Journal of Statistical Planning and Inference 231 (2024) 106139

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Designs for half-diallel experiments with commutative orthogonal block structure

R.A. Bailey [a,b,*], Peter J. Cameron [a,b], Dário Ferreira [c], Sandra S. Ferreira [c], Célia Nunes [c]

[a] School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews, KY16 9SS, Fife, UK
[b] School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK
[c] Department of Mathematics and Center of Mathematics and Applications, University of Beira Interior, Avenida Marquês D'Ávila e Bolama, 6200, Covilhã, Portugal

ARTICLE INFO

MSC:
62K99
05B30
05E30

Keywords:
Half-diallel experiments
Commutative orthogonal block structure
Linear model
Triangular association scheme

ABSTRACT

In some experiments, the experimental units are all pairs of individuals who have to undertake a given task together. The set of such pairs forms a triangular association scheme. Appropriate randomization then gives two non-trivial strata. The design is said to have commutative orthogonal block structure (COBS) if the best linear unbiased estimators of treatment contrasts do not depend on the stratum variances. There are precisely three ways in which such a design can have COBS. We give a complete description of designs for which all treatment contrasts are in the same stratum. Then we give a very general construction for designs with COBS which have some treatment contrasts in each stratum.

1. Experiments with half-diallel structure

A diallel experiment generally estimates the variation of quantitative characters in genetic components. For example, they are used in plant breeding experiments to assess attributes like plant seed quality, and are also used in animal breeding experiments. The so-called *diallel cross* is the most balanced and systematic approach to investigate continuous variations among the genotypes of the individuals used in the study. In a full diallel experiment, the experimental units are all ordered crosses between m parental lines. Sometimes the self-crosses are excluded: see Yates (1947). Sometimes only a subset of these crosses is used, and the structure is called a *partial diallel cross*: see Curnow (1963), Fyfe and Gilbert (1963) and Kempthorne and Curnow (1961). In situations where the gender of the parent is irrelevant, it is efficient to use *half-diallel* experiments, in which the experimental units consist of all unordered crosses between m parental lines, excluding self-crosses: see Jones (1965).

The most common use of diallel experiments is to estimate properties of the genotypes. However, the half-diallel structure is also useful in some experiments where the experimental units naturally consist of unordered pairs and each treatment of interest is applied to some of these pairs: see Bailey (1991, 2003). These include experiments in human–computer interaction in which the computer is used for some task involving two people whose roles are the same. They may be conducting collaborative research, or simply having an online conversation: see Howes et al. (2009) and Özkan et al. (2021). This idea extends to any experiment where pairs of individuals are needed to complete a task, with both playing the same role. For example, the experiment might be conducted to compare different methods for researchers to collaborate when they are unable to meet face-to-face, such as email, online meetings, old-fashioned letters, and telephone calls with or without video.

* Corresponding author at: School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews, KY16 9SS, Fife, UK.
E-mail address: rab24@st-andrews.ac.uk (R.A. Bailey).

https://doi.org/10.1016/j.jspi.2023.106139
Received 15 March 2023; Received in revised form 23 November 2023; Accepted 18 December 2023
Available online 21 December 2023

	1	2	3	4	5
2					
3					
4					
5					
6					

(a) 15 elements

	1	2	3	4	5
2	1				
3	1	1			
4	1	1	1		
5	1	1	1	1	
6	1	1	1	1	1

(b) vector $\mathbf{v}_0$

	1	2	3	4	5
2	0				
3	0	0			
4	1	1	1		
5	0	0	0	1	
6	0	0	0	1	0

(c) vector $\mathbf{v}_4$

Fig. 1. The triangular association scheme $T(6)$.

In the half-diallel structure, the set Ω of experimental units consists of all unordered pairs from a set of m individuals, labelled $1, \ldots, m$. In Section 2 we remind readers that such a set can be described by a triangular association scheme, and give the relevant linear algebra for the benefit of those who are not very familiar with association schemes. Section 3 gives our assumed linear model, including how the process of randomization affects assumptions about the variance–covariance matrix. In Section 4 we remind readers about the property called *commutative orthogonal block structure*, and deduce what this means for the half-diallel structure. It turns out that there are three types of design with this property. These three types are fully described in the remainder of the paper.

2. The triangular association scheme

Association schemes were introduced in the 1950's to provide constructions for incomplete-block designs of sizes where no balanced incomplete-block designs are possible: see Bailey (2004), Bannai and Ito (1984), Bose and Mesner (1959), Bose and Nair (1939) and Bose and Shimamoto (1952). Here we can specialize immediately to association schemes with two associate classes.

Given a finite set Ξ of size N, let $\mathbf{I}$ be the $N \times N$ identity matrix whose rows and columns are labelled by the elements of Ξ, and let $\mathbf{J}$ be the $N \times N$ matrix with rows and columns labelled by the elements of Ξ with all elements equal to 1. The set of elements (ξ_1, ξ_2) of $\Xi \times \Xi$ with $\xi_1 \neq \xi_2$ is partitioned into two associate classes, in such a way that (ξ_1, ξ_2) is in the same class as (ξ_2, ξ_1). The *adjacency matrix* $\mathbf{A}$ for the first class is the $N \times N$ matrix whose (ξ_1, ξ_2)-entry is equal to 1 if (ξ_1, ξ_2) is in the first class; otherwise the entry is 0. The adjacency matrix for the second class is $\mathbf{J} - \mathbf{A} - \mathbf{I}$. This partition of $\Xi \times \Xi$ into the diagonal and two non-diagonal associate classes is defined to be an *association scheme* if and only if

$$\mathbf{A}^2 \text{ is a linear combination of } \mathbf{I}, \mathbf{A} \text{ and } \mathbf{J}. \tag{1}$$

If the coefficient of $\mathbf{J}$ in Eq. (1) is non-zero then the equation shows that $\mathbf{A}$ commutes with $\mathbf{J}$. Hence there is some constant c such that $\mathbf{A}$ has exactly c non-zero entries in each row and each column. Let $\mathbf{v}_0$ be the column vector of length N with all entries equal to 1. Then $\mathbf{A}\mathbf{v}_0 = c\mathbf{v}_0$. Let W_0 be the 1-dimensional subspace of $\mathbb{R}^N$ spanned by $\mathbf{v}_0$. If $\mathbf{v} \in \mathbb{R}^N \setminus W_0$ then some entries in $\mathbf{v}$ are different and so $\mathbf{A}\mathbf{v}$ cannot be equal to $c\mathbf{v}$. Hence W_0 is an eigenspace of $\mathbf{A}$ with dimension 1. The matrix $\mathbf{Q}_0$ of orthogonal projection onto W_0 is given by $\mathbf{Q}_0 = N^{-1}\mathbf{J}$.

If we restrict attention to vectors in the orthogonal complement $W_0^\perp$ of W_0, Eq. (1) shows that $\mathbf{A}$ satisfies a quadratic equation. Therefore $\mathbf{A}$ has two other eigenvalues λ_1 and λ_2, with their corresponding eigenspaces W_1 and W_2, and $\mathbb{R}^N$ is the orthogonal direct sum of W_0, W_1 and W_2.

For the triangular association scheme $T(m)$, which was first described by Bose and Shimamoto (1952), we let Ξ be the set Ω given in Section 1. Thus $N = m(m-1)/2$. We can picture the elements of Ω as the cells in a $m \times m$ square array with the diagonal missing and only the cells below the diagonal included. Fig. 1(a) shows this for the case that $m = 6$. Fig. 1(b) shows the vector $\mathbf{v}_0$ in this case.

For $i = 1, \ldots, m$, let $\mathbf{v}_i$ be the vector taking value 1 on each pair that contains individual i and the value 0 elsewhere. Fig. 1(c) shows the vector $\mathbf{v}_4$ when $m = 6$. Also, if $i \neq j$, let $\mathbf{v}_{ij}$ be the vector containing the value 1 on the pair $\{i, j\}$, with all other values 0. Thus

$$\sum_{i=1}^{m} \mathbf{v}_i = 2\mathbf{v}_0 \tag{2}$$

and

$$\sum_{j \neq i} \mathbf{v}_{ij} = \mathbf{v}_i. \tag{3}$$

Two distinct pairs in Ω are first associates if they have an individual in common; otherwise, they are second associates. In order to ensure that each pair has some second associates, we insist that $m \geq 4$.

