

DESIGNS WITH BLOCKS OF SIZE TWO AND APPLICATIONS TO MICROARRAY EXPERIMENTS

BY JANET GODOLPHIN

University of Surrey

Designs with blocks of size two have numerous applications. In experimental situations where observation loss is common, it is important for a design to be robust against breakdown. For designs with one treatment factor and a single blocking factor, with blocks of size two, conditions for connectivity and robustness are obtained using combinatorial arguments and results from graph theory. Lower bounds are given for the breakdown number in terms of design parameters. For designs with equal or near equal treatment replication, the concepts of treatment and block partitions, and of linking blocks, are used to obtain information on the number of blocks required to guarantee various levels of robustness. The results provide guidance for construction of designs with good robustness properties.

Robustness conditions are also established for row column designs in which one of the blocking factors involves blocks of size two. Such designs are particularly relevant for microarray experiments, where the high risk of observation loss makes robustness important. Disconnectivity in row column designs can be classified as three types. Techniques are given to assess design robustness according to each type, leading to lower bounds for the breakdown number. Guidance is given for robust design construction.

Cyclic designs and interwoven loop designs are shown to have good robustness properties.

1. Introduction. Designs incorporating a blocking factor with blocks of size two have received considerable attention. [Wu and Hamada \(2009\)](#) provide a description of the special role of blocks of size two which arises, in part, due to occurrences of blocks of this size in nature. Developmental studies on twins are well known and other relevant designs include those which have subjects as blocks, with the experimental units comprising individual shoes, for example. Incomplete block designs, with a single treatment factor and with treatments allocated to b blocks each of two experimental units, are the simplest such designs and the most commonly used; see, for example, [Bagchi and Cheng \(1993\)](#) and [Yang and Draper \(2003\)](#).

If a second blocking factor is included with two levels, so that, of the two experimental units contained in a block of size two, exactly one unit corresponds to

Received June 2017; revised September 2017.

MSC2010 subject classifications. Primary 62K10; secondary 62K05.

Key words and phrases. Breakdown number, connectivity, incomplete block design, microarray experiment, observation loss, robustness, row column design.

each level of the second blocking factor, then the design can be regarded as a row column design and represented by a $2 \times b$ array. Two-colour microarray experiments provide an important application of such designs. The technical aspects of microarray experiments are covered by [Nguyen et al. \(2002\)](#) and design issues are discussed by [Kerr and Churchill \(2001\)](#). A thorough description of the use of row column designs for microarray experiments is given by [Bailey \(2007\)](#). Other relevant work includes [Wit, Nobile and Khanin \(2005\)](#), [Gupta \(2006\)](#) and [Chai, Liao and Tsai \(2007\)](#).

The aim of an experiment in v treatments, using an incomplete block design or a row column design, is to determine whether the data collected depend on the specific treatment applied and to obtain estimates of treatment contrasts. A design for which these aims can be achieved, that is, for which all treatment contrasts are estimable, is described as *connected*. Otherwise, the design is said to be *disconnected*.

In practice, observations may be lost during experimentation. Therefore, as part of the process of planning an experiment, it is useful to take account, not only of the properties of the planned design, but also those of potential eventual designs that can result from a degree of observation loss. The extent of observation loss considered will depend on knowledge of the particular experimental situation. The loss of one observation from a block of size two is equivalent to loss of the block, since no information is obtained from a single observation in a block. Thus, for designs with blocks of size two, in any investigation of the effect of missing observations, attention can be restricted to the loss of entire blocks.

Criteria for assessing the robustness of incomplete block designs against observation loss are introduced by [Dey \(1993\)](#), following the work of [Ghosh \(1979, 1982\)](#). A design is robust against the loss of t blocks, according to Criterion 1, if all treatment contrasts are estimable from any eventual design resulting from the loss of t blocks. [Baksalary and Tabis \(1987\)](#), [Bhaumik and Whittinghill \(1991\)](#) and [Godolphin and Godolphin \(2015\)](#) provide conditions for assessing the robustness of incomplete block designs. The block breakdown number of a design 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; see, for example, [Mahbub Latif, Bretz and Brunner \(2009\)](#) and [Bailey, Schiffel and Hilgers \(2013\)](#). For designs with blocks of size two, the terms block breakdown number and breakdown number are interchangeable.

Little work has been done on the connectivity of designs involving more than one blocking factor, and even less on the robustness of such designs against observation loss. [Godolphin \(2013\)](#) gives a theoretical consideration of connectivity properties of designs with arbitrary numbers of treatment and blocking factors. Disconnectivity in a row column design is characterised by the existence of linear relationships between columns of the design matrix in addition to those that exist in connected designs. The forms of the additional relationships are covered by [Godolphin and Godolphin \(2001\)](#). A method to assess the robustness of row

column designs against the loss of a small number of observations is given by Godolphin (2004).

In this work, general conditions on connectivity and Criterion 1 robustness are developed, for incomplete block designs and row column designs with blocks of size two, that do not depend on the planned design having properties such as balance. The aim is to provide conditions and construction strategies which ensure that robustness properties of a design are appropriate for the anticipated level of observation loss. The main value of the work is in facilitating design construction that ensures a level of robustness, rather than in determining properties of individual designs. For a given incomplete block design, with blocks of size two, the approach of Godolphin and Warren (2014) can be used to obtain information on Criterion 1 robustness.

For models with one blocking factor, cyclic designs are shown to be optimally robust, subject to specified conditions. For two blocking factors, optimal robustness of many of the interwoven loop designs is established.

The paper is structured as follows. In Sections 2–4, the focus is on incomplete block designs with blocks of size two. The model and basic concepts are introduced in Section 2, together with elementary results on connectivity and robustness. In Section 3, representation of a design by its concurrence graph enables use of results from graph theory to provide conditions for connectivity and lower bounds for the breakdown number in terms of design parameters. Section 4 focuses on designs with treatment replications differing by one at most. An algorithm which generates lower bounds for the number of blocks required to guarantee various levels of robustness is developed. Robustness of designs with the additional complexity of a second blocking factor, that is, row-column designs with the blocking factors involving blocks of size two and size b , respectively, are covered in Sections 5–6. These designs are represented as digraphs, which parallels the graphical representation of designs in Section 3. Results leading to lower bounds for the breakdown number are derived and guidance is given for the construction of robust designs. It is demonstrated that *minimum breakdown designs* proposed in the literature do not all have the robustness properties claimed. A brief investigation of equireplicate designs suggests poor correlation between design robustness and an optimality criterion. Most proofs are located in the Supplementary Material [Godolphin (2017)].

2. Preliminaries. Let D denote an incomplete block design on v treatments applied to $2b$ experimental units arranged in b blocks of size two. Model 1 assumes that the $2b \times 1$ observation vector, Y , is specified by

$$(2.1) \quad E(Y) = \mu 1_{2b} + X_1 \beta_1 + X_2 \beta_2,$$

where μ is a scalar constant, the $n \times 1$ vector with all elements unity is denoted by 1_n , and β_1 and β_2 are $v \times 1$ and $b \times 1$ parameter vectors relating to treatments and blocks, respectively. The design matrix for Model 1 is given by $X = [1_{2b} \ X_1 \ X_2]$.

Here, X_1 and X_2 are $2b \times v$ and $2b \times b$ component matrices pertaining to treatments and blocks. Each row of X_1 has one element unity and $v - 1$ zeros and each row of X_2 has one element unity and $b - 1$ zeros. This structure leads to the relationship

$$(2.2) \quad X_1 1_v = X_2 1_b.$$

All designs are taken to be binary since no treatment information is gained from a block if the same treatment is applied to both experimental units. Further, it is assumed that no two blocks contain the same pair of treatments, so $b \leq v(v - 1)/2$. Treatment replication numbers are given in decreasing order by $r_{[1]}, \dots, r_{[v]}$, with

$$(2.3) \quad v - 1 \geq r_{[1]} \geq r_{[2]} \geq \dots \geq r_{[v]} \geq 1.$$

The condition $r_{[1]} \leq v - 1$ arises since the treatment pairs in blocks of D are distinct and, therefore, no treatment can occur in more than $v - 1$ blocks. Optimality criteria for D are functions of the nonzero eigenvalues of its information matrix, which is

$$(2.4) \quad C = X_1^T \left(I - \frac{1}{2} X_2 X_2^T \right) X_1;$$

see [John and Williams \(1995\)](#) for details.

If D is connected, all treatment contrasts are estimable. Conversely, if D is disconnected, some treatment contrasts are not estimable. Two equivalent features of treatment allocation to experimental units in a disconnected incomplete block design characterise the disconnectivity. Description of the first feature uses the concepts of partitions. A treatment partition, $\{\mathcal{V}_1, \mathcal{V}_2\}$, is an arrangement of the treatments into disjoint nonempty sets, \mathcal{V}_1 and \mathcal{V}_2 , of sizes u and $v - u$ with $u \in U$, where $U = \{2, \dots, [v/2]\}$ and $[x]$ denotes *the integer part of x* . A block partition, $\{\mathcal{B}_1, \mathcal{B}_2\}$, is a partitioning of the blocks into disjoint nonempty 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 $i = 1, 2$. For a pairwise treatment contrast to be estimable, both treatments must occur in the same \mathcal{V}_i . In a connected design, there will be at least one *linking block* for any treatment partition, that is, at least one block which contains a treatment from each of \mathcal{V}_1 and \mathcal{V}_2 . The notation $\Omega_1(D)$ is used for the smallest number of linking blocks over all treatment partitions in D . The notions of treatment and block partitions and of linking blocks are used extensively in this work.

An alternative and equivalent characterisation of a disconnected design is the existence of a relationship between the columns of X that is additional to and independent of (2.2). Thus, for a disconnected design there is a relationship of the form

$$(2.5) \quad X_1 \alpha_1 = X_2 \alpha_2,$$

where α_1 is not a multiple of 1_ν ; see [Godolphin and Godolphin \(2001\)](#) for further details.

An eventual design realized after the loss of one or more blocks from D is denoted by $D_\#$ and the class of eventual designs resulting from the loss of every set of p blocks is denoted by $D(p)$. The block breakdown number, b_0 , is the smallest number of blocks that needs 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 $b_0 > 0$. For $p < b_0$, every design in $D(p)$ is a connected design in ν treatments, but for $p \geq b_0$ there will be at least one member of $D(p)$ from which some treatment contrasts are 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. Thus, we have the following.

DEFINITION 2.1. The breakdown number of a design D is given by

$$(2.6) \quad b_0 = \min\{r_{[\nu]}, \Omega_1(D)\}.$$

Following [Ghosh \(1982\)](#), D is described as being maximally robust against the unavailability of data and with respect to estimability of treatment contrasts if $b_0 = r_{[\nu]}$.

Any treatment partition requires at least two treatments in each of \mathcal{V}_1 and \mathcal{V}_2 . It follows immediately that all designs with $\nu < 4$ are connected and, further, that the loss of any $p < r_{[\nu]}$ blocks yields a connected eventual design in ν treatments. Therefore, it is assumed throughout that $\nu \geq 4$.

Obtaining an upper bound for b_0 is straightforward, even before allocation of treatments to experimental units. Theorem 2.1 is given by [Bailey, Schiffel and Hilgers \(2013\)](#).

THEOREM 2.1. For any design, $b_0 \leq [2b/\nu]$.

PROOF. At least one treatment has replication $[2b/\nu]$ or smaller. Thus, $r_{[\nu]} \leq [2b/\nu]$ and the result follows immediately from (2.6). \square

Theorem 2.1 leads to definitions of levels of design robustness.

