Robustness of Diallel Cross Designs Against the Loss of Any Number of Observations (Crosses) in a Block

Krishan Lal and T. Jeisobers

*Indian Agricultural Statistics Research Institute, New Delhi*

(Received : January, 2002)

SUMMARY

The robustness of universal optimal binary balanced block (BBB) designs for diallel crosses has been investigated for the loss of any $t \geq 1$ observations (crosses) in a single block as per connectedness and A-efficiency criteria. It is observed that all the BBB designs for diallel crosses with the minimum eigenvalue strictly greater than 2 are robust as per connectedness criterion. Further A-efficiency of the residual design in comparison to original design, except the BBB design for diallel crosses with 4 lines, 3 blocks each of size 2, is also computed for the loss of any $t \geq 1$ crosses in a block. The designs which are not robust as per A-efficiency criterion are also identified.

Key words: Connectedness, Diallel cross, A-efficiency, Robustness, Universal optimality, Binary balanced block designs.

1. Introduction

The diallel cross is a type of mating design used to study the genetic properties of a set of inbred lines. Suppose there are $p$ inbred lines and it is desired to perform a diallel cross experiment involving $p(p - 1)/2$ crosses of type $(i \times j)$ for $i < j, i, j = 1, 2, \ldots, p$. This is a type IV mating design of Griffing [13]. The problem of obtaining the optimal block designs for complete diallel cross experiments have been studied by Gupta and Kageyama [11], Dey and Midha [5], Das et al. [2], Parsad et al. [17]. All these studies have been made for the situations where the experimenter is interested in estimating the general combining ability (gca) effects and the specific combining ability (sca) effects have been excluded from the model. A catalogue of block designs for diallel crosses has also been presented by Parsad et al. [17]. These optimal designs perform well under ideal conditions. However, disturbances may occur during the experimentation. The loss of observation(s) during the experimentation is one such aberration and because of this an optimal design may become
non-optimal design. Therefore, it is necessary for an experimenter to investigate whether the design adopted by him is robust for the loss of observation(s) or not. Following Ghosh [8], a block design for diallel crosses (d) is termed as robust against the loss of observations if the resulting design (d(t)) obtained after the loss of t (≥1) observation(s) remains connected. This criterion of robustness is called as connectedness criterion.

Many times design remains robust as per connectedness criterion but the A-efficiency of the residual design in comparison to original design may fall considerably. So, another criterion of robustness is A-efficiency criterion. A connected design is said to be robust as per A-efficiency criterion against the loss of any t (≥1) observation(s) in a block if the A-efficiency of residual design in comparison to original design is sufficiently large, i.e., it is at least 0.95. In A-efficiency the term A- is used because the efficiency is based on average variance of the design.

The robustness of block designs against missing data has been investigated by John [14], Baksalary and Tabis [1], Dey and Dhall [4], Kageyama [15], Gupta and Srivastava [12], Dey [3], Dey et al. [6], Srivastava et al. ([20], [21], [22]), Lal et al. [16], among others. In the context of block designs for diallel cross Ghosh and Desai ([9], [10]) investigated the robustness of complete diallel cross plans, obtained by taking all possible crosses of lines of a balanced incomplete block (BIB) designs and Singular Group Divisible designs, due to unavailability of a complete block. Dey et al. [7] studied the robustness of block designs for diallel cross against loss of one observation and one complete block for binary balanced block designs. No other work seems to have been done on the robustness of designs for diallel cross against missing data. In the present article we shall investigate the robustness of block designs for diallel cross for the loss of any t (≥1) observations in a block as per connectedness criterion in Section 2 and as per A-efficiency criterion in Section 3.

Throughout the present investigation, we shall denote an n-component vector of all unities by 1n, an identity matrix of order n by In, and an m × n matrix of all ones by Jm,n. Jm,n is simply denoted by Jm. Further, A’, ξ(A), A− and A+ will respectively denote the transpose, column space (range), a generalized inverse (g-inverse) and Moore-Penrose inverse of matrix A.

