

Variance and Concurrence in Block Designs, and Distance in the Corresponding Graphs

R. A. BAILEY

In memory of D. G. Higman

1. Introduction

When Fisher initially advocated partitioning the units in an experiment into blocks of similar units in [14, Sec. 48], he proposed that each treatment should occur on one unit in each block. (In statistical contexts, the points of a design are usually called *treatments*.) Such designs were eventually called *randomized block designs* or *complete-block designs*. However, natural blocks may not be large enough to contain every treatment and so, in [41], Yates introduced designs with incomplete blocks. He had the intuition to propose designs in which each pair of treatments occurs together in the same number of blocks. He called these *symmetrical incomplete randomized block arrangements*; nowadays, statisticians call them *balanced incomplete-block designs* while pure mathematicians call them *2-designs*. The last phrase was first used in print by D. R. Hughes [18], who told me in 2001 that it had been suggested by his then colleague D. G. Higman.

When a design is used for an experiment, there is one observation for each treatment in each block. These data are analyzed to estimate the relative effects of the different treatments. In this context, 2-designs have three clear advantages: (i) analysis of data from an experiment using such a design does not require matrix inversion (this consideration was important in pre-computer days); (ii) the variance of the estimator of the difference between the effect of treatment i and the effect of another treatment j is independent of the pair $\{i, j\}$; and (iii) the design minimizes the average value of these pairwise variances.

For a given practical experiment, there may not exist a 2-design with the required parameters. What design should one use then? In [42], Yates introduced *square lattice* designs for n^2 treatments in rn blocks of size n , where $2 \leq r \leq n + 1$. The treatments are the cells of an $n \times n$ square array; the blocks correspond to the rows and columns of the array and to the letters of $r - 2$ mutually orthogonal $n \times n$ Latin squares. In such a design, each pair of treatments occurs together in either one or zero blocks, and the variance of the estimator of their difference depends only on this number, being slightly smaller in the former case. Use of these designs led statisticians to believe that, in any incomplete-block design, the variance would always depend on the number of blocks containing a given pair

of treatments; this is called the *concurrency* of those two treatments. It was also believed that variance would always decrease as concurrency increases. See [22; 31; 34; 36] for clear examples of this. Both of these ideas are now known to be false.

To understand a third widespread misconception, consider experiments whose blocks have size 2. Then the design is just a graph, possibly with multiple edges. If the graph is a tree then the only estimator of the difference between the effects of treatments i and j is obtained from the unique path joining i to j . If the path is (i, a, b, \dots, j) then the difference between i and a is estimated from one block, the difference between a and b is estimated from the next block, and so on, and the sum of these differences estimates the difference between i and j . The variance of the sum is equal to the sum of the variances, so the variance of the estimator of the difference between the effects of treatments i and j is proportional to the length of the path. This leads to the idea that variance must increase with distance, even though this is no longer true in general graphs.

Even when the block size is greater than 2, each incomplete-block design defines a *concurrency graph*, which may have multiple edges but no loops (as will be explained in Section 2). If the design is used for an experiment, then the differences between treatments will be estimated from the ensuing data: the relative variances of those estimators can be deduced from the adjacency matrix of the graph. If there is no 2-design for the given parameters, then we need to decide (a) what is a good design to use and (b) how to allocate actual treatments to the abstract ones in the design.

Question (a) has prompted a large amount of research (see [32; 35]). One obvious criterion is the average pairwise variance: a design that minimizes this variance is said to be *A-optimal*. A design that minimizes the maximum pairwise variance is said to be “MV-optimal” in [20] and “ \tilde{E} -optimal” in [30]. Alternatively, any convex function of the pairwise variances could be minimized; see [23].

In [27], Kempthorne proved that the average pairwise variance is inversely proportional to the harmonic mean of the nontrivial eigenvalues of the information matrix, which will be defined in Section 2. Other popular optimality criteria are other concave functions of these eigenvalues, such as the geometric mean and the minimum; the latter is studied in [11; 19; 32]. Designs that minimize the last two are called *D-optimal* and *E-optimal*, respectively.

In [29], Kshirsagar showed that 2-designs are A- and D-optimal; in [33], Mote proved that they are E-optimal. This led to the idea that optimal designs were likely to have some degree of symmetry. “Symmetry” might mean a transitive automorphism group: in [24], John recommended cyclic designs, which are those possessing a cyclic group of automorphisms acting regularly on the treatments. It might also mean a nontrivial group of automorphisms; see [4]. It might mean that the concurrency graph is based on a strongly regular graph (see [1; 10; 12]). In one way of generalizing strongly regular graphs, the design is based on an association scheme; see Section 2. In another, a strongly regular graph is weakened to a regular graph (see [25]). At the very least, “symmetry” should imply that all

treatments occur in the same number of blocks—in other words, that the design is *equireplicate* or is a *1-design*.

Perhaps unsurprisingly, exhaustive computer searches have shown that many such naive ideas are not correct. In [25], John and Mitchell showed that the A-optimal design for ten treatments in thirty blocks of size 2 is not the complement of the Petersen graph. Jones and Eccleston investigated designs for v treatments in v blocks of size 2 for $v \leq 12$ in [26]; they were surprised to find that, in the A-optimal design for $v \geq 10$, most treatments occur in only one block.

These results were rediscovered more recently with the widespread use of blocks of size 2 in microarray experiments. For v treatments in v blocks, biologists—trained to compare everything with a control—proposed the *reference* design, which is the star $K_{1,v}$ with a standard treatment at the center. Statisticians, familiar with the idea that equal replication is desirable, counterproposed the *loop* design, which is a polygon with v edges. In this case the intuition that variance increases with distance proved to be correct: computer searches reported in [28; 40] showed that the loop design is A-optimal for $v \leq 8$ but that for higher values of v the A-optimal design consists of a star glued to a vertex of either a triangle or a quadrangle. The theoretical underpinning is in [3].

The purpose of this paper is to investigate question (b) under the assumption that the design is equireplicate. If we are particularly interested in the difference between the effects of treatments i and j , then we should ensure that the variance of the estimator of this difference is small. Thus we explore when properties of the variances can be deduced from an examination of the graph itself *without* calculating the generalized inverse of a matrix. In particular, when is it true that variance decreases as concurrence increases or that variance increases with distance?

Section 2 introduces the terminology and three instructive examples. Each of the subsequent sections shows that, for a certain class of block designs, a simple statement linking variance to concurrence or distance holds for all incomplete-block designs in that class.

2. Definitions and Notation

Let Δ be an incomplete-block design for $v(\Delta)$ points (treatments). Denote by $\mathcal{T}(\Delta)$ the set of treatments of Δ . Thus Δ is a family of $b(\Delta)$ subsets of $\mathcal{T}(\Delta)$, called *blocks*, each of size $k(\Delta)$. Each treatment occurs in $r(\Delta)$ blocks; the number $r(\Delta)$ is called the *replication*. Notation such as $v(\Delta)$ will be abbreviated to v where no confusion is likely. The *concurrence* $\lambda_{ij}(\Delta)$ of treatments i and j is the number of blocks in which i and j both occur: in particular, $\lambda_{ii}(\Delta) = r(\Delta)$ for all i in $\mathcal{T}(\Delta)$. The concurrence matrix $\Lambda(\Delta)$ is the $v \times v$ matrix with entries $\lambda_{ij}(\Delta)$.

Let $\mathcal{G}(\Delta)$ be the concurrence graph of Δ . The vertex set of $\mathcal{G}(\Delta)$ is $\mathcal{T}(\Delta)$, and the number of edges between vertices i and j is λ_{ij} if $i \neq j$; there are no loops. The block design is said to be *connected* if $\mathcal{G}(\Delta)$ is a connected graph. From now on, we assume that Δ is connected. Put $B(\Delta) = \Lambda(\Delta) - rI$, where I is the identity matrix; then $B(\Delta)$ is the adjacency matrix of $\mathcal{G}(\Delta)$.

