text{ even} .$$ (3.34b)

Proof. Without loss of generality, we consider the $p \times p$ PLS of the form (3.32) with the first row $(1, 2, \cdots, p)$. Let $\mu = \alpha_1 = 0$, $\tau_l = -\beta_l$ for any $l$ and $q_{ij} = \beta_j + \tau_l$ if the treatment $\tau_l$ is applied to the cell $(i,j)$. Also, let $h(1, \cdots, p)$ be such that $\theta(h) = 1$, and let

$$\sum_{i=1}^{p} \beta_j = 0. \quad \beta_j - \beta_{\theta(j)} = \pi, \text{ for all } j \text{ such that } \theta(j) \neq 1 .$$ (3.35)
Then the first row of the LS is zero and \( \tau_i = \tau_{\beta i-1(j)} = -\beta_{\beta i-1(j)} \) if the treatment \( \tau_i \) is applied to the cell \((i, j)\). Also, we have

\[
\beta_h - \beta_1 = -(\beta_1 - \beta_h) = -\sum_{s=1}^{p-1} [\beta_{\theta s-1(1)} - \beta_{\theta s}(1)] = -(p-1)\pi,
\]

since \( h = \theta^{p-1}(1) \). Furthermore, consider \( u, v \) with \( 1 \leq u, v \leq p \), such that for some \( r, s \), with \( 0 \leq r, s \leq p-1 \), we have

\[
\begin{align*}
    u &= \theta^r(1), \\
    v &= \theta^s(1);
\end{align*}
\]

then we get

\[
\beta_u - \beta_v = (s - r)\pi
\]

For \( s > r \), (3.38) follows directly from (3.35), and for \( s < r \), we have \( \beta_u - \beta_v = (\beta_1 - \beta_u) - (\beta_1 - \beta_v) \), and (3.35) again gives the result. Now, consider, the value of \( (\beta_j - \beta_g) \), where treatment \( g \) has been applied to the cell \((i, j)\) of the LS. (Indeed, because we have a PLS, we know that \( g = \theta^{i-1}(j) \)). Let \( 1 = \theta^w(j) \). Then, (3.38) gives

\[
\beta_j - \beta_g = (i - 1)\pi, \text{ if } w > i - 1
\]

\[
= -(p - i + 1)\pi, \text{ if } 1 \leq w \leq i - 2.
\]

Since \( \tau_i = \beta_i \), for all \( \ell \), it is clear that the value of \( (\beta_j + \tau_g) \) (for \( g = g(i, j) \)), the treatment applied to cell \((i, j)\), for \( j = 1, \ldots, p \) is either \((i - 1)\pi\) or \(-(p + i - 1)\pi\). Now, notice that for \( i > 1 \), we can not have \((g = j)\), and every entry in the \( i \)th row is nonzero, being one of the two quantities at (3.39). Let

\[
h_0 = 1 = h_p, h_i = \theta^i(1), \text{ for } 1 \leq i \leq p - 1, h_{p-1} = h.
\]

Then, it is clear that for \( i > 1 \), the value of \( (\beta_j + \tau_g) \) (for \( g \) in \((i, j)\)) is \((i - 1)\pi\), in the cells \((i, h_0), (i, h_1) \cdots, (i, h_{p-1}), \) and \((-p + i - 1)\pi\) in the remaining cells.

When \( p \) is odd \((p = 2k + 1)\), we choose

\[
\alpha_i = -(i - 1)\pi, \quad 2 \leq i \leq (k + 1)
\]

\[
= (p - i + 1)\pi, \quad (k + 2) \leq i \leq (2k + 1).
\]

This choice of the \( \alpha \)'s, \( \beta \)'s and \( \tau \)'s, will thus lead to \((i - 1)\) nonzero cells in row \(i\) (for \(2 \leq i \leq k + 1\)), and \((p - i + 1)\) nonzero cells in row \(i\) (with \(k + 2 \leq i \leq 2k + 1\)). The total number of nonzero cells thus produced is

\[
0 + (1 + 2 + \cdots + (k - 1) + k) + (k + (k - 1) + \cdots + 2 + 1) = k(k+1)
\]

When \( p \) is even \((p = 2k)\), we take

\[
\alpha_i = -(i - 1)\pi, 2 \leq i \leq k + 1
\]

\[
= (p - i + 1), k + 2 \leq i \leq 2k.
\]
The total number of nonzero cells produced is
\[ 0 + (1 + 2 + \cdots + (k - 1)) + k + ((k - 1) + \cdots + 2 + 1) = k^2 \]  \hspace{1cm} (3.44)

Expressions (3.43) and (3.44) respectively lead us to the bounds (3.34a,b). This completes the proof.

For the special case when \( p = 3 \), it is easy to check that every LS is a PLS; for this case, the right side of (3.34a) reduces to zero, which shows that no NA cells are identifiable, a result which was first proved in Srivastava (1993).

**Theorem 3.5** For any \((p \times p)\) LS, under Model I, we have
\[ \text{MNIC} \leq (p - 1) \]  \hspace{1cm} (3.45)

Proof Without loss of generality, consider treatments 1 and 2. Suppose treatment 1 occurs in cells \((1, j_{11}), (2, j_{12}), \ldots, (p, j_{1p})\), and treatment 2 occurs in the cells \((1, j_{21}), \ldots, (p, j_{2p})\). Let \( \sum_{u=1,2} \) indicate the sum over the \( p \) cells \((i, j)\) in which treatment \( u \) occurs. Then, we have
\[ \sum_1 [E(y_{ij})] - \sum_2 [E(y_{ij})] = p(\tau_1 - \tau_2) + \lambda_1 - \lambda_2, \]  \hspace{1cm} (3.46)

where \( \lambda_u(u = 1, 2) \) is the sum of the \( \delta_{ij} \) (nonadditivity parameters) for those cells \((i, j)\) which are NA, and in which treatment \( u \) is applied. By taking \( \tau_1 = \tau_2 \), the right side of (3.46) equals \( (\lambda_1 - \lambda_2) \) which is a linear function of the NA parameters of at most \( 2p \) cells. Hence, comparing with (2.4), it is clear that we have
\[ \text{MNIC} < [2p/2], \]  \hspace{1cm} (3.47)

which leads to the results. This completes the proof.

Notice that the last result is similar to theorem 3.1 which is under model II.

For further results on model II for the LS design, for Youden Squares, and Lattice Squares, etc, the reader is referred to Wang (1997). These results give upper and/or lower bounds on MNIC (and, frequently, its exact value). The method \( R_1 \), of course, works for the purpose of identifying the NA parameters, under both models I and II. But, in certain cases, simpler methods (based on transversals etc.) are presented. For similar results under model I, reference is made to Srivastava (1997).