Consider the pair $\{i,j\}$. It has $2(m-2)$ first associates. If $k \notin \{i,j\}$ then $\{i,j\}$ and $\{i,k\}$ have $(m-2)$ first associates in common; one of these is $\{j,k\}$ and the others are $\{i,\ell\}$ for $\ell \notin \{i,j,k\}$. If i, j, k and ℓ are all different, then the pairs $\{i,j\}$ and $\{k,\ell\}$ are second associates and they have four first associates in common (these are $\{i,k\}$, $\{i,\ell\}$, $\{j,k\}$ and $\{j,\ell\}$). Therefore

$$\mathbf{A}^2 = 2(m-2)\mathbf{I} + (m-2)\mathbf{A} + 4(\mathbf{J}-\mathbf{A}-\mathbf{I}) = (2m-8)\mathbf{I} + (m-6)\mathbf{A} + 4\mathbf{J}. \tag{4}$$

The coefficient of $\mathbf{J}$ in Eq. (4) is non-zero, so the previous argument shows that W_0 is an eigenspace of $\mathbf{A}$ with eigenvalue $2(m-2)$. If λ is an eigenvalue of $\mathbf{A}$ with eigenvectors in $W_0^\perp$ then Eq. (4) shows that

$$\lambda^2 = (2m-8) + (m-6)\lambda.$$

Therefore

$$(\lambda+2)(\lambda-(m-4)) = \lambda^2 - (m-6)\lambda - (2m-8) = 0,$$

and so $\lambda_1 = -2$ and $\lambda_2 = m-4$. Let d_1 and d_2 be the dimensions of W_1 and W_2. Then $d_1 + d_2 = N-1$. Also, $\mathrm{Tr}(\mathbf{A}) = 0$, because all the diagonal elements of $\mathbf{A}$ are zero, and so

$$2(m-2) + d_1(-2) + (m(m-1)/2 - 1 - d_1)(m-4) = 0.$$

Hence $d_1 = m(m-3)/2$. It follows that $d_2 = m-1$.

For any i in $\{1,\dots,m\}$, let us consider $\mathbf{A}\mathbf{v}_i$. If $j \neq i$ then the pair $\{i,j\}$ has exactly $m-2$ first associates which involve i. Thus the coefficient of $\mathbf{v}_{ij}$ in $\mathbf{A}\mathbf{v}_i$ is equal to $m-2$. If i, j and k are all different then the pair $\{j,k\}$ has exactly 2 first associates which involve i. Thus the coefficient of $\mathbf{v}_{jk}$ in $\mathbf{A}\mathbf{v}_i$ is equal to 2. Therefore

$$\mathbf{A}\mathbf{v}_i = (m-2)\mathbf{v}_i + 2(\mathbf{v}_0 - \mathbf{v}_i) = (m-4)\mathbf{v}_i + 2\mathbf{v}_0.$$

Therefore, if $j \neq i$ then $\mathbf{A}(\mathbf{v}_i - \mathbf{v}_j) = (m-4)(\mathbf{v}_i - \mathbf{v}_j)$, and so $\mathbf{v}_i - \mathbf{v}_j \in W_2$. Vectors of the form $\mathbf{v}_i - \mathbf{v}_j$ span a subspace of dimension $m-1$, and so this is the whole of W_2. The eigenspace W_1 is the orthogonal complement of $W_0 \oplus W_2$, which is spanned by vectors of the form $\mathbf{v}_{ij} + \mathbf{v}_{k\ell} - \mathbf{v}_{ik} - \mathbf{v}_{j\ell}$ with i, j, k and ℓ all different.

Let $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be the matrices of orthogonal projection onto W_1 and W_2 respectively. Then

$$\mathbf{I} = \mathbf{Q}_0 + \mathbf{Q}_1 + \mathbf{Q}_2$$

and

$$\mathbf{A} = 2(m-2)\mathbf{Q}_0 - 2\mathbf{Q}_1 + (m-4)\mathbf{Q}_2. \tag{5}$$

Therefore

$$\mathbf{A} + 2\mathbf{I} = 2(m-1)\mathbf{Q}_0 + (m-2)\mathbf{Q}_2.$$

It follows that

$$(m-2)\mathbf{Q}_2 = 2\mathbf{I} + \mathbf{A} - \frac{2(m-1)}{N}\mathbf{J} = 2\mathbf{I} + \mathbf{A} - \frac{4}{m}\mathbf{J},$$

and so

$$\mathbf{Q}_2 = \frac{2}{m-2}\mathbf{I} + \frac{1}{m-2}\mathbf{A} - \frac{4}{m(m-2)}\mathbf{J}.$$

Likewise,

$$(m-4)\mathbf{I} - \mathbf{A} = -m\mathbf{Q}_0 + (m-2)\mathbf{Q}_1,$$

and so

$$(m-2)\mathbf{Q}_1 = (m-4)\mathbf{I} - \mathbf{A} + m\mathbf{Q}_0 = (m-4)\mathbf{I} - \mathbf{A} + \frac{2}{m-1}\mathbf{J}.$$

Therefore

$$\mathbf{Q}_1 = \frac{m-4}{m-2}\mathbf{I} - \frac{1}{m-2}\mathbf{A} + \frac{2}{(m-1)(m-2)}\mathbf{J}.$$

3. Design, linear model, and randomization

Let $\mathcal{T}$ be the set of t treatments in the experiment (for example, various methods of remote communication). We always assume that $t \geq 2$. The *design* is a function $f : \Omega \to \mathcal{T}$ allocating treatment $f(\omega)$ to the pair ω in Ω. This can be summarized in the $N \times t$ *design matrix* $\mathbf{X}$. Its rows are indexed by elements of Ω, its columns are indexed by the treatments. In row ω there is a 1 in column $f(\omega)$, and all other entries are zero.

Denote by Y_ω the response on pair ω. We assume that, for each treatment A, there is a constant τ_A such that

$$Y_\omega = \tau_{f(\omega)} + \varepsilon_\omega,$$

where ε_ω is a normally distributed random variable with expectation zero. If we assemble the entries Y_ω and ε_ω into column vectors $\mathbf{Y}$ and $\boldsymbol{\varepsilon}$ of length N, and the entries τ_A into a column vector $\boldsymbol{\tau}$ of length t, we can write this as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\tau} + \boldsymbol{\varepsilon}. \tag{6}$$

The expectation $\mathbb{E}(\mathbf{Y})$ of $\mathbf{Y}$ is equal to $\mathbf{X}\boldsymbol{\tau}$.

Eq. (6) is an example of a linear model. Such models have been studied since the time of Gauss (1777–1855), as explained by Sprott (1978). See Gauss.

The design should be randomized by applying an appropriate permutation to the elements of Ω. In our case, where Ω consists of all unordered pairs from $\{1,\ldots,m\}$, the appropriate group of permutations is the symmetric group S_m of all permutations of the m individuals. One of these permutations should be chosen at random.

Let $\mathbf{V}$ be the variance–covariance matrix of $\mathbf{Y}$. Following Grundy and Healy (1950), Bailey (1981, 1991) argued that the entries $\mathbf{V}(\omega_1,\omega_2)$ and $\mathbf{V}(\omega_3,\omega_4)$ should be the same (albeit not known in advance) if and only if there is at least one of the permutations that might be chosen for randomization that takes (ω_1,ω_2) to (ω_3,ω_4) or to (ω_4,ω_3).

In our case, this means that $\mathbf{V}$ is a linear combination of $\mathbf{I}$, $\mathbf{A}$ and $\mathbf{J}-\mathbf{A}-\mathbf{I}$. The entries on the diagonal are variances, while the others are covariances, so we can write this as

$$\mathbf{V} = \sigma^2\mathbf{I} + \rho_1\sigma^2\mathbf{A} + \rho_2\sigma^2(\mathbf{J}-\mathbf{A}-\mathbf{I}). \tag{7}$$

Here σ^2 is the common variance, ρ_1 is the correlation between responses on pairs which are first associates, while ρ_2 is the correlation between responses on pairs which are second associates.

The results in Section 2 show that Eq. (7) can be rewritten as

$$\mathbf{V} = \gamma_0\mathbf{Q}_0 + \gamma_1\mathbf{Q}_1 + \gamma_2\mathbf{Q}_2, \tag{8}$$

where γ_0, γ_1 and γ_2 must all be non-negative but there are no other constraints on the values of γ_0, γ_1 and γ_2.

In fact, Eq. (7) can also be rewritten as

$$\mathbf{V} = \sigma^2(1-\rho_2)\mathbf{I} + \sigma^2(\rho_1-\rho_2)\mathbf{A} + \rho_2\sigma^2\mathbf{J}.$$

Then Eq. (5) gives

$$\begin{aligned}\mathbf{V} &= \sigma^2(1-\rho_2)(\mathbf{Q}_0+\mathbf{Q}_1+\mathbf{Q}_2)\\ &\quad + \sigma^2(\rho_1-\rho_2)[2(m-2)\mathbf{Q}_0 - 2\mathbf{Q}_1 + (m-4)\mathbf{Q}_2]\\ &\quad + \sigma^2\rho_2\frac{m(m-1)}{2}\mathbf{Q}_0.\end{aligned}$$

Comparing this with Eq. (8) shows that

$$\begin{aligned}\gamma_0 &= \sigma^2\left[1+2(m-2)\rho_1+\frac{(m-2)(m-3)}{2}\rho_2\right],\\ \gamma_1 &= \sigma^2[1-2\rho_1+\rho_2],\\ \gamma_2 &= \sigma^2[1+(m-4)\rho_1-(m-3)\rho_2].\end{aligned}$$

Houtman and Speed (1983) defined an experiment to have *orthogonal block structure*, OBS, if the variance–covariance matrix $\mathbf{V}$ of the model has a representation of the form

$$\mathbf{V} = \sum_{i=0}^{n}\gamma_i\mathbf{Q}_i,$$

where $\gamma_0,\ldots,\gamma_n$ are the eigenvalues of $\mathbf{V}$, there are no constraints, other than non-negativity, on the values of $\gamma_0,\ldots,\gamma_n$, and $\mathbf{Q}_0,\ldots,\mathbf{Q}_n$ are known symmetric, idempotent and pairwise orthogonal matrices, summing to $\mathbf{I}$.