Let $M = I - (rk)^{-1}\Lambda$. Statisticians call rM the *information matrix* of the design. Since \mathcal{G} is regular with degree $r(k-1)$, graph theorists know the matrix rkM as the *Laplacian* of \mathcal{G} .

Since M has zero row sums, 0 is an eigenvalue of M . Call this the *trivial* eigenvalue. Because Δ is a connected design, this eigenvalue has multiplicity 1, and its corresponding eigenspace is spanned by the all-1 vector. Let Φ be the unique generalized inverse of M that is symmetric and has zero row sums. It is obtained by expressing M in spectral form as a linear combination of its eigenprojectors and then replacing each nonzero eigenvalue coefficient by its reciprocal.

It is usually assumed that each observation in the experiment is the sum of an effect for the relevant treatment, an effect for the relevant block, and a random error with expectation zero; furthermore, the errors are mutually independent and each has variance σ^2 . Then the variance of the estimator of the difference between the effects of treatment i and treatment j is equal to $(\Phi_{ii} + \Phi_{jj} - 2\Phi_{ij})\sigma^2/r$ (see [2, Chap. 4]). In a complete-block design with the same replication, this variance is $2\sigma^2/r$. Define the *pseudovariance* ψ_{ij} for the difference between treatments i and j to be the ratio between these two variances, so that $\psi_{ij} = (\Phi_{ii} + \Phi_{jj} - 2\Phi_{ij})/2$. Thus $\psi_{ij} \geq 1$, with equality if and only if $\lambda_{ij} = r$.

If the design is a 2-design then $\lambda_{ij} = \lambda = r(k-1)/(v-1)$ for all i and j with $i \neq j$. In this case Λ is completely symmetric—that is, a linear combination of I and the all-1 matrix J . It follows that Φ is also completely symmetric and hence that ψ_{ij} has a constant value ψ for all i and j with $i \neq j$. In fact, $\psi = k(v-1)/v(k-1)$.

After the square lattice designs, the first nonbalanced incomplete-block designs to be studied were the partially balanced designs introduced by Bose and Nair in [6]. These are defined in terms of association schemes. Square lattice designs are a special case.

An *association scheme* on \mathcal{T} with s associate classes is a partition of $\mathcal{T} \times \mathcal{T}$ into $s+1$ classes satisfying some conditions that can be most easily expressed in terms of the $v \times v$ indicator matrices A_0, A_1, \dots, A_s for the associate classes. Here the (i, j) -element of A_x is equal to 1 if i and j are x th associates, and otherwise it is equal to 0. The conditions are (i) $A_0 = I$, (ii) each of A_0, A_1, \dots, A_s is symmetric, and (iii) each product $A_x A_y$ is in the set \mathcal{A} of all real linear combinations of A_0, \dots, A_s . These conditions ensure that \mathcal{A} is a commutative algebra; see [2, Chap. 2].

In the original definition, implicit in [6] and made explicit in [8], the associate classes $1, \dots, s$ were essentially defined on unordered pairs of treatments, so condition (ii) is natural. It is sensible because Λ is a symmetric matrix. In his investigations of finite permutation groups, Higman introduced the concept of *homogeneous coherent configuration* in [17]. This generalizes an association scheme by weakening condition (ii) to the condition that if A_x is an indicator matrix then so is its transpose. The theory of coherent configurations is fruitful, but in general it cannot assume the commutativity of the algebra \mathcal{A} , which we need to use in Section 4.

A block design is *partially balanced* with respect to a given association scheme if the concurrence λ_{ij} depends only on the associate class containing (i, j) . Thus Λ has the form $\sum_{x=0}^s \lambda_x A_x$, which is in \mathcal{A} . Because \mathcal{A} is an algebra, Φ is also in \mathcal{A} , so there are scalars $\phi_0, \phi_1, \dots, \phi_s$ such that $\Phi = \sum_{x=0}^s \phi_x A_x$. Moreover, if s is small then it is easy to calculate Φ by hand. If (i, j) is in the associate class x then $\psi_{ij} = \psi_x$ with $\psi_x = \phi_0 - \phi_x$. Thus both concurrence and pseudovariance are functions of associate class. However, it is possible to have $\lambda_x = \lambda_y$ for $x \neq y$ or, independently, to have $\psi_x = \psi_y$ for $x \neq y$.

If $s = 2$ then the pairs in each class form the edges of a strongly regular graph (possibly not connected). Now if $\lambda_1 = \lambda_2$ then the design is balanced and so $\psi_1 = \psi_2$ as well. Hence pseudovariance is a function of concurrence for partially balanced incomplete-block designs with two associate classes. We shall show in Sections 3 and 6 that, for such designs, $\lambda_1 > \lambda_2$ implies $\psi_1 < \psi_2$.

EXAMPLE 1. The following array shows a block design Γ for the ten treatments A, \dots, J in six blocks of size 5 (blocks are columns).

<i>D</i>	<i>E</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>F</i>
<i>E</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>H</i>
<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>J</i>
<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>	<i>G</i>
<i>I</i>	<i>J</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>

Figure 1 shows the treatments as the vertices of the Petersen graph. This graph is strongly regular and so defines an association scheme on $\mathcal{T}(\Gamma)$ with two associate classes, the edges and nonedges of the Petersen graph. We find that $\lambda_{ij}(\Gamma) = 2$ if $\{i, j\}$ is an edge and $\lambda_{ij}(\Gamma) = 1$ if $\{i, j\}$ is a nonedge. Moreover, $\psi_{ij}(\Gamma) = 13/12$ if $\{i, j\}$ is an edge and $\psi_{ij}(\Gamma) = 7/6$ if $\{i, j\}$ is a nonedge.

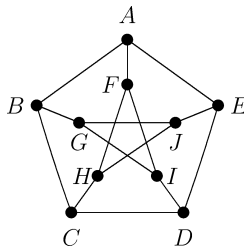


Figure 1 The Petersen graph used in Example 1

It was probably their experience from balanced designs and from partially balanced designs with two associate classes—as well as intuition—that once led statisticians to think that pseudovariance should be a (preferably decreasing) function of concurrence. The following examples show that this is not true in general.

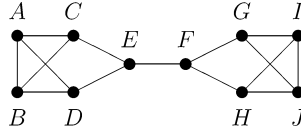


Figure 2 The design in Example 2: $v = 10, r = 3, k = 2,$ and $b = 15$

EXAMPLE 2. Let Δ be the design for ten treatments whose fifteen blocks of size 2 are the edges of the graph in Figure 2. Then

$$\Phi = \begin{bmatrix} A & B & C & D & E & F & G & H & I & J \\ 4.83 & 3.33 & 2.88 & 2.88 & 1.08 & -1.92 & -3.12 & -3.12 & -3.42 & -3.42 \\ 3.33 & 4.83 & 2.88 & 2.88 & 1.08 & -1.92 & -3.12 & -3.12 & -3.42 & -3.42 \\ 2.88 & 2.88 & 4.18 & 2.18 & 1.38 & -1.62 & -2.82 & -2.82 & -3.12 & -3.12 \\ 2.88 & 2.88 & 2.18 & 4.18 & 1.38 & -1.62 & -2.82 & -2.82 & -3.12 & -3.12 \\ 1.08 & 1.08 & 1.38 & 1.38 & 2.58 & -0.42 & -1.62 & -1.62 & -1.92 & -1.92 \\ -1.92 & -1.92 & -1.62 & -1.62 & -0.42 & 2.58 & 1.38 & 1.38 & 1.08 & 1.08 \\ -3.12 & -3.12 & -2.82 & -2.82 & -1.62 & 1.38 & 4.18 & 2.18 & 2.88 & 2.88 \\ -3.12 & -3.12 & -2.82 & -2.82 & -1.62 & 1.38 & 2.18 & 4.18 & 2.88 & 2.88 \\ -3.42 & -3.42 & -3.12 & -3.12 & -1.92 & 1.08 & 2.88 & 2.88 & 4.83 & 3.33 \\ -3.42 & -3.42 & -3.12 & -3.12 & -1.92 & 1.08 & 2.88 & 2.88 & 3.33 & 4.83 \end{bmatrix}.$$

