Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Conditions for connectivity of incomplete block designs

J.D.Godolphin

[1]Department of Mathematics, University of Surrey, Guildford, UK

Correspondence
Email: j.godolphin@surrey.ac.uk

Abstract

In experimental situations where observation loss is common, it is important for a design to be robust against breakdown. For incomplete block designs, with one treatment factor and a single blocking factor, conditions for connectivity and robustness are developed using the concepts of treatment and block partitions, and of linking blocks. Lower bounds are given for the block breakdown number in terms of parameters of the design and its support. The results provide guidance for construction of designs with good robustness properties.

KEYWORDS:
block breakdown number, observation loss, robustness

1 | INTRODUCTION

Incomplete block designs, with v treatments allocated to b blocks of size k, are widely used in industrial experiments. The aim of such an experiment is to determine whether the data depend on the specific treatment applied and to obtain estimates of treatment contrasts. A design which enables these aims to be realised, that is, a design which yields estimates of all treatment contrasts, is described as *connected*. Otherwise, the design is *disconnected*. See Butz (1982)[4] for early work on design connectivity.

It is common for observations to be be lost during experimentation. The extent of observation loss will depend on the particular situation. Thus, when planning an experiment, in addition to checking properties of the planned design, it is prudent to investigate the properties of potential eventual designs that can arise in the event of some observation loss.

There has been considerable investigation on the robustness of incomplete block designs against missing observations. For early work, see Ghosh (1979[8], 1982[9]). Dey (1993)[5] introduces criteria for assessing design robustness. In particular, Dey defines a design as being robust against the loss of $b_\dagger$ blocks, according to Criterion 1, if all treatment contrasts are estimable from any eventual design resulting from the loss of $b_\dagger$ blocks. Conditions for assessing design robustness, based on treatment replicate numbers, block sizes and treatment concurrences, have been developed by Baksalary and Tabis (1987)[2], Godolphin and Warren

(2011)[12] and Godolphin and Godolphin (2015)[11], amongst others. A key concept is the block breakdown number, which is the smallest number of blocks that can be removed so there is at least one eventual design from which not all treatment contrasts are estimable. The breakdown number is similarly defined and relates to the loss of individual observations, rather than whole blocks. See Latif *et al.*[15] (2009), Bailey *et al.* (2013)[1] and Tsai and Liao(2013)[16].

When planning an experiment, practical constraints can have the consequence that incomplete block designs with advantageous properties such as balance, high efficiency and equal replication are not available. For example, limited resources may mean that b is too small to accommodate designs known to have good properties. Alternatively, factors such as geographic location or technical expertise varying between sites might make the grouping of some treatments in a block more economical or practical than others. This can result in designs with $b < \binom{v}{k}$ having repeated blocks, or designs with $b \geq \binom{v}{k}$ in which some sets of k treatments do not appear together in a block. The set of distinct blocks in a design is defined as the *support*, and the number of distinct blocks as the *support size*. The literature on designs with repeated blocks focuses mainly on balanced designs. Foody and Hedayat (1977)[7] and Hedayat and Li (1979)[13] give algorithms for constructing balanced incomplete block designs with common b, v and k but with different support sizes. Hedayat and Pesotan (1985)[14] provide a general study on balanced incomplete designs and the support. Recent work on balanced incomplete designs includes Dobcsányi *et al.* (2007)[6], who derive bounds for the multiplicities of repeated blocks and give constructions for designs with repeated blocks. Other work focuses on t-designs. See for example Behbahani *et al.* (2008)[3]. There appears to be no work investigating the relationship between Criterion 1 robustness and properties of the design support for balanced designs or indeed for incomplete block designs in general.

The motivation for this work is to develop conditions on connectivity and Criterion 1 robustness for incomplete block designs, that do not depend on the planned design having properties such as balance. The pattern of observation loss focussed on is the loss of whole blocks. However, the bounds and conditions developed for robustness in the event of the loss of whole blocks also apply if the pattern of observation loss involves individual observations: for example any connected design which is robust against the loss of $b_\dagger$ blocks, according to Criterion 1 will also be robust against the loss of $b_\dagger$ individual observations. Designs with blocks of size two are a special case when considering alternative types of observation loss. For such designs the loss of a single observation is equivalent to the loss of the whole block since no information can be gained from a single observation in a block. Bailey *et al.* (2013)[1] and Godolphin (2018)[10] use methods from graph theory to investigate designs with $k = 2$. Other work on designs with $k = 2$ is found in Tsai and Liao (2013)[16].

The paper is structured as follows. The model is introduced in Section 2 along with notation and concepts on connectivity and robustness. In Section 3, lower bounds for the block breakdown number are obtained in terms of parameters of the design and its support. Sufficient conditions are given for connectivity, and lower bounds are provided for the block breakdown number. The bounds on the block breakdown number in Section 3 also apply to the breakdown number. The results are illustrated by

examples to aid clarity. However, the primary aim of the work is to provide guidance for the construction of designs with good robustness properties, rather than to obtain properties of specific designs.