The main conclusion from this section is that under both models, the value of MNIC is quite small relative to the total number of cells in the design. Because of this, even if \( \sigma^2 = 0 \), confounding of NA parameters is possible in a relatively large number of cases. When \( \sigma^2 > 0 \), which is the case in real life, the situation is much worse, as we shall see later.
4 CORRECT IDENTIFICATION OF NA CELLS

We now study the case $\sigma^2 > 0$, which is the real life situation. All discussion is under model I. We consider a $(p \times p)$ LS. Suppose the LS has $m$ NA cells, where $m$ is defined by (1.23). Let the NA cells be $(i_1, j_1), \ldots, (i_m, j_m)$, and let $\delta(i_u, j_u) (u = 1, \ldots, m)$ be the corresponding NA parameters. For simplicity, we shall write $\delta_u$ for $\delta(i_u, j_u)$, and $\theta_u = \delta_u / \sigma$, for $u = 1, \ldots, m$.

The question now is this. Given the LS, the NA cells $(i_u, j_u)$, and the corresponding NA parameters $\theta_u (u = 1, \ldots, m)$, can we correctly identify all or some of the NA cells. Well, before answering this, one might wonder what is the need of identifying the NA cells. On the other hand, this latter point should now be quite clear. From the remark 1.1, and the discussion around (1.14), it follows that the estimate of $(\tau'_k - \tau'_l)$ is biased by the amount $(\delta'_k - \delta'_l)$. In view of this, the reader can easily construct examples to verify that this bias will cause quite a distortion in the results, and will usually defeat the purpose of the experiment. The estimate of $\sigma^2$ will also be inflated, and results otherwise significant will fail to appear to be so. Some studies on the effect of this are reported in the next section.

So, the idea behind the need of identifying the NA cells is simple. Notice that a NA cell, particularly one with a large value of $\delta / \sigma$, will tend to heavily distort the results; however, it otherwise gives no information on the $\tau$'s. At the best, from such a cell, we may estimate the corresponding $\delta$ in case the values of the corresponding $\alpha, \beta$, and $\tau$ (which correspond to the cell) is already estimable from other data. Thus, it is clear that if we could identify the NA cells, then we could ignore the observations in these cells, and get accurate estimates of the parameters of interest from the remaining cells.

Consider therefore, the problem of identification of the $m$ NA cells. In the last section, we gave some upper bounds on the number $m$, so that $m$ NA cells can be identified. Thus, we know that for a $4 \times 4$ LS, we have $\text{MNIC} = 1$ This means that if $\sigma^2 = 0$, and if $m = 1$, then we can identify the NA cell with certainty, and also estimate the various parameters of interest with variance zero. (As we shall presently see, when $\sigma^2 > 0$, the probability that the NA cell will be identified correctly, is less than 1. Of course, the estimates of various parameters, will have positive variance.) On the other hand, if $m = 2$, the Remark 2.2 applies. This means there do exist certain pairs of values of the parameters ($\alpha$'s, $\beta$'s, $\tau$'s, $\delta$'s), which will be confounded; as a result, if the $y$ is such that one of these parameter values is the correct one, then (even if $\sigma^2 = 0$), there will be more than one parameter value corresponding to this $y$, and the identification of the correct parameter value will not be possible. However, this confounding will arise, only if the correct parameter value happens to be a value which is confounded with some other set of parameter value; in case this does not happen, the NA cells will be correctly identified (with probability one (if $\sigma^2 = 0$) even if $m > \text{MNIC}$.

In this section, for $\sigma^2 > 0$, we consider the probability of correct identification of the NA cells. Such studies were first conducted for $p = 4$ and 5, with $m = 1$ in a totally theoretical manner (using the method $R_1$). The investigation uses ideas similar to that in Srivastava (1991), and in Shirakura et al (1997). Basically, one has to evaluate an intergal whose region
of integration is the intersection of a large number of quadratic regions; this makes the work exceedingly messy. Such theoretical studies are reported in Wang (1997). For identifying the NA cells, Wang (1995) compared $R_1$ with other methods available in the literature, and found $R_1$ to be best.

The above study (for $m = 1, p = 4, 5$, using $R_1$) was also conducted using simulation. We generated $n'$ samples; for each sample, the data was analysed using $R_1$. Then, the proportion of samples in which $R_1$ led to correct identification, was noted. The value of $n'$ was increased, until the theoretical and the simulation results matched very closely (for a variety of values of $\delta$). Thus, the (minimal) value of $n'$ which gives a reasonably close approximation to the theoretical value was determined. Later, in those cases (larger $p$, and/or larger $m$) where theoretical study was not feasible, we used sample sizes (suitably) larger then this minimal value of $n'$, so as to obtain stable results.

The results are reported in the following three tables (I, II, III), which deal respectively with $m = 1, 2$ and $3$. In each case, a cyclic LS (in which the first two rows are $(1, 2, \cdots, p)$, and $(p, 1, 2, \cdots, p - 2, p - 1)$) was used. The entries inside the tables are probabilities of correct identification of the NA cells.

Table I, $m = 1$

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$p = 4$</th>
<th>$p = 5$</th>
<th>$p = 6$</th>
<th>$p = 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.0962</td>
<td>0.0971</td>
<td>0.0980</td>
<td>0.0989</td>
</tr>
<tr>
<td>2.0</td>
<td>0.1972</td>
<td>0.2012</td>
<td>0.2048</td>
<td>0.2085</td>
</tr>
<tr>
<td>3.0</td>
<td>0.3558</td>
<td>0.3976</td>
<td>0.4254</td>
<td>0.4328</td>
</tr>
<tr>
<td>4.0</td>
<td>0.5430</td>
<td>0.6271</td>
<td>0.6721</td>
<td>0.6971</td>
</tr>
<tr>
<td>5.0</td>
<td>0.7178</td>
<td>0.8415</td>
<td>0.8626</td>
<td>0.8906</td>
</tr>
<tr>
<td>6.0</td>
<td>0.8497</td>
<td>0.9307</td>
<td>0.9568</td>
<td>0.9672</td>
</tr>
<tr>
<td>7.0</td>
<td>0.9311</td>
<td>0.9791</td>
<td>0.9910</td>
<td>0.9929</td>
</tr>
<tr>
<td>8.0</td>
<td>0.9728</td>
<td>0.9952</td>
<td>0.9981</td>
<td>0.9992</td>
</tr>
<tr>
<td>9.0</td>
<td>0.9907</td>
<td>0.9992</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
<tr>
<td>10.0</td>
<td>0.9973</td>
<td>0.9999</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
Table II, $m = 2$

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>$p = 4$</th>
<th>$p = 5$</th>
<th>$p = 6$</th>
<th>$p = 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.0, 1.5)</td>
<td>0.0138</td>
<td>0.0155</td>
<td>0.0164</td>
<td>0.0170</td>
</tr>
<tr>
<td>(2.0, 3.0)</td>
<td>0.0436</td>
<td>0.0925</td>
<td>0.1037</td>
<td>0.1055</td>
</tr>
<tr>
<td>(3.0, 4.5)</td>
<td>0.0992</td>
<td>0.2913</td>
<td>0.3490</td>
<td>0.3530</td>
</tr>
<tr>
<td>(4.0, 6.0)</td>
<td>0.1928</td>
<td>0.5638</td>
<td>0.6445</td>
<td>0.6790</td>
</tr>
<tr>
<td>(5.0, 7.5)</td>
<td>0.3134</td>
<td>0.7937</td>
<td>0.8582</td>
<td>0.8730</td>
</tr>
<tr>
<td>(6.0, 9.0)</td>
<td>0.4584</td>
<td>0.9174</td>
<td>0.9538</td>
<td>0.9640</td>
</tr>
<tr>
<td>(7.0, 10.5)</td>
<td>0.5704</td>
<td>0.9753</td>
<td>0.9908</td>
<td>0.9918</td>
</tr>
</tbody>
</table>