Now some pairs have the same concurrence but different pseudovariance; for example, $\lambda_{AB} = \lambda_{AD} = 1$ but $\psi_{AB} = 3/2$ and $\psi_{AD} = 13/8$. Other pairs have the same pseudovariance but different concurrence; for example, $\psi_{CD} = \psi_{CE} = 2$ but $\lambda_{CD} = 0$ and $\lambda_{CE} = 1$. Even worse, $\lambda_{EF} = 1 > 0 = \lambda_{BE}$ but $\psi_{EF} = 3 > 21/8 = \psi_{BE}$.

EXAMPLE 3. The rectangular association scheme on the $2 \times n$ rectangle has three associate classes: pairs in the same row are in class 1; those in the same column in class 2; and other pairs in class 3. Consider the design with $n(n + 1)$ blocks of size 2 in which each column occurs as two blocks and there is also one block containing each pair of treatments in the same row. Then $\Lambda = (n + 1)I + A_1 + 2A_2$ and

$$\Phi = \frac{n + 1}{2n(n + 4)} \left[(5n + 8)I + nA_1 + 8A_2 - \frac{n^2 + 4n + 16}{2n} J \right].$$

Hence $\psi_1 = (4n + 8)(n + 1)/2n(n + 4)$ and $\psi_2 = 5n(n + 1)/2n(n + 4)$. If $n > 8$, then $\psi_2 > \psi_1$ even though $\lambda_2 > \lambda_1$.

Thus, even in a partially balanced design with three associate classes, variance may not decrease as concurrence increases.

3. Graphs with Two Nontrivial Eigenvalues

In [21], James and Wilkinson defined the *canonical efficiency factors* of a connected incomplete-block design Δ to be the nontrivial eigenvalues of M . Evidently, these

are linear combinations of the eigenvalues of B . Less evidently, they all lie in the interval $(0, 1]$.

James and Wilkinson also defined Δ to have n th-order balance if there are n distinct values among the canonical efficiency factors. Thus, first-order balance is equivalent to M being completely symmetric. In this case, all canonical efficiency factors are equal to $v(k - 1)/k(v - 1)$. In other words, if any one of concurrences, pseudovariances, and canonical efficiency factors takes just a single value, then so do the other two. However, if all take more than one value then there is no straightforward relationship in general. The design in Example 2 has two different concurrences, ten different pseudovariances, and seven different canonical efficiency factors.

The following theorem shows that there is a simple relationship between variance and concurrence if there are only two different canonical efficiency factors. It is proved in [2, Sec. 5.3].

THEOREM 1. *Let Δ be a connected incomplete-block design with block size k , replication r , concurrences λ_{ij} , and pseudovariances ψ_{ij} . If Δ has second-order balance with canonical efficiency factors e and f , then*

$$\psi_{ij} = c - d\lambda_{ij}$$

for all pairs of distinct treatments i and j , where

$$c = \frac{k(e + f - 1) + 1}{efk}, \tag{1}$$

$$d = \frac{1}{efrk}. \tag{2}$$

COROLLARY 1.1. *If a block design has second-order balance, then pseudovariance is a monotonic decreasing function of concurrence.*

This corollary applies to three important families of designs. The first consists of partially balanced incomplete-block designs with two associate classes. The other two are defined in terms of duals.

The dual Δ^* of a design Δ is obtained from Δ by interchanging the roles of blocks and treatments. The canonical efficiency factors of Δ^* are the same as those of Δ , including multiplicities, apart from $|b - v|$ canonical efficiency factors equal to 1. Thus if $v = b$ then Δ and Δ^* have the same order of balance, whereas if $v \neq b$ then they have the same order of balance if and only if they both have some canonical efficiency factor equal to 1. In particular, if Δ is a nonsymmetric 2-design then Δ^* has second-order balance.

EXAMPLE 4. Let Δ consist of all 3-subsets of a 6-set. The single canonical efficiency factor for Δ is $4/5$. Thus Δ^* has second-order balance with canonical efficiency factors $4/5$ (with multiplicity 5) and 1 (with multiplicity 14). Also, $r(\Delta^*) = 3$ and $k(\Delta^*) = 10$, so Theorem 1 shows that $\psi_{ij}(\Delta^*) = (27 - \lambda_{ij}(\Delta^*))/24$. In fact, Δ^* is partially balanced with respect to the Johnson association scheme $J(6, 3)$,

which has three associate classes; see [2, Chap. 1]. Its concurrences are 0, 1, and 2, so its corresponding pseudovariances are $9/8$, $13/12$, and $25/24$.

EXAMPLE 5. The dual Γ^* of the design in Example 1 is a 2-design for six treatments in ten blocks of size 3. Hence Γ has second-order balance with canonical efficiency factors $4/5$ and 1. Theorem 1 gives the values of $\psi_{ij}(\Gamma)$ reported in the last sentence of Example 1.

For the third family, take Δ to be a group-divisible design in which every block contains the same number of treatments from each group, so that the between-group canonical efficiency factor is equal to 1. Since Δ is partially balanced with two associate classes, Δ^* also has second-order balance if it is not a 2-design. Affine-resolved designs belong to this family: they are resolved designs in which the intersection of any two blocks from different parallel classes has the same size. Theorem 1 is given for affine-resolved designs in [5, Thm. 3.6].

COROLLARY 1.2. *If Δ and Γ are designs with second-order balance and with the same values of v , r , k , e , and f , then pseudovariance is the same monotonic decreasing function of concurrence for both Δ and Γ .*

Most optimality criteria are functions of the canonical efficiency factors. Suppose that Δ and Γ do have the same values of v , r , k , e , and f . The multiplicities p and q of e and f satisfy

$$\begin{aligned} 1 + p + q &= v, \\ pe + qf &= \frac{v(k-1)}{k}, \end{aligned}$$

so Δ and Γ also have the same values of p and q and hence of optimality criteria such as A, D, and E. It would then be sensible to choose between Δ and Γ by minimizing $\max \psi_{ij}$. Theorem 1 and Corollary 1.2 show that $\max \psi_{ij}(\Delta) < \max \psi_{ij}(\Gamma)$ if and only if $\min \lambda_{ij}(\Delta) > \min \lambda_{ij}(\Gamma)$. So we should choose the design with the larger minimal concurrence.

EXAMPLE 6. Let Ω be the 2-design for six treatments in twenty blocks of size 3 consisting of two copies of Γ^* , where Γ is the design in Example 1. Then Ω^* has the same canonical efficiency factors and other parameters as the design Δ^* in Example 4. However, Ω^* has concurrences 1, 2, and 3, whereas Δ^* has concurrences 0, 1, and 2; thus Δ^* has the larger value of $\max \psi_{ij}$ and so Ω^* should be preferred.

Similarly, affine-resolved designs for the same parameters do not all have the same value of $\max \psi_{ij}$ (see [5]).