The eigenspaces of $\mathbf{V}$ are often called *strata*. In this case, the matrices $\mathbf{Q}_i$ are called *stratum projectors*.

The definition of OBS was originally given by Nelder (1965a,b) for structures defined by factors. Nelder's definition is used in Bailey (1981, 1991). Other definitions of OBS may be found in Bailey (1994), Bailey and Brien (2016), Caliński and Kageyama (2000) and Ferreira et al. (2013). In this paper we use the Houtman–Speed definition.

Thus we have shown that, under the definition of half-diallel structures in Section 1, the linear model (6) and appropriate randomization, designs for half-diallel structures have OBS in the Houtman–Speed sense.

4. Commutative orthogonal block structure

Define the $N\times N$ matrix $\mathbf{T}$ by $\mathbf{T} = \mathbf{X}\left(\mathbf{X}^\top\mathbf{X}\right)^{-1}\mathbf{X}^\top$. This is the matrix of orthogonal projection onto the *treatment subspace* V_T, which is the subspace of $\mathbb{R}^N$ consisting of vectors which have constant entries on each treatment.

If $\mathbf{T}$ commutes with the stratum projector $\mathbf{Q}_i$ then $\mathbf{T}\mathbf{Q}_i$ is the matrix of orthogonal projection onto the intersection of V_T and stratum W_i. In the special case that $\mathbf{T}\mathbf{Q}_i = \mathbf{Q}_i\mathbf{T} = \mathbf{0}$, this intersection is the zero subspace, containing only the zero vector. Thus if $\mathbf{T}$ commutes with all stratum projectors then V_T is the orthogonal direct sum of all such intersections that are non-zero. In this case the best linear unbiased estimators of treatment contrasts are obtained by the standard method of ordinary least squares, with no need to know the values of the eigenvalues of $\mathbf{V}$. This was pointed out by Kruskal (1968).

The consequences of the combination of OBS with the property that $\mathbf{T}$ commutes with all stratum projectors were investigated by Isotalo et al. (2008), Markewicz et al. (2010), Puntanen and Styan (1989) and Zmyślony (1980). Fonseca et al. (2008) named this combination of properties *commutative orthogonal block structure*. It is now called COBS for short. See also Bailey et al. (2016), Carvalho et al. (2015), Ferreira et al. (2010, 2013), Fonseca et al. (2010), Nunes et al. (2008) and Santos et al. (2017, 2020). It has also been called *equivalent estimation* in Macharia and Goos (2010), Parker et al. (2007) and Vining et al. (2005).

Because $\mathbf{T}$ commutes with both $\mathbf{I}$ and $\mathbf{J}$, a half-diallel design has COBS if and only if $\mathbf{TA} = \mathbf{AT}$. Thus there are only three ways in which a half-diallel design can have COBS. We name the three types as follows.

Type I: the treatment subspace V_T is contained in $W_0 \oplus W_1$.
Type II: the treatment subspace V_T is contained in $W_0 \oplus W_2$.
Type III: the treatment subspace V_T has non-zero intersections with both W_1 and W_2.

To understand these three types, readers may like to compare this situation to an experiment which has 48 experimental units grouped into 12 blocks of size 4. Given a treatment factor F with four levels, one possibility is to allocate each level of F to one experimental unit per block, thus giving a complete-block design. This is analogous to Type I. Given a different treatment factor G with four levels which cannot easily be applied to small units, another possibility is to allocate each level of G to three whole blocks, thus giving a design with repeated measures. This is analogous to Type II. The third possibility is to make a factorial design with treatment factors F and G allocated as above, thus giving what is often called a split-plot design. This is analogous to Type III.

Not all designs for half-diallel experiments have COBS. Bailey (2003) gives three designs for triangular association schemes; none of them has COBS. Like the designs in Bailey (2005), they are *balanced* in the sense that all contrasts between two treatments have the same variances, irrespective of the values of γ_1 and γ_2.

In the remainder of this paper, we give a complete description of half-diallel designs which have COBS and Types I or II, and a very general construction for such designs with Type III.

5. Designs of Type I

For each treatment A and each individual i, denote by p_{Ai} the number of pairs on which treatment A occurs and which include individual i. For example, in the design in Fig. 2(b) we have $p_{Ai} = 2$ for $i = 1, \ldots, 8$ and $p_{Di} = 1$ for $i = 1, \ldots, 8$.

Let $\mathbf{v}(A)$ be the vector which has entry 1 on each pair where treatment A occurs and entry 0 elsewhere. (This notation is a little different from $\mathbf{v}_i$ because we have some examples where an individual and a treatment may have the same label.) Thus $p_{Ai} = \mathbf{v}(A)\cdot\mathbf{v}_i$, where $\cdot$ denotes the inner product.

Theorem 1. *A half-diallel design has Type I if and only if, for every treatment A and all pairs of individuals i and j, we have $p_{Ai} = p_{Aj}$.*

Proof. The design has Type I if and only if V_T is contained in $W_0 \oplus W_1$. Equivalently, a design has Type I if and only if V_T is orthogonal to W_2. The subspace V_T is spanned by the vectors $\mathbf{v}(A)$ for all treatments A, while W_2 is spanned by the vectors $\mathbf{v}_i - \mathbf{v}_j$ for all pairs of individuals i and j. Since $\mathbf{v}(A) \cdot (\mathbf{v}_i - \mathbf{v}_j) = p_{Ai} - p_{Aj}$, it follows that the design has Type I if and only if $p_{Ai} = p_{Aj}$ for every treatment A and all pairs of individuals i and j. $\square$

Thus, for a Type I design, we can write p_{Ai} simply as p_A. For half-diallel design, the overall replication r_A of treatment A is equal to $\left(\sum_{i=1}^{m} p_{Ai}\right)/2$. Therefore $r_A = mp_A/2$. Thus we have this immediate corollary.

Corollary 1. *In a design of Type I, every product mp_A is even.*

Theorem 1 gives us two combinatorial ways of thinking about a design of Type I. In one, the individuals label the vertices of a complete graph. The edge between vertices i and j is labelled by A if treatment A occurs on $\{i, j\}$. Denote by Γ_A the graph whose edges are labelled by A. Theorem 1 shows that Γ_A is regular with degree p_A for each A.

In the other way, create an $m \times m$ square array. For $i \neq j$, put treatment $f(\{i, j\})$ in cells (i, j) and (j, i). Put an unused treatment ∞ in every cell on the diagonal. If $p_A = 1$ for every treatment A, the result is a symmetric Latin square with a constant diagonal. If any $p_A > 1$, the ratios of treatment occurrences in each row and column are the same, so the square is called a *frequency square* or *F-square*. These were introduced by Finney (1945, 1946a,b). See also Freeman (1966), Addelman (1967), and Hedayat and Seiden (1970).

If $p_A = 1$ for all treatments A, Corollary 1 shows that m must be even. Moreover, all treatments have replication $m/2$, and so $t = m-1$. From the context of graphs, such a design is called a *one-factorization* of the complete graph. Here we give two constructions for such designs.

Construction 1. *Identify the treatments with the integers modulo t. If $i \neq j$, allocate treatment $i + j \mod t$ to the pair $\{i, j\}$. If $i \neq m$, put treatment $2i \mod t$ on the pair $\{i, m\}$.*

Example 1. When $t = 7$, Construction 1 gives the design in Fig. 2(a).

	1	2	3	4	5	6	7
2	3						
3	4	5					
4	5	6	7				
5	6	7	1	2			
6	7	1	2	3	4		
7	1	2	3	4	5	6	
8	2	4	6	1	3	5	7

(a) Example 1

	1	2	3	4	5	6	7
2	B						
3	B	C					
4	C	C	D				
5	C	D	A	A			
6	D	A	A	B	B		
7	A	A	B	B	C	C	
8	A	B	C	A	B	C	D

(b) Example 4

Fig. 2. Designs with $m = 8$ in Examples 1 and 4.