2 | PRELIMINARIES

Let D denote a binary incomplete block design on υ treatments applied to the experimental units arranged in b blocks of size k. An additive model is assumed with the $bk \times 1$ observation vector Y specified by

$$E(Y) = \mu 1_{bk} + X_1\beta_1 + X_2\beta_2, \tag{1}$$

where μ is a scalar constant, the $bk \times 1$ vector with all elements unity is denoted 1_{bk}, and β_1 and β_2 are $\upsilon \times 1$ and $b \times 1$ parameter vectors relating to treatments and blocks respectively. The design matrix for (1) is given by $X = [1_{bk}\ X_1\ X_2]$. Here X_1 and X_2 are the $bk \times \upsilon$ and $bk \times b$ components of the design matrix pertaining to treatments and blocks. Each row of X_1 has one element unity and $\upsilon - 1$ zeros and each row of X_2 has one element unity and $b - 1$ zeros. Treatment replication numbers are given in decreasing order by $r_{[1]}, \dots r_{[\upsilon]}$, with

$$b \geq r_{[1]} \geq r_{[2]} \geq \cdots \geq r_{[\upsilon]} \geq 1.$$

The number of distinct blocks in D, the support size, is denoted by d, thus $d \leq b$. The number of copies of each of block in D is given by

$$n_{[1]} \geq n_{[2]} \geq \cdots \geq n_{[d]},$$

where $\Sigma_{i=1}^{d} n_{[i]} = b$. The sub-design comprising a copy of each of the d distinct blocks of D is denoted by D^{sup}. Treatment replication numbers in D^{sup} are $s_{[1]}, \dots s_{[\upsilon]}$, with

$$\min\left\{\binom{\upsilon-1}{k-1}, d\right\} \geq s_{[1]} \geq s_{[2]} \geq \cdots \geq s_{[\upsilon]} \geq 1. \tag{2}$$

The condition $s_{[1]} \leq \binom{\upsilon-1}{k-1}$ is justified as follows. Consider any treatment: the sets of $k-1$ treatments occurring with this treatment in the blocks of D^{sup} are all distinct sets from the remaining $\upsilon-1$ treatments. The number of such sets is $\binom{\upsilon-1}{k-1}$. Thus no treatment can occur in more than $\binom{\upsilon-1}{k-1}$ blocks of D^{sup}. If D has no repeated blocks then $b = d \leq \binom{\upsilon}{k}$ and $n_{[i]} = 1$, for $i = 1, \dots, b$, and $r_{[j]} = s_{[j]}$ for $j = 1, \dots, \upsilon$. Otherwise, $b > d$ and $r_{[j]} \geq s_{[j]}$ for all $j \in \{1, \dots, \upsilon\}$ with $r_{[j]} > s_{[j]}$ for at least one $j \in \{1, \dots, \upsilon\}$.

If D is connected, all treatment contrasts are estimable. Conversely, if D is disconnected, some treatment contrasts are not estimable. The feature of treatment allocation that results in a disconnected incomplete block design can be described using the concepts of treatment and block partitions. For a design satisfying model (1) and also having $\upsilon \geq 2k$, a treatment partition, $\{\mathcal{V}_1, \mathcal{V}_2\}$, is defined as an arrangement of the treatments into disjoint non-empty sets, $\mathcal{V}_1$ and $\mathcal{V}_2$, of sizes u and $\upsilon - u$ with $u \in U$, where $U = \{k, \dots, [\upsilon/2]\}$ and [.] denotes *the integer part of*. A block partition, $\{\mathcal{B}_1, \mathcal{B}_2\}$, is a partitioning of the blocks into

disjoint non-empty sets $\mathcal{B}_1$ and $\mathcal{B}_2$. A disconnected design is characterised by the existence of consistent treatment and block partitions so that all replicates of treatments in $\mathcal{V}_i$ occur in blocks in $\mathcal{B}_i$, for $i = 1, 2$. By comparison, a connected design will contain at least one *linking block* for any treatment partition, that is, at least one block which contains treatments from both $\mathcal{V}_1$ and $\mathcal{V}_2$. The notation $\Omega(D)$ is used for the smallest number of linking blocks for a treatment partition in design D. The notions of treatment and block partitions and of linking blocks are useful in obtaining conditions on connectivity and bounds for the block breakdown number. For a design with $v < 2k$, there are no treatment partitions since any treatment partition would need to contain at least k treatments in each of disjoint sets $\mathcal{V}_1$ and $\mathcal{V}_2$. It follows immediately that all designs with $v < 2k$ are connected.

Example 1 Design $D1$ is given by

$$D1 = \begin{array}{cccccccccccc} 1 & 1 & 1 & 1 & 3 & 3 & 3 & 5 & 2 & 2 & 2 & 4 \\ 3 & 5 & 8 & 8 & 5 & 7 & 8 & 7 & 4 & 4 & 4 & 6 \\ 7 & 7 & 10 & 10 & 10 & 8 & 10 & 10 & 6 & 6 & 9 & 9 \end{array}$$

where, as with all designs displayed in this work, columns correspond to blocks. This design has $v = 10$, $k = 3$, $b = 12$ and $d = 10$. Design $D1$ is disconnected: the treatment partition, $\mathcal{V}_1 = \{2, 4, 6, 9\}$, $\mathcal{V}_2 = \{1, 3, 5, 7, 8, 10\}$ is consistent with the block partition $\mathcal{B}_1$, $\mathcal{B}_2$, where the last four blocks are contained in $\mathcal{B}_1$ and the first eight in $\mathcal{B}_2$. Contrasts in treatment effects involving treatments in $\mathcal{V}_1$ are estimable as are those in $\mathcal{V}_2$. However, a pairwise treatment difference involving a treatment from each of $\mathcal{V}_1$ and $\mathcal{V}_2$, such as the difference in effect of treatments 1 and 2, is inestimable.