DEFINITION 2.2. A design with $b_0 = [2b/\nu]$ is described as having *optimal breakdown number*.

DEFINITION 2.3. A design, D , with $b_0 = [2b/\nu]$ and for which there are exactly $\nu - \eta$ members of $D([2b/\nu])$ from which not all treatment contrasts are estimable, where $\eta = 2b - [2b/\nu]\nu$, is *optimally robust*.

Designs with optimal breakdown number or optimal robustness have $r_{[v]} = [2b/v]$. Optimal breakdown number implies that $\Omega_1(D) \geq [2b/v]$. Optimal robustness implies that treatment replications differ at most by one, so $r_{[1]} \leq [2b/v] + 1$, and that $\Omega_1(D) > [2b/v]$. For an optimally robust design, each member of $D([2b/v])$ for which not all treatment contrasts are estimable is an eventual design in which all replicates of a treatment have been lost. The properties of being optimally robust, of having optimal breakdown number and of being maximally robust are hierarchical. Any optimally robust design has optimal breakdown number and any design with optimal breakdown number is maximally robust.

For $p < r_{[v]}$, an eventual design $D_\#$ in $D(p)$ has v treatments and $b - p$ blocks. The design matrix, $X_\# = [1_{2(b-p)} \ X_{1\#} \ X_{2\#}]$, is obtained by deleting $2p$ rows of X and p columns of X_2 . As with the planned design, $D_\#$ is disconnected iff there is a relationship between the columns of $X_\#$ of the form $X_{1\#\alpha_1} = X_{2\#\alpha_2}$, where α_1 is not a multiple of 1_v . This is equivalent to the existence of consistent treatment and block partitions in $D_\#$.

The first condition for connectivity depends only on v and b .

THEOREM 2.2. *A sufficient condition for D to be connected is given by*

$$(2.7) \quad b > \binom{v-2}{2} + 1.$$

All designs in a class with given v and b are connected iff Theorem 2.2 is satisfied. For a design, D , in a class satisfying Theorem 2.2, there are at least $b - (v - 2)(v - 3)/2 - 1$ linking blocks for any treatment partition. This gives a lower bound for $\Omega_1(D)$.

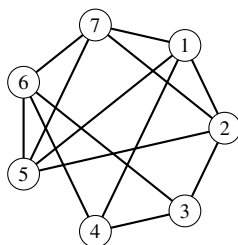
COROLLARY 2.1. *Let D be such that $b > (v - 2)(v - 3)/2 + 1$. Then*

$$\Omega_1(D) \geq b - (v - 2)(v - 3)/2 - 1.$$

EXAMPLE 1. Design $D1$ is given by

$$D1 = \begin{matrix} 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 5 & 5 & 6 \\ 2 & 4 & 5 & 7 & 3 & 5 & 7 & 4 & 6 & 6 & 6 & 7 & 7 \end{matrix}$$

where, as with all designs in this work, columns correspond to blocks of size two. For designs of Model 1, the order of the blocks is immaterial, as is the order of treatments within a block. Design $D1$ is in the class of designs with $v = 7$ and $b = 13$. By Theorem 2.2, all these designs are connected and, by Corollary 2.1, all have $\Omega_1(D1) \geq 2$. Since $D1$ has $r_{[7]} = 3$, by (2.6) $b_0 \geq 2$. Note that not all designs in the class have $b_0 \geq 2$: any design with $r_{[7]} = 1$ will have $b_0 = 1$.

FIG. 1. Graphical representation of $D1$.

3. Robustness conditions from graph theory. A design of Model 1 can be represented by a graph, G_D , termed the concurrence graph of D , as in Bailey and Cameron (2009). Treatments and blocks of D correspond to vertices and edges in G_D . Thus, a treatment with replication $r_{[i]}$ in D corresponds to a vertex of degree $r_{[i]}$ in G_D . The concurrence graph of $D1$ is given in Figure 1. In this section, results from graph theory yield sufficient conditions for connectivity and lower bounds for b_0 in terms of b , v and $r_{[i]}$. Graph theory terminology used is consistent with Bondy and Murty (2008).

The concurrence graph of a Model 1 design will be simple and finite, and only graphs with these properties are considered. Note that a design is connected iff its concurrence graph is connected. A connected graph with v vertices must contain at least $v - 1$ edges: a connected graph with exactly $v - 1$ edges has no cycles and is termed a tree. Thus, every tree in v vertices corresponds to a connected design in $v - 1$ blocks. However, such designs necessarily contain at least two treatments with replication one and so have poor robustness.

In general, a graph G is k -edge-connected if the removal of any set of fewer than k edges gives a connected graph. Thus, every connected graph is 1-edge-connected. The maximum value of k for which G is k -edge-connected is the *edge connectivity*, $\kappa'(G)$. Note that $\kappa'(G)$ cannot exceed the minimum vertex degree, which is $r_{[v]}$ for G_D . It follows that a design has $b_0 = \kappa'(G_D)$. A sufficient condition for G to be k -edge-connected and, therefore, for the corresponding design to have $b_0 \geq k$, is given by Theorem 4.1 of Bauer et al. (2009). The result is reproduced here without proof, first as a result in graph theory as given by Bauer et al. (2009), and then as the design equivalent.

THEOREM 3.1 [Bauer et al. (2009)]. *Let the graph G have v vertices, with vertex degrees $r_{[1]}, \dots, r_{[v]}$ of (2.3), and let $k \in \{1, \dots, r_{[v]}\}$. If $r_{[v+k-i]} \leq i - 1$ and $r_{[v+1-i]} \leq i + k - 2$ implies $r_{[1]} \geq v + k - 1 - i$, for $i = k + 1, \dots, \lfloor v/2 \rfloor$ then $\kappa'(G) \geq k$.*

THEOREM 3.2. *Let the treatment replicates of design D be such that $r_{[v+k-i]} \leq i - 1$ and $r_{[v+1-i]} \leq i + k - 2$ implies $r_{[1]} \geq v + k - 1 - i$, for $i = k + 1, \dots, \lfloor v/2 \rfloor$, where $k \in \{1, \dots, r_{[v]}\}$. Then $b_0 \geq k$.*

Theorem 3.2 uses information on the treatment replicates not used in Theorem 2.2 and can therefore give improved lower bounds for b_0 . This is demonstrated by returning to Example 1:

EXAMPLE 1 (Revisited). Design $D1$ has $r_{[j]} = 4$ for $j = 1, \dots, 5$ and $r_{[6]} = r_{[7]} = 3$. Use of Theorem 3.2 with $k = 3$ establishes that $b_0 = 3$.

Theorem 3.2 can be used to find a lower bound for b_0 for all designs with the same set of treatment replications.

EXAMPLE 2. Consider the design class with $v = 9$ and treatment replications: $r_{[1]} = 6, r_{[2]} = \dots = r_{[7]} = 4, r_{[8]} = r_{[9]} = 3$. Note that, with $b = 18$, Theorem 2.2 is not satisfied. By Theorem 3.2 with $k = 2$, all designs in the class have $b_0 \geq 2$. The condition of Theorem 3.2 is not satisfied for $k = 3$, and so it cannot be concluded that all designs have $b_0 = 3$.

The bound cannot be improved since some designs in the class do have $b_0 = 2$. Such a design is $D2$, which is given below and which has concurrence graph in Figure 2. Other designs in the class, such as $D3$, also below and with concurrence graph in Figure 3, are maximally robust with $b_0 = 3$.

$$D2 = \begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 6 & 6 & 6 & 7 & 7 & 8 \\ 2 & 3 & 4 & 5 & 6 & 7 & 3 & 4 & 5 & 4 & 5 & 5 & 7 & 8 & 9 & 8 & 9 & 9' \end{matrix}$$

$$D3 = \begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 6 & 6 \\ 2 & 3 & 4 & 5 & 6 & 8 & 3 & 5 & 7 & 7 & 9 & 5 & 6 & 8 & 7 & 9 & 8 & 9' \end{matrix}$$

The design in $D2(2)$ resulting from the loss of the fifth and sixth blocks is disconnected. From Figure 2, it is easily seen that removal of the edges in G_{D2} corresponding to these blocks gives rise to a disconnected graph, demonstrating that $\kappa'(G_{D2}) = 2$. It is less easy to see, but there is no pair of edges in G_{D3} whose removal gives rise to a disconnected graph. The removal of the edges incident to any of the vertices of degree three does, of course, yield a disconnected graph demonstrating that $\kappa'(G_{D3}) = 3$.

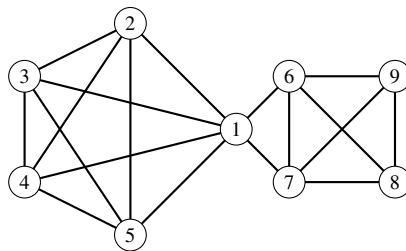
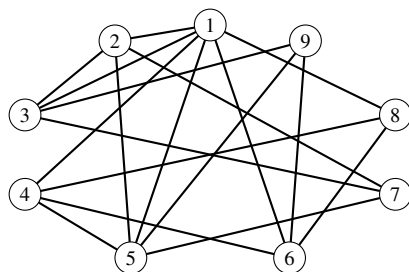


FIG. 2. G_{D2} of Example 2.

FIG. 3. G_{D_3} of Example 2.

Straightforward sufficient conditions for design connectivity are given as corollaries to Theorem 3.2. The first, obtained through setting $k = 1$, was originally given in graph equivalent form by Bondy (1969).

COROLLARY 3.1. *If $r_{[v+1-i]} \leq i - 1$ implies $r_{[1]} \geq v - i$, for $i = 2, \dots, \lfloor v/2 \rfloor$ then D is connected.*

COROLLARY 3.2. *If $r_{[1]} \geq v - 2$, then D is connected.*

Use of $k = r_{[v]}$ in Theorem 3.2 provides a sufficient condition for D to have $b_0 = r_{[v]}$ and, therefore, to be maximally robust. This result is given by Chartrand (1966) in graph equivalent form.

COROLLARY 3.3. *If $r_{[v]} > \lfloor v/2 \rfloor - 1$, then D is maximally robust.*

EXAMPLE 3. Consider the class of designs with $v = 9$ and $r_{[v]} = 4$. By Corollary 3.3, every design in the class is maximally robust.

A graph is described as being *super-edge-connected* if $\kappa'(G) = r_{[v]}$ and the number of sets of $r_{[v]}$ edges whose removal results in a disconnected graph is the same as the number of vertices with degree $r_{[v]}$. Such graphs correspond to designs with $\Omega_1(D) > r_{[v]}$. Kel'mans (1972) gives a sufficient condition for a graph to be super-edge-connected.

THEOREM 3.3. *Let G be a graph on v vertices, with vertex degrees given by $r_{[1]}, \dots, r_{[v]}$ of (2.3). If $r_{[v]} \geq (v + 1)/2$, then G is super-edge-connected.*

Theorem 3.3 and Definition 2.3 lead to a condition for optimal robustness.

COROLLARY 3.4. *Let D have $r_{[v]} = \lfloor 2b/v \rfloor \geq (v + 1)/2$ and $r_{[1]} \leq \lfloor 2b/v \rfloor + 1$. Then D is optimally robust.*

EXAMPLE 4. Consider design classes with $v = 10$ and treatment replicates:

- (i) $r_{[1]} = r_{[2]} = 7, r_{[3]} = \cdots = r_{[10]} = 6$;
- (ii) $r_{[1]} = r_{[2]} = 6, r_{[3]} = \cdots = r_{[10]} = 5$.