	1	2	3	4	5	6	7	8	9
2	3								
3	2	1							
4	7	9	8						
5	9	8	7	6					
6	8	7	9	5	4				
7	4	6	5	1	3	2			
8	6	5	4	3	2	1	9		
9	5	4	6	2	1	3	8	7	
10	1	2	3	4	5	6	7	8	9

(a) Example 2

	1	2	3	4	5	6	7	8	9
2	A								
3	A	A							
4	C	C	C						
5	C	C	C	B					
6	C	C	C	B	B				
7	B	B	B	A	A	A			
8	B	B	B	A	A	A	C		
9	B	B	B	A	A	A	C	C	
10	A	A	A	B	B	B	C	C	C

(b) Example 3

Fig. 3. Designs with $m = 10$ in Examples 2 and 3.

From the point of view of graph theory, Wallis (1973) gave Construction 1. Moreover, he showed that, up to isomorphism, this is the only possibility if $m = 4$ or $m = 6$, but, for larger even values of m, there are non-isomorphic possibilities. See also Cameron (1994).

Construction 2. *If $t \geq 3$ and $t \equiv 1 \mod 6$ or $t \equiv 3 \mod 6$ then there is a Steiner triple system with t points. This is a balanced incomplete-block design for t treatments in $t(t-1)/6$ blocks of size 3 in which every pair of treatments concur in exactly one block. Label the individuals and the treatments by the points of such a Steiner triple system, with one extra individual labelled m. If i, j and m are all different and $\{i, j, k\}$ is a block of the Steiner triple system then put treatment k on the pair $\{i, j\}$. If $i \neq m$, put treatment i on the pair $\{i, m\}$.*

Example 2. When $t = 9$ there is a Steiner triple system with these blocks.

$$\begin{array}{cccccc} \{1,2,3\} & \{4,5,6\} & \{7,8,9\} & \{1,4,7\} & \{2,5,8\} & \{3,6,9\} \\ \{1,5,9\} & \{2,6,7\} & \{3,4,8\} & \{1,6,8\} & \{2,4,9\} & \{3,5,7\} \end{array}$$

Using this in Construction 2 gives the design in Fig. 3(a).

Constructions 1 and 2 produce large numbers (more than exponentially many) of different examples for large even values of m: see Cameron (1976).

If some treatment A has $p_A > 1$ then we can make a design from one of the previous constructions by simply choosing p_A treatments and replacing them all by A. The variance of the best linear unbiased estimator of $\tau_A - \tau_B$ is equal to

$$\left(\frac{1}{r_A} + \frac{1}{r_B}\right)\gamma_1.$$

The sum of these variances, over all pairs of treatments, is minimized when all replications are as equal as possible. This is equivalent to making the p_A values as equal as possible.

	1	2	3	4	5	6	7	8	9	10	11	12
2	1											
3	2	1										
4	3	2	1									
5	4	3	2	1								
6	5	4	3	2	1							
7	6	5	4	3	2	1						
8	6	6	5	4	3	2	1					
9	5	6	6	5	4	3	2	1				
10	4	5	6	6	5	4	3	2	1			
11	3	4	5	6	6	5	4	3	2	1		
12	2	3	4	5	6	6	5	4	3	2	1	
13	1	2	3	4	5	6	6	5	4	3	2	1

Fig. 4. Design with $m = 13$ in Example 5.

If t divides $m - 1$ then we can put $p_A = (m - 1)/t$ for every treatment A. Otherwise, we should ensure that $|p_A - p_B| \leq 1$ whenever A and B are different treatments.

Example 3. Suppose that $t = 3$ and $m = 10$. Then we put $p_A = 3$ for every treatment A. We can start with the design in Fig. 3(a) and merge the nine original treatments into three lots of three in any way at all. One possibility is shown in Fig. 3(b).

Example 4. Suppose that $t = 4$ and $m = 8$. Then we can put $p_A = p_B = p_C = 2$ and $p_D = 1$. Starting from the design in Fig. 2(a), one possibility is the design shown in Fig. 2(b).

Now we turn attention to Type I designs with m odd. Corollary 1 shows that p_A is even for all treatments A. We begin with two constructions for the case that $p_A = 2$ for all treatments, so that $m = 2t + 1$.

Construction 3. *Consider the individuals as the vertices of a polygon, numbered in order. If i and j are different vertices, then there are two distances between them using the edges of the polygon. Because m is odd, these two distances are not the same. Use the smaller one to label the treatment applied to the pair $\{i, j\}$. Equivalently, this label is whichever is smaller of the differences $i - j$ and $j - i$ mod m.*

Example 5. When $m = 13$, Construction 3 gives the design in Fig. 4.

If Construction 3 is used when m is prime, the edges labelled by each treatment give a single m-gon. Such an m-gon is called a *Hamiltonian cycle*. If m is not prime then the edges labelled by any difference which divides m do not form a single cycle.

A decomposition of the complete graph K_m for odd m into Hamiltonian cycles is attributed to Walecki by Édouard Lucas in Volume 2 of Lucas (1882–1894). This is called the answer to the *problème de ronde*. The construction is as follows.

Construction 4. *Put $2q = m - 1$. Label the individuals by the integers modulo $2q$, with another special one labelled m. The first treatment is put on the edges of the following Hamiltonian cycle:*

$$(m, 1, 2, 2q, 3, 2q - 1, \ldots, q, q + 2, q + 1).$$

The cycles for the remaining treatments are obtained by adding $1, 2, \ldots, q - 1$ (mod $2q$), with m remaining fixed.

Example 6. When $m = 9$, Construction 4 gives the design in Fig. 5(a). It is not possible to obtain this using Construction 3.

If m is odd and $m > 2t + 1$ then there must be some treatment A for which $p_A > 2$. As in the case for even m, we should ensure that, for different treatments A and B, the replications are as equal as possible. Since p_A and p_B must both be even, we should have $|p_A - p_B| \leq 2$. Now we can start with a design made from Construction 3 or Construction 4 and then merge treatments in the same way that we did for even m.

6. Designs of Type II

Theorem 2. *A half-diallel design has Type II if and only if, for every treatment A, the vector $\mathbf{v}(A)$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.*

	1	2	3	4	5	6	7	8
2	1							
3	2	2						
4	2	3	3					
5	3	3	4	4				
6	3	4	4	1	1			
7	4	4	1	1	2	2		
8	4	1	1	2	2	3	3	
9	1	2	3	4	1	2	3	4

(a) Example 6

	1	2	3	4	5	6	7	8
2	A							
3	A	B						
4	A	B	B					
5	A	B	B	B				
6	A	B	B	B	B			
7	A	B	B	B	B	B		
8	A	B	B	B	B	B	B	
9	A	B	B	B	B	B	B	B

(b) Example 7

Fig. 5. Designs with $m = 9$ in Examples 6 and 7.

Proof. The design has Type II if and only if the treatment subspace V_T is contained in $W_0 \oplus W_2$. The vectors $\mathbf{v}_1, \dots, \mathbf{v}_m$ form a basis for $W_0 \oplus W_2$. □

Theorem 3. *Suppose that $t = 2$ in a half-diallel design. If there is some individual i such that treatment A is applied to all pairs containing i and treatment B is applied to all other pairs, then the design has Type II.*

Proof. Here $\mathbf{v}(A) = \mathbf{v}_i$ and $\mathbf{v}(B) = \mathbf{v}_0 - \mathbf{v}_i$. Eq. (2) shows that $\mathbf{v}(B) = (\mathbf{v}_1 + \cdots + \mathbf{v}_m)/2 - \mathbf{v}_i$, so both $\mathbf{v}(A)$ and $\mathbf{v}(B)$ are in $W_0 \oplus W_2$. □

Theorem 4. *There are no other half-diallel designs of Type II.*

Proof. Let A be any treatment. Theorem 2 shows that if the design has Type II then there are constants $a_1, a_2, \dots, a_m$ such that

$$\mathbf{v}(A) = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m. \tag{9}$$

Hence Eq. (3) shows that

$$\mathbf{v}(A) = \sum_{i=1}^{m-1}\sum_{j=i+1}^{m}(a_i + a_j)\mathbf{v}_{ij}. \tag{10}$$

All the coefficients in Eq. (10) must be 1 or 0, and so we have $a_i + a_j \in \{0, 1\}$ whenever $i \neq j$.

Suppose that i, j and k are all different. The only way that all three of $a_i + a_j$, $a_i + a_k$ and $a_j + a_k$ can be zero is if $a_i = a_j = a_k = 0$. However, the coefficients in Eq. (9) cannot all be zero. Without loss of generality, let us assume that $a_i + a_j = 1$.

If $a_i + a_k = 1$ then $a_j = a_k$ and so either $a_j = a_k = 0$ and $a_i = 1$ or $a_j = a_k = 1/2$ and $a_i = 1/2$. On the other hand, if $a_i + a_k = 0$ then $a_j = a_k + 1$ and so $a_j + a_k = 2a_k + 1$. Thus either $a_k = 0$, in which case $a_i = 0$ and $a_j = 1$, or $a_k = -1/2$, in which case $a_i = a_j = 1/2$.