The values of c and d in Theorem 1 depend on e and f . It is possible to find designs with the same values of v , r , and k but with different values of e and f , so that they have different linear functions giving pseudovariance in terms of concurrence.

EXAMPLE 7. Let Θ be the dual of the affine plane with nine treatments. Thus $v(\Theta) = 12$, $b(\Theta) = 9$, $r(\Theta) = 3$, and $k(\Theta) = 4$. This design is group-divisible,

with four groups of size 3 corresponding to parallel classes in Θ^* . It has within-group concurrence 0 and between-group concurrence 1. Its canonical efficiency factors are $3/4$ (with multiplicity 8) and 1 (with multiplicity 3). By Theorem 1, $\psi_{ij}(\Theta) = (12 - \lambda_{ij}(\Theta))/9$ and so Θ has pseudovariances $4/3$ and $11/9$.

Consider the complete bipartite graph $K_{3,3}$ as a block design for six treatments in nine blocks of size 2. Obtain the block design Ξ from this by replacing each treatment with a pair of treatments. Then Ξ is partially balanced with respect to the hierarchical group-divisible association scheme $\underline{2}/\underline{3}/\underline{2}$, which has three associate classes (see [2, Chap. 3]). It has the same values of $v, b, r,$ and k as does Θ . It has concurrences 3 (within pairs), 0 (between pairs within parts of $K_{3,3}$), and 1 (between parts). Its canonical efficiency factors are $1/2$ (with multiplicity 4) and 1 (with multiplicity 7). By Theorem 1, $\psi_{ij}(\Xi) = (9 - \lambda_{ij}(\Xi))/6$ and so Ξ has pseudovariances $3/2, 4/3,$ and 1.

Now we prove the converse of Theorem 1 and hence give a partial converse to Corollary 1.1.

THEOREM 2. *Let Δ be a connected incomplete-block design with block size $k,$ replication $r,$ concurrences $\lambda_{ij},$ and pseudovariances $\psi_{ij}.$ If there are at least two different concurrences and if there are constants c and d such that $\psi_{ij} = c - d\lambda_{ij}$ for all i and j with $i \neq j,$ then Δ has second-order balance and its canonical efficiency factors e and f are the solutions of equations (1) and (2).*

Proof. By definition of $\psi_{ij},$ we have

$$c - d\lambda_{ij} = \frac{\Phi_{ii} + \Phi_{jj} - 2\Phi_{ij}}{2}, \tag{3}$$

where Φ is a symmetric generalized inverse of M and $\sum_{j=1}^v \Phi_{ij} = 0$ for all $i.$ Summing equation (3) over $j \neq i$ gives

$$2(v - 1)c - 2r(k - 1)d = v\Phi_{ii} + \sum_{j=1}^v \Phi_{ij};$$

hence Φ_{ii} is independent of i and so is equal to $[(v - 1)c - r(k - 1)d]/v.$ Substituting in equation (3) gives $\Phi_{ij} = [-c - r(k - 1)d + v d\lambda_{ij}]/v$ for $i \neq j.$ Thus

$$\begin{aligned} v\Phi &= [(v - 1)c - r(k - 1)d]I - [c + r(k - 1)d](J - I) + vd(\Lambda - rI) \\ &= v[c + r(k - 1)d]\tilde{I} - vrkdM, \end{aligned}$$

where $\tilde{I} = I - v^{-1}J,$ which is the identity for symmetric matrices such as M and Φ with zero row sums. Since Φ is a generalized inverse of $M,$ this shows that

$$[(c + r(k - 1)d)\tilde{I} - rkdM]M = \tilde{I}. \tag{4}$$

Thus any eigenvalue e of M on the image of \tilde{I} satisfies

$$[c + r(k - 1)d]e - rkde^2 = 1. \tag{5}$$

Hence there are at most two such eigenvalues. Because the concurrences are not all the same, equation (5) has exactly two solutions, e and f , and so Δ has second-order balance. The usual theory of quadratic equations shows that

$$ef = \frac{1}{rkd}$$

and

$$e + f = \frac{c + r(k - 1)d}{rkd} = \frac{c - rd}{rkd} + 1,$$

whence e and f satisfy equations (1) and (2). □

COROLLARY 2.1. *If the block design Δ has two different concurrences and if its pseudovariances are a function of concurrence, then Δ is partially balanced with two associate classes.*

Proof. Let the two concurrences be λ_1 and λ_2 . Then the two pseudovariances are $h(\lambda_1)$ and $h(\lambda_2)$ for some real function h . Since h is specified at only two points, h has the form

$$h(x) = c - dx$$

for some constants c and d , and Theorem 2 applies.

Let A be the $v \times v$ matrix with $A_{ij} = 1$ if $\lambda_{ij} = \lambda_1$ and $i \neq j$ and with $A_{ij} = 0$ otherwise. Then

$$\Lambda = rI + \lambda_1 A + \lambda_2 (J - I - A).$$

Now equation (4) shows that A^2 is a linear combination of A , I , and J . This is precisely condition (iii) for the classes corresponding to I , A , and $J - I - A$ to form an association scheme. □

4. Amorphic Association Schemes

In a *square* association scheme there are n^2 treatments in a square array. A set of t mutually orthogonal $n \times n$ Latin squares is given. Distinct treatments are first associates if they are in the same row, second associates if they are in the same column. For $x = 3, \dots, t + 2$, they are x th associates if they have the same letter in the $(x - 2)$ th square. If $t = n - 1$ then all pairs of treatments are accounted for; otherwise, the remaining pairs form the final class.

Square association schemes are *amorphic* in the sense that any fusion of the associate classes yields another association scheme. Of course, all association schemes with two associate classes are amorphic. We shall show that amorphic association schemes with three or more classes have the property that variance decreases as concurrence increases. The proof depends upon the fact that, for an association scheme, the algebra \mathcal{A} is commutative and so has a basis consisting of the matrices of orthogonal projection onto the mutual eigenspaces of matrices in \mathcal{A} . These are called the *minimal idempotents* of the association scheme. The trivial minimal idempotent is the projector P_0 onto the space spanned by the all-1 vector.

THEOREM 3. *If $s \geq 3$ then the nontrivial minimal idempotents of an amorphic association scheme with s associate classes can be numbered as P_1, \dots, P_s in such*

a way that $P_x = \alpha_x I + \beta A_x + \gamma_x J$ for $x = 1, \dots, s$, where α_x, β , and γ_x are real numbers.

Proof. Let \mathcal{P} be the set of nontrivial idempotents and \mathcal{Q} the set of indicator matrices other than I . Both sets have size s . Since the association scheme is amorphic it follows that, given any partition π of \mathcal{Q} with p parts, the matrices formed by summing the matrices in each part (together with I) form an association scheme with p associate classes. The nontrivial minimal idempotents of this must be sums of the parts of a partition π' of \mathcal{P} , also with p parts. Because the indicator matrices of a scheme span the same algebra as its minimal idempotents, the map $\pi \mapsto \pi'$ is a bijection between partitions of \mathcal{Q} and partitions of \mathcal{P} , which obviously preserves refinement.

A partition of \mathcal{Q} with parts $\{A_x\}$ and $\mathcal{Q} \setminus \{A_x\}$ has $2^{s-2} - 1$ refinements with three parts. A partition of \mathcal{P} into parts of size m and $s - m$ has $(2^{m-1} - 1) + (2^{s-m-1} - 1)$ refinements with three parts. The first number is odd, while the second is even unless $m = 1$ or $m = s - 1$. Hence the bijection must take $\{\{A_x\}, \mathcal{Q} \setminus \{A_x\}\}$ to a partition of \mathcal{P} of the form $\{\{P_y\}, \mathcal{P} \setminus \{P_y\}\}$. Because $s \geq 3$, $\{P_y\}$ is the only singleton in this partition. Therefore we can unambiguously label the minimal idempotents in such a way that the minimal idempotents of the scheme with indicator matrices I, A_x , and $J - I - A_x$ are P_0, P_x , and $I - P_0 - P_x$.