Designs in class (i) are optimally robust by Corollary 3.4. Thus, for any design, D , in the class, $b_0 = 6$ and there are exactly eight members of $D(6)$ from which not all treatment contrasts are estimable. Each of these is formed by the loss of all replicates of a treatment with replication six. Designs in class (ii) are not established as optimally robust by Corollary 3.4, but are maximally robust by Corollary 3.3 and have optimal breakdown number, since $r_{[10]} = \lfloor 2b/v \rfloor$.

4. Designs with equal, or near equal, replication. In many situations, there is no practical or economic advantage in replicating some treatments more than others. It should be noted that for small b , many optimal designs with block size two do not exhibit the equal or near equal treatment replication typically observed in optimal incomplete block designs. For example, for design classes with $v > 8$ and $b = v + 1$, D - and E -optimal designs have $r_{[v]} = 1$. By contrast, A -optimal designs do have treatments replicated as equally as possible. See Morgan (2015) for a summary of relevant work on optimality. However, amongst equireplicate designs with block size two and replication at least three, designs which perform well with regards to the A -criterion tend to be the same as those that perform well with regards to the D -criterion. Further, by investigation of equireplicate connected designs with $v = 8$ and $b = 12$, Bailey (2007) demonstrates that the designs with $b_0 = 3$ tend to have better A - and D - criteria measures than those with $b_0 = 2$, which in turn generally have better A - and D - criteria measures than those with $b_0 = 1$. Results from Bagchi and Cheng (1993) indicate that in many classes with specific v and b , where $v|2b$, equireplicate designs are E -optimal. Thus, optimality considerations provide some motivation for giving particular attention to designs in which treatments are replicated as equally as possible. Also, of course, such designs maximise $r_{[v]}$ and, therefore, have the potential for high breakdown number.

In this section, designs have equal, or near equal, replication. There are $v - \eta$ treatments with replication $r = \lfloor 2b/v \rfloor$ and η treatments with replication $r + 1$, where $\eta = 2b - \lfloor 2b/v \rfloor v$. With this additional constraint, the properties of being maximally robust and of having optimal breakdown number coincide.

4.1. *Conditions depending only on the number of blocks.* Conditions on the number of blocks that guarantee connectivity, optimal breakdown number and optimal robustness, are now derived.

LEMMA 4.1. *A sufficient condition for D to be connected is given by*

$$(4.1) \quad b \geq \frac{2(v-1)^2 + 7 + (-1)^{v+1}}{8}.$$

The bound of (4.1) cannot be improved on since disconnected designs exist with $b = \{2(v-1)^2 - 1 + (-1)^{v+1}\}/8$. To demonstrate this, v even and odd are dealt with separately. Suppose v is even and $b = v(v-2)/4$. Then each treatment has replication $(v-2)/2$. Treatments can be allocated to blocks so there are consistent treatment and block partitions with $u = v/2$. This occurs if \mathcal{B}_1 contains $v(v-2)/8$ blocks comprising all pairs of the $v/2$ treatments in \mathcal{V}_1 and \mathcal{B}_2 similarly contains $v(v-2)/8$ blocks comprising all pairs of the $v/2$ treatments in \mathcal{V}_2 . Such a design is clearly disconnected. Now suppose v is odd and $b = (v-1)^2/4$. The design has $r = (v-3)/2$ and $\eta = v - r - 1$. Thus, $(v-1)/2$ treatments have replication $(v-3)/2$ and the remaining $(v+1)/2$ treatments each have replication $(v-1)/2$. Arrangement of all pairs of the $(v-1)/2$ lower replicated treatments into $(v-1)(v-3)/8$ blocks and of all pairs of the $(v+1)/2$ higher replicated treatments into $(v-1)(v+1)/8$ blocks gives a disconnected design.

EXAMPLE 5. Designs with $v = 9$ and $b > 16$ are connected by Lemma 4.1.

An algorithm is now developed to determine values of b required to guarantee optimal breakdown number and optimal robustness for a design in v treatments. The algorithm evaluates the smallest number of linking blocks for treatment partitions with every possible cardinality for \mathcal{V}_1 . As will be demonstrated, the bounds for b_0 achieved improve on those obtained from Theorem 3.2 in many cases. A preliminary result is needed.

LEMMA 4.2. Let D satisfy the condition of Lemma 4.1. A lower bound for the number of linking blocks for a treatment partition such that \mathcal{V}_1 has cardinality u is given by

$$(4.2) \quad b_u = \min_{\theta \in \Theta_u} \{ \max\{u(r-u+1) + \theta, (v-u)(r-v+u+1) + \eta - \theta\} \},$$

where $\Theta_u = \{\max\{0, \eta - v + u\}, \dots, \min\{u, \eta\}\}$.

THEOREM 4.1. Let D satisfy the condition of Lemma 4.1. Define $\omega_1 = \min_{u \in U} \{b_u\}$. Then $\Omega_1(D) \geq \omega_1$.

PROOF. For a treatment partition such that \mathcal{V}_1 has cardinality u , Lemma 4.2 establishes that b_u is a lower bound for the number of linking blocks between \mathcal{V}_1 and \mathcal{V}_2 . Thus, $\omega_1 = \min_{u \in U} \{b_u\}$ is a lower bound for the number of linking blocks for any treatment partition and the result is established. \square

The number of blocks required to guarantee optimal breakdown number and optimal robustness can be identified from Theorem 4.1. Designs with $\omega_1 \geq r$ have optimal breakdown number and those with $\omega_1 > r$ have optimal robustness. Table 1 gives values of ω_1 and of $b_{\dagger} = \min\{r, \omega_1\}$ for $v = 4, \dots, 15$. It follows that

TABLE 1
 Values of $b_{\dagger} = \min\{r, \omega_1\}$ for $4 \leq v \leq 15$. The smallest numbers of blocks to guarantee optimal breakdown number and optimal robustness are labelled *ob* and *or*, respectively

v	b	r	ω_1	b_{\dagger}	v	b	r	ω_1	b_{\dagger}	v	b	r	ω_1	b_{\dagger}
4	3	1	1	1 <i>ob</i>	9	20	4	4	4	13	38	5	4	4
4	4	2	2	2	9	21	4	5	4 <i>or</i>	13	39	6	6	6 <i>ob</i>
4	5	2	3	2 <i>or</i>	10	21	4	1	1	13	40	6	6	6
5	5	2	2	2 <i>ob</i>	10	22	4	2	2	13	41	6	6	6
5	6	2	2	2	10	23	4	3	3	13	42	6	6	6
5	7	2	3	2 <i>or</i>	10	24	4	4	4 <i>ob</i>	13	43	6	7	6 <i>or</i>
6	7	2	1	1	10	25	5	5	5	14	43	6	1	1
6	8	2	2	2 <i>ob</i>	10	26	5	6	5 <i>or</i>	14	44	6	2	2
6	9	3	3	3	11	26	4	2	2	14	45	6	3	3
6	10	3	4	3 <i>or</i>	11	27	4	4	4 <i>ob</i>	14	46	6	4	4
7	10	2	2	2 <i>ob</i>	11	28	5	5	5	14	47	6	5	5
7	11	3	3	3	11	29	5	5	5	14	48	6	6	6 <i>ob</i>
7	12	3	3	3	11	30	5	5	5	14	49	7	7	7
7	13	3	4	3 <i>or</i>	11	31	5	6	5 <i>or</i>	14	50	7	8	7 <i>or</i>
8	13	3	1	1	12	31	5	1	1	15	50	6	2	2
8	14	3	2	2	12	32	5	2	2	15	51	6	4	4
8	15	3	3	3 <i>ob</i>	12	33	5	3	3	15	52	6	6	6 <i>ob</i>
8	16	4	4	4	12	34	5	4	4	15	53	7	7	7
8	17	4	5	4 <i>or</i>	12	35	5	5	5 <i>ob</i>	15	54	7	7	7
9	17	3	2	2	12	36	6	6	6	15	55	7	7	7
9	18	4	4	4 <i>ob</i>	12	37	6	7	6 <i>or</i>	15	56	7	7	7
9	19	4	4	4	13	37	5	2	2	15	57	7	8	7 <i>or</i>

$b_{\dagger} \leq \min\{r, \Omega_1(D)\} \leq b_0$. For each value of v , the number of blocks ranges from that required to guarantee connectivity by Lemma 4.1, to the minimum to ensure optimal robustness. For example, for $v = 10$, any design with at least 21 blocks is connected, and the numbers of blocks to guarantee optimal breakdown number and optimal robustness are 24 and 26, respectively.

In many cases, lower bounds for b_0 achieved by Theorem 4.1 improve on those of Section 3. For example, for $v = 10$ and $b = 24$, Theorem 4.1 gives $\omega_1 \geq 4 = r$, showing that all such designs have optimal breakdown number $b_0 = 4$. With the same treatment replications, Theorem 3.2 only establishes that $b_0 \geq 3$. Likewise, for $v = 10$ and $b = 26$, Theorem 4.1 gives $\omega_1 \geq 6 > r$ and all the designs are optimally robust. This is not established from Corollary 3.4, as seen from Example 4(ii).

For large v the number of blocks required by Theorem 4.1 to guarantee desirable robustness properties might exceed resources. If this is the case and if candidate designs are available, robustness properties can be determined using the approach of Godolphin and Warren (2014). This essentially involves scanning potential designs to identify linking blocks for each treatment partition. Alternatively,

a cyclic construction, with all treatments having even replication $r \geq 4$, yields optimally robust designs.

4.2. *Cyclic designs.* A class of cyclic designs with good robustness properties is now described.

DEFINITION 4.1. Let $\delta < v/2$. A set of v blocks containing treatment pairs i and $i + \delta$, for $i = 1, \dots, v$, where $i + \delta \equiv i + \delta - v$ if $i + \delta > v$ is termed a *loop* with jump δ .

The blocks of a loop for which δ and v are coprime form a cycle which includes all v treatments. Such a loop is a connected design in v blocks and has $b_0 = 2$. Note that the term *loop* is sometimes used to specify a set of v blocks in a row-column design; see, for example, Kerr and Churchill (2001), Wit, Nobile and Khanin (2005). Such arrangements will be termed *directed loops* in this work and are covered in Section 6.2.

THEOREM 4.2. Let D comprise $p \geq 2$ loops L_1, \dots, L_p with distinct jumps $\delta_1, \dots, \delta_p$, such that $\delta_1, \dots, \delta_q$ are coprime to v for some q with $1 \leq q \leq p$. Then $\Omega_1(D) \geq 2p + q$ and D is optimally robust.

EXAMPLE 6. Consider the two loop design for $v = 10$ with $\delta_1 = 1$ and $\delta_2 = 3$:

$$D4 = \begin{array}{cccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 1 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 1 & 2 & 3 \end{array}.$$

Design $D4$ has $p = q = 2$. By Theorem 4.2, $\Omega_1(D4) \geq 6$ and $D4$ is optimally robust.