Thus every triple of coefficients in Eq. (9) must be one of the multi-sets $\{0, 0, 0\}$, $\{1, 0, 0\}$, $\{1/2, 1/2, 1/2\}$ and $\{1/2, 1/2, -1/2\}$. They cannot all be $\{0, 0, 0\}$, as already explained. Moreover, they cannot all be $\{1/2, 1/2, 1/2\}$, because that implies that $\mathbf{v}(A) = \mathbf{v}_0$, in which case there are no other treatments.

Thus at most one of the a_i is 1, and if this occurs then $a_j = 0$ whenever $j \neq i$. In this case, $\mathbf{v}(A) = \mathbf{v}_i$. If there is another treatment B with $\mathbf{v}(B) = \mathbf{v}_j$ then treatments A and B both occur on pair $\{i, j\}$. This cannot happen, and so there is at most one treatment whose triple is $\{1, 0, 0\}$.

If none of the coefficients is 1, then the only possibility is that one is equal to $-1/2$ and the rest equal to $1/2$. Suppose that $a_i = -1/2$. Then $\mathbf{v}(A) = \mathbf{v}_0 - \mathbf{v}_i$. Suppose that there is another treatment B with $\mathbf{v}(B) = \mathbf{v}_0 - \mathbf{v}_j$. Because $m \geq 4$, there are two different individuals k and ℓ which are both different from i and j. Thus treatments A and B both occur on pair $\{k, \ell\}$. This cannot happen, and so there is at most one treatment whose triple is $\{-1/2, 1/2, 1/2\}$.

Thus there are precisely two treatments. One of these occurs on all pairs containing some designated individual i, and the other occurs on all other pairs. □

Example 7. Fig. 5(b) shows a design of Type II with $m = 9$ and $t = 2$.

7. Designs of Type III

Theorem 5. *If every treatment has replication 1 then the design has Type III.*

Proof. If every treatment has replication 1 then $\mathbf{T} = \mathbf{I} = \mathbf{Q}_0 + \mathbf{Q}_1 + \mathbf{Q}_2$ and the treatment subspace V_T is the whole of $\mathbb{R}^N$, which contains W_0, W_1 and W_2. Therefore $\mathbf{T}\mathbf{Q}_i = \mathbf{Q}_i\mathbf{T} = \mathbf{Q}_i$ for $i = 0$, 1 and 2. □

A design with no replicated treatments is probably of no practical use, but this theorem does at least show that there are designs of Type III.

We will now give a very general construction for half-diallel designs with COBS. Since those of Types I and II have been identified by Theorems 1, 2, 3 and 4, all the remaining ones have Type III.

It turns out that there is a strong connection between the COBS property for half-diallel structures and a combinatorial concept called *equitable partitions* of graphs. See Bailey et al. (2019) and Gavrilyuk and Goryainov (2013), for example. Let Γ_m be the graph whose vertices are the pairs in Ω, with an edge between two pairs if they have an individual in common. Thus the adjacency matrix of Γ_m (in the graph-theoretical sense) is precisely the adjacency matrix $\mathbf{A}$ defined in Section 2.

For each treatment A, let Δ_A be the set of pairs ω for which $f(\omega) = A$. The subsets Δ_A, Δ_B, … give a *partition* $\mathcal{P}$ of Ω, which means a set of non-empty disjoint subsets whose union is Ω. The size of Δ_A is the replication r_A of treatment A. The treatment subspace V_T is precisely the subspace of $\mathbb{R}^N$ consisting of vectors which are constant on each part of $\mathcal{P}$, and the matrix of orthogonal projection onto this subspace is $\mathbf{T}$.

Given any graph Γ and any partition $\mathcal{P}$ of its set of vertices into t parts, labelled Δ_A, Δ_B, …, $\mathcal{P}$ is defined to be *equitable* with respect to Γ if there is a $t \times t$ matrix $C = (c_{AB})$ such that, for $\omega \in \Delta_A$, the number of vertices in Δ_B joined to ω is c_{AB}, so depending only on A and B and not on the choice of ω in Δ_A.

Theorem 6. *The treatment partition $\mathcal{P}$ is equitable with respect to the graph Γ_m if and only if $\mathbf{T}$ commutes with $\mathbf{A}$.*

Proof. First assume that the partition $\mathcal{P}$ is equitable. Let $\{i, j\}$ be a vertex in Δ_A, where A is any treatment. Then

$$\mathbf{A}\mathbf{v}_{ij} = \sum_{k \notin \{i,j\}} \mathbf{v}_{ik} + \sum_{k \notin \{i,j\}} \mathbf{v}_{jk}.$$

Of the neighbours listed in this sum, precisely c_{AB} belong to Δ_B for each treatment B. Then applying $\mathbf{T}$ averages these vectors over each part Δ_B, which has size r_B, so that the vector $\mathbf{v}(B)$ has coefficient c_{AB}/r_B. Thus

$$\mathbf{T}\mathbf{A}\mathbf{v}_{ij} = \sum_{B=1}^{t} c_{AB}\mathbf{v}(B)/r_B.$$

On the other hand, applying $\mathbf{T}$ to $\mathbf{v}_{ij}$ averages it over Δ_A, so that $\mathbf{T}\mathbf{v}_{ij} = \mathbf{v}(A)/r_A$. Now, every vertex in Δ_B has c_{BA} neighbours in Δ_A; so

$$\mathbf{A}\mathbf{T}\mathbf{v}_{ij} = \frac{1}{r_A}\sum_{B=1}^{t} c_{BA}\mathbf{v}(B).$$

However, double counting edges between Δ_A and Δ_B shows that

$$r_A c_{AB} = r_B c_{BA},$$

so $\mathbf{T}\mathbf{A}\mathbf{v}_{ij} = \mathbf{A}\mathbf{T}\mathbf{v}_{ij}$ for all $\{i, j\}$ in Ω, and hence $\mathbf{T}\mathbf{A} = \mathbf{A}\mathbf{T}$.

Conversely, suppose that $\mathbf{T}\mathbf{A} = \mathbf{A}\mathbf{T}$. Then $\mathbf{A}\mathbf{T}\mathbf{v}$ lies in the image of $\mathbf{T}$ for any vector $\mathbf{v}$, and so is constant each part of $\mathcal{P}$. If this constant is b_{AB} on Δ_B for a vector $\mathbf{v}_{ij}$ with $\{i, j\}$ in Δ_A, then each vertex in Δ_B has $r_A b_{AB}$ neighbours in Δ_A. So the partition $\mathcal{P}$ is equitable, with $c_{BA} = r_A b_{AB}$. □

Previous work on equitable partitions by Bailey et al. (2019), Gavrilyuk and Goryainov (2013) and Gavrilyuk and Metsch (2014) for graphs whose adjacency matrices have three or four eigenspaces has been quite successful when only one non-trivial eigenspace intersects the subspace defined by the partition (so, essentially like our Types I and II). However, in the other cases (like our Type III) it seems much harder to give a complete classification. So we will give one, rather general, method of constructing half-diallel designs with COBS of Type III, but we cannot guarantee that all designs of Type III can be obtained by this method.

Recall that the set Ω of experimental units consists of all 2-element subsets (called "pairs") of the set $\{1, \dots, m\}$ of individuals. Section 2 described the triangular association scheme $T(m)$ on Ω; two pairs are first or second associates according as they intersect in 1 or 0 individuals. Said otherwise, the pairs of the association scheme are the edges of the complete graph K_m. To avoid confusion with the graph Γ_m introduced at the start of this section, we call the edges of K_m "lines" and its vertices "points".

We now give a fairly general construction for an equitable partition of Γ_m. Each part is thus a set of lines of a graph on the point-set $\{1, \dots, m\}$, and we describe the sets in the partition as graphs.

Construction 5. *Here are the steps.*

1. *Partition $\{1, \dots, m\}$ into n disjoint sets S_1, …, S_n called sorts. Let s_i be the size of S_i.*
2. *For each i, let $\mathcal{T}_i$ be a set of regular graphs on the point-set S_i whose line-sets partition the set S_{ii} of unordered pairs from S_i. If $s_i = 1$ then $S_{ii} = \emptyset$ and $\mathcal{T}_i = \emptyset$.*
3. *For $i < j$, let $\mathcal{T}_{ij}$ be a set of semiregular bipartite graphs with parts S_i and S_j, whose line-sets partition the set S_{ij} of pairs with one element of each of these two sorts. (The valency of a point in each graph should depend only on whether the point is in S_i or S_j.) (Sometimes it is convenient to write the set S_{ij} as S_{ji}.)*

Theorem 7. *The line-sets of the graphs in the sets $\mathcal{T}_i$ ($1 \le i \le n$, with $s_i > 1$) and $\mathcal{T}_{ij}$ ($1 \le i < j \le n$) form an equitable partition of Γ_m.*

Proof. It is clear from the construction that this really is a partition of the set of all pairs.