Table III, $m = 3$

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2, \theta_3)$</th>
<th>$p = 5$</th>
<th>$p = 6$</th>
<th>$p = 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.0, 1.5, 2.0)</td>
<td>0.0010</td>
<td>0.0020</td>
<td>0.0050</td>
</tr>
<tr>
<td>(2.0, 3.0, 4.0)</td>
<td>0.0530</td>
<td>0.0820</td>
<td>0.0930</td>
</tr>
<tr>
<td>(3.0, 4.5, 6.0)</td>
<td>0.1640</td>
<td>0.2880</td>
<td>0.3300</td>
</tr>
<tr>
<td>(4.0, 6.0, 8.0)</td>
<td>0.3420</td>
<td>0.6020</td>
<td>0.6360</td>
</tr>
<tr>
<td>(5.0, 7.5, 10.0)</td>
<td>0.5530</td>
<td>0.8500</td>
<td>0.8640</td>
</tr>
<tr>
<td>(6.0, 9.0, 12.0)</td>
<td>0.6480</td>
<td>0.9220</td>
<td>0.9560</td>
</tr>
</tbody>
</table>

To appreciate the results in the above tables, consider the case $m = 1, p = 4$. For the cases $\theta = 2$ and 3, the probabilities are respectively 0.197 and 0.356. Now, for a normal distribution, the probability of exceeding the mean by $\theta$ units is 95.5% and 99.7% for $\theta = 2$ and 3. Thus, it is clear that for correct identification, the probabilities jump from about 0.955 to 0.197, and from 0.997 to 0.356. This is quite discouraging. Indeed, the chance of correct identification is about 45% only when $\theta$ takes the excessively large value of about 7.5.

The probabilities increase as $p$ increases, but not very much. Indeed, this comparison is somewhat misleading, since as $p$ increases, it is almost certain that $m$ may increase rapidly, as a result of which the probability that the NA cells will be correctly identified will decrease rapidly. To see this, consider the average of the absolute value of the $\theta$'s, for $p = 6$. When this average is 4.0, for $m = 1$, the probability is 0.672. For $m = 2$, and the average equal
to 3.75, the probability is 0.349, and for $m = 3$ (with the average = 4.5), the probability is 0.288. This shows a rapid decreasing trend.

The moral from these tables is that unless the NA parameters are excessively large (and the value of $m$ is relatively small), there is not much hope of correctly identifying the NA cells.

We shall now look into a few examples of real life data from published books. The purpose is to see if nonadditivity really occurs often in real life, or we (the authors) are unnecessarily making a big issue out of it. We present the results of this study in Table IV below. The first column of this table gives the name of the author(s) of the books, which altogether are six in number. Except for Montgomery, from where we obtained four examples, only one example is taken from each other book.

For each example, the (published) data was re-analysed to see if nonadditivity is really present. From Tables I-III, it is clear that unless NA parameters are excessively large, it would be difficult to identify the NA cells. So, before the analysis of this published data, we (the authors) were not too hopeful that we would be able to demonstrate the existence of nonadditivity. Nethertheless, the analysis was conducted, and to our dismay, as we shall presently see, we found nonadditivity occurring too abundantly.

The methodology for the re-analysis of the data will be explained in detail with reference to the first example (from Bliss). The basic idea (from Srivastava (1996)) is as follows.

Let $U$ be a set of experimental units which are included in a given experiment under a design $D$, and suppose that the data from a unit $u$ are considered to obey a (linear) model $M$, for all $u \in U_0$, where $U_0$ is some subset of $U$, and the number of units in $(U - U_0)$ (which is the set of all units in $U$ which are not in $U_0$) is at most $m$, where $m$ is some known positive integer. Let $U_0^*$ be a subset of $U$ such that $|U - U_0^*|$ is a minimum (say, $m'$, where $0 \leq m' \leq m$). Note, that $U_0^*$ need not be unique. Let $Y$ denote the (whole) data from $U$. As explained in remark 2.1, if some of the units in $U$ do not obey $M$, then in general, confounding may occur. But we must emphasize that such confounding occurs with respect to specific values of parameters in $M$; and with respect to a given data set $Y$, there may be no such confounding, and $U_0^*$ may be unique and may be identifiable.

The procedure is as follows. We analyse $Y$ under $M$ using all units $U$, and obtain $s_e^2$, the (mean) sum of squares due to error. Next, we analyse $Y_{om'}$, where $Y_{om'}$ is the portion of $Y$ coming from the units in $U_{om'}$, where $U_{om'}$ is such that $|U - U_{om'}| = m'$, for $1 \leq m' \leq m$. The analysis of $Y_{om'}$ is, of course, done under $M$; let $s_e^2(U_{om'})$ denote the (mean) sum of squares due to error obtained from this analysis. Let $U_{om'}^*$ be such that

$$s_e^2(U_{om'}^*) = \min_{U_{om'}}(s_e^2(U_{om'})) \quad 1 \leq m' \leq m. \quad (4.1)$$

Define, for $1 \leq m' \leq m$,

$$\pi(m') = s_e^2(U)/s_e^2(U_{om'}) \quad (4.2)$$

28
Now, it is clear that if NA is present, then $s_e^2$ will be inflated. This is so, because the $\delta_{ij}$ are confounded with the $e_{ij}$. Thus, if $m'$ NA cells are present, then by removing these $m'$ cells, the error mean square should decrease appreciably. Hence, a large value of the ratio $\pi(m')$ is indicative of the fact that $U_{om'}^*$ contains NA cells.

The question now is this. For a given $m'$, how can we decide whether the observed ratio $\pi(m')$ is "too large". Let $\Pi(m')$ be the distribution function of $\pi(m')$, under the hypothesis $H_0$ that $m = 0$, i.e., there are no NA cells. Obviously, we need to know $\Pi(m')$. Let $\pi_0(m')$ be the upper 100$\alpha$ percent cut off point in the distribution $\Pi(m')$, i.e., for $0 \leq \alpha \leq 1$, we have

$$\text{Prob}\{\pi(m') > \pi_0(m') \mid H_0\} = \alpha$$

(4.3)

Then, we can reject $H_0$ at the 100$\alpha$% level, if $\pi(m') > \pi_0(m')$.

Unfortunately, the distribution of $\pi(m')$ would, in general, be difficult to obtain theoretically, since it will involve an integral in which the region of integration is the intersection of $\binom{n}{m'}$ quadratic regions, where $n$ is the total number of cells in the design. Thus, we have to take recourse to simulation.