5. A second blocking factor and applications to microarray experiments.

Detailed descriptions of two-colour microarray experiments are given by Kerr and Churchill (2001) and Nguyen et al. (2002). To summarise, and to relate notation used in this work to microarray experiments: the treatments comprise v biological samples which are applied to b microarray slides, with two samples applied to each slide. This justifies modelling each slide as a block of size two. One sample on each slide is labelled with a red fluorescent dye and the other with a green fluorescent dye. The two dye aspect of the experiment is accommodated by the inclusion of a second blocking factor in the model, and is the motivation for treating microarray experiments as row column designs, as in Bailey (2007). Such a design can be represented by a $2 \times b$ array, with *columns* corresponding to the blocks of size 2 and *rows* to the additional blocking factor comprising two blocks of size b . As an alternative and useful representation, which parallels the use of concurrence graphs in Section 3, the designs can be depicted as digraphs on v vertices: a block with treatments v_1 and v_2 in rows 1 and 2 corresponds to an arc, that is, directed

edge, from vertex v_1 to vertex v_2 . Kerr and Churchill (2001), Wit, Nobile and Khanin (2005), Mahbub Latif, Bretz and Brunner (2009), amongst others, include dye colour in graphical portrayals of microarray designs in this way.

In this section and in Section 6, row column designs are investigated with emphasis on properties of eventual designs resulting from the loss of blocks of size two, and on the construction of designs which are robust to observation loss. This is particularly relevant for microarray experiments since, as described by Mahbub Latif, Bretz and Brunner (2009), it is common for these to experience observations loss.

5.1. *Preliminaries for row-column designs.* Model 1 is amended to Model 2 which incorporates a two level blocking factor, that is, *dye colour* for microarray experiments. For Model 2, the observation vector Y is given by

$$(5.1) \quad E(Y) = \mu 1_{2b} + X_1\beta_1 + X_2\beta_2 + X_3\beta_3,$$

where β_3 is the 2×1 parameter vector relating to the second blocking factor and X_3 is the corresponding $2b \times 2$ component of the design matrix, with each row containing one element unity and one element zero. Other terms are as in (2.1). As with Model 1 designs, no block of size two contains two replicates of the same treatment, and no two blocks contain the same pair of treatments. The design matrix is $X = [1_{2b} \ X_1 \ X_2 \ X_3]$ and the information matrix is

$$(5.2) \quad C = X_1^T \left(I - \frac{1}{2} X_2 X_2^T - \frac{1}{b} X_3 X_3^T + \frac{1}{2b} 1_{2b} 1_{2b}^T \right) X_1.$$

For such designs, (2.2) holds and there is a parallel relationship between columns of X_1 and X_3 :

$$(5.3) \quad X_1 1_v = X_3 1_2.$$

In Sections 5–6, D is assumed to be a design following Model 2. A design is disconnected iff there is a relationship between the columns of X_1 and the columns of the blocking component matrices, which is independent of (2.2) and (5.3). That is, there is a relationship of the form

$$(5.4) \quad X_1 \alpha_1 = X_2 \alpha_2 + X_3 \alpha_3,$$

where α_1 is not a multiple of 1_v . If such a relationship exists, not all treatment contrasts will be estimable and the aims of the experiment will not be achieved. Unlike the situation with Model 1, a disconnected Model 2 design is not necessarily characterised by the existence of a consistent partition between treatments and a single blocking factor. Relationships covered by (5.4) can be categorised as three types and these are now described briefly; see Godolphin and Godolphin (2001) for a more thorough treatment.

Type 1: α_2 is not a multiple of 1_b , α_3 is a multiple of 1_2

In this case, the additional relationship is identical in nature to that observed in a

disconnected design of Model 1 and can be given as (2.5). The disconnectivity is characterised by the existence of a treatment partition which is consistent with a partition of the blocks of size two.

Type 2: α_2 is a multiple of 1_b , α_3 is not a multiple of 1_2
The additional relationship can be expressed as

$$X_1\alpha_1 = X_3\alpha_3.$$

Here, the treatments are partitioned, with all replicates of a subset of the treatments contained in one row and all replicates of the remaining treatments in the other row.

Type 3: α_2 is not a multiple of 1_b and α_3 is not a multiple of 1_2
Such a relationship is more complex than types 1 and 2 and cannot be simplified from (5.4). It is characterised by a partition of the treatments into sets $\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_z$, where $3 \leq z \leq v$, which is consistent with a joint partition of the columns into sets $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_{z-1}$, and of the rows, such that, treatments in row j of blocks in \mathcal{B}_i are from \mathcal{W}_{i+j-1} for $i = 1, \dots, z$ and $j = 1, 2$. After possible rearrangement of the blocks of size two, but not of the position of treatments within blocks, the relationship is represented diagrammatically as a *staircase* structure, with treatments from the same \mathcal{W}_i lying in a North East to South West diagonal:

$$(5.5) \quad \begin{array}{c|ccccc} & \mathcal{B}_1 & \mathcal{B}_2 & \mathcal{B}_3 & \cdots & \mathcal{B}_{z-1} \\ \hline \text{row 1} & \mathcal{W}_1 & \mathcal{W}_2 & \mathcal{W}_3 & \cdots & \mathcal{W}_{z-1} \\ \text{row 2} & \mathcal{W}_2 & \mathcal{W}_3 & \mathcal{W}_4 & \cdots & \mathcal{W}_z \end{array}$$

For a pairwise treatment contrast to be estimable, both treatments must be in the same set.

As with Model 1 designs, $D(p)$ denotes the class of eventual designs in which p blocks of size two are lost from D , and the breakdown number, b_0 , is the smallest number of blocks of size two that need to be lost before the possibility of an eventual design from which not all treatment contrasts are estimable. In this situation, all replicates of one or more treatments will have been lost or the eventual design will be disconnected. The notation $\Omega_1(D)$, introduced in Section 2, is extended for Model 2 designs, so that $\Omega_i(D)$ is the smallest number of blocks of size two that need to be lost before there is at least one $D_\#$ in which there is a relationship of type i , for $i \in \{1, 2, 3\}$. Thus, we have the following.

DEFINITION 5.1. The breakdown number of D is given by

$$(5.6) \quad b_0 = \min\{r_{[v]}, \Omega_1(D), \Omega_2(D), \Omega_3(D)\}.$$

The main aim is to develop a construction approach to ensure that b_0 is sufficiently large that breakdown will not occur, given anticipated levels of observation loss. A secondary aim is the determination of a lower bound for b_0 for a given

design. For $p < r_{[v]}$, an eventual design, $D_{\#}$, in $D(p)$ follows Model 2 and has design matrix $X_{\#} = [1_{2(b-p)} \ X_{1\#} \ X_{2\#} \ X_{3\#}]$, obtained by removal of $2p$ rows from X and p columns from X_2 . All treatment contrasts are estimable in $D_{\#}$ unless a relationship of type i exists for at least one $i \in \{1, 2, 3\}$. Determining b_0 for a specific design is technically straightforward, since a Model 2 design is connected iff the rank of the design matrix is $v + b$. However, even for moderate values of b and $r_{[v]}$, the process of checking the design matrix rank for every member of $D(p)$, for $p < r_{[v]}$, can be computationally demanding due to the large number of eventual designs. Thus, derivation of lower bounds for each $\Omega_i(D)$ in terms of design properties is useful towards both aims. As with Model 1 designs, Model 2 designs are described as having optimal breakdown number or being optimally robust with regards to robustness in the event of the loss of blocks of size two.

5.2. *Robustness of row-column designs.* It is stressed that consideration of $\Omega_i(D)$ for all relationship types is important to avoid the use of designs with smaller b_0 and, therefore, greater vulnerability to breakdown, than anticipated. In particular, the possibility of a type 3 relationship for members of $D(p)$ for small values of p should not be overlooked. This is illustrated by reference to Tsai and Liao (2013) who investigate designs with regards to robustness against loss of blocks of size two. For given v and b , all designs following Model 2 are identified. For each design, a computationally intensive procedure seeks to determine $M\{D\}_p$, the number of members of $D(p)$ for which not all treatment contrasts are estimable, for $1 \leq p \leq b - v + 1$. For two designs, DI and DII say, DI is said to have less breakdown than DII if there is an integer q such that $M\{DI\}_j = M\{DII\}_j$ for $j < q$ and $M\{DI\}_q < M\{DII\}_q$. If no design has less breakdown than DI , then DI is a *minimum breakdown design* (MBD). Tsai and Liao describe the procedure as computationally infeasible for large b or v and only generate results for $4 \leq v \leq 9$. However, the process of design assessment appears to overlook type 3 relationships, as the following example shows.

EXAMPLE 7.

$$DTL = \begin{array}{c|cccccccccccc} \text{row 1} & 1 & 1 & 7 & 2 & 2 & 7 & 5 & 6 & 3 & 5 & 6 & 4 \\ \text{row 2} & 5 & 6 & 1 & 5 & 6 & 2 & 3 & 3 & 7 & 4 & 4 & 7 \end{array}$$

Design DTL has $r_{[1]} = r_{[2]} = r_{[3]} = 4, r_{[4]} = \dots = r_{[7]} = 3$. A digraph representation of DTL is given in Figure 4. Tsai and Liao (2013) report DTL as a MBD with $b_0 = 3$ and $M\{DTL\}_3 = 4$, suggesting that DTL is optimally robust. This would be correct if DTL was a Model 1 design. However, when viewed as a row column design according to Model 2, DTL is far more vulnerable to observation loss. In fact, the correct value of $M\{DTL\}_3$ is 28 and, of even more concern, $M\{DTL\}_2 = 2$ with two members of $DTL(2)$ being disconnected via type 3 relationships. To demonstrate this, consider the loss of columns 9 and 12. The treatments and remaining blocks can be partitioned as $\mathcal{W}_1 = \{\text{treatment 7}\}$,

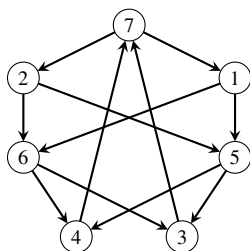


FIG. 4. The digraph corresponding to DTL of Example 7.

$\mathcal{W}_2 = \{\text{treatments } 1, 2\}$, $\mathcal{W}_3 = \{\text{treatments } 5, 6\}$, $\mathcal{W}_4 = \{\text{treatments } 3, 4\}$ and $\mathcal{B}_1 = \{\text{blocks } 3, 6\}$, $\mathcal{B}_2 = \{\text{blocks } 1, 2, 4, 5\}$, $\mathcal{B}_3 = \{\text{blocks } 7, 8, 10, 11\}$. A consequence of the treatment groupings of the type 3 relationship is that only three of the 21 pairwise treatment contrasts are estimable. Arranging the ten remaining columns according to the joint partition illustrates the type 3 relationship:

	\mathcal{B}_1		\mathcal{B}_2				\mathcal{B}_3			
row 1	7	7	1	1	2	2	5	6	5	6
row 2	1	2	5	6	5	6	3	3	4	4

Thus, $\Omega_3(\text{DTL}) \leq 2$ and, from (5.6), $b_0 \leq 2$.

Attention is now given to the general problem of obtaining lower bounds for each $\Omega_i(D)$ and in gaining an understanding of design features which promote larger $\Omega_i(D)$ values.

$\Omega_1(D)$. Results of Sections 3–4 can be used to determine lower bounds for $\Omega_1(D)$, for a design of Model 2, and to provide information on design properties to guarantee desired levels of robustness against type 1 relationships. Alternatively, for a given candidate design, the algorithm of Godolphin and Warren (2014) can identify $\min\{r_{\lceil v \rceil}, \Omega_1(D)\}$.

$\Omega_2(D)$. A relationship of type 2 exists in D or $D_\#$ iff rows 1 and 2 have no common treatments. Let $r_{i,j}$ be the number of replicates of treatment i in row j of D . Then at least $\min\{r_{i,1}, r_{i,2}\}$ of these need to be lost before the possibility of a type 2 relationship. This leads to a lower bound for $\Omega_2(D)$.