Example 2 Design $D2$ has support $D2^{sup}$:

$$D2^{sup} = \begin{array}{cccccccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 3 & 4 & 4 \\ 2 & 3 & 3 & 3 & 3 & 5 & 5 & 6 & 3 & 6 & 5 & 6 \\ 6 & 5 & 5 & 6 & 9 & 7 & 7 & 9 & 9 & 7 & 7 & 9 \\ 10 & 8 & 9 & 8 & 10 & 9 & 10 & 10 & 10 & 9 & 9 & 10 \end{array}.$$

Design $D2$ has parameters $v = 10$, $k = 4$ and $d = 12$ and is connected. No treatment partition is consistent with a block partition. Thus, for any treatment partition $\{\mathcal{V}_1, \mathcal{V}_2\}$, there is at least one linking block in $D2$ (and also in $D2^{sup}$), containing treatments from both $\mathcal{V}_1$ and $\mathcal{V}_2$. Design $D2$ will be returned to in a subsequent example.

An eventual design realized after the loss of one or more blocks from a planned design, D, is denoted $D_\#$ and the class of eventual designs resulting from the loss of every set of p blocks is denoted $D(p)$. The block breakdown number, b_0, gives the smallest number of blocks that need to be lost before the possibility of an eventual design from which not all treatment contrasts are estimable. Thus, a disconnected design has $b_0 = 0$ and a connected design has $b_0 > 0$. Every design in $D(p)$, for $p < b_0$, is a connected design in v treatments, but for $p \geq b_0$ there will be at least one member of $D(p)$ from which some treatment contrasts will be inestimable. If some treatment contrasts are not estimable in $D_\#$ then one or both of the following must be true: all replicates of a treatment are missing; at least one treatment partition exists in $D_\#$ for which there are no linking blocks. It

follows immediately that all designs with $v < 2k$ have $b_0 = r_{[v]}$. Henceforth, it will be assumed that $v \geq 2k$. The two causes of inestimable treatment constrasts leads to:

Definition 1. The block breakdown number of design D is given by

$$b_0 = \min\{r_{[v]}, \Omega(D)\}. \tag{3}$$

In Example 1, the blocks of $D1$ are ordered, so that the disconnected nature of the design can be easily recognised, with blocks of $\mathcal{B}_1$ arranged together and similarly blocks of $\mathcal{B}_2$ arranged together. In practice, it can be challenging recognise that a design is disconnected, and even more challenging to determine if a connected design has a low value of b_0. Obtaining an upper bound for b_0 is straightforward, even before allocation of treatments to experimental units. Theorem 1 is given by Bailey *et al.* (2013)[1].

Theorem 1. For any design, $b_0 \leq [bk/v]$.

Proof. At least one treatment has replication $[bk/v]$ or smaller. Thus $r_{[v]} \leq [bk/v]$ and the result follows directly from (3). □

The problem of planning a design to ensure that b_0 exceeds a desirable value is challenging. For $p < r_{[v]}$, an eventual design $D_\#$ in $D(p)$ is an incomplete block design in v treatments and with $b - p$ blocks of size k. The design matrix of $D_\#$ is denoted by $X_\# = [1_{(b-p)k}\ X_{1\#}\ X_{2\#}]$ and is obtained by removal of kp rows from X and p columns from X_2. As with the planned design, $D_\#$ is disconnected iff there is a partition of the treatments into $\mathcal{V}_1$ and $\mathcal{V}_2$ which is consistent with a partitioning of the blocks of $D_\#$ into sets $\mathcal{B}_{\#1}$ and $\mathcal{B}_{\#2}$, so that all replicates of treatments in $\mathcal{V}_i$ occur in blocks in $\mathcal{B}_{\#i}$ for $i = 1, 2$.

3 | CONNECTIVITY AND ROBUSTNESS CONDITIONS IN TERMS OF SUPPORT PARAMETERS

In this section, sufficient conditions are developed to guarantee design connectivity. A second focus is the derivation of lower bounds for $\Omega(D)$, and hence b_0. The conditions and bounds are based on various levels of information on D and are given in terms of the design parameters: $v, k, d, n_{[i]}$ and $s_{[i]}$. Results are illustrated by reference to specific examples. However, the main value of the results is to provide guidance in planning an experiment, to facilitate construction of designs that have a level of robustness appropriate to the perceived risk of observation loss associated with the experimental situation. The first result depends only on parameters v, k, d.