Below, we give the results of the analysis of the data from nine different ($p \times p$) LS designs. Table IV gives the summary of the results. The first column gives the name of the author(s) from whose book the data are taken. The next six columns give the value of $p$ in each case, the value of $m$, the set of cells $U_{om}^*$, $s_e^2$, $s_e^2(U_{om}^*)$ and $\pi(m)$. The last column of the table gives an upper bound to the probability $\alpha_0$, where $\alpha_0(=\alpha_0(\pi(m)))$ is the probability that, under the distribution function $\Pi(m)$, the quantity $\pi(m)$ will be exceeded with probability $\alpha_0$.

**Table IV Analysis of Published Data**

<table>
<thead>
<tr>
<th>Author</th>
<th>$p$</th>
<th>$m$</th>
<th>$U_{om}^*$</th>
<th>$s_e^2$</th>
<th>$s_e^2(U_{om}^*)$</th>
<th>$\pi(m)$</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.I. Bliss</td>
<td>4</td>
<td>3</td>
<td>(1,1)(1,4)(4,4)</td>
<td>35.67</td>
<td>0.93</td>
<td>38.22</td>
<td>3.73 $\times$ 10^{-7}</td>
</tr>
<tr>
<td>Cochran &amp; Cox</td>
<td>6</td>
<td>4</td>
<td>(1,2)(4,6)(6,5)(6,6)</td>
<td>3.33</td>
<td>0.95</td>
<td>3.51</td>
<td>0.1252</td>
</tr>
<tr>
<td>D.R. Cox</td>
<td>6</td>
<td>3</td>
<td>(3,4)(5,3)(6,2)</td>
<td>1.62</td>
<td>0.71</td>
<td>2.28</td>
<td>0.4677</td>
</tr>
<tr>
<td>Evans &amp; Alldredge</td>
<td>6</td>
<td>3</td>
<td>(2,5)(5,4)(6,2)</td>
<td>0.66</td>
<td>0.29</td>
<td>2.28</td>
<td>0.4677</td>
</tr>
<tr>
<td>Montgomery, p157</td>
<td>5</td>
<td>3</td>
<td>(2,5)(3,2)(5,4)</td>
<td>10.67</td>
<td>2.83</td>
<td>3.77</td>
<td>0.1065</td>
</tr>
<tr>
<td>Montgomery, p173</td>
<td>4</td>
<td>3</td>
<td>(1,1)(3,2)(4,3)</td>
<td>1.75</td>
<td>0.17</td>
<td>10.29</td>
<td>5.14 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>Montgomery, p173</td>
<td>5</td>
<td>4</td>
<td>(1,1)(3,1)(3,3)(4,3)</td>
<td>3.13</td>
<td>0.13</td>
<td>24.08</td>
<td>9.39 $\times$ 10^{-6}</td>
</tr>
<tr>
<td>Montgomery, p466</td>
<td>6</td>
<td>3</td>
<td>(2,1)(3,4)(4,2)</td>
<td>9.90</td>
<td>5.06</td>
<td>1.96</td>
<td>0.6422</td>
</tr>
<tr>
<td>Stephen Senn</td>
<td>4</td>
<td>3</td>
<td>(1,2)(1,3)(4,4)</td>
<td>0.073</td>
<td>0.001</td>
<td>73.00</td>
<td>1.02 $\times$ 10^{-11}</td>
</tr>
</tbody>
</table>

How did we decide on the chosen value of $m$? This was done informally. We computed $s_e^2(U_{om})^*$, for $m' = 1, 2, 3$ and 4, and stopped when the value of $s_e^2(U_{om})^*$, ceased to have an
appreciable decrease by increasing $m'$. Values of $m' > 4$ were not tried, since the amount of computation was prohibitive.

As mentioned above, simulation was used to approximate $II(m)$ to a high degree of accuracy. Many studies were first made on the sample size needed for this purpose. It was found that a sample size of 4000 provided stable results. Below, in Table V, we give the $II(m)$ obtained by simulation for the $4 \times 4$ LS of Bliss, when $m = 3$, and the NA cells are $(1,1)$, $(1,4)$, and $(4,4)$. Columns 1 and 3 of the table indicate the value of $g$, and columns 2 and 4, the probability under $II(m)$ for the interval $(0,g)$. The values of $g$ increase in increments of 0.25. For $g > 9.75$, the change in the probability for intervals of this size was too small; they were, therefore, pooled.

Table V. The Distribution $II(m)$ for $p = 4$, and three NA Cells

<table>
<thead>
<tr>
<th>$g$</th>
<th>Prob</th>
<th>$g$</th>
<th>Prob</th>
<th>$g$</th>
<th>Prob</th>
<th>$g$</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>.0183</td>
<td>2.75</td>
<td>.5568</td>
<td>5.25</td>
<td>.8270</td>
<td>7.75</td>
<td>.9463</td>
</tr>
<tr>
<td>0.50</td>
<td>.0428</td>
<td>3.00</td>
<td>.6058</td>
<td>5.50</td>
<td>.8438</td>
<td>8.00</td>
<td>.9533</td>
</tr>
<tr>
<td>0.75</td>
<td>.0683</td>
<td>3.25</td>
<td>.6430</td>
<td>5.75</td>
<td>.8635</td>
<td>8.25</td>
<td>.9613</td>
</tr>
<tr>
<td>1.00</td>
<td>.1035</td>
<td>3.50</td>
<td>.6748</td>
<td>6.00</td>
<td>.8780</td>
<td>8.50</td>
<td>.9673</td>
</tr>
<tr>
<td>1.25</td>
<td>.1478</td>
<td>3.75</td>
<td>.7035</td>
<td>6.25</td>
<td>.8938</td>
<td>8.75</td>
<td>.9725</td>
</tr>
<tr>
<td>1.50</td>
<td>.2040</td>
<td>4.00</td>
<td>.7278</td>
<td>6.50</td>
<td>.9025</td>
<td>9.00</td>
<td>.9770</td>
</tr>
<tr>
<td>1.75</td>
<td>.2755</td>
<td>4.25</td>
<td>.7505</td>
<td>6.75</td>
<td>.9123</td>
<td>9.25</td>
<td>.9798</td>
</tr>
<tr>
<td>2.00</td>
<td>.3578</td>
<td>4.50</td>
<td>.7725</td>
<td>7.00</td>
<td>.9233</td>
<td>9.50</td>
<td>.9820</td>
</tr>
<tr>
<td>2.25</td>
<td>.4323</td>
<td>4.75</td>
<td>.7905</td>
<td>7.25</td>
<td>.9323</td>
<td>9.75</td>
<td>.9838</td>
</tr>
<tr>
<td>2.50</td>
<td>.4990</td>
<td>5.00</td>
<td>.8105</td>
<td>7.50</td>
<td>.9388</td>
<td>$\infty$</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The value of $\pi(m)$ for the Bliss data is 38.22, which is much beyond the range of the intervals in Table V. The corresponding upper bound ($3.73 \times 10^{-7}$) to the probability of exceeding this value was obtained using Tschebyscheff's inequality. The same procedure was adopted in many other cases where the probability is quite small.