THEOREM 5.1. For design D , let $\omega_2 = \lceil \sum_{i=1}^v \min\{r_{i,1}, r_{i,2}\} / 2 \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling. Then $\Omega_2(D) \geq \omega_2$.

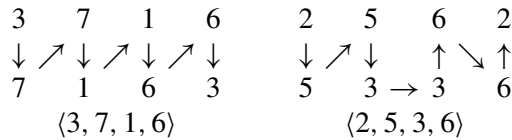
When constructing a design, monitoring the distribution of treatment replicates between the rows will ensure that $\omega_2 > r_{\lceil v \rceil}$, so that b_0 is not limited by type 2 relationships. For microarray experiments, this strategy is consistent with recommendations of Kerr and Churchill (2001) and Bailey (2007), amongst others, who advocate allocating dye colour as equally as possible for each treatment.

Work to obtain a lower bound for $\Omega_3(D)$, which follows, also provides an alternative bound for $\Omega_2(D)$.

$\Omega_3(D)$. Approaches to determine a lower bound for $\Omega_3(D)$ for a specific design, and to construct designs which are not vulnerable to type 3 relationships, use concepts similar to those of Butz (1982).

DEFINITION 5.2. Let v_1, \dots, v_n be a subset of $n \geq 3$ treatments and let b_1, \dots, b_n be a subset of n blocks of size two from D , such that the treatments in b_i are v_i and v_{i+1} , for $i = 1, \dots, n - 1$, and those in b_n are v_1 and v_n . Such an arrangement is termed an n -circuit, or circuit of length n and denoted $\langle v_1, \dots, v_n \rangle$.

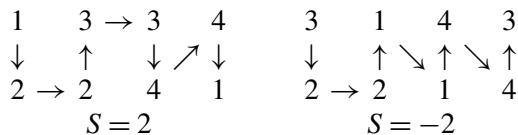
Design DTL of Example 7 contains many 4-circuits and 6-circuits. For example, columns 2, 3, 8 and 9 comprise the 4-circuit, $\langle 3, 7, 1, 6 \rangle$. Another 4-circuit, $\langle 2, 5, 3, 6 \rangle$, is formed from columns 4, 5, 7 and 8. These circuits can be seen in Figure 4 and are also depicted as



DEFINITION 5.3. The sum of an n -circuit is $S = \sum_{i=1}^n s_i$, where $s_1 = 1$ and, for $i = 2, \dots, n$, $s_i = s_{i-1}$ if treatments common to blocks i and $i - 1$ are in different rows, and $s_i = -s_{i-1}$ otherwise.

DEFINITION 5.4. An n -circuit with $S = n$ is termed an n -cycle.

For the two circuits of DTL already referred to: $\langle 3, 7, 1, 6 \rangle$ has $S = 4$ and is a 4-cycle; $\langle 2, 5, 3, 6 \rangle$ has $S = 0$. Note that the sign of a circuit sum relates to the direction of travel around the circuit and can depend on the initial block. For example, the following 4-circuits are equivalent, being formed from the same set of four blocks, but have sums $S = 2$ and $S = -2$, respectively:



The blocks of the circuits are displayed in Figure 5. Traversing the circuit in a clockwise direction gives $S = 2$ and in an anti-clockwise direction gives $S = -2$. In general, traversing an arc in the direction of the arc contributes $+1$ to S and traversing in the opposite direction contributes -1 . The initial block fixes the direction of travel. Note that circuits of odd length cannot have $S = 0$.

Lower bounds for $\Omega_2(D)$ and $\Omega_3(D)$ are now obtained in terms of disjoint circuits in D .

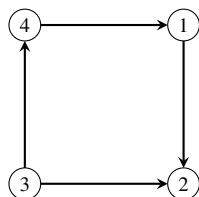


FIG. 5. A 4-circuit.

THEOREM 5.2. *Let a subset of the blocks of D be allocated to ω_3 disjoint sets, for $j = 1, \dots, \omega_3$, with blocks from the j th set forming a circuit with nonzero sum, S_j . Then $\Omega_3(D) \geq \omega_3$ and $\Omega_2(D) \geq \omega_2 \geq \sum_{i=1}^{\omega_3} |S_i|/2$.*

When the circuits of Theorem 5.2 are all cycles, the bound for $\Omega_2(D)$ from Theorem 5.2 is maximised.

COROLLARY 5.1. *Let the nonzero sum circuits of Theorem 5.2 be cycles of lengths n_1, \dots, n_{ω_3} . Then $\Omega_2(D) \geq \omega_2 \geq \sum_{i=1}^{\omega_3} n_i/2$.*

The use of disjoint cycles to provide lower bounds for $\Omega_2(D)$ and $\Omega_3(D)$ is illustrated by returning to Example 7.

EXAMPLE 7 (Revisited). A subset of the blocks of DTL can be arranged into two disjoint cycles in various ways. For example, the twelve blocks can be arranged into two 4-cycles, each with $S = 4$, namely $(3, 7, 1, 6)$ and $(4, 7, 2, 5)$, and four remaining blocks:

$$\text{DTL} = \begin{array}{|c|c|c|c|} \hline 3 & 7 & 1 & 6 \\ \hline \downarrow \nearrow & \downarrow \nearrow & \downarrow \nearrow & \downarrow \nearrow \\ \hline 7 & 1 & 6 & 3 \\ \hline \end{array} \quad \begin{array}{|c|c|c|c|} \hline 4 & 7 & 2 & 5 \\ \hline \downarrow \nearrow & \downarrow \nearrow & \downarrow \nearrow & \downarrow \nearrow \\ \hline 7 & 2 & 5 & 4 \\ \hline \end{array} \quad \begin{array}{|c|c|c|c|} \hline 1 & 2 & 5 & 6 \\ \hline & & & \\ \hline 5 & 6 & 3 & 4 \\ \hline \end{array}.$$

With $r_{[1]} = 4$ and $r_{[7]} = 3$, the results of Section 4 apply: from Table 1, $\Omega_1(\text{DTL}) \geq 3$. Theorems 5.1 and 5.2 give $\Omega_2(\text{DTL}) \geq 4$ and $\Omega_3(\text{DTL}) \geq \omega_3 = 2$. Thus, $b_0 \geq 2$ by (5.6). Removal of the blocks containing treatment pairs 3 and 7, and 4 and 7 destroys both cycles and gives the disconnected member of $\text{DTL}(2)$ displayed earlier. Removal of the blocks containing treatment pairs 1 and 7, and 2 and 7 gives the other disconnected member of $\text{DTL}(2)$. Breaking both cycles by the removal of other pairs of blocks does not lead to a $D_{\#}$ containing a type 3 relationship since, in each case, the eventual design still contains at least one circuit with nonzero sum.

Using the understanding gained of design features that promote robustness against each relationship type, an alternative to DTL is proposed, with the same treatment replicates, but with improved robustness properties.

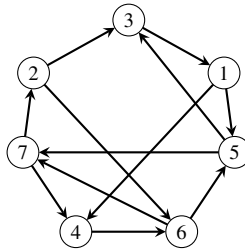


FIG. 6. The digraph corresponding to DTLI of Example 8.

EXAMPLE 8. The treatment replicates of Example 7 are arranged as three disjoint cycles and two additional blocks, to give DTLI:

$$\text{DTLI} = \begin{array}{|c|c|c|} \hline 1 & 5 & 3 \\ \hline \downarrow & \nearrow & \downarrow \\ \hline 5 & 3 & 1 \\ \hline \end{array} \begin{array}{|c|c|c|} \hline 2 & 6 & 7 \\ \hline \downarrow & \nearrow & \downarrow \\ \hline 6 & 7 & 2 \\ \hline \end{array} \begin{array}{|c|c|c|c|} \hline 4 & 6 & 5 & 7 \\ \hline \downarrow & \nearrow & \downarrow & \nearrow \\ \hline 6 & 5 & 7 & 4 \\ \hline \end{array} \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 4 & 3 \\ \hline \end{array}$$

As with DTL, Table 1 gives $\omega_1 = 3$, so $\Omega_1(\text{DTLI}) \geq 3$. By Theorems 5.1 and 5.2, $\Omega_2(\text{DTLI}) \geq 5$ and $\Omega_3(\text{DTLI}) \geq 3$. Thus, $b_0 = 3$ by (5.6), which establishes that DTLI has optimal breakdown number. Further, a check involving the removal of each set of three blocks indicates that DTLI is optimally robust. The digraph representing DTLI is given in Figure 6.

5.3. Robust design construction and optimality conflict. Strategies for constructing designs with good robustness properties are suggested by the definition of b_0 in (5.6) and results in Section 4 and Section 5. If all treatments can be replicated $r \geq 4$ times where r is even, then cyclic designs should be considered. These are covered in Section 6.2. Otherwise, having treatments replicated as equally as possible has the advantages that $r_{[v]}$ is maximised for the resources used, and that Theorem 4.1 applies. If possible, b should be such that ω_1 of Theorem 4.1 is large enough to provide appropriate robustness against type 1 relationships, given the anticipated level of observation loss for the experiment. Second, Theorem 5.2 indicates that constructing designs with as many disjoint cycles as possible gives good robustness against type 2 and 3 relationships.

There are two issues with this combined approach. For large v , the number of blocks required by Theorem 4.1 to guarantee desired levels of robustness may exceed the resources available. In this case, a candidate design, D , can be constructed just taking ω_2 and ω_3 into consideration. By then treating D as a design of Model 1, $\min\{r_{[v]}, \Omega_1(D)\}$ can be determined by the Godolphin and Warren (2014) scanning procedure to check that b_0 is of sufficient size. The other issue relates to a conflict between design features which maximise ω_3 and those which promote optimality, and is now discussed.

The quality of a connected design in v qualitative treatments is typically measured by the A -criterion:

$$(5.7) \quad \Phi_A(D) = (v - 1) \left(\sum_{i=1}^{v-1} \mu_i^{-1} \right)^{-1},$$

where μ_1, \dots, μ_{v-1} are the nonzero eigenvalues of the information matrix (5.2). Within a design class, D is A -optimal if no design has A -criterion higher than $\Phi_A(D)$. Paterson (1983) conjectures that designs with minimal numbers of short circuits tend to have higher A -criterion than those with greater numbers of short circuits. The implication from Theorem 5.2 that constructing designs with as many disjoint cycles as possible, to achieve good robustness against type 2 and 3 relationships, favours designs with short cycles. Thus, there is some incompatibility between designs which guarantee good robustness properties and designs with good A -criterion. This is most problematic if v and $2b/v$ are both relatively small, when even achieving a small number of disjoint cycles may require 3-cycles.

The conflict is illustrated by Examples 6 and 7. The robustness properties of DTLI have been demonstrated to be superior to those of DTL. However, whilst the shortest circuits in DTL are 4-circuits, DTLI contains several 3-circuits, as noted by inspection of the digraphs in Figures 4 and 6. The A -criterion values reflect these design features, with $\Phi_A(\text{DTL}) = 1.7809$ and $\Phi_A(\text{DTLI}) = 1.7107$, indicating that DTL has better A -criterion.

A suggested approach to construct designs which are robust against type 2 and 3 relationships is to arrange treatment replicates to form up to $r_{[v]} + 1$ disjoint cycles, and to make the shortest cycle as long as possible. In practice, with $r_{[v]} \geq 4$ and $v \geq 10$ it has been straightforward to construct cycles without compromising the A -criterion unduly.