Theorem 2. Consider the class of designs with given v, k and d, where $d \leq \binom{v}{k}$. Then a sufficient condition for all designs in the class to be connected is given by

$$d > \binom{v-k}{k} + 1. \tag{4}$$

Proof. Assume that the design class contains a disconnected design, D, and consider the sub-design D^{sup}. A treatment partition which is consistent with a block partition exists in D, *iff* the same treatment partition is consistent with a block partition in D^{sup}. Therefore D^{sup} is disconnected and there is a treatment partition $\{\mathcal{V}_1, \mathcal{V}_2\}$, where $\mathcal{V}_1$ contains u treatments with $k \leq u \leq \upsilon/2$ and $\mathcal{V}_2$ contains the remaining $\upsilon - u$ treatments, which is consistent with a partition, $\{\mathcal{B}_{1S}, \mathcal{B}_{2S}\}$, of the d blocks of D^{sup}. Thus, in D^{sup} all replicates of treatments from $\mathcal{V}_i$ occur in blocks in $\mathcal{B}_{iS}$ for $i = 1, 2$. Since all blocks in $\mathcal{B}_{1S}$ are distinct and each contains k members of the u treatments of $\mathcal{V}_1$, set $\mathcal{B}_{1S}$ contains at most $\binom{u}{k}$ blocks. Similarly, $\mathcal{B}_{2S}$ contains at most $\binom{\upsilon-u}{k}$ blocks, giving

$$d \leq \binom{u}{k} + \binom{\upsilon - u}{k}. \tag{5}$$

A straightforward algebraic argument establishes that the right hand side of (5) is maximised when $u = k$. Thus the assumption that D is disconnected implies that $d \leq 1 + \binom{\upsilon-k}{k}$. Therefore the result follows by contraposition. □

Within the class of designs for given υ, k and d, Theorem 2 establishes whether or not all designs in the class are connected. For a design with no repeated blocks, Theorem 2 is expressed as follows.

Corollary 1. Consider the sub-class of designs with given υ, k and d, where $d = b \leq \binom{\upsilon}{k}$. A sufficient condition for all designs in the sub-class to be connected is given by

$$b > \binom{\upsilon - k}{k} + 1.$$

Example 3 Design $D3$ is given by

$$D3 = \begin{matrix} 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 5 & 5 \\ 2 & 4 & 5 & 7 & 3 & 5 & 7 & 7 & 4 & 6 & 6 & 6 & 6 & 6 & 7 \end{matrix}$$

blocks. This design has $\upsilon = 7$, $k = 2$ and $d = 12$. By Theorem 2, any design with these parameters is connected.

Example 4 Consider the class of designs with $\upsilon = 8$, $k = 4$, $d = 10$. By Theorem 2, all designs in the class are connected. The class of designs of Example 4 will be returned to once further results have been established.

For design classes, specified by υ, k, d, satisfying Theorem 2, a lower bound for $\Omega(D^{sup})$ is given by $\alpha = d - 1 - \binom{\upsilon-k}{k}$. Thus, for any treatment partition, there are at least α linking blocks in D^{sup}. A lower bound for $\Omega(D)$ can be determined with knowledge of the $n_{[i]}$ values.

Theorem 3. Let design D satisfy the condition of Theorem 2. Let $\alpha = d - 1 - \binom{\upsilon-k}{k}$ and $\omega_\alpha = \Sigma_{i=d-\alpha+1}^{d} n_{[i]}$. Then $\Omega(D) \geq \omega_\alpha$.

Proof. Design D is connected by Theorem 2. In general, a set of $\Sigma_{i=d-x+1}^{d} n_{[i]}$ blocks of D is the smallest set that can contain all copies of x distinct blocks. A lower bound for $\Omega(D^{sup})$ is given by α. The smallest number of blocks of D that can contain all copies of α distinct blocks is $\Sigma_{i=d-\alpha+1}^{d} n_{[i]}$. Thus, $\Omega(D) \geq \omega_\alpha$ as required. □

For a design with no repeated blocks, Theorem 3 simplifies to:

Corollary 2. Let D be such that $d = b$ and $d > 1 + \binom{v-k}{k}$. Then $\Omega(D) \geq \alpha$.

Example 4 revisited All designs in the class with $v = 8$, $k = 4$ and $d = 10$ have $\alpha = 8$ and $\omega_\alpha = \Sigma_{i=3}^{10} n_{[i]}$ by Theorem 3. Thus, for any design in the class and any treatment partition $\{\mathcal{V}_1, \mathcal{V}_2\}$, the design support contains at least eight linking blocks and the design itself contains at least $\Sigma_{i=3}^{10} n_{[i]}$ linking blocks. Hence, $b_0 \geq \min\{r_{[v]}, \Sigma_{i=3}^{10} n_{[i]}\}$.

The bound on d to guarantee connectivity of D and the bound on $\Omega(D)$, achieved with Theorems 2 and 3, give useful guidance in design planning for many v, k combinations. In particular, the bound on d is moderate in size for designs with v close to or equal to $2k$: for a design with $v = 2k$, by Theorem 2 $d \geq 3$ is sufficient to guarantee connectivity. However, unless v is close to $2k$, the value of d required to ensure connectivity from Theorem 2 can be prohibitively large as v and k increase.