Now there are constants α_x, β_x , and γ_x such that $P_x = \alpha_x I + \beta_x A_x + \gamma_x J$ for $x = 1, \dots, s$. Summing these equations gives $I - P_0 = (\sum \alpha_x)I + \sum \beta_x A_x + (\sum \gamma_x)J$. The only linear combination of A_1, \dots, A_s that is also a linear combination of I and J is their sum, so there is a constant β such that $\beta_x = \beta$ for $x = 1, \dots, s$. \square

Theorem 3 can be used to prove that the number of elements in an amorphic association scheme with more than two classes has the form n^2 for some positive integer n . Moreover, there are only two possibilities. In one, every indicator matrix A_x satisfies the equation

$$A_x^2 = c_x(n - 1)I + [(c_x - 1)(c_x - 2) + n - 2]A_x + c_x(c_x - 1)(J - I - A_x)$$

for some positive integer c_x , as happens for fusions of a square scheme, including the association scheme of a square lattice design. In the other possibility, every indicator matrix A_x satisfies

$$A_x^2 = c_x(n + 1)I + [(c_x + 1)(c_x + 2) - n - 2]A_x + c_x(c_x + 1)(J - I - A_x)$$

for some positive integer c_x . The corresponding strongly regular graphs are said to have *Latin-square* type and *negative Latin-square* type, respectively. A different proof of this was given in [16].

THEOREM 4. *If a block design is partially balanced with respect to an amorphic association scheme with more than two classes, then pseudovariance is a monotonic decreasing function of concurrence.*

Proof. Let the concurrences be $\lambda_1, \dots, \lambda_s$. Then $rkM = r(k - 1)I - \sum_{x \neq 0} \lambda_x A_x$. Theorem 3 shows that there is some constant κ_0 such that

$$rkM = \sum_{x \neq 0} \left(\kappa_0 - \frac{\lambda_x}{\beta} \right) P_x.$$

Therefore

$$\begin{aligned} \Phi &= rk \sum_{x \neq 0} \left(\frac{\beta}{\beta\kappa_0 - \lambda_x} \right) P_x \\ &= rk \left[\kappa_1 I + \kappa_2 J + \sum_{x \neq 0} \left(\frac{\beta^2}{\beta\kappa_0 - \lambda_x} \right) A_x \right] \end{aligned}$$

for some constants κ_1 and κ_2 . Thus

$$\psi_x = rk \left[\kappa_1 - \frac{\beta^2}{\beta\kappa_0 - \lambda_x} \right],$$

which is a monotonic decreasing function of λ_x . □

5. Walk-Regular Graphs

For $t = 1, \dots, v - 1$, let $\lambda_{ij}^{[t]}$ be the number of walks of length t from i to j in \mathcal{G} . Thus $\lambda_{ij}^{[1]} = \lambda_{ij}$ if $i \neq j$ and $\lambda_{ii}^{[1]} = 0$. For completeness, put $\lambda_{ii}^{[0]} = 1$ and $\lambda_{ij}^{[0]} = 0$ if $i \neq j$. Now $\lambda_{ij}^{[t]}$ is the (i, j) -entry of B^t .

Authors such as Paterson [34] and Mead [31] have suggested that, if pseudovariance is not simply a function of concurrence, then it should be related to the numbers of walks of lengths 2, 3, ..., in \mathcal{G} . We shall prove a generalization of Theorems 1 and 2 giving pseudovariance as a linear combination of the $\lambda_{ij}^{[t]}$. However, the proofs of Theorems 1 and 2 involve a step showing that Φ_{ii} is independent of i . As Example 2 shows, this is not true in general.

Extending the definition given in [15] for simple graphs, we define the graph \mathcal{G} to be *walk-regular* if $\lambda_{ii}^{[t]}$ is independent of i for all t . If Δ has n th-order balance with canonical efficiency factors e_1, \dots, e_n , then

$$(M - e_1 \tilde{I})(M - e_2 \tilde{I}) \cdots (M - e_n \tilde{I}) = 0,$$

so B^n is a linear combination of $I, B, B^2, \dots, B^{n-1}$, and J ; hence \mathcal{G} is walk-regular if and only if $\lambda_{ii}^{[t]}$ is independent of i for $0 \leq t \leq n - 1$.

In [36], Sinha defined a binary equireplicate design to be *simple* if the multiset $(\lambda_{ij})_{j \neq i}$ is independent of i —that is, if each row of Λ is a permutation of the first row. This is yet another way in which a design can be considered to have symmetry. All partially balanced designs are simple, as are all regular-graph designs, which were defined in [25] to be those designs in which all concurrences differ by at most 1. Many designs used in practice are simple.

LEMMA 1. *The concurrence graph $\mathcal{G}(\Delta)$ is walk-regular if the block design Δ satisfies any of the following conditions:*

- (a) Δ has a group of automorphisms that is transitive on the treatments;
- (b) Δ is partially balanced;
- (c) Δ has second-order balance;
- (d) Δ is simple and has third-order balance.

Proof. (a) Let i and j be different treatments. If Δ has a group of automorphisms that is transitive on the treatments, then there is an automorphism of Δ that takes i to j . For every t , such an automorphism also permutes the rows and columns of B^t accordingly and so $\lambda_{ii}^{[t]} = \lambda_{jj}^{[t]}$.

(b) If Δ is partially balanced with respect to the association scheme with indicator matrices A_0, A_1, \dots, A_s , then there are scalars b_0, b_1, \dots, b_s such that $B^t = \sum_{x=0}^s b_x A_x$. Hence all diagonal elements of B^t are equal to b_0 .

(c) By definition, $\lambda_{ii}^{[0]} = 1$ and $\lambda_{ii}^{[1]} = 0$ for all i . If Δ has second-order balance, this ensures that $\mathcal{G}(\Delta)$ is walk-regular.

(d) For each i , $\lambda_{ii}^{[2]}$ is the sum of the squares of the entries in the i th row of B . If Δ is simple, then these rows have the same multiset of entries and so $\lambda_{ii}^{[2]}$ is independent of i . Since Δ has third-order balance, this ensures that $\mathcal{G}(\Delta)$ is walk-regular. \square

The special case of Lemma 1(d) for which $\mathcal{G}(\Delta)$ has no multiple edges was proved in [13].

THEOREM 5. *Let Δ be a connected incomplete-block design with concurrences λ_{ij} and pseudovariances ψ_{ij} such that \mathcal{G} is walk-regular. If Δ has n th-order balance then there are constants c_0, c_1, \dots, c_{n-1} such that*

$$\psi_{ij} = c_0 + c_1 \lambda_{ij}^{[1]} + \dots + c_{n-1} \lambda_{ij}^{[n-1]}$$

for all pairs of distinct treatments i and j .