6. Row-column designs with equal replication. In this section, D is assumed to follow Model 2 and to have all treatments equally replicated for some $r \in \{3, \dots, v - 1\}$, with applications divided as evenly as possible between the rows. That is, for r even, each treatment is replicated $r/2$ times in each row and, for r odd and v even, each treatment has $[r/2]$ replicates in one row and $[r/2] + 1$ in the other. Designs of these forms include the interwoven loop designs investigated by Wit, Nobile and Khanin (2005), where r is even, and the designs for $r = 3, 4, 5, 6$ with up to 30 blocks given in Bailey (2007).

With every treatment occurring in both rows, $\Omega_2(D) > 0$ and $\Omega_3(D) > 0$. Therefore, (4.1) can be used as a condition for connectivity. To obtain information on the robustness guaranteed by D , the relationship types are considered separately. For robustness against type 1 relationships, Theorem 4.1 applies. If resources do not enable b to be large enough for Theorem 4.1 to guarantee the desired level of robustness, the approach of Godolphin and Warren (2014) can be

used to assess candidate designs for connectivity and robustness against type 1 relationships. For robustness against type 2 relationships, Theorem 5.1 gives

$$(6.1) \quad \Omega_2(D) \geq \omega_2 = \left\lceil \frac{v}{2} \left\lceil \frac{r}{2} \right\rceil \right\rceil.$$

From (6.1), $\Omega_2(D) > r$, except for the case $v = 4, r = 3$, for which $\Omega_2(D) \geq 2$.

6.1. *Robustness against type 3 relationships.* Determination of robustness against type 3 relationships is more challenging. Consideration of possible arrangements of the treatment replicates, taking into account the binary nature of the blocks of size two and the property that each pair of treatments appears together in no more than one such block, enables identification of configurations involving the largest subset of the $vr/2$ blocks that can exhibit a type 3 relationship. These are now described and illustrated for r even and odd.

r even. The largest number of blocks that can exhibit a type 3 relationship is $vr/2 - r^2/4$. This occurs when $|\mathcal{W}_2|, \dots, |\mathcal{W}_{z-1}|$ are all $r/2$, where $|\mathcal{W}_i|$ denotes the cardinality of \mathcal{W}_i , and $|\mathcal{W}_1|$ and $|\mathcal{W}_z|$ have values in $\{1, \dots, r/2\}$, such that the sum of the z cardinalities is v . The $r^2/4$ omitted blocks must contain: $r/2$ replicates of each treatment of \mathcal{W}_z and $r/2 - |\mathcal{W}_z|$ replicates of each treatment of \mathcal{W}_{z-1} in row 1; $r/2$ replicates of each treatment of \mathcal{W}_1 and $r/2 - |\mathcal{W}_1|$ replicates of each treatment of \mathcal{W}_2 in row 2. Any other type 3 arrangement contains fewer blocks. For illustration, configurations for $v = 10, r = 4$ are displayed, with details of the omitted treatment replicates.

Configuration 1. $\mathcal{W}_1 = \{1\}, \mathcal{W}_2 = \{2, 3\}, \mathcal{W}_3 = \{4, 5\}, \mathcal{W}_4 = \{6, 7\}, \mathcal{W}_5 = \{8, 9\}, \mathcal{W}_6 = \{10\}$. Omitted replicates are 8, 9, 10, 10 in row 1 and 1, 1, 2, 3 in row 2.

	\mathcal{B}_1	\mathcal{B}_2	\mathcal{B}_3	\mathcal{B}_4	\mathcal{B}_5
row 1	1 1	2 2 3 3	4 4 5 5	6 6 7 7	8 9
row 2	2 3	4 5 4 5	6 7 6 7	8 9 8 9	10 10

Configuration 2. $\mathcal{W}_1 = \{1, 2\}, \mathcal{W}_2 = \{3, 4\}, \mathcal{W}_3 = \{5, 6\}, \mathcal{W}_4 = \{7, 8\}, \mathcal{W}_5 = \{9, 10\}$. Omitted replicates are 9, 9, 10, 10 in row 1 and 1, 1, 2, 2 in row 2.

	\mathcal{B}_1	\mathcal{B}_2	\mathcal{B}_3	\mathcal{B}_4
row 1	1 1 2 2	3 3 4 4	5 5 6 6	7 7 8 8
row 2	3 4 3 4	5 6 5 6	7 8 7 8	9 10 9 10

r odd. The largest number of blocks demonstrating a type 3 relationship is $vr/2 - [r/2]^2$. This occurs when $|\mathcal{W}_1| = [r/2], |\mathcal{W}_2| = |\mathcal{W}_3| = [r/2] + 1, |\mathcal{W}_z| = [r/2]$ and the staircase is formed from a Basic Structure and arbitrary numbers of Structure A and/or Structure B arrangements positioned between \mathcal{W}_3 and \mathcal{W}_z , such that the sum of all cardinalities is v . Structure A arrangements contain treatment

sets \mathcal{W}_i and \mathcal{W}_{i+1} , with $|\mathcal{W}_i| = |\mathcal{W}_{i+1}| = [r/2] + 1$, and Structure B arrangements contain sets $\mathcal{W}_j, \mathcal{W}_{j+1}, \mathcal{W}_{j+2}$ and \mathcal{W}_{j+3} with $|\mathcal{W}_j| = |\mathcal{W}_{j+1}| = [r/2]$ and $|\mathcal{W}_{j+2}| = |\mathcal{W}_{j+3}| = [r/2] + 1$. Within any given treatment set, all the treatments have higher replication in the same row. The arrangements are depicted in (6.2)–(6.4), with the row containing $[r/2] + 1$ replicates indicated by (h) for each treatment set. If there are no Structure A or B arrangements, then $z = 4$.

(6.2) Basic Structure

	\mathcal{B}_1	\mathcal{B}_2	\mathcal{B}_3	\cdots	\mathcal{B}_{z-1}
row 1	$\mathcal{W}_1(h)$	$\mathcal{W}_2(h)$	\mathcal{W}_3	\cdots	\cdots
row 2	\mathcal{W}_2	$\mathcal{W}_3(h)$	\cdots	\cdots	$\mathcal{W}_z(h)$

(6.3) Structure A

	\mathcal{B}_{i-1}	\mathcal{B}_i	\mathcal{B}_{i+1}
row 1	\cdots	$\mathcal{W}_i(h)$	\mathcal{W}_{i+1}
row 2	\mathcal{W}_i	$\mathcal{W}_{i+1}(h)$	\cdots

(6.4) Structure B

	\mathcal{B}_{j-1}	\mathcal{B}_j	\mathcal{B}_{j+1}	\mathcal{B}_{j+2}	\mathcal{B}_{j+3}
row 1	\cdots	\mathcal{W}_j	$\mathcal{W}_{j+1}(h)$	$\mathcal{W}_{j+2}(h)$	\mathcal{W}_{j+3}
row 2	$\mathcal{W}_j(h)$	\mathcal{W}_{j+1}	\mathcal{W}_{j+2}	$\mathcal{W}_{j+3}(h)$	\cdots

The $[r/2]^2$ omitted blocks contain: $[r/2]$ replicates of each treatment of \mathcal{W}_z in row 1 and $[r/2]$ replicates of each treatment of \mathcal{W}_1 in row 2. Any other type 3 arrangement contains fewer blocks. For illustration, the configurations for $v = 10, r = 3$ and $v = 10, r = 5$ are displayed.

Configuration for $v = 10, r = 3$. $\mathcal{W}_1 = \{1\}, \mathcal{W}_2 = \{2, 3\}, \mathcal{W}_3 = \{4, 5\}, \mathcal{W}_4 = \{6, 7\}, \mathcal{W}_5 = \{8, 9\}, \mathcal{W}_6 = \{10\}$. Omitted replicates are 10 in row 1 and 1 in row 2. Note that \mathcal{W}_4 and \mathcal{W}_5 are contained in a Structure A arrangement.

	\mathcal{B}_1		\mathcal{B}_2			\mathcal{B}_3		\mathcal{B}_4			\mathcal{B}_5			
row 1	1	1	2	2	3	3	4	5	6	6	7	7	8	9
row 2	2	3	4	5	4	5	6	7	8	9	8	9	10	10

Configuration for $v = 10, r = 5$. $\mathcal{W}_1 = \{1, 2\}, \mathcal{W}_2 = \{3, 4, 5\}, \mathcal{W}_3 = \{6, 7, 8\}, \mathcal{W}_4 = \{9, 10\}$. Omitted replicates are 9, 9, 10, 10 in row 1 and 1, 1, 2, 2 in row 2. The arrangement is a Basic Structure.

	\mathcal{B}_1					\mathcal{B}_2					\mathcal{B}_3										
row 1	1	1	1	2	2	2	3	3	3	4	4	4	5	5	5	6	6	7	7	8	8
row 2	3	4	5	3	4	5	6	7	8	6	7	8	6	7	8	9	10	9	10	9	10

The results of the combinatorial investigation are expressed succinctly as the following.

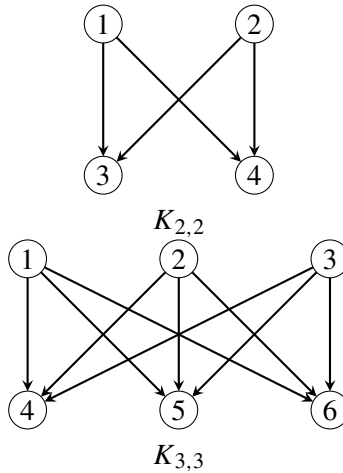


FIG. 7. Complete bipartite digraphs.

PROPOSITION 6.1. *Let all treatments in D have replication r and be distributed as evenly as possible between the rows. Then $\Omega_3(D) \geq \lceil r/2 \rceil^2$.*

By Proposition 6.1, if $r \geq 6$, then $\Omega_3(D) > r$. Thus, since $\Omega_2(D) > r$ from (6.1), it follows that $b_0 = \min\{r, \Omega_1(D)\}$. However, for $r = 3, 4, 5$, Proposition 6.1 indicates that $\Omega_3(D)$ can be as small as 1, 4, 4, respectively. From the configurations which minimise $\Omega_3(D)$, it is seen that the corresponding digraphs contain copies of the complete bipartite digraph $K_{t,t}$, where $t = \lceil r/2 \rceil + 1$. The digraphs $K_{2,2}$ and $K_{3,3}$ are displayed in Figure 7. This feature can be used as a means of avoiding designs which have lowest robustness against type 3 relationships for $r = 3, 4, 5$ and is now investigated:

Robustness against type 3 relationships: $r = 3$. Using the result of Proposition 6.1, $\Omega_3(D) \geq 1$. By inspection of (6.2)–(6.4), the digraph of a design with $\Omega_3(D) = 1$ will contain one $K_{2,2}$ sub-digraph for the Basic Structure and one for each Structure A and Structure B arrangement. The Basic Structure involves six treatments and A and B structures involve four and six treatments, respectively. Thus, for the possibility of $\Omega_3(D) = 1$, the number of treatments must be of the form $v = 6 + 4n_A + 6n_B$ where $n_A, n_B \in \{0, \mathbb{Z}^+\}$. This implies that any design with $v = 4$ or $v = 8$ must have $\Omega_3(D) \geq 2$. For $v \neq 4, 8$, let $v = 6 + 4n_{A1} + 6n_{B1}$ be the representation of v in the required form that minimises n_A . For $\Omega_3(D) = 1$, there must be at least $1 + n_{A1} + n_{B1}$ copies of $K_{2,2}$ on disjoint sets of four treatments. Therefore, ensuring that the digraph of a design does not have this property ensures that $\Omega_3(D) \geq 2$.