Additional knowledge of D promotes improved conditions. In particular, knowledge of the treatment replication numbers in D^{sup}, that is $s_{[1]}, \ldots, s_{[v]}$, enables a lower bound to be established for u, the cardinality of $\mathcal{V}_1$, for a treatment partition characterising a disconnected design. A set of conditions is now developed, as an alternative to Theorem 2, which guarantee a connected design without the need for (4) to be satisfied. These conditions are based on $s_{[1]}, \ldots, s_{[v]}$. To achieve this aim, a function is defined:

Definition 2. For integers γ and δ, with $\gamma \geq 1$ and $\delta \geq 2$, define $\Phi(\gamma, \delta)$ be the positive integer such that:

$$\binom{\Phi(\gamma,\delta)-2}{\delta-1} < \gamma \leq \binom{\Phi(\gamma,\delta)-1}{\delta-1}. \tag{6}$$

Values for $\Phi(\gamma, \delta)$, for $\gamma \leq 30$ and $2 \leq \delta \leq 10$ are given in Table 1 in the Appendix.

Lemma 1. Let t be a treatment in D with replication s in D^{sup}. If D is disconnected then, in any treatment partition which is consistent with a block partition in D^{sup}, treatment t is contained in a treatment set of cardinality at least $\Phi(s, k)$.

Proof. Consider the s blocks of D^{sup} which contain treatment t. Each of these blocks contains $k - 1$ treatments in addition to t. Let the number of treatments occuring in at least one block with t be m. Then, m satisfies

$$s \leq \binom{m}{k-1}.$$

From (6),

$$\binom{\Phi(s,k)-2}{k-1} < s \leq \binom{\Phi(s,k)-1}{k-1}.$$

Thus, $\Phi(s, k) - 2 < m$ and $\Phi(s, k) - 1 \leq m$. It follows that $\Phi(s, k) - 1$ is the smallest number of treatments, which can occur with t in blocks of D^{sup}. Hence, in a disconnected design, a lower bound for the cardinality of the treatment set containing t in any treatment partition consistent with a block partition in D or D^{sup} is $\Phi(s, k)$. □

Lemma 1 prompts a range of conditions for connectivity.

Theorem 4. The inequalities listed each provide a sufficient condition for D to be connected:

$$(i) \quad \Phi(s_{[1]}, k) > v - k; \tag{7}$$

$$(ii) \quad \Phi(s_{[v+1-k]}, k) > \left[\frac{v}{2}\right]; \tag{8}$$

$$(iii) \quad \Phi(s_{[1]}, k) + \Phi(s_{[v+1-k]}, k) > v. \tag{9}$$

Proof. Assume D is disconnected and let $\{\mathcal{V}_1, \mathcal{V}_2\}$ be any treatment partition which is consistent with a block partition in D. The conditions are considered in turn.

(i) Let (7) hold. Consider a treatment with replication $s_{[1]}$ in D^{sup}. By Lemma 1 this treatment occurs in a treatment set with cardinality at least $\Phi(s_{[1]}, k)$, that is, at least $v - k + 1$. Since the cardinalities of $\mathcal{V}_1$ and $\mathcal{V}_2$ are bounded below by k, the sum of the cardinalities of the two treatment sets is at least $(v - k + 1) + k > v$, giving a contradiction and proving that D is connected.

(ii) Let (8) hold. For any set of w treatments, $\mathcal{W}$ say, the largest replication of a treatment from $\mathcal{W}$ in D^{sup} must be at least $s_{[v+1-w]}$. The cardinalities of $\mathcal{V}_1$ and $\mathcal{V}_2$ are bounded below by k. Thus each set contains a treatment with replication at least $s_{[v+1-k]}$ in D^{sup}. Then, by Lemma 1, it follows that $\mathcal{V}_1$ and $\mathcal{V}_2$ each have cardinality at least $\Phi(s_{[v+1-k]}, k)$. But $2\Phi(s_{[v+1-k]}, k) > v$, again providing a contradiction and establishing that D is connected.

(iii) Let (9) hold. In D^{sup} at least one of $\mathcal{V}_1$, $\mathcal{V}_2$ contains a treatment with replication $s_{[1]}$ and both sets contain a treatment with replication at least $s_{[v+1-k]}$. Thus a lower bound for the sum of the set cardinalities is $\Phi(s_{[1]}, k) + \Phi(s_{[v+1-k]}, k) > v$. This gives a contradiction and therefore D is connected. □

The conditions of Theorem 4 provide guidance in the construction of designs to ensure connectivity. Note that if condition (iii) is satisfied, then condition (i) is automatically also satisfied since $\Phi(s_{[v+1-k]}, k) \geq k$. However, condition (i) is useful in its own right because it depends only on $s_{[1]}$.

Example 5 A design, $D4$, is proposed for an experiment with $v = 9$, $k = 3$ and $d = 18$ with support treatment replication numbers:

$$s_{[1]} = 10, \; s_{[2]} = s_{[3]} = s_{[4]} = s_{[5]} = s_{[6]} = s_{[7]} = 6, \; s_{[8]} = s_{[9]} = 1.$$

Connectivity of $D4$ is not guaranteed by Theorem 2, which requires $d \geq 21$ for designs with $v = 9$, $k = 3$. Using Table 1, the conditions of Theorem 4 are considered in turn:

$$(i) \quad \Phi(s_{[1]}, k) = \Phi(10, 3) = 6 \not> v - k = 6$$

$$(ii) \quad \Phi(s_{[v+1-k]}, k) = \Phi(s_{[7]}, k) = \Phi(6, 3) = 5 > \left[\frac{v}{2}\right] = 4$$

$$(iii) \quad \Phi(s_{[1]}, k) + \Phi(s_{[v+1-k]}, k) = \Phi(10, 3) + \Phi(6, 3) = 11 > v = 9.$$

Thus, $D4$ is connected by conditions (ii) and (iii).