2. Robustness as per Connectedness Criterion

Let d be a block design for a diallel cross experiment of the type mentioned in Section 1 involving p inbred lines, b blocks such that the jth block is of size kj, j = 1, ..., b. For the data obtained from the design d, we postulate the model

\[ Y = μ1_n + Δ' g + D' β + e \]  
(2.1)
where Y is n x 1 vector of observations, \( \mu \) is general mean effect, \( \mathbf{g} \) and \( \beta \) are vector of p gca effects and b block effects, respectively. \( \Delta' \) and \( \mathbf{D}' \) are the corresponding n x p and n x b design matrices, respectively i.e. the (s, t)\(^{th} \) element of \( \Delta' \) is 1 if the s\(^{th} \) observation pertaining to t\(^{th} \) line and is zero otherwise. Also, the (s, t)\(^{th} \) element of \( \mathbf{D}' \) is 1 if the s\(^{th} \) observation pertaining to t\(^{th} \) block and is zero otherwise, \( \mathbf{e} \) is the random vector follows \( \mathcal{N}(0, \sigma^2 \mathbf{I}_n) \).

\[
\mathbf{n} = \sum_{j=1}^{b} k_j \text{ the total number of experimental units in the design.}
\]

Suppose n observations are so ordered that the first \( k_1 \) come from the first block, the next \( k_2 \) come from the 2\(^{nd} \) block, and so on, the last \( k_b \) come from the b\(^{th} \) block. The n x b matrix \( \mathbf{D}' \) can be partitioned as \( [\mathbf{D}_1, \ldots, \mathbf{D}_b]' \) where \( \mathbf{D}_j' \) is a \( k_j \times b \) matrix with j\(^{th} \) column of all one's, other columns 0's and \( \mathbf{D}' = \sum_{j=1}^{b} \mathbf{J}_{k_j} \). Using the similar partitioning we can write the matrix \( \Delta' = [\Delta_1, \ldots, \Delta_b]' \), where \( \Delta_j \) is a \( k_j \times p \) matrix.

Under the model 2.1, the C-matrix (coefficient matrix) for reduced normal equations for estimating linear functions of gca effects using design \( \mathbf{d} \) is

\[
\mathbf{C} = \Delta' (\mathbf{I} - \mathbf{D}' \mathbf{D}^{-1}) \mathbf{D}' \Delta = \mathbf{G} - \mathbf{N} \mathbf{K}^{-1} \mathbf{N}' \]

where \( \mathbf{G} \) is the matrix of order \( p \times p \) with diagonal elements as replication of the p-inbred lines and off-diagonal elements are the number of replications of the crosses in the design \( \mathbf{d} \), \( n_{ij} \) is the number of times line i occurs in block j, \( \mathbf{K} = \text{diag}(k_1, \ldots, k_0) \). For an universal optimal binary balanced block design, the information matrix simplifies to

\[
\mathbf{C}_p = \theta \left( \mathbf{I}_p - \mathbf{p}^{-1} \mathbf{J}_p \right) \tag{2.2}
\]

where \( \theta = 2 \left( p - 1 \right)^{-1} (n - b) \) is the unique non-zero eigenvalue of \( \mathbf{C}_p \).

Let \( \mathbf{d}_v \) be the connected and universal optimal binary balanced block design for diallel cross, then its C-matrix is as given in (2.2) and Rank (\( \mathbf{C}_p \)) = \( p - 1 \). Suppose any t observations are lost in the design \( \mathbf{d}_v \) then the model of t missing observations will be \( \mathbf{A} \mathbf{Y} = \mu \mathbf{A} \mathbf{I}_n + \Delta \mathbf{A} \mathbf{g} + \mathbf{A} \mathbf{D}' \mathbf{\beta} + \mathbf{A} \mathbf{e} \) \tag{2.3}

where \( \mathbf{A} = \mathbf{I}_n - \mathbf{U} \) is an idempotent matrix as \( \mathbf{U} \) is a diagonal matrix of order \( n \) with 1 at the i\(^{th} \) diagonal position if i\(^{th} \) observation is missing and 0, otherwise. The C-matrix under (2.3) will be