Proof. Since Δ has n th-order balance, it follows that B^n is a linear combination of $I, B, B^2, \dots, B^{n-1}$, and J ; hence these $n + 1$ matrices span a commutative algebra \mathcal{B} . Now, $M \in \mathcal{B}$ and so $\Phi \in \mathcal{B}$. Thus there are constants $\phi_0, \phi_1, \dots, \phi_{n-1}, \phi_\infty$ such that $\Phi = \sum_{t=0}^{n-1} \phi_t B^t + \phi_\infty J$. Hence

$$\begin{aligned} 2\psi_{ij} &= \Phi_{ii} + \Phi_{jj} - 2\Phi_{ij} \\ &= \sum_{t=0}^{n-1} \phi_t (\lambda_{ii}^{[t]} + \lambda_{jj}^{[t]} - 2\lambda_{ij}^{[t]}) \\ &= 2\phi_0 + \sum_{t=1}^{n-1} \phi_t (\lambda_{ii}^{[t]} + \lambda_{jj}^{[t]} - 2\lambda_{ij}^{[t]}). \end{aligned}$$

Because \mathcal{G} is walk-regular, $\lambda_{ii}^{[t]}$ is a constant $\lambda^{[t]}$ independent of i , so

$$\psi_{ij} = c_0 + \sum_{t=1}^{n-1} c_t \lambda_{ij}^{[t]},$$

where $c_0 = \phi_0 + \sum_{t=1}^{n-1} \phi_t \lambda^{[t]}$ and $c_t = -\phi_t$ for $1 \leq t \leq n - 1$. \square

Theorem 5 can be applied to *rectangular lattice designs*, which are resolved designs for $k(k + 1)$ treatments in blocks of size k constructed from $r - 2$ mutually orthogonal $(k + 1) \times (k + 1)$ Latin squares with a common transversal; see [2, Chap. 5]. The treatments are the cells of the square array that are not in the

Table 1 The rectangular lattice design Δ in Example 8:
 $v = 20, r = 3, k = 4,$ and $b = 15$ (blocks are columns)

A	E	I	M	Q	E	A	B	C	D	H	B	D	A	C
B	F	J	N	R	I	J	F	G	H	K	I	G	E	F
C	G	K	O	S	M	N	O	K	L	O	P	N	L	J
D	H	L	P	T	Q	R	S	T	P	R	T	Q	S	M

Table 2 The rectangular lattice design Γ in Example 8:
 $v = 20, r = 3, k = 4,$ and $b = 15$ (blocks are columns)

A	E	I	M	Q	E	A	B	C	D	F	A	B	C	D
B	F	J	N	R	I	J	F	G	H	J	G	E	H	K
C	G	K	O	S	M	N	O	K	L	P	I	L	N	M
D	H	L	P	T	Q	R	S	T	P	T	O	R	Q	S

transversal. The blocks of the first parallel class are the rows of the array; those of the second are the columns. For $x = 3, \dots, r,$ the blocks of the x th parallel class are the letters of the $(x - 2)$ th square.

All concurrences are 0 or 1, so these are regular-graph designs and hence simple. If $r = k + 1$ then they have second-order balance; otherwise, they have third-order balance with canonical efficiency factors 1, $(k + 1)(r - 1)/kr,$ and $(kr - k - 1)/kr.$ Many of them do not satisfy (a) or (b), yet Lemma 1 shows that their concurrence graphs are walk-regular.

EXAMPLE 8. Any single 5×5 Latin square with a transversal gives a rectangular lattice design for twenty treatments in fifteen blocks of size 4. Here $r = 3$ and $k = 4,$ so $B = 9I - 12M.$ The canonical efficiency factors are $5/6, 7/12,$ and 1, so

$$\left(M - \frac{5}{6}\tilde{I}\right)\left(M - \frac{7}{12}\tilde{I}\right)(M - \tilde{I}) = 0,$$

which expands as

$$M^3 - \frac{29}{12}M^2 + \frac{137}{72}M - \frac{35}{72}\tilde{I} = 0.$$

Hence

$$\Phi = \frac{72}{35}\left[M^2 - \frac{29}{12}M + \frac{137}{72}\tilde{I}\right] = \frac{1}{70}\left[94I + 11B + B^2 - \frac{274}{20}J\right].$$

The diagonal elements of B^2 are all equal to 9, so

$$\psi_{ij} = \frac{1}{70}(103 - 11\lambda_{ij} - \lambda_{ij}^{[2]}). \tag{6}$$

Tables 1 and 2 show different rectangular lattice designs for these parameters, using the treatment array in Table 4. Design Δ is constructed from the cyclic Latin square in Table 3, while design Γ is constructed from the noncyclic Latin square

Table 3 Cyclic Latin square

α	β	γ	δ	ε
β	γ	δ	ε	α
γ	δ	ε	α	β
δ	ε	α	β	γ
ε	α	β	γ	δ

Table 4 Treatment array

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>E</i>		<i>F</i>	<i>G</i>	<i>H</i>
<i>I</i>	<i>J</i>		<i>K</i>	<i>L</i>
<i>M</i>	<i>N</i>	<i>O</i>		<i>P</i>
<i>Q</i>	<i>R</i>	<i>S</i>	<i>T</i>	

Table 5 Noncyclic Latin square

α	β	γ	δ	ε
γ	ε	α	β	δ
β	α	δ	ε	γ
ε	δ	β	γ	α
δ	γ	ε	α	β

Table 6 Walks of lengths 1 and 2 in the designs Δ and Γ in Example 8, and corresponding pseudovariances

$\lambda_{ij}^{[0]}$	λ_{ij}	$\lambda_{ij}^{[2]}$	$70\psi_{ij}$	Design Δ		Design Γ	
				Number of pairs	Example pair i, j	Number of pairs	Example pair i, j
1	0	9		20	<i>A, A</i>	20	<i>A, A</i>
0	1	4	88	80	<i>A, D</i>	72	<i>A, B</i>
0	1	3	89	80	<i>A, B</i>	96	<i>A, I</i>
0	1	2	90	20	<i>A, E</i>	12	<i>A, D</i>
0	0	5	98	40	<i>A, F</i>	48	<i>A, K</i>
0	0	4	99	160	<i>A, G</i>	144	<i>A, H</i>
0	0	3	100	0		8	<i>G, R</i>

in Table 5. The values of λ_{ij} and $\lambda_{ij}^{[2]}$ that occur in these designs are shown in Table 6, together with the pseudovariances given by equation (6).

Neither Δ nor Γ is partially balanced. However, Δ does have a transitive automorphism group; Γ does not because the rows of $(B(\Gamma))^2$ are not all permutations

of each other. The largest pseudovariance for Γ does not occur for any pair in Δ . This shows that rectangular lattice designs with the same parameters do not necessarily have the same pseudovariances. In this case, Γ is MV-worse than Δ .

Williams [37] showed that, for a rectangular lattice design, Φ is a linear combination of I , Λ , and Λ^2 with explicit coefficients, but he did not go on to work out the individual pseudovariances explicitly. For $u = 0, 1$, Williams calculated the average value of the ψ_{ij} for which $\lambda_{ij} = u$; the figures given in Table 6 are consistent with his results.

Theorem 5 also has a partial converse, as follows.

THEOREM 6. *Let Δ be a connected incomplete-block design with concurrences λ_{ij} and pseudovariances ψ_{ij} such that \mathcal{G} is walk-regular. If there are constants c_0, c_1, \dots, c_{n-1} such that*

$$\psi_{ij} = c_0 + \sum_{t=1}^{n-1} c_t \lambda_{ij}^{[t]}$$

for all pairs of distinct treatments i and j , and if n is minimal subject to this, then Δ has n th-order balance.