We now consider the results in Table IV. It is to be noted first that the value of $\pi(m)$ obtained in the table would usually be an underestimate, which would become an excessively gross underestimate as the number of NA cells, and the magnitude of the NA parameters increased. The reason is this. Consider a $p \times p$ LS, and suppose the number of NA cells (under model I) is $m$, i.e., in (2.4), all elements of $\hat{\delta}$ are exactly zero except $m$ elements. It is clear that the quantity $\pi(m)$, as described above in this section, correspond precisely to this situation.

However, the rows and columns of the LS represent nuisance factors. In real life, there is no reason why the precise situation in the last paragraph (with exactly $m$ nonzero elements in $\hat{\delta}$) will hold. Indeed, the reverse should be true most of the time. In other words, all elements
of \( \delta \) would generally be nonzero. This will inflate both \( s^2_e \) and \( s^2_e(\overline{U}_{om}) \), so that the ratio would be expected to decrease, since most of the time, we shall have \( s^2_{0e} > s^2_{0e}(\overline{U}_{om}) \), where the subscript 0 indicates that the quantities correspond to the situation where \( m \) is strictly zero, i.e., where \( \delta = 0 \). (The reason for this that \( s^2_e \approx s^2_{0e} + s^2_{se} \) and \( s^2_e(\overline{U}_{om}) \approx s^2_{0e}(\overline{U}_{om}) + s^2_{se} \), where \( \approx \) denotes "approximately", and \( s^2_{se} \) is the contribution due to the \( \delta_{ij} \).)

Now, Table IV shows that in four cases out of nine, the probability is quite small, and thus a heavy amount of nonadditivity is present. In two out of the remaining five cases, the probability is of the order of about 11%. On the surface, this seems to indicate an additive model is present. However, as discussed above, the value of \( \pi(m) \) is probably too small, because of the large number of nonzero \( \delta \)'s. This may be true in the remaining three cases as well, where the LS is large and values of \( m > 4 \) were not tried. To elaborate, a 6 \times 6 LS may have 7 NA cells with relatively large \( \delta \)'s, and other cells with small \( \delta \)'s. Our method of detecting nonadditivity through the use of \( \pi(m) \) may then fail if we go only up to \( m = 4 \). The reason is that with \( m = 4 \), there are still three large \( \delta \)'s (and other moderate \( \delta \)'s) left, which would inflate both \( s^2_e \) and \( s^2_e(\overline{U}_{om}) \), resulting in a small value of \( \pi(m) \).

It is, therefore, clear that a small value of \( \pi(m) \) with correspondingly large probability of being exceeded, does not necessarily imply lack of nonadditivity. Thus, the situation in Table IV is that in four out of nine cases an excessive amount of nonadditivity is present, and in two cases, probably a moderate amount of the same exists. In the other three cases, we can not make a rational judgement based on the given data.

We do not wish to enter into a discussion here of the results of Table IV in the light of "Extreme Value Theory". But, readers familiar with the latter, may intuitively judge that a considerable amount of nonadditivity probably exists in most cases where cross-classified nuisance factors are present. (The reasoning behind this intuitive guess would run somewhat like this: If we go to a community, and pick nine people randomly, and four of them turn out to be multimillionaires, then probably it is safe to conclude that, in general, people in this community are quite rich.)

5 EFFECT OF NON-ADDITIVITY ON THE POWER FUNCTION

We now study the power function of the \( F \)-test (on treatment effects) under non-additivity. The following lemma is easy to check. We continue using the notation of sections 1 and 2.

Lemma 5.1 For any row-column design, under the notations of Section 1 and 2, we have

\[
\text{Exp}[S^2_e] = \begin{cases} 
(n - \nu_1)\sigma^2, & \text{if } \delta = 0 \\
(n - \nu_1)\sigma^2 + \delta'(I_n - A(A'A)^{-1}A')\delta, & \text{if } \delta \neq 0
\end{cases}
\]  

(5.1)
From this lemma, we see that non-additivity causes $S_e^2/(n - \nu_1)$ to be a biased estimate of $\sigma^2$, with the bias $\delta'(I_n - A' A)^{-1}A' \delta$. Of course, nonadditivity will grossly distort the estimate of treatment contrasts. Thus, there will be a great loss in the sensitivity of the experiment. We now consider the power function of the F-test.

Consider the hypothesis

$$H_0 : \quad C \xi = 0$$

(5.2)

where $C$ is a $(k \times \nu_1)$ matrix with $\text{Rank}(C) = k < \nu_1$. Let

$$S_e^2 = y' S_e^* y = \text{the sum of squares due to error},$$

$$S_h^2 = y' S_h^* y = \text{the sum of squares due to the hypothesis } H_0,$$

(5.3)

where $S_e^* = I_n - A'(A'A)^{-1}A'$ and $S_h^* = A'(A'A)^{-1}C'(C(A'A)^{-1}C')^{-1}C(A'A)^{-1}A'$. Then the F-statistic for testing $H_0$ is

$$F = \frac{S_h^2/k}{S_e^2/(n - \nu_1)}.$$  

(5.4)

**Lemma 5.2** For any row-column design, if $\delta \neq 0$, then

(1) $S_e^2 \sim \chi^2(n - \nu_1, \lambda_e)$,

(2) $S_h^2 \sim \chi^2(k, \lambda_h)$,

(3) $S_e^2$ and $S_h^2$ are independent.

where $\chi^2(\cdot, \cdot)$ is a non-central $\chi^2$ distribution, and

$$\lambda_e = \frac{1}{2} \delta' S_e^* \delta, \quad \lambda_h = \frac{1}{2} (A \xi + \delta)' S_h^* (A \xi + \delta).$$

(5.5)

Specifically, under $H_0$, we have

$$\lambda_h = \frac{1}{2} \delta' S_h^* \delta.$$  

(5.6)

**Proof** (1) and (2) can be proved by general linear model theory. To prove (3), we can directly check that $S_e^* S_h^* = 0$ (zero matrix).

**Theorem 5.1** For any row-column design,

if $\delta \neq 0$, then $F$ has a doubly non-central F distribution, and the power of the F-test is given by $P(F_0)$, where

$$P(F_0) = \text{Prob}\{F > F_0(\alpha, k, n - \nu_1)\},$$

(5.7)
where the quantity \( F_0(\alpha, k, n - \nu_1) \) is the upper \( \alpha \) probability point ((1 - \( \alpha \)) percentile) of a central \( F \) distribution with degrees of freedom \( k \) and \((n - \nu_1)\), and \( \alpha \) is the probability of Type I error. Also, we have

\[
P(F_0) = \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} \frac{(-\frac{1}{2} \lambda_h)^s}{s! t!} \left( \sum_{i=0}^{s} (-1)^{i+j} \binom{s}{i} \binom{t}{j} I_{u_0} \left[ k+j, \frac{1}{2} (n-\nu_1)+i \right] \right), \tag{5.8}
\]

where \( I_x(a,b) = \frac{1}{B(a,b)} \int_0^x u^{a-1}(1-u)^{b-1}du \) and where \( B(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1}du \) to be the Beta function.