We have to show that, given two graphs from the sets constructed above, the number of lines of the second graph meeting a line e of the first — in other words, the number of lines of the second graph through either point v or point w, where $e = \{v, w\}$ (but not equal to e) — does not depend on the choice of e. This just involves some (rather straightforward) checking.

Consider first the case where the two graphs are the same, say $\mathcal{G}$. Then the number of lines of $\mathcal{G}$ meeting a given line $e = \{v, w\}$ is the sum of the valencies of v and w, minus 2; these numbers are independent of e (since each graph is either regular or semiregular bipartite).

So suppose that the first graph is $\mathcal{G}_1$ and the second is $\mathcal{G}_2$. In the cases where $\mathcal{G}_1 \in \mathcal{T}_i$ and $\mathcal{G}_2 \in \mathcal{T}_j$ with $i \ne j$, or $\mathcal{G}_1 \in \mathcal{T}_i$ and $\mathcal{G}_2 \in \mathcal{T}_{jk}$ with $i \notin \{j, k\}$ or *vice versa*, or $\mathcal{G}_1 \in \mathcal{T}_{ij}$ and $\mathcal{G}_2 \in \mathcal{T}_{kl}$ with $\{i, j\} \cap \{k, l\} = \emptyset$, the number is zero. This leaves a few cases to check. Let e be a line of $\mathcal{G}_1$. The number of lines of $\mathcal{G}_2$ meeting e is as follows.

- $\mathcal{G}_1, \mathcal{G}_2 \in \mathcal{T}_i$: twice the valency of $\mathcal{G}_2$.
- $\mathcal{G}_1 \in \mathcal{T}_i$, $\mathcal{G}_2 \in \mathcal{T}_{ij}$: twice the valency of $\mathcal{G}_2$ associated with points in $\mathcal{S}_i$.
- $\mathcal{G}_1 \in \mathcal{T}_{ij}$, $\mathcal{G}_2 \in \mathcal{T}_i$: the valency of $\mathcal{G}_2$.
- $\mathcal{G}_1, \mathcal{G}_2 \in \mathcal{T}_{ij}$: the sum of the two valencies of $\mathcal{G}_2$.
- $\mathcal{G}_1 \in \mathcal{T}_{ij}$, $\mathcal{G}_2 \in \mathcal{T}_{ik}$: the valency of $\mathcal{G}_2$ associated with points in $\mathcal{S}_i$.

This covers all cases. □

We remark that, by Theorem 6, this shows that (translating back from lines of K_m to pairs of the triangular scheme $T(m)$) we have a COBS.

Here is another, equivalent, way of describing Construction 5 once the sets $\mathcal{S}_i$ have been specified. Recall that $\mathcal{S}_i$ is a set of s_i individuals, and that $\mathcal{S}_{ii}$ is the set of pairs of individuals of sort i. If $s_i > 1$, put a design of Type I on pairs of individuals of sort i, using t_i treatments forming a set $\mathcal{T}_i$. If $s_i = 2$ then $\mathcal{T}_i$ has a single treatment with replication one, so this case should be avoided. If $s_i = 3$ then the only way to avoid replication one is to have $t_i = 1$.

Suppose that $i < j$. Recall that $\mathcal{S}_{ij}$ is the set of pairs of individuals where one individual is of sort i and the other is of sort j. Because our figures use the cells below the main diagonal, we can picture the set $\mathcal{S}_{ij}$ as a rectangle with s_j rows and s_i columns. For example, in Fig. 6(b), $\mathcal{S}_{12}$ is the rectangle formed by rows 4–9 and columns 1–3. Let t_{ij} be any common divisor of s_i and s_j. Make a set $\mathcal{T}_{ij}$ of t_{ij} treatments and allocate them to pairs of individuals in $\mathcal{S}_{ij}$ in such a way that each treatment occurs s_j/t_{ij} times with each individual of sort i and s_i/t_{ij} times with each individual of sort j. If $s_i = s_j = 1$ then $t_{ij} = 1$ and the single treatment in $\mathcal{T}_{ij}$ has replication one, so this should be avoided.

Example 8. Suppose that $m = 9$. If $n = 1$ then Construction 5 gives a design of Type I. One possibility is shown in Fig. 5(a).

If $n = 2$ and no treatment has replication one then the only possibilities for $\{s_1, s_2\}$ are $\{1, 8\}$, $\{3, 6\}$ and $\{4, 5\}$. The first possibility can give the design of Type II in Fig. 5(b), but we can also put a non-trivial Type I design on $\mathcal{T}_2$, as shown in Fig. 6(a), where treatment A occurs throughout $\mathcal{S}_2 \times \mathcal{S}_1$ while the other treatments occur in $\mathcal{S}_{22}$. The second and third are shown in Fig. 6(b) and (c), with all treatment sets $\mathcal{T}_i$ and $\mathcal{T}_{ij}$ as large as possible. In any of these cases, any treatments within the same treatment set may be merged without violating the COBS condition.

If $n = 3$ and no treatment has replication one then the only possibilities for $\{s_1, s_2, s_3\}$ (considered as a multiset) are $\{3, 3, 3\}$, $\{1, 4, 4\}$ and $\{1, 3, 5\}$. These are shown in Fig. 6(d), (e) and (f), again with all treatment sets as large as possible. As above, any treatments within the same treatment set may be merged.

It is not possible to have $n \ge 4$ without having some treatment with replication one.

The proof of the following theorem shows that we can be explicit about which treatment contrasts are in W_1 and which are in W_2 when a design is made using Construction 5.

Theorem 8. *Suppose that a half-diallel design with t treatments is constructed using Construction 5 with the individuals partitioned into n sorts. Then $\dim(V_T \cap W_1) = t - n$ and $\dim(V_T \cap W_2) = n - 1$.*

Proof. Since the subsets $\mathcal{S}_1, \ldots, \mathcal{S}_n$ are considered to give n sorts of individual, then the non-empty subsets among $\mathcal{S}_{11}, \ldots, \mathcal{S}_{nn}$, $\mathcal{S}_{12}, \ldots, \mathcal{S}_{n-1,n}$ might be called *sort-pairs*. Let V_1 be the subspace of treatment contrasts whose coefficients sum to zero within each of $\mathcal{T}_1, \ldots, \mathcal{T}_n, \mathcal{T}_{12}, \ldots, \mathcal{T}_{n-1,n}$. Then the coefficients sum to zero within each sort-pair, and so the contrasts in V_1 could be called *contrasts within sort-pairs*. Any treatment contrast that is entirely within one of the sort-pairs is orthogonal to the characteristic vectors of all individuals, and so it is in W_1. Therefore $V_1 \le W_1$.

Let z be the number of individuals i which have $s_i = 1$. Then the number of non-empty sort-pairs is $n(n+1)/2 - z$. Hence $\dim(V_1) = t - n(n+1)/2 + z$.

Let V_2 be the subspace of V_T consisting of vectors which are constant on each sort-pair and sum to zero on each sort. This is orthogonal to V_1, and its vectors could be called *contrasts between sort-pairs within sorts*.

For $i \ne j$, let $\mathbf{w}_{ij}$ be the vector with entry $1/s_i s_j$ on $\mathcal{S}_{ij}$ and 0 elsewhere. (This is like a scaled version of the vector $\mathbf{v}_{ij}$ in Section 2: its inner product with any vector $\mathbf{v}$ averages the entries of $\mathbf{v}$ within $\mathcal{S}_{ij}$.) If $s_i > 1$, let $\mathbf{w}_{ii}$ be the vector with entry $2/s_i(s_i - 1)$ on $\mathcal{S}_{ii}$ and 0 elsewhere. These vectors are all in $V_T \cap V_1^{\perp}$.

	1	2	3	4	5	6	7	8
2	A							
3	A	B						
4	A	C	D					
5	A	D	C	B				
6	A	E	F	G	H			
7	A	F	E	H	G	B		
8	A	G	H	E	F	C	D	
9	A	H	G	F	E	D	C	B

(a) $s_1 = 1,\ s_2 = 8$

	1	2	3	4	5	6	7	8
2	A							
3	A	A						
4	B	C	D					
5	B	C	D	E				
6	D	B	C	F	I			
7	D	B	C	G	H	E		
8	C	D	B	H	F	G	I	
9	C	D	B	I	G	H	F	E

(b) $s_1 = 3,\ s_2 = 6$

	1	2	3	4	5	6	7	8
2	A							
3	B	C						
4	C	B	A					
5	D	D	D	D				
6	D	D	D	D	E			
7	D	D	D	D	F	E		
8	D	D	D	D	F	F	E	
9	D	D	D	D	E	F	F	E

(c) $s_1 = 4,\ s_2 = 5$

	1	2	3	4	5	6	7	8
2	A							
3	A	A						
4	B	C	D					
5	D	B	C	E				
6	C	D	B	E	E			
7	F	G	H	I	J	K		
8	G	H	F	K	I	J	L	
9	H	F	G	J	K	I	L	L

(d) $s_1 = 3,\ s_2 = 3,\ s_3 = 3$

	1	2	3	4	5	6	7	8
2	A							
3	A	B						
4	A	C	D					
5	A	D	C	B				
6	E	F	G	H	I			
7	E	G	H	I	F	J		
8	E	H	I	F	G	K	L	
9	E	I	F	G	H	L	K	J

(e) $s_1 = 1,\ s_2 = 4,\ s_3 = 4$

	1	2	3	4	5	6	7	8
2	A							
3	A	B						
4	A	B	B					
5	C	D	D	D				
6	C	D	D	D	E			
7	C	D	D	D	F	E		
8	C	D	D	D	F	F	E	
9	C	D	D	D	E	F	F	E

(f) $s_1 = 1,\ s_2 = 3,\ s_3 = 5$

Fig. 6. Designs with $m = 9$ in Example 8.