It is useful to define two further functions. These will facilitate identification of design parameters that lead to one or more conditions of Theorem 4 being satisfied and enable some measure of the robustness of a design satisfying Theorem 4 to be obtained.

Definition 3. For $q \geq k$, define $\Theta_1(q, \delta)$ to be the smallest integer γ_1 such that $\Phi(\gamma_1, \delta) = q$. For $q \geq 2k$, define $\Theta_2(q, \delta)$ be the set of ordered integer pairs $\langle \gamma_1, \gamma_2 \rangle$ with $\gamma_1 \geq \gamma_2$ such that

$$\Phi(\gamma_1, \delta) + \Phi(\gamma_2, \delta) = q \text{ and}$$

$$\Phi(\gamma_1 - 1, \delta) + \Phi(\gamma_2, \delta) < q \text{ and}$$

$$\Phi(\gamma_1, \delta) + \Phi(\gamma_2 - 1, \delta) < q.$$

For example, using Table 1, $\Theta_1(6, 3) = 7$ and $\Theta_2(11, 3) = \{\langle 16, 1 \rangle \langle 11, 2 \rangle \langle 7, 4 \rangle\}$. The functions $\Theta_1(q, \delta)$ and $\Theta_2(q, \delta)$ aid in the identification of design parameters that satisfy Theorem 4 and in assessing design robustness. For example, the smallest value of $s_{[1]}$ that satisfies condition (i) is $\Theta_1(v + 1 - k, k)$. Similarly, the set $\Theta_2(v + 1, k)$ contains pairs $\langle s_{[1]}, s_{[v+1-k]} \rangle$ that satisfy condition (iii), and for which the condition is not satisfied if either replication number is reduced. The function $\Theta_1(q, \delta)$ leads to bounds for $\Omega(D)$ for designs satisfying condition (i) or (ii) of Theorem 4. Likewise, $\Theta_2(q, \delta)$ gives rise to a bound for $\Omega(D)$ for designs satisfying condition (iii). These are combined in an obvious way in the following result.

Theorem 5. Let design D be connected by Theorem 4. Then let

$$\beta = \max \left\{ s_{[1]} - \Theta_1(v + 1 - k, k), s_{[v+1-k]} - \Theta_1\left(\left[\frac{v}{2}\right] + 1, k\right), \max_{\langle \gamma_1, \gamma_2 \rangle \in \Theta_2(v+1,k)} \left\{ \min\{ s_{[1]} - \gamma_1, s_{[v+1-k]} - \gamma_2 \} \right\} \right\}$$

and $\omega_\beta = \Sigma_{i=d-\beta+1}^{d} n_{[i]}$. Then $\Omega(D) \geq \omega_\beta$.

Example 5 revisited Design $D4$ has been established as being connected by conditions (ii) and (iii) of Theorem 4. A bound on $\Omega(D4)$ can be obtained from Theorem 5. For $D4$:

$$\beta = \max \left\{ s_{[1]} - \Theta_1(7, 3), s_{[7]} - \Theta_1(5, 3), \max_{\langle \gamma_1, \gamma_2 \rangle \in \Theta_2(10,3)} \left\{ \min\{ s_{[1]} - \gamma_1, s_{[7]} - \gamma_2 \} \right\} \right\}.$$

From Table 1, $\Theta_1(7, 3) = 11$, $\Theta_1(5, 3) = 4$ and $\Theta_2(10, 3) = \{\langle 11, 1 \rangle, \langle 7, 2 \rangle, \langle 4, 4 \rangle\}$. This gives

$$\beta = \max \{-1, 2, \max\{\min\{-1, 5\}, \min\{3, 4\}, \min\{6, 2\}\}\} = \max \{-1, 2, 3\} = 3.$$

Thus, for every treatment partition $\{\mathcal{V}_1, \mathcal{V}_2\}$, the support design $D4^{sup}$ contains at least three linking blocks. Hence, $\Omega(D4) \geq \Sigma_{i=16}^{18} n_{[i]}$.

The use of the functions $\Theta_1(q, \delta)$ and $\Theta_2(q, \delta)$ in identifying design parameters that guarantee connectivity is demonstrated in the next example.

Example 6 A design is required with $v = 10$ and $k = 4$. Support treatment replication numbers can be identified which satisfy the conditions of Theorem 4. For condition (i), $\Theta_1(v + 1 - k, k) = \Theta_1(7, 4) = 11$. Thus designs with $s_{[1]} \geq 11$ satisfy condition (i). For condition (ii), $\Theta_1([v/2] + 1, k) = \Theta_1(6, 4) = 5$. Thus, any design with $s_{[7]} \geq 5$ will satisfy condition (ii). Finally, for condition (iii), the set $\Theta_2(v + 1, k) = \Theta_2(11, 4) = \{\langle 11, 1\rangle, \langle 5, 2\rangle\}$ is informative. From this set, it follows that a design with $s_{[1]} \geq 11$ and $s_{[7]} \geq 1$ or with $s_{[1]} \geq 5$ and $s_{[7]} \geq 2$ will satisfy condition (iii).