**Proof** The result follows from Tiku (1974), page 142.

The function \( P(F_0) \) depends on \( \lambda_h \) and \( \lambda_k \), and hence on the number and locations of NA cells. Wang (1995) studied the \( 5 \times 5 \) cyclic LS design with one NA cell (1,1).

The LS considered is given by

\[
\begin{bmatrix}
1 & 2 & 3 & 4 & 5 \\
2 & 3 & 4 & 5 & 1 \\
3 & 4 & 5 & 1 & 2 \\
4 & 5 & 1 & 2 & 3 \\
5 & 1 & 2 & 3 & 4 \\
\end{bmatrix}
\tag{5.9}
\]

Two values of \( k(= 1, 2) \) were tried; the corresponding linear functions \( C\xi \), and the parameter values tried are indicated below.

For \( k = 1 \), \( C\xi = \tau_1 - \tau_2 \), and \( \tau_2 = -1.2 \) \tag{5.10a}

For \( k = 2 \), \( C\xi = \begin{bmatrix} \tau_1 - \tau_2 \\ \tau_3 - \tau_4 \end{bmatrix} \), and \( \begin{bmatrix} \tau_2 \\ \tau_4 \end{bmatrix} = \begin{bmatrix} -1.2 \\ 2.8 \end{bmatrix} \) \tag{5.10b}

Since there is only one NA cell, there is also one nonadditivity parameter, which we denote by \( \delta \). Let \( \theta = \delta / \sigma \). Values of the power are presented in Table VI below for 11 different values of \( \theta \), and two different values of \( C\xi \), for each value of \( k \) at (5.10).
\[
\begin{array}{|c|c|c|c|c|}
\hline
C_k & (-1.5) & (-2.0) & (-1.5,1.5) & (-2.0,2.0) \\
\hline
\theta = 0 & 0.3666 & 0.5601 & 0.4861 & 0.7310 \\
\theta = 1 & 0.2915 & 0.4732 & 0.4270 & 0.6805 \\
\theta = 2 & 0.2160 & 0.3746 & 0.3571 & 0.6098 \\
\theta = 3 & 0.1482 & 0.2741 & 0.2828 & 0.5222 \\
\theta = 4 & 0.0943 & 0.1831 & 0.2117 & 0.4242 \\
\theta = 5 & 0.0564 & 0.1110 & 0.1500 & 0.3251 \\
\theta = 6 & 0.0326 & 0.0612 & 0.1012 & 0.2344 \\
\theta = 7 & 0.0191 & 0.0310 & 0.0655 & 0.1591 \\
\theta = 8 & 0.0120 & 0.0152 & 0.0412 & 0.1020 \\
\theta = 9 & 0.0109 & 0.0105 & 0.0253 & 0.0625 \\
\theta = 10 & 0.0091 & 0.0067 & 0.0161 & 0.0397 \\
\hline
\end{array}
\]

The row for \((\theta = 0)\) in the above table corresponds to the case where there is no nonadditivity; this row shows the power that the experiment is expecting under the assumption of nonadditivity. The above table shows that if \(\theta\) is quite large, then even one NA cell can reduce the power to a negligible value. Even a moderate value of \(\theta\) \((= 3)\) causes the power to drop to less than 50\% when \(k = 1\). For \(k = 2\), we took both components of \(C_k\) to be equal in magnitude; even then the power drops to about 57\% and 75\% in the two cases (when \(\theta = 3\)). Using (5.8), the reader can compute the power for various other combinations of \(k\), values of \(C_k\), and the number of NA cells, and their location. The general picture emerging is quite dismal.

6 HAZARDS ASSOCIATED WITH ROW-COLUMN DESIGNS

In the previous sections we have investigated the effect of nonadditivity from many angles. In Section 1, we noted that the estimate of \((\tau_k - \tau_l)\) is biased by the amount \((\delta_k^* - \delta_l^*)\), which is an unknown quantity. Indeed, we would not know even the sign of \((\delta_k^* - \delta_l^*)\), which means that we would not know whether the estimates are biased upward or downward. Thus, the estimates of the various effects will be distorted, the amount and the direction of distortion being unknown. On the other hand, from Section 4, it appears that in most real life situations a great deal of nonadditivity might be present. Also, the last section shows that even one nonadditive cell could reduce the power of the \(F\)-test by a considerable amount. These results, therefore, establish the fact that in most real situations where a row column design is used, much of the inferences drawn might be misleading.

It is also clear that if the situation is generally additive, except for a set of \(m\) cells, then one way to handle it will be to, if possible, identify the NA cells, and ignore them while doing
the analysis of the remaining data. However, for this to be possible, the results of Section 3 tell us that \( m \) should be relatively quite small. This difficulty is even further compounded by the message from the first part of Section 4, which tells us that unless the NA parameters are quite large, the probability of identifying the NA cells is pathetically small.

In scientific work, cross-classified nuisance factors arise quite often. Because of this, the field of row-column designs has been active since its inception in the 1930's. The possibility of the presence of nonadditivity, particularly in large Latin Squares, was apprehended by Fisher and Yates (1948), who pointed out that this may invalidate the analysis of the experimental data, which is based on the additive linear model. This remark was picked up by one of the authors (Srivastava), who (in a conversation with R.C. Bose in October 1959) deemphasized large sized Latin Squares from the statistical angle. The same point of view was taken for general row-column designs when the paper of Kiefer (1975) was being processed. This also led to the emphasis on, and some theory for, the nested multidimensional block designs, in Srivastava (1977, 1981, 1986). In 1981, Srivastava and Kiefer planned to systematically attack the nonadditivity problem, but the project did not start because of Kiefer's sudden death. Only in 1990, Srivastava began the investigation, which were published in Srivastava (1993). Some alarming conclusions were reached in this study, leading to the present investigation. During the three decades (1960-90), though the idea of nonadditivity and its dangers was there, it never occurred that it will turn out to be so significant.

With respect to nonadditivity, the situation in a row-column design can be categorized into six cases, threat of the distortions getting progressively worse as we move from one case to the next. We briefly describe these.

The first category is where the model is totally additive. We believe that the probability that such a model arises in nature is very small. Also, recall from section 1, that even under this model, the estimate can be biased. The next two cases correspond to the situation where the \( \delta \)'s are all small, except for \( m \) NA cells (which correspond to large \( \delta \)'s) For case II, \( m \) is small, and for case III, \( m \) is moderate. Cases IV and V are similar to the last two except that the \( \delta \)'s of the \( m \) "bigger" NA cells are not quite so large. Case VI is where there is a large number of \( \delta \)'s of moderate or large sizes.

In cases I, II and III, identification is possible with a large probability, which decreases significantly in case III. In cases IV, V and VI, identification is not possible, and \( \sigma^2 \gets \) inflated, the situation getting progressively worse, as we go from IV to V to VI.