Robustness against type 3 relationships: $r = 4$. By the result of Proposition 6.1, $\Omega_3(D) \geq 4$. Thus, $b_0 = \min\{4, \Omega_1(D)\}$. Also, the configuration necessary for $\Omega_3(D) = 4$ means that the digraph for such a design contains at least $\lfloor \frac{v-2}{2} \rfloor$ copies of $K_{2,2}$.

Robustness against type 3 relationships: $r = 5$. Using the result of Proposition 6.1, $\Omega_3(D) \geq 4$. The digraph of a design with $\Omega_3(D) = 4$ contains copies of $K_{3,3}$: one copy for the Basic Structure and one for each A or B structure. The Basic Structure involves ten treatments and A and B structures involve six and ten treatments, respectively. For $\Omega_3(D) = 4$ to be a possibility, v must be of the form $10 + 6n_A + 10n_B$ where $n_A, n_B \in \{0, \mathbb{Z}^+\}$. Thus, if $v < 10$ or $v \in \{12, 14, 18, 24\}$, no such configuration is possible and $\Omega_3(D) \geq 5$. Suppose v can be expressed in the required form and let the representation which minimises n_A be $v = 10 + 6n_{A1} + 10n_{B1}$. For $\Omega_3(D) = 4$, there must be at least $1 + n_{A1} + n_{B1}$ copies of $K_{3,3}$ on disjoint sets of six treatments. Otherwise, $\Omega_3(D) \geq 5$.

Thus, for $r = 4, 5$ building designs with $\Omega_3(D) \geq 5$ is straightforward. However, for $r = 3$, constructing designs with good robustness against type 3 relationships can be problematic. Avoidance of the configuration required for $\Omega_3(D) = 1$ still gives the possibility that $\Omega_3(D) = 2$. The approach of Section 5 would be to include three or four nonzero sum disjoint circuits in a design, which can be challenging since no treatment can appear in more than one circuit of such a set. It may not be possible to construct a design containing a set of three or four disjoint circuits if v is small, or, taking optimality into consideration, it may be possible but undesirable.

6.2. *Interwoven loop designs.* The interwoven loop designs investigated by Wit, Nobile and Khanin (2005) can be considered as the Model 2 version of the cyclic designs of Section 4.2. The specification of a *loop* given in Definition 4.1 is amended to accommodate the second blocking factor of Model 2 designs.

DEFINITION 6.1. A set of v blocks containing treatment pairs i in row 1 and $i + \gamma$ in row 2, for $i = 1, \dots, v$, where $1 \leq \gamma \leq v - 1$ and $i + \gamma \equiv i + \gamma - v$ if $i + \gamma > v$ is termed a directed loop with jump γ .

The interwoven loop designs comprise $p \geq 2$ directed loops with jumps $\gamma_1, \dots, \gamma_p$ such that $i \neq j$ implies $\gamma_i \neq \gamma_j$ and $\gamma_i \neq v - \gamma_j$. These designs have the advantages of being simple to construct and of having the good robustness against type 1 relationships of the cyclic designs. More general robustness properties are now established.

THEOREM 6.1. Let D be an interwoven loop design such that $\gamma_1, \dots, \gamma_q$ are coprime to v for some q with $1 \leq q \leq p$. Then D is optimally robust if: (i) $p \geq 3$; or (ii) $p = 2$ and $\gamma_1 = 1, \gamma_2 \neq v/2 + 1$.

PROOF. Each treatment is replicated $r = 2p$ times in D . By Theorem 4.2, $\Omega_1(D) \geq 2p + q > r$. By Theorem 5.1, $\Omega_2(D) > r$. If $p \geq 3$, then $r \geq 6$ and, by Proposition 6.1, $\Omega_3(D) > r$ and (i) holds. If $p = 2$, then $r = 4$ and, by Proposition 6.1, $\Omega_3(D) \geq 4$. The configuration for $\Omega_3(D) = 4$ requires D to contain

copies of $K_{2,2}$. Suppose D has $\gamma_1 = 1$ and $\gamma_2 \neq v/2 + 1$ and does contain $K_{2,2}$ and let v_0 be any treatment involved in row 1 of such an arrangement, then the following four blocks of D form $K_{2,2}$:

row 1	v_0	v_0	$v_0 + \gamma_2 - 1$	$v_0 + \gamma_2 - 1$
row 2	$v_0 + 1$	$v_0 + \gamma_2$	$v_0 + \gamma_2$	$v_0 + 2\gamma_2 - 1$

where $v_0 + 2\gamma_2 - 1 \equiv v_0 + 1$. But this implies that $\gamma_2 = v/2 + 1$ and gives a contradiction. Thus, $\Omega_3(D) > 4 = r$ and (ii) holds. \square

EXAMPLE 9. Consider the interwoven loop designs, $D5$ and $D6$, with $v = 8$ and $r = 4$:

$D5 =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td style="border-right: 1px solid black; padding: 2px;">row 1</td> <td style="padding: 2px;">1</td><td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td> <td style="border-left: 1px solid black; padding: 2px;">1</td><td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">row 2</td> <td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td><td style="padding: 2px;">1</td> <td style="border-left: 1px solid black; padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td><td style="padding: 2px;">1</td><td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">5</td> </tr> </table>	row 1	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	row 2	2	3	4	5	6	7	8	1	4	5	6	7	8	1	2	3	5
row 1	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8																				
row 2	2	3	4	5	6	7	8	1	4	5	6	7	8	1	2	3	5																			
$D6 =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td style="border-right: 1px solid black; padding: 2px;">row 1</td> <td style="padding: 2px;">1</td><td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td> <td style="border-left: 1px solid black; padding: 2px;">1</td><td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">row 2</td> <td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td><td style="padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td><td style="padding: 2px;">1</td> <td style="border-left: 1px solid black; padding: 2px;">6</td><td style="padding: 2px;">7</td><td style="padding: 2px;">8</td><td style="padding: 2px;">1</td><td style="padding: 2px;">2</td><td style="padding: 2px;">3</td><td style="padding: 2px;">4</td><td style="padding: 2px;">5</td> </tr> </table>	row 1	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	row 2	2	3	4	5	6	7	8	1	6	7	8	1	2	3	4	5	
row 1	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8																				
row 2	2	3	4	5	6	7	8	1	6	7	8	1	2	3	4	5																				

The dashed vertical lines are used for clarity to separate the blocks of each directed loop. Design $D5$ has $\gamma_1 = 1$ and $\gamma_2 = 3$ and is optimally robust by Theorem 6.1. Design $D6$ has $\gamma_1 = 1$ and $\gamma_2 = 5 = v/2 + 1$ and is not established as optimally robust by Theorem 6.1. The blocks of $D6$ can be rearranged to highlight that the design contains four copies of $K_{2,2}$:

$D6 =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td style="border-right: 1px solid black; padding: 2px;">row 1</td> <td style="padding: 2px;">1</td><td style="padding: 2px;">1</td><td style="padding: 2px;">5</td><td style="padding: 2px;">5</td> <td style="border-left: 1px solid black; padding: 2px;">2</td><td style="padding: 2px;">2</td><td style="padding: 2px;">6</td><td style="padding: 2px;">6</td> <td style="border-left: 1px solid black; padding: 2px;">3</td><td style="padding: 2px;">3</td><td style="padding: 2px;">7</td><td style="padding: 2px;">7</td> <td style="border-left: 1px solid black; padding: 2px;">4</td><td style="padding: 2px;">4</td><td style="padding: 2px;">8</td><td style="padding: 2px;">8</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px;">row 2</td> <td style="padding: 2px;">2</td><td style="padding: 2px;">6</td><td style="padding: 2px;">2</td><td style="padding: 2px;">6</td> <td style="border-left: 1px solid black; padding: 2px;">3</td><td style="padding: 2px;">7</td><td style="padding: 2px;">3</td><td style="padding: 2px;">7</td> <td style="border-left: 1px solid black; padding: 2px;">4</td><td style="padding: 2px;">8</td><td style="padding: 2px;">4</td><td style="padding: 2px;">8</td> <td style="border-left: 1px solid black; padding: 2px;">1</td><td style="padding: 2px;">5</td><td style="padding: 2px;">1</td><td style="padding: 2px;">5</td> </tr> </table>	row 1	1	1	5	5	2	2	6	6	3	3	7	7	4	4	8	8	row 2	2	6	2	6	3	7	3	7	4	8	4	8	1	5	1	5
row 1	1	1	5	5	2	2	6	6	3	3	7	7	4	4	8	8																			
row 2	2	6	2	6	3	7	3	7	4	8	4	8	1	5	1	5																			

Dashed vertical lines are again used for clarity, this time to separate the blocks of each $K_{2,2}$. There are eight ways of removing four blocks to give a $D_\#$ with a type 3 relationship. Thus, $D6$ is not optimally robust.

6.3. *Designs recommended for $r < 7$.* Bailey (2007) recommends designs for $r = 3, 4, 5, 6$. These are selected as having optimal or near-optimal A-criterion values and no robustness properties are claimed. For those designs without repeated blocks and which have $rv > 2v + 4r$, so the degrees of freedom available do not preclude the possibility of optimal robustness, it is interesting to determine b_0 . All these designs have $\Omega_1(D) > r$ by Corollary 3.4, Theorem 4.1, Theorem 6.1 or by use of Godolphin and Warren (2014). All also have $\Omega_2(D) > r$, by Theorem 6.1. Thus, all the designs have $b_0 = \min\{r, \Omega_3(D)\}$: any design with $\Omega_3(D) = r$ has optimal breakdown number and any with $\Omega_3(D) > r$ is optimally robust.

By Proposition 6.1, the designs with $r = 6$ have $\Omega_3(D) \geq 9$ and are therefore optimally robust. Of the designs with $r = 5$, none of the corresponding digraphs have properties necessary for $\Omega_3(D) = 4$, indicating that all have optimal breakdown number. In fact, in each case removal of every set of five blocks confirms that $\Omega_3(D) > 5$ and so the designs are optimally robust. For the case $r = 4$, only the

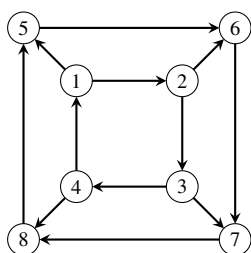
design with $v = 7$ has digraph containing as many as $\lfloor \frac{v-2}{2} \rfloor$ copies of $K_{2,2}$. This design is not optimally robust since $\Omega_3(D) = 4$: there are two ways of removing four blocks to give eventual designs which are disconnected due to type 3 relationships. The other designs with $r = 4$ are optimally robust. Several are interwoven loop designs.

None of the digraphs for the designs with $r = 3$ contain the configuration for $\Omega_3(D) = 1$. However, designs for $v = 8, 12$ have $\Omega_3(D) = 2$ and so do not have optimal breakdown number, whilst those for $v = 10, 18$ have $\Omega_3(D) = 3$ and so have optimal breakdown number but are not optimally robust. Alternatives to the design for $v = 8$ are now investigated.

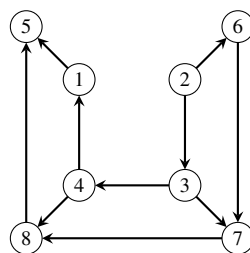
EXAMPLE 10. The design given by Bailey for $v = 8, r = 3$, denoted here by DB , contains two disjoint nonzero sum circuits, in the form of the 4-cycles $\langle 1, 2, 3, 4 \rangle$ and $\langle 5, 6, 7, 8 \rangle$, and four additional blocks. There are other nonzero sum circuits, for example, the 6-circuit $\langle 1, 5, 8, 4, 3, 2 \rangle$ has $S = -4$, but no other pairs of disjoint nonzero sum circuits. The digraph for DB is given in Figure 8(i). Four designs in $DB(2)$ are disconnected via type 3 relationships. One such example, formed by the loss of blocks with treatment pairs 1, 2 and 5, 6, has a digraph in Figure 8(ii). The relevant feature of the digraph is that all circuits have zero sum. For completeness, the staircase is given below.