This information is useful in planning an experiment. For example, design $D2$ of Example 2, with $v = 10$ and $k = 4$, has support treatment replication numbers

$$s_{[1]} = s_{[2]} = 8,\ s_{[3]} = s_{[4]} = 6,\ s_{[5]} = s_{[6]} = 5,\ s_{[7]} = 4,\ s_{[8]} = s_{[9]} = s_{[10]} = 2.$$

Thus, $D2$ is connected by condition (iii) of Theorem 4. Also, by Theorem 5, the support contains at least two linking blocks for every treatment partition and so $\Omega(D2) \geq \Sigma_{i=11}^{12} n_{[i]}$ and $b_0 \geq \min\{\Omega(D2), r_{[10]}\}$. With only knowledge of the support of $D2$, it follows that $\Omega(D2) \geq 2$ and $r_{[10]} \geq 2$, giving $b_0 \geq 2$.

For experimental situations where there is confidence that observation loss is highly unlikely, the conditions for connectivity from Theorems 2 and 4 can be improved on.

Lemma 2. Let $\phi_0 = k$ and define

$$\phi_{i+1} = \Phi(s_{[v+1-\phi_i]}, k) \text{ for } i = 0, \ldots. \tag{10}$$

Then $\phi_0, \phi_1, \ldots$ is a non-decreasing sequence of integers which terminates with stop value ϕ_*, where $1 \leq \phi_* \leq \Phi(s_{[1]}, k)$.

Proof. Using (10), $\phi_1 = \Phi(s_{[v+1-k]}, k) \geq k = \phi_0$. Then $v + 1 - \phi_1 \leq v + 1 - k$, and so $s_{[v+1-\phi_1]} \geq s_{[v+1-k]}$ giving $\phi_2 = \Phi(s_{[v+1-\phi_1]}, k) \geq \phi_1$. A recursive argument establishes that ϕ_0, ϕ_1, ϕ_2, … is a non-decreasing sequence of positive integers. The sequence is bounded below by k and above by v. A stop value is attained if there is a value for m such that $\phi_{m+1} = \phi_m$; in this case set $\phi_* = \phi_m$. Otherwise, the stop value is $\phi_* = \Phi(s_{[1]}, k)$. □

Lemma 3. Let D be a disconnected design. Then any treatment partition consistent with a block partition has $u \geq \phi_*$.

Proof. Let $\{\mathcal{V}_1, \mathcal{V}_2\}$ be any treatment partition which is consistent with a block partition in D. Then $\{\mathcal{V}_1, \mathcal{V}_2\}$ is also consistent with a block partition in D^{sup}. Let $\mathcal{B}_{1S}$ be the blocks of D^{sup} that contain all replicates of the u treatments from $\mathcal{V}_1$. Since $u \geq k$, the set $\mathcal{V}_1$ contains a treatment with replication at least $s_{[v+1-k]}$ in D^{sup}. From Lemma 1, $\mathcal{V}_1$ has cardinality at least $\Phi(s_{[v+1-k]}, k) = \phi_1$, where ϕ_1 is defined by Lemma 2. Therefore, $\mathcal{V}_1$ contains a treatment of replication at least $s_{[v+1-\phi_1]}$ and, again using Lemma 1, $\mathcal{V}_1$ has cardinality at least $\Phi(s_{[v+1-\phi_1]}, k) = \phi_2$. The argument proceeds in the same manner to conclude that $\mathcal{V}_1$ contains at least ϕ_* treatments, that is, $u \geq \phi_*$. □

The lower bound for u provided by ϕ_* is used to improve on conditions for connectivity of Theorems 2 and 4. The proofs mirror those of the earlier results and are not included:

Theorem 6. The inequalities listed each provide a sufficient condition for D to be connected:

$$\begin{aligned} &(i) \quad d > \binom{v-\phi_*}{k} + 1; \\ &(ii) \quad \phi_* > \left[\frac{v}{2}\right]; \\ &(iii) \quad \Phi(s_{[1]}, k) + \phi_* > v. \end{aligned}$$

Example 7 Design $D5$ is proposed for an experiment with $v = 9$, $k = 3$ and $d = 10$. The support design, $D5^{sup}$, has treatment replication numbers:

$$s_{[1]} = s_{[2]} = s_{[3]} = s_{[4]} = s_{[5]} = s_{[6]} = 4; s_{[7]} = s_{[8]} = s_{[9]} = 2.$$

Neither Theorem 2 or Theorem 4 guarantee the connectivity of $D5$. The value of ϕ_* is obtained via Lemma 2: $\phi_0 = k = 3$; $\phi_1 = \Phi(s_{[7]}, 3) = \Phi(2, 3) = 4$; $\phi_2 = \Phi(s_{[6]}, 3) = \Phi(4, 3) = 5$; $\phi_3 = \Phi(s_{[5]}, 3) = \Phi(4, 3) = 5 = \phi_2$. Thus $\phi_* = 5$. The conditions of Theorem 6 are considered in turn:

$$\begin{aligned} &(i) \quad d = 10 > 5 = \binom{4}{3} + 1 = \binom{v-\phi_*}{k} + 1; \\ &(ii) \quad \phi_* = 5 > 4 = \left[\frac{v}{2}\right]; \\ &(iii) \quad \Phi(s_{[1]}, k) + \phi_* = \Phi(4, 3) + 5 = 10 > 9 = v. \end{aligned}$$

Design $D5$ is established as connected by all three conditions of Theorem 6.