Table 1
Skeleton analysis-of-variance tables for some half-diallel designs with COBS which use nine individuals, showing degrees of freedom.

Figure		5(a)	5(b)	6(a)	6(b)	6(c)	6(d)	6(e)	6(f)
n sorts		1	2	2	2	2	3	3	3
$s_1 + \cdots + s_n$		9	1+8	1+8	3+6	4+5	3+3+3	1+4+4	1+3+5
z		0	1	1	0	0	0	1	1
t treatments		4	2	8	9	6	12	12	6
W_0		1	1	1	1	1	1	1	1
W_1	V_1	3	–	6	6	3	6	7	1
	V_2	–	–	–	1	1	3	2	2
	residual	24	27	21	20	23	18	18	24
	total	27	27	27	27	27	27	27	27
W_2	V_3	–	1	1	1	1	2	2	2
	residual	8	7	7	7	7	6	6	6
	total	8	8	8	8	8	8	8	8

If i, j, k and ℓ are all different then $\mathbf{w}_{ij}+\mathbf{w}_{k\ell}-\mathbf{w}_{ik}-\mathbf{w}_{j\ell} \in W_1$. The argument is similar to that for $\mathbf{v}_{ij}+\mathbf{v}_{k\ell}-\mathbf{v}_{ik}-\mathbf{v}_{j\ell}$ in Section 2, but slightly more complicated, as follows. If u is an individual whose sort is none of i, j, k or ℓ then the coefficient of $\mathbf{v}_u$ throughout $\mathcal{S}_{ij}$, $\mathcal{S}_{k\ell}$, $\mathcal{S}_{ik}$ and $\mathcal{S}_{j\ell}$ is zero. If u is an individual of sort i then $\mathbf{v}_u \cdot \mathbf{w}_{ij} = s_j/s_is_j$, $\mathbf{v}_u \cdot \mathbf{w}_{ik} = s_k/s_is_k$, and $\mathbf{v}_u \cdot \mathbf{w}_{k\ell} = \mathbf{v}_u \cdot \mathbf{w}_{j\ell} = 0$. The dimension of the subspace of V_2 which is spanned by these vectors is equal to $n(n-3)/2$ if $n \geq 3$ (the calculation is similar to the calculation of d_1 in Section 2), and zero if $n = 2$.

If $s_i \geq 2$ and i, j and k are all different then $\mathbf{w}_{ii}+\mathbf{w}_{jk}-\mathbf{w}_{ij}-\mathbf{w}_{ik} \in W_1$. The argument is as above if the sort of u is not i. If u has sort i then

$$\mathbf{v}_u \cdot (\mathbf{w}_{ii}+\mathbf{w}_{jk}-\mathbf{w}_{ij}-\mathbf{w}_{ik}) = (s_i-1)\frac{2}{s_i(s_i-1)} + 0 - s_j\frac{1}{s_is_j} - s_k\frac{1}{s_is_k},$$

which is zero. These contrasts, for different values of i with $s_i \geq 2$, are all linearly independent of each other, and of those given immediately before, so, if $n \geq 3$, then the dimension of V_2 is at least

$$\frac{n(n-3)}{2} + (n-z) = \frac{n(n-1)}{2} - z.$$

For $i = 1, \ldots, n$, let $\mathbf{w}_i$ be the vector with entry 2 on $\mathcal{S}_{ii}$, entry 1 on $\mathcal{S}_{ij}$ for $j \neq i$, and zeros elsewhere. These vectors are linearly independent of each other, and span the subspace $V_T \cap (W_0 \oplus W_2)$, so this subspace has dimension n. The subspace of treatment contrasts within this has dimension $n-1$. Call this V_3. It could be called the subspace of *contrasts between sorts*.

We now have

$$\dim(V_1 \oplus V_3 \oplus W_0) = t - \frac{n(n+1)}{2} + z + n = t - \frac{n(n-1)}{2} + z,$$

and so $\dim(V_2) \leq n(n-1)/2 - z$. Combining the two inequalities for $\dim(V_2)$ shows that $\dim(V_2) = n(n-1)/2 - z$ and that V_T is equal to the orthogonal direct sum $W_0 \oplus V_1 \oplus V_2 \oplus V_3$. This gives an alternative proof that the design has COBS.

Finally, we deal with the case that $n = 2$. Since $s_1 + s_2 = m \geq 4$, at most one of s_1 and s_2 is equal to 1. Thus the arguments about V_1 and V_3 are still valid, giving $\dim(V_1) = t-3+z$ and $\dim(V_3) = 1$. If $z = 1$ then $\dim(V_1) = t-2$ and $\dim(V_3) = 1$, so $V_1 \oplus V_3 = V_T \cap W_0^\perp$, with no need for V_2.

If $z = 0$ then $s_1 > 1$, $s_2 > 1$ and $\mathbf{w}_{11}+\mathbf{w}_{22}-2\mathbf{w}_{12} \in V_2$. This vector is also in W_1, because if u is an individual of sort 1 then

$$\mathbf{v}_u \cdot (\mathbf{w}_{11}+\mathbf{w}_{22}-2\mathbf{w}_{12}) = (s_1-1)\frac{2}{s_1(s_1-1)} + 0 - 2s_2\frac{1}{s_1s_2} = 0.$$

Hence $\dim(V_1) = t-3$ and $\dim(V_2) = \dim(V_3) = 1$. $\square$

8. Analysis of variance

The statistical properties of a half-diallel design with COBS can be summarized in a skeleton analysis-of-variance table: see Bailey (2008). Each stratum W_i gives a subtable, with one row for each relevant treatment subspace in W_i and another for $W_i \cap V_T^\perp$, which is usually called *residual*. Each row shows the dimension of the relevant subspace, which is equal to the relevant degrees of freedom. If there is more than one subspace then it is helpful to have another row giving the total degrees of freedom for that stratum. If $W_i \cap V_T^\perp$ is non-zero then the mean square for this subspace gives an unbiased estimator for γ_i; otherwise, there is no such estimator.

Table 1 shows these skeleton analysis-of-variance tables for the designs with nine treatments in Figs. 5 and 6.

9. Use of these designs in practice

In Sections 5–7 we have tried to give very general constructions for half-diallel designs with COBS. In real half-diallel experiments, some of these designs may be preferred over others.

The description of the correlations ρ_1 and ρ_2 in Section 3 suggests that $\rho_1 > \rho_2$. The derivation of the eigenvalues γ_0, γ_1 and γ_2 later in Section 3 shows that $\gamma_2 - \gamma_1 = \sigma^2(m-2)(\rho_1 - \rho_2)$, which suggests that $\gamma_2 > \gamma_1$. In this case, designs of Type I are preferred if they are possible. However, the replications may be larger than is desired.

Designs of Type II permit only two treatments. Moreover, they have very unequal replication when $m \geq 6$, so they are probably not a good choice.

Designs of Type III give us a compromise, as it is possible to have smaller replications than in designs of Type I but less unequal replications than in designs of Type II. In practical experiments, just as the blocks in a block design usually all have the same size, it would probably be good to have $s_1 = s_2 = \cdots = s_n = s$; that is, each sort consists of s individuals. For example, in the experiment on methods of remote collaboration described in Section 1, the individuals might be located in n different time-zones, with s individuals in each time-zone. The design in Fig. 6(d) is like this, with $n = 3$ and $s = 3$. When $s = 3$ then all treatments have replication 3. In the design in Fig. 6(d), the treatment contrast between sorts 1 and 2 has the following coefficients

$$\begin{array}{cccccccccccc} A & B & C & D & E & F & G & H & I & J & K & L \\ 2 & 0 & 0 & 0 & -2 & 1 & 1 & 1 & -1 & -1 & -1 & 0 \end{array}$$

and the treatment contrast between sorts 1 and 3 has the following coefficients

$$\begin{array}{cccccccccccc} A & B & C & D & E & F & G & H & I & J & K & L \\ 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -2 \end{array}.$$

These contrasts are both in W_2. All treatment contrasts orthogonal to these are in W_1.