Also, we believe that in nature, cases IV, V, VI occur frequently, II and III occasionally, and I only rarely.

Leaving case I aside, only in case II, do we have a reasonably high chance of identifying the NA cells, which would lead to them being ignored in the analysis, thus freeing the results adequately from bias. In case III, the above may be possible only one-half or one-third of the time. It is difficult to estimate how often these cases arise in nature, since no studies are available. Our intuitive guess is that the proportion of cases where identification of "large"
NA cells is possible, would not exceed 20% in LS designs of size \( p \times p \), with \( p \neq 4 \). For \( p = 4 \), the proportion may be somewhat larger.

We conclude, therefore, that in general, the use of a row-column design is hazardous in the sense that the results may get too distorted. In the next section, we discuss possible remedies against this trouble.

7 EXPERIMENTAL DESIGNS UNDER TWO OR MORE NUISANCE FACTORS

We shall limit our discussion to two \((p \times q)\) cross-classified nuisance factors. However, the ideas presented can be easily generalized to larger numbers.

There are \( pq \) cells, corresponding to the different level-combination of the nuisance factors. Two cases arise according as (I) it is possible to take two or more observations in each cell, and (II) we can take only one observation per cell.

In case I, the situation is much easier to handle. In general, each cell can be considered as a block, and the usual incomplete block designs (with a simple block system) can be used. Below, we discuss in more detail, the subcase Ia, where in each cell exactly two observations are taken, and \( p = q \neq 6 \).

By the famous Bose-Shrikhande-Parker theorem, two m.o.l.s. (mutually orthogonal Latin squares) of size \( p \times p \) exist (if \( p > 2, p \neq 6 \)). We take the two squares and superimpose them, thus creating 2 observations per cell. There will be exactly \( p \) cells, in which the two treatments assigned are identical. One may ignore these \( p \) cells, which will give a design with \( r = 2(p-1) \), where \( r \) is the number of replications. We prefer this to the full design with \( r = 2p \) if there are no treatment-cell interactions, since \( r \) will be somewhat smaller in the former case.

Let \( L \) and \( L^* \) be the two m.o.l.s., and let \( y_{ij} \) and \( y_{ij}^* \) respectively be the observations in the \((i,j)\) cells in the two squares, satisfying the model (1.6), so that, in an obvious notation, we have

\[
\begin{align*}
y_{ij} &= \mu + \alpha_i + \beta_j + \tau_k + \delta_{ij} + \epsilon_{ij} \\
y_{ij}^* &= \mu + \alpha_i + \beta_j + \tau_{k^*} + \delta_{ij} + \epsilon_{ij}^* \tag{7.1}
\end{align*}
\]

where treatments \( k \) and \( k^* \) are applied in the \((i,j)\) cell respectively in the two squares, and where \( \epsilon_{ij}, \epsilon_{ij}^* \sim \text{i.i.d. } N(0, \sigma^2) \). Let

\[
\begin{align*}
z_{ij} &= y_{ij} - y_{ij}^* \\
w_{ij} &= y_{ij} + y_{ij}^* \tag{7.2}
\end{align*}
\]

\[h_{ij} = \epsilon_{ij} + \epsilon_{ij}^*\]
so that \( g_{ij} \) and \( h_{ij} \) are i.i.d. \( N(0, 2\sigma^2) \) variables. Then, it is clear that for all \((i, j)\), we have
\[
E(z_{ij}) = \tau_k - \tau_{k^*},
\]
(7.3)
which is free of the \( \delta \)'s as well as \( \mu, \alpha \)'s, and \( \beta \)'s. Also,
\[
E(w_{ij}) = 2(\mu + \alpha_i + \beta_j + \delta_{ij}) + (\tau_k + \tau_{k^*}).
\]
(7.4)
Recall that we have
\[
\sum \alpha_i = \sum \beta_j = \sum \tau_k = 0.
\]
(7.5)
Let \( w_i, w_j, \) and \( w_\cdot \) respectively be the means of \( w_{ij} \) with respect to \( j, i \), and both \( j \) and \( i \), and let
\[
w_{ijo} = w_{ij} - w_i - w_j + w_\cdot.
\]
(7.6)
Then
\[
E(w_{ijo}) = 2\delta_{ij} + (\tau_k + \tau_{k^*}),
\]
(7.7)
since \( L_1 \) and \( L_2 \) are Latin Squares. Now
\[
\tau_k = \frac{1}{p} \sum_{k' = 1}^{p} (\tau_k - \tau_{k'}) ,
\]
(7.8)
so that \( \tau_k \) and \( \tau_{k^*} \) are estimable by the \( z \)'s. Hence, using (7.3) and (7.7), an unbiased estimate of the \( \delta_{ij} \) can be obtained. The \( z \)'s provide \( (p^2 - p + 1) \) degrees of freedom for error when \( r = 2p \), and \( (p^2 - p - p + 1) \) d.f. when \( r = 2(p - 1) \). Using these, \( \sigma^2 \) can be estimated accurately.

When a good estimate \( (\hat{\sigma}^2, \mathrm{say}) \) of \( \sigma^2 \) is obtained (from the \( z \)'s) one may proceed to analyse the \( \delta_{ij} \), which are the estimates of the \( \delta_{ij} \) obtained as above. The problem is to find out what is the magnitude of the \( \delta \)'s relative to \( \sigma \). We advocate probability plotting and similar procedures, the details of which still need to be worked out. If all the \( \delta \)'s are of a similar magnitude, further analysis can be done using appropriate random effect models, where the models relate to all cells. In case some \( \delta \)'s are large, the corresponding cells may be ignored, and a random effects model considered for the remaining cells which are more "homogeneous" with respect to the \( \delta \)'s. Finally, after fitting such models further information about the \( \tau \)'s may be retrieved. These types of investigation will be reported elsewhere.

It should be emphasized that using the random effects model is just one approach for extracting whatever information the \( w \)'s may have regarding the \( \tau \)'s. Indeed, much further research is needed to elaborate in what situations there would be any appreciable gain in the information on \( \tau \)'s by using the \( w \)'s.

When more than 2 observations per cell are possible, one may consider simply imposing more than two m.o.l.s. Again, we do not necessarily need to consider all cells. The procedure would be similar to the above, and we shall not discuss the same here.
The question arises whether or not one should use incomplete blocks all the time under cases IV, V and VI. The answer seem to depend on the relative values of the \( \delta \)'s, and needs further investigation. Case I is already clear. In cases II and III, we can work with one observation per cell, because it is usually possible to identify and hence ignore the NA cells.

We now discuss the situation where we are able to take only one observation in any cell. Suppose, through prior studies or otherwise, we are convinced that we are in cases I, II, or III. In our opinion, case I will be very rare, and when we really believe we are in case I, we may actually be under case II, and similarly case III may actually seem like case II in practice. For this situation, we should use the ordinary row-column design, and after the data are collected, analyse it with an effort to identify the NA cells. Of course, this work will be done under the Search Linear Model (SLM). Below, we discuss model I given by (2.4).