References

1. Bailey, R.A., Schiffl, K. and Hilgers, R-D. A note on the robustness of D-optimal block designs for two-colour microarray experiments. *J. Statist. Plann. Inference.* 2013; **143**: 1195-1202.

2. Baksalary, J.K., Tabis, Z. Conditions for the robustness of block designs against the unavailability of data. *J. Statist. Plann. Inference.* 1987; **16**: 49-54.

3. Behbahani, M., Khosrovshahi, G.B., Tayfeh-Rezaie, B. On the spectrum of simple T(2,3,v) trades. *J. Statist. Plann. Inference.* 2008; **138**: 2236-2242.

4. Butz, L. *Connectivity in multi-factor designs*. Berlin: Heldermann Verlag. 1982.

5. Dey, A. Robustness of block designs against missing data. *Statist. Sinica.* 1993; **3**: 219-231.

6. Dobcsányi, P., Preece, D.A. and Soicher, L.H. On balanced incomplete-block designs with repeated blocks. *European Journal of Combinatorics* 2007; **28**: 1955-1970.

7. Foody, W., Hedayat, A. On theory and applications of BIB designs with repeated blocks. *Ann. Statist.* 1977; **5**: 932-945.

8. Ghosh, S. On robustness of block designs against incomplete data. *Sankhya.* 1979; **40**: 204-208.

9. Ghosh, S. Robustness of BIBD against the unavailability of data. *J. Statist. Plann. Inference.* 1982; **6**: 29-32.

10. Godolphin, J.D. Designs with blocks of size two and applications to microarray experiments. *Ann. Statist.* 2018; **46**: 2775-2805.

11. Godolphin, J.D., Godolphin, E.J. The use of treatment concurrences to assess robustness of binary block designs against the loss of whole blocks. *Aust. N. Z. J. Stat.* 2015; **57**: 225-239.

12. Godolphin, J.D., Warren, H.R. Improved conditions for the robustness of binary block designs against the loss of whole blocks. *J. Statist. Plann. Inference.* 2011; **141**: 3498-3505.

13. Hedayat, A., Li, S-Y.R. The trade-off method in the construction of BIB designs with variable support sizes. *Ann. Statist.* 1979; **7**: 1277-1287.

14. Hedayat, A., Pesotan, H. A study of BIB designs through support matrices. *J. Statist. Plann. Inference.* 1985; **11**: 363-372.

15. Latif, A.H.M.M., Bretz, F. and Brunner, E. Robustness considerations in selecting efficient two-color microarray designs. *Bioinformatics.* 2009; **25**: 2355-2361.

16. Tsai, S-F., Liao, C-T. Minimum breakdown designs in blocks of size two. *J. Statist. Plann. Inference.* 2013; **143**: 202-208.

□

APPENDIX

TABLE 1 Values for $\Phi(\gamma, \delta)$, for $\gamma \leq 30$ and $2 \leq \delta \leq 10$

	δ								
γ	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
2	3	4	5	6	7	8	9	10	11
3	4	4	5	6	7	8	9	10	11
4	5	5	5	6	7	8	9	10	11
5	6	5	6	6	7	8	9	10	11
6	7	5	6	7	7	8	9	10	11
7	8	6	6	7	8	8	9	10	11
8	9	6	6	7	8	9	9	10	11
9	10	6	6	7	8	9	10	10	11
10	11	6	6	7	8	9	10	11	11
11	12	7	7	7	8	9	10	11	12
12	13	7	7	7	8	9	10	11	12
13	14	7	7	7	8	9	10	11	12
14	15	7	7	7	8	9	10	11	12
15	16	7	7	7	8	9	10	11	12

	δ								
γ	2	3	4	5	6	7	8	9	10
16	17	8	7	8	8	9	10	11	12
17	18	8	7	8	8	9	10	11	12
18	19	8	7	8	8	9	10	11	12
19	20	8	7	8	8	9	10	11	12
20	21	8	7	8	8	9	10	11	12
21	22	8	8	8	8	9	10	11	12
22	23	9	8	8	9	9	10	11	12
23	24	9	8	8	9	9	10	11	12
24	25	9	8	8	9	9	10	11	12
25	26	9	8	8	9	9	10	11	12
26	27	9	8	8	9	9	10	11	12
27	28	9	8	8	9	9	10	11	12
28	29	9	8	8	9	9	10	11	12
29	30	10	8	8	9	9	10	11	12
30	31	10	8	8	9	9	10	11	12