	B_1	B_2	B_3	B_4
row 1	2 2	3 3 6	4 4 7	1 8
row 2	3 6	4 7 7	1 8 8	5 5

With $v = 8, r = 3$, it is not possible to construct a design with more than two disjoint circuits. Thirteen designs with $v = 8, r = 3$, which also contain the cycles $\langle 1, 2, 3, 4 \rangle$ and $\langle 5, 6, 7, 8 \rangle$, are considered as potential alternatives to DB . For each design, the four blocks not included in these cycles are given in Table 2 together with robustness properties and A -criterion. Design 14 in the table is DB .



(i) Digraph for DB



(ii) Digraph for a disconnected member of $DB(2)$

FIG. 8. Digraphs corresponding to Example 10.

TABLE 2
Properties of designs with $v = 8$ and $r = 3$

	Blocks 9–12	b_0	$M\{D\}_{b_0}$	$\Phi_A(D)$	Robustness rank	Optimality rank
1	1 6 8 4 5 2 3 7	3	8	1.4388	1=	1
2	1 6 3 8 5 2 7 4	3	8	1.4359	1=	2=
3	1 2 8 7 5 6 3 4	3	8	1.4328	1=	4
4	1 6 3 4 5 2 8 7	3	8	1.4119	1=	7=
5	1 2 3 4 5 6 8 7	3	8	1.3243	1=	12
6	1 2 3 4 5 8 7 6	3	8	1.3125	1=	13=
7	1 2 7 6 5 8 3 4	3	10	1.4189	7=	5=
8	1 2 8 4 5 6 3 7	3	10	1.4119	7=	7=
9	1 2 7 8 5 6 3 4	3	12	1.4189	9=	5=
10	1 2 3 7 5 6 8 4	3	12	1.4119	9=	7=
11	1 2 3 6 5 8 7 4	3	14	1.4000	11	10=
12	1 6 3 4 5 2 7 8	3	18	1.4000	12	10=
13	1 8 3 6 5 2 7 4	2	2	1.4359	13	2=
14	1 2 3 4 5 6 7 8	2	4	1.3125	14	13=

There is a striking lack of correlation between the design ranking according to robustness against breakdown and the ranking by A -criterion. Six of the 14 designs are optimally robust and these have A -criterion ranks ranging from first to thirteenth. The 14 designs do only form a subset of those with $v = 8, r = 3$: even so, this lack of correlation contrasts with the findings of Bailey (2007) discussed in Section 4 for Model 1 designs also with $v = 8, r = 3$.

7. Concluding comments. Combinatorial arguments and known results from graph theory can aid in construction of robust Model 1 designs, by providing in-

formation on robustness based only on v , b and treatment replicates. For Model 2 designs, robustness considerations are more complex and guidance is given for design construction. Much of the previous work on selecting Model 2 designs for microarray experiments involves starting with a Model 1 design with good properties, and assigning dye colour to experimental units in blocks to maximise dye balance over treatments; see, for example, Sanchez and Glonek (2009). This approach yields designs which tend to be robust against type 1 and 2 relationships, but not necessarily against type 3 relationships. Further, investigation in Example 10 suggests a lack of correlation between A -criterion and propensity for design breakdown via a type 3 relationship. Thus, the predominant message is that care should be taken to avoid designs with small $\Omega_3(D)$.

Amongst equireplicate designs with treatments distributed as evenly as possible between rows, designs with $r \geq 6$ are robust against type 3 relationships and those with $r = 3$ are particularly vulnerable. This supports the suggestion from Wit, Nobile and Khanin (2005) that only designs with $r \geq 4$ should be used.

For both models, cyclic constructions are shown to yield designs which are optimally robust for even r .

Graphical representation of Model 1 designs is particularly appealing since an eventual design in which breakdown has occurred corresponds to a disconnected graph and vice versa. The relationship between the properties of Model 2 designs and those of the corresponding digraphs are more complicated. The digraph for an eventual design, which is disconnected by a type 1 relationship or which has lost all replicates of one or more treatments, will be disconnected. By comparison, the digraph for an eventual design which is disconnected by a type 2 or type 3 relationship, but contains no type 1 relationship, is connected, but not strongly connected. However, digraphs for connected eventual designs can be connected or strongly connected. Therefore, it is not evident that results on connectivity properties of digraphs will provide information on robustness of Model 2 designs above that provided by results on graphs.

Acknowledgement. The author gives grateful thanks to Jan van den Heuvel for discussion and advice regarding graph theory.

SUPPLEMENTARY MATERIAL

Supplement to “Designs with blocks of size two and applications to microarray experiments” (DOI: [10.1214/17-AOS1638SUPP](https://doi.org/10.1214/17-AOS1638SUPP); .pdf). Contains proofs of: Theorem 2.2, Lemma 4.1, Lemma 4.2, Theorem 4.2 and Theorem 5.2.

REFERENCES

- BAGCHI, S. and CHENG, C.-S. (1993). Some optimal designs of block size two. *J. Statist. Plann. Inference* **37** 245–253. [MR1243802](https://doi.org/10.1080/01621459308838802)
- BAILEY, R. A. (2007). Designs for two-colour microarray experiments. *J. Roy. Statist. Soc. Ser. C* **56** 365–394. [MR2409757](https://doi.org/10.1111/j.1467-9892.2007.00575.x)

- BAILEY, R. A. and CAMERON, P. J. (2009). Combinatorics of optimal designs. In *Surveys in Combinatorics 2009* (S. Huczynska, J. D. Mitchell and C. M. Roney-Dougal, eds.). *London Mathematical Society Lecture Notes* **365** 19–73. Cambridge Univ. Press, Cambridge.
- BAILEY, R. A., SCHIFFL, K. and HILGERS, R.-D. (2013). A note on robustness of D-optimal block designs for two-colour microarray experiments. *J. Statist. Plann. Inference* **143** 1195–1202. [MR3049621](#)
- BAKSALARY, J. K. and TABIS, Z. (1987). Conditions for the robustness of block designs against the unavailability of data. *J. Statist. Plann. Inference* **16** 49–54.
- BAUER, D., HAKIMI, S. L., KAHL, N. and SCHMEICHEL, E. (2009). Sufficient degree conditions for k -edge-connectedness of a graph. *Networks* **54** 95–98.
- BHAUMIK, D. K. and WHITTINGHILL, D. C. (1991). Optimality and robustness to the unavailability of blocks in block designs. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **53** 399–407.
- BONDY, J. A. (1969). Properties of graphs with constraints on degrees. *Studia Sci. Math. Hungar.* **4** 473–475.
- BONDY, J. A. and MURTY, U. S. R. (2008). *Graph Theory*. Springer, Berlin.
- BUTZ, L. (1982). *Connectivity in Multifactor Designs*. Heldermann Verlag, Berlin. [MR0671251](#)
- CHAI, F.-S., LIAO, C.-T. and TSAI, S.-F. (2007). Statistical designs for two-color spotted microarray experiments. *Biom. J.* **49** 259–271. [MR2364336](#)
- CHARTRAND, G. (1966). A graph-theoretic approach to a communications problem. *SIAM J. Appl. Math.* **14** 778–781.
- DEY, A. (1993). Robustness of block designs against missing data. *Statist. Sinica* **3** 219–231.
- GHOSH, S. (1979). On robustness of designs against incomplete data. *Sankhya* **40** 204–208.
- GHOSH, S. (1982). Robustness of BIBD against the unavailability of data. *J. Statist. Plann. Inference* **6** 29–32.
- GODOLPHIN, J. D. (2004). Simple pilot procedures for the avoidance of disconnected experimental designs. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **53** 133–147.
- GODOLPHIN, J. D. (2013). On the connectivity problem for m -way designs. *J. Stat. Theory Pract.* **7** 732–744. [MR3196631](#)
- GODOLPHIN, J. (2018). Supplement to “Designs with blocks of size two and applications to microarray experiments.” DOI:[10.1214/17-AOS1638SUPP](#).
- GODOLPHIN, J. D. and GODOLPHIN, E. J. (2001). On the connectivity of row-column designs. *Util. Math.* **60** 51–65.
- GODOLPHIN, J. D. and GODOLPHIN, E. J. (2015). The use of treatment concurrences to assess robustness of binary block designs against the loss of whole blocks. *Aust. N. Z. J. Stat.* **57** 225–239.
- GODOLPHIN, J. D. and WARREN, H. R. (2014). An efficient procedure for the avoidance of disconnected incomplete block designs. *Comput. Statist. Data Anal.* **71** 1134–1146.
- GUPTA, S. (2006). Balanced factorial designs for cDNA microarray experiments. *Comm. Statist. Theory Methods* **35** 1469–1476.
- JOHN, J. A. and WILLIAMS, E. R. (1995). *Cyclic and Computer Generated Designs*, 2nd ed. Chapman & Hall, London. [MR1382127](#)
- KEL’MANS, A. K. (1972). Asymptotic formulas for the probability of k -connectedness of random graphs. *Theory Probab. Appl.* **17** 243–254.
- KERR, M. K. (2006). Efficient 2^k factorial designs for blocks of size 2 with microarray applications. *J. Qual. Technol.* **38** 309–318.
- KERR, M. K. and CHURCHILL, G. A. (2001). Experimental design for gene expression microarrays. *Biostatistics* **2** 183–201.
- MAHBUB LATIF, A. H. M., BRETZ, F. and BRUNNER, E. (2009). Robustness considerations in selecting efficient two-color microarray designs. *Bioinformatics* **25** 2355–2361.

- MORGAN, J. P. (2015). *Blocking with Independent Responses* (A. Dean, M. Morris, J. Stufken and D. Bingham, eds.). *Handbook of Design and Analysis of Experiments*. CRC Press, Boca Raton, FL.
- NGUYEN, D. V., ARPAT, A. B., WANG, N. and CARROLL, R. J. (2002). DNA microarray experiments: Biological and technological aspects. *Biometrics* **58** 701–717.
- PATERSON, L. (1983). Circuits and efficiency in incomplete block designs. *Biometrika* **70** 215–225.
- SANCHEZ, P. S. and GLONEK, G. F. V. (2009). Optimal designs for 2-colour microarray experiments. *Biostatistics* **10** 561–574.
- TSAI, S.-F. and LIAO, C.-T. (2013). Minimum breakdown designs in blocks of size two. *J. Statist. Plann. Inference* **143** 202–208.
- WIT, E., NOBILE, A. and KHANIN, R. (2005). Near-optimal designs for dual channel microarray studies. *J. Roy. Statist. Soc. Ser. C* **54** 817–830. [MR2209033](#)
- WU, C. F. J. and HAMADA, M. S. (2009). *Experiments: Planning, Analysis, and Optimization*, 2nd ed. Wiley, Hoboken, NJ. [MR2583259](#)
- YANG, Y. J. and DRAPER, N. R. (2003). Two-level factorial and fractional factorial designs in blocks of size two. *J. Qual. Technol.* **35** 294–305.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SURREY
GUILDFORD
SURREY GU2 7XH
UNITED KINGDOM
E-MAIL: j.godolphin@surrey.ac.uk