\[
\mathbf{C}_{p(t)} = \Delta \mathbf{A} \left[ \mathbf{I} - \mathbf{X}(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \right] \mathbf{A} \Delta' \tag{2.4}
\]

where \( \mathbf{X} = [\mathbf{A} \mathbf{I}_n \quad \mathbf{A} \mathbf{D}' \quad \mathbf{A} \mathbf{X}] \) and \( (\mathbf{X}' \mathbf{X})^{-} = \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & (\mathbf{DAD}')^{-1} \end{bmatrix} \)
On solving, we get
\[ C_p(t) = \Delta A \Delta' - \Delta A D' (D A D')^{-1} D A \Delta' \]
because \( A = A' = A A \) \hspace{1cm} (2.5)

Now we consider the model by devoting an extra parameter to each missing observation
\[ Y = \mu I + \Delta' g + D' \beta + U \gamma + \epsilon \]
where \( \gamma \) is an \( n \times 1 \) vector of unknown parameters. The usual C-matrix under this model is
\[ C_s = \Delta [I - X(X' X)^{-1} X'] \Delta' \]
where \( X = [1 D' U] \). By applying \( U = U' = U U \), it simplifies to
\[ C_s = \Delta A \Delta' - \Delta A D' (D A D')^{-1} D A \Delta' \]
(2.7)

Thus the C-matrices under model of \( t \) missing observations and the model (2.6) is same. Henceforth, we shall use the C-matrix under model (2.6) for the \( t \) missing observations in the design. Now we derive the relation between the C-matrix of the original design \( d_p \) and the design of \( t \) missing observation i.e. \( d_{p(t)} \).

**Lemma 2.1.** \[ C_p = C_a + Q C_a Q' = C_{p(0)} + Q C_a Q' \]
where \( Q = \Delta U, C_a = U \phi U \) and \( \phi = I - D' K^{-1} D = \text{diag} (\phi_1, ..., \phi_b) \), and \( \phi_j = I_{k_j} - k_j^{-1} J_{k_j} \), \( j = 1, ..., b \).

**Proof:** The proof can easily be derived after having
\[
\begin{bmatrix}
K & DU \\
UD' & U
\end{bmatrix}^{-1} = \begin{bmatrix}
K^{-1} + K^{-1} D U D^{-1} & K^{-1} D U C_a C_a^{-1} \\
K^{-1} D U C_a C_a^{-1} & C_a^{-1}
\end{bmatrix}
\]

\[ C_s = \Delta \phi A' - \Delta \phi U C_a U \phi A' \] and \( C_p = \Delta \phi A' \)

**Theorem 2.1.** The design \( d_p \) under model (2.1) is robust as per connectedness criterion against the loss of any \( t(\geq 1) \) observations iff \( I_n - C_a^{-1/2} Q' (C_p)_{+}^{-1} Q C_a^{-1/2} \) is positive definite.

**Proof:** From Lemma 2.1, \( C_p = C_{p(0)} + Q C_a Q' \) and \( \xi(Q) \subset \xi(C_p) \). Using Theorem 1 of Lal et al. [16], we have Rank \( (C_p) \) = Rank \( (C_{p(0)}) \) iff \( I_n - C_a^{-1/2} Q' (C_p)_{+}^{-1} Q C_a^{-1/2} \) is positive definite and also we know that Rank \( (C_p) \) = \( p - 1 \). Hence the theorem is proved.

Since it is not easy to confirm the positive definiteness of a matrix so we shall derive the sufficient condition for the robustness which is easy for the search of robust designs.