Proof. For $i \neq j$ we have

$$c_0 + \sum_{t=1}^{n-1} c_t \lambda_{ij}^{[t]} = \frac{\Phi_{ii} + \Phi_{jj} - 2\Phi_{ij}}{2}. \tag{7}$$

Since $BJ = r(k-1)J$ we have $B^t J = r^t(k-1)^t J$ for $t \geq 1$, so that $\sum_{j=1}^v \lambda_{ij}^{[t]}$ is independent of i . Since \mathcal{G} is walk-regular, $\lambda_{ii}^{[t]}$ is also independent of i and so $\sum_{j \neq i} \lambda_{ij}^{[t]}$ is independent of i . As in the proof of Theorem 2, summing equation (7) over $j \neq i$ shows that Φ_{ii} has a constant value ϕ independent of i . Hence

$$\Phi_{ij} = \phi - c_0 - \sum_{t=1}^{n-1} c_t \lambda_{ij}^{[t]}$$

for $i \neq j$. Thus Φ is a linear combination of $I, B, B^2, \dots, B^{n-1}$, and J . Because Φ has zero row sums, it is a linear combination of $\tilde{I}, M, M^2, \dots, M^{n-1}$. But $\Phi M = \tilde{I}$, so M satisfies a polynomial equation of degree n on the image of \tilde{I} ; hence M has at most n eigenvalues on the image of \tilde{I} . If it has fewer than n eigenvalues then Δ has a smaller order of balance than n , so Theorem 5 contradicts the minimality of n . Hence M has exactly n eigenvalues on the image of \tilde{I} and so Δ has n th-order balance. \square

6. Distance-Regular Graphs

If there are only two different concurrences, then pseudovariances may be calculated from knowledge of the regular simple graph whose vertices are the treatments and whose edges are the pairs with the higher concurrence. It seems plausible that pseudovariance should be a monotonic increasing function of distance in this graph, but Example 2 shows that this is not true in general.

A connected simple graph is *distance-regular* if its distance classes form an association scheme. Label the associate classes so that class x consists of all pairs of vertices at distance x . Let s be the number of associate classes, which is equal to the diameter of the graph. The graph must be regular (of valency d , say), and there are nonnegative integers b_0, \dots, b_s and c_0, \dots, c_s such that $c_0 = b_s = 0$, $b_0 = d$, $c_1 = 1$, $0 < b_x + c_x \leq d$, and

$$A_x A_1 = b_{x-1} A_{x-1} + (d - b_x - c_x) A_x + c_{x+1} A_{x+1}$$

for $x = 0, \dots, s$; see [9, Chap. 4]. Moreover, if n_x is the number of x th associates of each treatment, then $n_1 = d$ and $n_x b_x = n_{x+1} c_{x+1}$ for $0 \leq x \leq s - 1$.

THEOREM 7. *Let Δ be a connected incomplete-block design with block size k , replication r , and two concurrences λ and $\lambda + \mu$, where $\mu > 0$. If the pairs of treatments with concurrence $\lambda + \mu$ form the edges of a distance-regular graph, then pseudovariance is a monotonic increasing function of distance.*

Proof. We have

$$\Lambda = rI + \lambda(J - I) + \mu A_1$$

and

$$r(k - 1) = \lambda(v - 1) + \mu d,$$

where $v = v(\Delta)$ and d is the valency of the graph. Hence

$$rkM = (\lambda v + \mu d)I - \mu A_1 - \lambda J.$$

If s is the diameter of the graph then the distance classes form an association scheme with s associate classes, so there are scalars ϕ_0, \dots, ϕ_s such that $\Phi = \phi_0 I + \phi_1 A_1 + \dots + \phi_s A_s$. Therefore, if treatments i and j are at distance x then $\psi_{ij} = \psi_x = \phi_0 - \phi_x$.

The row sums of Φ are all zero, so $\Phi J = 0$. Also, $\Phi M = \tilde{I}$, so

$$(\lambda v + \mu d)\phi_x - \mu\phi_{x-1}c_x - \mu\phi_x(d - b_x - c_x) - \mu\phi_{x+1}b_x = -rk/v$$

for $x = 1, \dots, s$. This can be rewritten as

$$\lambda v n_y \phi_y + \mu n_y c_y (\phi_y - \phi_{y-1}) + \mu n_y b_y (\phi_y - \phi_{y+1}) = -rk n_y / v. \tag{8}$$

Summing equations (8) from $y = x$ to $y = s$ gives

$$\lambda v (n_x \phi_x + \dots + n_s \phi_s) + \mu n_x c_x (\phi_x - \phi_{x-1}) = -rk (n_x + \dots + n_s) / v.$$

This shows that $\lambda v^2 (n_x \phi_x + \dots + n_s \phi_s) + rk (n_x + \dots + n_s)$ and $\phi_x - \phi_{x-1}$ have opposite signs. Thus, if $\lambda v^2 \phi_s \leq -rk$ then reverse induction yields $\phi_x \leq \phi_{x+1} \leq \dots \leq \phi_{s-1} \leq \phi_s$, so that $\lambda v^2 (n_x \phi_x + \dots + n_s \phi_s) \leq -rk (n_x + \dots + n_s)$ and $\phi_x \geq \phi_{x-1}$. Consequently, $\phi_0 \leq \phi_1$, which contradicts the fact that $\phi_0 - \phi_1 = \psi_1 \geq 1$.

This contradiction shows that $\lambda v^2 \phi_s > -rk$. Now a similar reverse induction gives $\phi_x > \phi_{x+1} > \dots > \phi_{s-1} > \phi_s$ and so $\lambda v^2 (n_x \phi_x + \dots + n_s \phi_s) > -rk (n_x + \dots + n_s)$ and $\phi_x < \phi_{x-1}$. Hence $\psi_x = \phi_0 - \phi_x > \phi_0 - \phi_{x-1} = \psi_{x-1}$. □

EXAMPLE 9. Let Γ be a projective plane of order q , so that $v(\Gamma) = b(\Gamma) = q^2 + q + 1$, $k(\Gamma) = r(\Gamma) = q + 1$, and $\lambda(\Gamma) = 1$. The following construction for a

resolved design Δ with $v(\Delta) = v(\Gamma)k(\Gamma)$, $k(\Delta) = k(\Gamma)$, and $r(\Delta) = 2$ was recommended by Bose and Nair in [7]. Take $\mathcal{T}(\Delta)$ to be the set of incident point-line pairs in Γ . In the first parallel class of Δ , two treatments are in the same block if they have the same line in Γ ; in the second parallel class of Δ , two treatments are in the same block if they have the same point in Γ . Such a design is called a *generalized hexagon* in [9, Sec. 6.5].

The concurrences in Δ are all equal to 0 or 1. The pairs with concurrence 1 form the edges of a distance-regular graph with diameter 3 that has $n_0 = 1$, $d = n_1 = 2q$, $n_2 = 2q^2$, and $n_3 = q^3$. Also,

$$\begin{aligned} A_1^2 &= 2qI + (q-1)A_1 + A_2, \\ A_2A_1 &= qA_1 + (q-1)A_2 + 2A_3, \\ A_3A_1 &= qA_2 + 2(q-1)A_3. \end{aligned}$$

Since $\Lambda = 2I + A_1$ and $k = q + 1$, we have $2(q+1)M = 2qI - A_1$. Some calculation gives

$$\Phi = \frac{1}{q^2 + q + 1} [(q^2 + 3q + 5)I + (q + 3)A_1 + A_2 + \kappa J],$$

where $\kappa = -(5q^2 + 9q + 5)/v$. Thus $\psi_1 = (q^2 + 2q + 2)/(q^2 + q + 1)$, $\psi_2 = (q^2 + 3q + 4)/(q^2 + q + 1)$, and $\psi_3 = (q^2 + 3q + 5)/(q^2 + q + 1)$.

It was shown in [38; 39] that this design is A-optimal among resolved designs.

Theorem 7 can be generalized to disconnected graphs whose components are distance-regular. In this case, we need $\lambda > 0$ for a connected design.