Under the SLM, if there are \( m \) NA parameters, then we need to estimate the parameters in \( \xi \) plus \( m \) parameters out of \( \delta \). Also, we do not know which \( m \) elements of \( \delta \) are nonzero. Consider a fixed set of \( m \) elements of \( \delta \); under D-optimality, we need to estimate these \( m \) elements plus the \( v_1 \) elements of \( \xi \), through a design which maximizes the information matrix, corresponding to these \( (v_1 + m) \) parameters. But, under the SLM, the set of \( m \) elements of \( \delta \) is not fixed. For this case, the concept of ordinary D-optimality is not applicable, and Srivastava (1977) introduced AD-optimality. This criterion, denoted by \( K \) below equals the average of the determinants of the \( \binom{n}{m} \) information matrices, each of size \( (v_1 + m) \times (v_1 + m) \), where these matrices correspond to the different possible choices of the \( m \) parameters out of the \( n \) parameters in \( \delta \).

Below, we shall restrict ourselves to the special case when all the \( \binom{n}{m} \) choices are assumed equally likely. For this case, under model (2.1), Srivastava (1977) proved the following result.

**Theorem 7.1.** (Srivastava, 1977). For any design \( T \), under the search linear model (2.1), we have

\[
K = c_1 c_2(t),
\]

\[
c_1 = |A_1'A_1| \tag{7.9}
\]

\[
c_2(t) = (-1)^t \sum_{l_1, \ldots, l_t} \frac{(-1)^{b_1 + \cdots + b_t}}{(l_1!)(2b_1!l_2!) \cdots (t^b l_t!)} \prod_{i=1}^t \text{trace}(GA_2A_2^t)^{l_i} \tag{7.10}
\]

\[
G_1 = I_n - A_1(A_1'A_1)^{-1}A_1' \tag{7.11}
\]

and where \( t \) is a general positive number which equals to \( 2m \) under the criteria of AD-optimality and \( (l_1, l_2, \ldots, l_t) \) satisfies

\[
l_1 + 2l_2 + \cdots + tl_t = t, \quad l_i \geq 0 \quad (i = 1, \ldots, t) \tag{7.12}
\]
We now apply the above theorem to the case where, in (2.1), we have

\[ n = v_2, A_2 = I_n \]  

(7.14)

**Theorem 7.2** Under (2.1) and (7.12), if a design is D-optimal when \( m = 0 \), then it is \( AD \)-optimal for all \( m \).

**Proof** Let \( D_1 \) and \( D_2 \) be two designs, and let \( A_{11}, A_{12}, A_{21}, A_{22} \) be the matrices \( A_1, A_2 \) under the model (2.1) for these two designs. Suppose \( D_1 \) is D-optimal when \( m = 0 \). Then

\[ |A'_{11}A_{11}| \geq |A'_{21}A_{21}| \]  

(7.15)

Note that since (7.14) holds, we have

\[ G_1A_{12}A'_{12} = G_1, \quad G_2A_{22}A'_{22} = G_2 \]  

(7.16)

where \( G_1 \) and \( G_2 \) are the \( G \)-matrices under \( D_1 \) and \( D_2 \). Also, we have (for all \( i \)),

\[ G_i^2 = G_i \]  

(7.17)

since \( G_i \) is idempotent. Hence, for all \( i \),

\[ \text{trace } G_i^2 = \text{trace } G_i = (n - v_i) \]  

(7.18)

Similarly, for all \( i \),

\[ \text{trace } G_i^2 = (n - v_i) \]  

(7.19)

Thus, in (7.11), the quantity \( c_2(t) \) is the same for the two designs \( D_1 \) and \( D_2 \). Hence, using (7.9) and (7.15), we find that the value of \( K \) for \( D_1 \) is larger than or equal to the value of \( K \) for \( D_2 \), where strict inequality holds if and only if it holds in (7.15). This completes the proof of the theorem.

**Theorem 7.3** Consider an experiment with two cross-classified nuisance factors each at \( p \) levels. Then, under (2.4), the \( (p \times p) \) LS design is \( AD \)-optimal for all permissible values of \( m \).

**Proof** We know that when \( m = 0 \), i.e. under the usual additive model, the LS design is D-optimal, a result proved first by Wald. Then, the result follows in view of the fact that we are working under the model (2.4) to which theorem 7.2 is applicable.

The above result shows that when \( p = q \) under cases I and II, and often under III, we shall be in good shape if we use the LS design.

We now consider cases IV, V and VI, under the assumption that only one observation per cell is allowed. This is the most difficult situation. For this, two approaches, \( A \) and \( B \), will be offered.
Approach A suggests the usage of "smaller" sized rectangles. Lattice Squares, or more generally, the NMBD (Nested Multidimensional Block Design) (Srivastava (1977), (1981), Srivastava and Beaver (1986)), are designs which belong to this category. The word "smaller" is, of course, relative. Thus, instead of using one $(9 \times 9)$ LS, one could use nine $(3 \times 3)$ squares. One $(9 \times 9)$ LS uses $(8+8 =)16$ parameters corresponding to the rows and columns, compared with $((2+2) \times 9 =)36$ parameters used by nine $3 \times 3$ squares. Thus, the number of parameters gets increased by $(36 - 16 =)20$, which would tend to absorb a lot of variability in the $\delta$'s if the $(9 \times 9)$ square is used. In other words, if we use the smaller squares, the $\delta$'s corresponding to these may be quite small in general. If there are a few large $\delta$'s, the others being small, the SLM approach could be used for identification of NA cells.

It may be noted here that going from a large rectangle to smaller ones (using an NMBD) has another big advantage. It is that in the latter cases, the number of analyses to be done (using $R_i$), is much smaller, and is in the practical range, compared to that in the former (which may be practically prohibitive, even though theoretically possible).

Approach B asks for studies prior to the conduct of the main experiment. Such a study may be called a Uniformity Trial, in which only one treatment is used on all units. Again, two situations (say, $b_1$ and $b_2$) arise accordingly as the set of nuisance factors have respectively fixed or random effects. Now, a uniformity trial would give an estimate of the $\delta_{ij}$ for each cell $(i, j)$. It would, of course, show what case we are under. Then, if $b_2$ holds, all this tells us is that such is the set of values of the $\delta$'s this time, without throwing light on what may happen next time (when the main experiment is done). If $b_1$ holds, then, of course, we know which cells correspond to large $\delta$'s, so that in the main experiment, we may ignore these. The authors believe that, even under $b_1$, the value of the $\delta$ (for any given cell) will change from one experiment to the next, so that the uniformity trial will tell us mainly which ones are the large NA cells.

Nevertheless, uniformity trails are a must, and should be performed whenever possible. They will give insight into which of the cases I-VI, we may fall under.

On the whole, when only one observation per cell is possible, the use of the NMBD's appears to be the safest possible strategy. This appears to be more prudent than the idea of using an exorbitantly large number of replications with the hope that eventually, the $\delta$'s may cancel out to some extent. Being smaller, the former is less expensive. It is also more practical, and as $s$ consequence, will entail fewer other problems (whose occurrence increases with the number of units).

References