This theorem can also be stated as the design is robust iff all the eigenvalues of \( C_a^{-1/2} Q' C_p^{-} Q C_a^{-1/2} \) are strictly less than 1 or if \( \lambda_{\max} (C_a^{-1/2} Q' C_p^{-} Q C_a^{-1/2}) \) is less than one. \( \lambda_{\max} (A) \) is the maximum eigenvalue of matrix \( A \).
Thus, $\lambda_{\text{max}} \left( C_{p}^{-1/2} Q' C_{p}^{-} Q C_{p}^{-1/2} \right) = \lambda_{\text{max}} \left( C_{p}^{-1/2} Q' C_{p}^{+} Q C_{p}^{-1/2} \right)$

$$= \lambda_{\text{max}} \left( Q C_{p}^{+} Q' C_{p}^{+} \right) < 1$$ \hspace{1cm} (2.8)

If $A$ and $B$ are non-negative definite matrix, then it is known (Marshal and Olkin [18]) that $\lambda_{\text{max}} (A B) \leq \lambda_{\text{max}} (A) \lambda_{\text{max}} (B)$ \hspace{1cm} (2.9)

So, from (2.8) and (2.9)

$$\lambda_{\text{max}} \left( Q C_{p}^{+} Q' C_{p}^{+} \right) \leq \lambda_{\text{max}} (C_{p}^{+}) \lambda_{\text{max}} (Q C_{p}^{+} Q')$$

Also $\lambda_{\text{max}} (C_{p}^{+}) = \{\lambda_{1} (C_{p})\}^{-1}$ where $\lambda_{1} (A)$ is the minimum eigenvalue of $A$. From this we have the following result:

Theorem 2.2. The design $d_{p}$ under model (2.1) is robust as per connectedness criterion against the loss of any $t \geq 1$ observations if the minimum eigenvalue of information matrix of the original design $d_{p}$ is strictly larger than maximum eigenvalue of $Q C_{p}^{+} Q'$.

Thus the eigenvalue of $Q C_{p}^{+} Q'$ are needed for investigating the robustness against the loss of any $t \geq 1$ observation(s) in a block of the design $d_{p}$, which we have considered a binary design. For this, the following cases arises:

**Case 1.** When $t$ missing observations are less than the block size i.e. $t < k_{1}$

Without any loss of generality, let the missing observations pertain to the first $t$ crosses and the $j^{th}$ block, in which these observations are lost, is the first block of size $k_{1}$. The $n \times n$ matrix $U$ can be rewritten as $U = \text{diag}(U_{1}, \ldots, U_{b})$ where $U_{j}$ is a $k_{1} \times k_{1}$ matrix with 1 at the diagonal position corresponding to the missing observation, is zero otherwise.

Thus, $U_{j} = \begin{bmatrix} 1 & \ldots & 0 \\ \ldots & \ddots & \ldots \\ 0 & \ldots & 1 \end{bmatrix}, U_{j} = 0 \forall j = 2, \ldots, b \ (t < k_{1})$ \hspace{1cm} (2.10)

$$C_{*} = U \phi U = \text{diag} (U_{1} \phi_{1} U_{1}, \ldots, U_{b} \phi_{b} U_{b})$$

and $$Q = (\Delta_{1} \phi_{1} U_{1}, \ldots, \Delta_{b} \phi_{b} U_{b})$$

where $\Delta_{i}, i = 1, \ldots, p$ and $\phi_{j}, j = 1, \ldots, b$ are as above. Because of (2.10) $C_{*} = U \phi U = \text{diag} (U_{1} \phi_{1} U_{1}, 0, \ldots, 0)$ where $k_{1} \times k_{1}$ matrix
\[
U_1 \phi_i U_1 = \begin{bmatrix}
I_t - k_i^{-1}J_t & 0 \\
0 & 0 
\end{bmatrix}
\]

So, \( C^{-1} = (U \phi_i U)^{-1} = \text{diag}([[U_1 \phi_i U_1]^{-1}, 0, \ldots, 0]) \) where

\[
(U_1 \phi_i U_1)^{-1} = \begin{bmatrix}
I_t + (k_i - t)^{-1}J_t & 0 \\
0 & 0 
\end{bmatrix}
\]