THEOREM 8. *Let Δ be a connected incomplete-block design with block size k , replication r , and two concurrences λ and $\lambda + \mu$, where $\mu > 0$. If the pairs of treatments with concurrence $\lambda + \mu$ form the edges of a disconnected graph whose components are distance-regular graphs with the same parameters, then pseudo-variance is a monotonic increasing function of distance, where the distance between vertices in different components is deemed to be ∞ .*

Proof. Let A_∞ be the indicator matrix for pairs of treatments in different components of the graph. Then $A_1A_\infty = n_1A_\infty$. Following the proof of Theorem 7, we obtain the extra equation $(\lambda v + \mu d - \mu n_1)\phi_\infty = -rk/v$. Hence $\phi_\infty = -rk/\lambda v^2$.

As in the previous proof, if $\phi_s \leq \phi_\infty$ then induction gives a contradiction, while if $\phi_s > \phi_\infty$ then induction shows that $\phi_0 > \phi_1 > \dots > \phi_s > \phi_\infty$. \square

References

- [1] B. Bagchi and S. Bagchi, *Optimality of partial geometric designs*, Ann. Statist. 29 (2001), 577–594.
- [2] R. A. Bailey, *Association schemes. Designed experiments, algebra and combinatorics*, Cambridge Stud. Adv. Math., 84, Cambridge Univ. Press, Cambridge, 2004.
- [3] ———, *Designs for two-colour microarray experiments*, J. Roy. Statist. Soc. Ser. C 56 (2007), 365–394.

- [4] R. A. Bailey, P. J. Cameron, P. Dobcsányi, J. P. Morgan, and L. H. Soicher, *Designs on the web*, Discrete Math. 306 (2006), 3014–3027.
- [5] R. A. Bailey, H. Monod, and J. P. Morgan, *Construction and optimality of affine-resolvable designs*, Biometrika 82 (1995), 187–200.
- [6] R. C. Bose and K. R. Nair, *Partially balanced incomplete block designs*, Sankhyā 4 (1939), 337–372.
- [7] ———, *Resolvable incomplete block designs with two replications*, Sankhyā Ser. A 29 (1962), 9–24.
- [8] R. C. Bose and T. Shimamoto, *Classification and analysis of partially balanced incomplete block designs with two associate classes*, J. Amer. Statist. Assoc. 47 (1952), 151–184.
- [9] A. E. Brouwer, A. M. Cohen, and A. Neumaier, *Distance-regular graphs*, Ergeb. Math. Grenzgeb. (3), 18, Springer-Verlag, Berlin, 1989.
- [10] C.-S. Cheng, *Optimality of certain asymmetrical experimental designs*, Ann. Statist. 6 (1978), 1239–1261.
- [11] ———, *On the E-optimality of some block designs*, J. Roy. Statist. Soc. Ser. B 42 (1980), 199–204.
- [12] C.-S. Cheng and R. A. Bailey, *Optimality of some two-associate-class partially balanced incomplete-block designs*, Ann. Statist. 19 (1991), 1667–1671.
- [13] E. R. van Dam, *Regular graphs with four eigenvalues*, Linear Algebra Appl. 226/228 (1995), 139–162.
- [14] R. A. Fisher, *Statistical methods for research workers*, Oliver & Boyd, Edinburgh, 1925.
- [15] C. D. Godsil and B. D. McKay, *Feasibility conditions for the existence of walk-regular graphs*, Linear Algebra Appl. 30 (1980), 51–61.
- [16] Ja. Ju. Gol’fand, A. V. Ivanov, and M. H. Klin, *Amorphic cellular rings*, Investigations in algebraic theory of combinatorial objects (I. A. Faradžev, A. A. Ivanov, M. H. Klin, A. J. Woldar, eds.), Math. Appl. (Soviet Ser.), 84, pp. 167–187, Kluwer, Dordrecht, 1994.
- [17] D. G. Higman, *Combinatorial considerations about permutation groups*, Mathematical Institute, Oxford, 1971.
- [18] D. R. Hughes, *Combinatorial analysis: t-designs and permutation groups*, Proc. Sympos. Pure Math., 6, pp. 39–41, Amer. Math. Soc., Providence, RI, 1962.
- [19] M. Jacroux, *On the E-optimality of regular graph designs*, J. Roy. Statist. Soc. Ser. B 42 (1980), 205–209.
- [20] ———, *Some minimum variance block designs for estimating treatment differences*, J. Roy. Statist. Soc. Ser. B 45 (1983), 70–76.
- [21] A. T. James and G. N. Wilkinson, *Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance*, Biometrika 58 (1971), 279–294.
- [22] R. G. Jarrett, *Definitions and properties for m-concurrence designs*, J. Roy. Statist. Soc. Ser. B 45 (1983), 1–10.
- [23] B. Jin and J. P. Morgan, *Optimal saturated block designs when observations are correlated*, J. Statist. Plann. Inference 38 (2008), 3299–3308.
- [24] J. A. John, *Cyclic designs*, Monogr. Statist. Appl. Probab., Chapman & Hall, London, 1987.
- [25] J. A. John and T. Mitchell, *Optimal incomplete block designs*, J. Roy. Statist. Soc. Ser. B 39 (1977), 39–43.
- [26] B. Jones and J. A. Eccleston, *Exchange and interchange procedures to search for optimal designs*, J. Roy. Statist. Soc. Ser. B 42 (1980), 238–243.

- [27] O. Kempthorne, *The efficiency factor of an incomplete-block design*, Ann. Math. Statist. 27 (1956), 846–849.
- [28] M. K. Kerr and G. A. Churchill, *Experimental design for gene expression microarrays*, Biostatistics 2 (2001), 183–201.
- [29] A. M. Kshirsagar, *A note on incomplete block designs*, Ann. Math. Statist. 29 (1958), 907–910.
- [30] J. Kunert, *Optimal experimental design when the errors are assumed to be correlated*, Statist. Decisions, suppl. 2 (1985), 287–298.
- [31] R. Mead, *The non-orthogonal design of experiments*, J. Roy. Statist. Soc. Ser. A 153 (1990), 151–178.
- [32] J. P. Morgan, *Optimal incomplete block designs*, J. Amer. Statist. Assoc. 102 (2007), 655–663.
- [33] V. L. Mote, *On a minimax property of a balanced incomplete block design*, Ann. Math. Statist. 29 (1958), 910–913.
- [34] L. J. Paterson, *Circuits and efficiency in incomplete block designs*, Biometrika 70 (1983), 215–225.
- [35] K. R. Shah and B. K. Sinha, *Theory of optimal designs*, Lecture Notes in Statist., 54, Springer-Verlag, New York, 1989.
- [36] B. K. Sinha, *Some aspects of simplicity in the analysis of block designs*, J. Statist. Plann. Inference 6 (1982), 165–172.
- [37] E. R. Williams, *A note on rectangular lattice designs*, Biometrics 33 (1977), 410–414.
- [38] E. R. Williams, H. D. Patterson, and J. A. John, *Resolvable designs with two replications*, J. Roy. Statist. Soc. Ser. B 38 (1976), 296–301.
- [39] ———, *Efficient two-replicate resolvable designs*, Biometrics 33 (1977), 713–717.
- [40] E. Wit, A. Nobile, and R. Khanin, *Near-optimal designs for dual channel microarray studies*, J. Roy. Statist. Soc. Ser. C 54 (2005), 817–830.
- [41] F. Yates, *Incomplete randomized blocks*, Ann. Eugenics 7 (1936), 121–140.
- [42] ———, *A new method for arranging variety trials involving a large number of varieties*, J. Agric. Sci. 26 (1936), 424–455.

School of Mathematical Sciences
Queen Mary, University of London
London E1 4NS
United Kingdom

r.a.bailey@qmul.ac.uk