Also, \( Q = (\Delta_i \phi_i U_t, \ldots, \Delta_0 \phi_0 U_b) \). As the design \( d \) a binary design, so, \( \Delta'_i = [I_{k_i} \otimes I_{j}^* : 0] \) because \( p \geq 2 k_j \forall j = 1, 2, \ldots, b \)

Here \( A \otimes B \) is the kronecker product of matrix \( A \) and \( B \). Since \( \phi_j = I_{k_j} - k_j^{-1}J_{k_j} \)

So, \( p \times k_i \) matrix \( \Delta_i \phi_i U_1 = \begin{bmatrix} H \otimes I_2 \\ 0 \end{bmatrix} \)

where \( H = \begin{bmatrix}
I_t - k_i^{-1}J_t & 0 \\
-k_i^{-1}J_{k_i - t, t} & 0 
\end{bmatrix} \)

Thus the \( p \times p \) matrix \( Q C^{-1} Q' = \Delta_i \phi_i U_1 (U_1 \phi_i U_1)^{-1} U_1 \phi_i \Delta_i' \)

\[
= \begin{bmatrix} M \otimes J_2 & 0 \\ 0 & 0 \end{bmatrix}
\]

where \( k_1 \times k_1 \) matrix \( M = \begin{bmatrix}
I_t - k_i^{-1}J_t & -k_i^{-1}J_{k_i - t, t} \\
-k_i^{-1}J_{k_i - t, t} & (k_i(k_i - t))^{-1}J_{k_i - t} 
\end{bmatrix} \)

The matrix \( M \) is an idempotent matrix and \( \text{Rank}(M) = t \), \( t < k_1 \).

Case 2. When \( t \) missing observations are equal to block size i.e. \( t = k_1 \)

Here \( U_1 = I_{k_1} U_1 \phi_i U_1 = \phi_i, (U_1 \phi_i U_1)^{-1} = \phi_i \)

So \( p \times p \) matrix \( Q C^{-1} Q' = \Delta_i \phi_i U_1 (U_1 \phi_i U_1)^{-1} U_1 \phi_i \Delta_i' \)

\[
= \begin{bmatrix} \phi_i \otimes J_2 & 0 \\ 0 & 0 \end{bmatrix}
\]

Also, the matrix \( \phi_i \) is an idempotent matrix and \( \text{Rank}(\phi_i) = k_1 - 1 \). From Searle [19], we have the following result:
Corollary 2.1: The eigenvalues of \( A \otimes B \) are the product of eigenvalues of matrix \( A \) with those of \( B \).

It is known that eigenvalues of an idempotent matrix are \( 1 \) with multiplicity equal to its rank and there are only two eigenvalues of \( J_k \) i.e. \( 2 \) and \( 0 \). Now combining both the cases \( 1, 2 \) and Corollary 2.1 we have the following result:

**Corollary 2.2.** The positive eigenvalues of \( Q \ C \cdot Q' \) are

- \( 2 \) with multiplicity \( t \) if \( 1 \leq t \leq k_1 - 1 \)
- \( 2 \) with multiplicity \( k_1 - 1 \) if \( t = k_1 \)
- \( 0 \) otherwise

Thus the maximum eigenvalue of \( Q \ C \cdot Q' \) for the loss of any number of observations is 2. We have the following result:

**Theorem 2.3.** The design \( d \) is criterion 1 robust against the loss of any \( t \) \((1 \leq t \leq k_1)\) observations in a block if the smallest eigenvalue of information matrix \( C_p \) of design \( d_p \) is strictly larger than 2.

**Corollary 2.3.** The 102 universally optimal binary block designs for diallel cross given by Gupta and Kageyama [11], Dey and Midha [5], Das et al. [2], Parsad et al. [17] for block size up to 15 and inbred lines up to 30 are studied. All the designs are found to be robust as per connectedness criterion except the design with \( p = 4, b = 3, k = 2, n = 6 \).

### 3. Robustness as per A-efficiency Criterion

In Section 2, we have concentrated on robustness in terms of connectedness of residual design. However, even if a design is robust according to connectedness criterion, the design may not be robust as per A-efficiency criterion because of poor efficiency of residual design in comparison to original design. It is therefore, necessary to investigate the robustness as per A-efficiency of residual design in comparison to original design. Let \( d_p \) be the design as defined in Section 2 and \( d_{p(t)} \) be the residual design when any \( t \) observations are lost in any of the block of \( d_p \). Then the A-efficiency of residual design with respect to original design is given by:

\[
E = \frac{\text{Sum of reciprocals of the non-zero eigen-values of } C_p}{\text{trace}[C_p]} - \frac{\text{trace}[C_p^*]}{\text{Sum of reciprocals of the non-zero eigen-values of } C_{p(t)}^*\text{trace}[C_{p(t)}^*]}
\]

where \( C_p \) (\( C_{p(t)} \)) is the information matrix of \( d_p \) (\( d_{p(t)} \)).

The result given by Lal et al. [16] in Section 4.1 for binary balanced block design can be used for the model (2.1) and is as follows:
Theorem 3.1. The A-efficiency $E$ of the residual design $d_{p(t)}$ when any $t$ observations are lost in comparison to original design $d_p$ is

$$E = [1 + (p - 1)^{-1} \sum_{i=1}^{m} e_i (\theta - e_i)^{t-1}]^{-1}$$

where $0 > e_1 \geq e_2 \geq \ldots \geq e_m$ be the $m(\leq t)$ positive eigenvalues of matrix $Q \Sigma^T$, $Q'$, $\theta$ is the unique non-zero eigenvalue of $C_p$ with multiplicity $p - 1$ as given in (2.2). Now for $m = t$, when $t < k_1$, the A-efficiency

$$E_1 = [1 + 2t \{ (p - 1) (\theta - 2) \}^{-1}]^{-1}$$

On solving, we get

$$E_1 = (n - b - p + t + 1)^{-1} (n - b - p + 1)$$

because $\theta = 2 (n - b) (p - 1)^{-1}$

Again when $t = k_1$, the multiplicity of eigenvalues is $k_1 - 1$.

So A-efficiency will be

$$E_2 = (n - b - p + k_1)^{-1} (n - b - p + 1)$$

The A-efficiency for all the universally optimal binary balanced block designs for the loss of $t = 2, 3$ and $4$ observations in a block, in comparison to original design, have been worked out and the designs for which A-efficiency is less than 0.95 i.e. designs which are not robust, are presented in Table-I. The A-efficiency for the loss of a single observation and a complete block is available in Dey et al. [7].

| Table 1. List of designs for which A-efficiency is less than 95% |
|----------------------|----------------------|----------------------|----------------------|
| $p$ | $b$ | $k$ | $n$ | Efficiency | Factor | $t = 2$ | $t = 3$ | $t = 4$ |
| 6 | 5 | 3 | 15 | 0.7143 | - | - |
| 6 | 15 | 3 | 45 | 0.9259 | - | - |
| 7 | 7 | 3 | 21 | 0.8000 | - | - |
| 8 | 7 | 4 | 28 | 0.8750 | 0.8235 | - |
| 9 | 9 | 4 | 36 | 0.9048 | 0.8636 | - |
| 10 | 9 | 5 | 45 | 0.9310 | 0.9000 | 0.8710 |
| 11 | 11 | 5 | 55 | 0.9444 | 0.9189 | 0.8947 |
| 12 | 11 | 6 | 66 | * | 0.9362 | 0.9167 |
| 13 | 13 | 6 | 78 | * | 0.9464 | 0.9298 |
| 14 | 13 | 7 | 91 | * | * | 0.9420 |

Note: Efficiency for the loss of $t = 1$ and $t = k$ are not shown here as these are available in Table-2 of Dey et al. [7].

* denotes that the efficiency is $\geq 0.95$. 
ACKNOWLEDGEMENT

The authors are grateful to the referee for valuable suggestions that have led to considerable improvement in the presentation.

REFERENCES