10. Historical remark

We note here the influential thesis of Delsarte (1973), which contains an investigation of subsets in arbitrary association schemes whose characteristic vectors are orthogonal to some strata, or common eigenspaces, of the scheme. In particular, Delsarte defined the notion of Q-polynomial schemes, in which the strata have a natural order W_0, W_1, W_2, …. He obtained a combinatorial characterization of those sets whose characteristic function is orthogonal to $W_1 \oplus \cdots \oplus W_n$ for some n. Although Delsarte was concerned with applications to coding theory rather than statistics, his results can be used to derive our analysis for Type I, although our proof is much more elementary.

Acknowledgements

The authors would like to thank the anonymous referees for their useful comments and suggestions. This work was partially supported by national funds of the Portuguese FCT-Foundation for Science and Technology under grant UIDB/MAT/00212/2020.

References

Addelman, S., 1967. Equal and proportional frequency squares. J. Amer. Statist. Assoc. 62, 226–240.
Bailey, R.A., 1981. A unified approach to design of experiments. J. R. Stat. Soc. Ser. A 144, 214–223.
Bailey, R.A., 1991. Strata for randomized experiments (with discussion). J. R. Stat. Soc. Ser. B Stat. Methodol. 53, 27–78.
Bailey, R.A., 1994. General balance: artificial theory or practical relevance?. In: Caliński, T., Kala, R. (Eds.), Proceedings of the International Conference on Linear Statistical Inference, LINSTAT '93. Kluwer, Dordrecht, Netherlands, pp. 171–184.
Bailey, R.A., 2003. Designs on association schemes. In: Goldstein, D.R. (Ed.), Science and Statistics: A Festschrift for Terry Speed. In: Institute of Mathematical Statistics Lecture Notes – Monograph Series, Vol. 40, IMS, Beachwood, Ohio, pp. 79–102.
Bailey, R.A., 2004. Association Schemes: Designed Experiments, Algebra and Combinatorics. Cambridge University Press.
Bailey, R.A., 2005. Balanced colourings of strongly regular graphs. Discrete Math. 293, 73–90.
Bailey, R.A., 2008. Design of Comparative Experiments. Cambridge University Press, Cambridge.
Bailey, R.A., Brien, C.J., 2016. Randomization-based models for multitiered experiments: I. a chain of randomizations. Ann. Stat. 44, 1131–1164.
Bailey, R.A., Cameron, P.J., Gavriluk, A.L., Goryainov, S.V., 2019. Equitable partitions of Latin-square graphs. J. Comb. Des. 27, 142–160.
Bailey, R.A., Ferreira, S.S., Ferreira, D., Nunes, C., 2016. Estimability of variance components when all model matrices commute. Linear Algebra Appl. 492, 144–160.
Bannai, E., Ito, T., 1984. Algebraic Combinatorics I: Association Schemes. Benjamin/Cummings, Menlo Park, CA.
Bose, R.C., Mesner, D.M., 1959. On linear associative algebras corresponding to association schemes of partially balanced designs. Ann. Math. Stat. 30, 21–38.
Bose, R.C., Nair, K.R., 1939. Partially balanced incomplete block designs. Sankhyā 4, 337–372.
Bose, R.C., Shimamoto, T., 1952. Classification and analysis of partially balanced incomplete block designs with two associate classes. J. Amer. Statist. Assoc. 47, 151–184.
Caliński, T., Kageyama, S., 2000. Block Ddesigns: A Randomization Approach. Vol. I: Analysis. Lecture Notes in Statistics, Vol. 150, Springer-Verlag, New York.
Cameron, P.J., 1976. Parallelisms of Complete Designs. London Mathematical Society Lecture Notes Series, Vol. 23, Cambridge University Press, Cambridge.
Cameron, P.J., 1994. Combinatorics: Topics, Techniques, Algorithms. Cambridge University Press, Cambridge.
Carvalho, F., Mexia, J.T., Santos, C., Nunes, C., 2015. Inference for types and structured families of commutative orthogonal block structures. Metrika 78, 337–372.
Curnow, R.N., 1963. Sampling the diallel cross. Biometrics 19, 287–306.
Delsarte, Ph., 1973. An algebraic approach to the association schemes of coding theory. Philips Research Reports Supplements 10.
Ferreira, S.S., Ferreira, D., Nunes, C., Mexia, J.T., 2010. Crossing segregated models with commutative orthogonal block structure. In: Proceedings of the 8th International Conference on Numerical Analysis and Applied Mathematics, ICNAAM 2010, AIP Conference Proceedings. Vol. 1281, pp. 1229–1232.
Ferreira, S.S., Ferreira, D., Nunes, C., Mexia, J.T., 2013. Estimation of variance components in linear mixed models with commutative orthogonal block structure. Rev. Colombiana Estadíst 36, 261–271.
Finney, D.J., 1945. Some orthogonal properties of the 4 × 4 and 6 × 6 latin squares. Ann. Eugen. 12, 213–219.

Finney, D.J., 1946a. Orthogonal partitions of the 5 × 5 latin squares. Ann. Eugen. 13, 1–3.
Finney, D.J., 1946b. Orthogonal partitions of the 6 × 6 latin squares. Ann. Eugen. 13, 184–196.
Fonseca, M., Mexia, J.T., Zmyślony, R., 2008. Inference in normal models with commutative orthogonal block structure. Acta Comment. Univ. Tartu. Math. 12, 3–16.
Fonseca, M., Mexia, J.T., Zmyślony, R., 2010. Least squares and generalized least squares in models with orthogonal block structure. J. Statist. Plann. Inference 140, 1346–1352.
Freeman, G.H., 1966. Some non-orthogonal partitions of 4 × 4, 5 × 5, and 6 × 6 latin squares. Ann. Math. Stat. 37, 661–681.
Fyfe, J.L., Gilbert, N., 1963. Partial diallel crosses. Biometrics 19, 278–286.
Gauss, C.F., 1821 1823, 1826. Theoria Combinationis Observationum Erroribus Minimus Obnoxiae, Parts 1 2 and Suppl. Werke 4 1–108.
Gavrilyuk, A.L., Goryainov, S.V., 2013. On perfect 2-colorings of Johnson graphs $J(v,3)$. J. Comb. Des. 21, 232–252.
Gavrilyuk, A.L., Metsch, K., 2014. A modular equality for Cameron–Liebler line classes. J. Comb. Theory Ser. A 127, 224–242.
Grundy, P.M., Healy, M.J.R., 1950. Restricted randomization and quasi-Latin squares. J. R. Stat. Soc. Ser. B Stat. Methodol. 12, 286–291.
Hedayat, A., Seiden, E., 1970. F-square and orthogonal F-squares design: A generalization of Latin square and orthogonal Latin squares design. Ann. Math. Stat. 41, 2035–2044.
Houtman, A.M., Speed, T.P., 1983. Balance in designed experiments with orthogonal block structure. Ann. Stat. 11, 1069–1085.
Howes, C., Healey, P.G.T., Mills, G.J., 2009. A: An experimental investigation into … B: … split utterances. In: Proceedings of SIGDIAL 2009: The 10th Annual Meeting of the Special Interest Group in Discourse and Dialogue. pp. 79–86.
Isotalo, J., Puntanen, S., Styan, G.P.H., 2008. The BLUE's covariance matrix revisited: a review. J. Statist. Plann. Inference 138, 2722–2737.
Jones, R.M., 1965. Analysis of variance of the half diallel table. Heredity 20, 117–121.
Kempthorne, O., Curnow, R.N., 1961. The partial diallel cross. Biometrics 17, 229–250.
Kruskal, W., 1968. When are Gauss–Markov and least squares estimators identical? A coordinate-free approach. Ann. Math. Stat. 39, 70–75.
Lucas, É., 1882–1894. Récréations Mathématiques (4 volumes). Gauthier-Villars, Paris.
Macharia, H., Goos, P., 2010. D-optimal and D-efficient equivalent-estimation second-order split-plot designs. J. Qual. Technol. 42, 358–372.
Markewicz, A., Puntanen, S., Styan, G.P.H., 2010. A note on the interpretation of the equality of OLSE and BLUE. Pakistan J. Stat. 26, 127–134.
Nelder, J.A., 1965a. The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. Proc. R. Soc. Lond. Ser. A 283, 147–162.
Nelder, J.A., 1965b. The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. Proc. R. Soc. Lond. Ser. A 283, 163–178.
Nunes, C., Santos, C., Mexia, J.T., 2008. Relevant statistics for models with commutative orthogonal block structure and unbiased estimator for variance components. J. Interdiscip. Math. 11, 553–564.
Özkan, E., Gurion, T., Hough, J., Healy, P.G.T., Jamone, L., 2021. Specific hand motion patterns correlate to miscommunications during dyadic conversations. In: 2021 IEEE International Conference on Development and Learning (ICDL). pp. 1–6.
Parker, P., Kowalski, S.M., Vining, G.G., 2007. Construction of balanced equivalent-estimation second-order split-plot designs. Technometrics 49, 56–65.
Puntanen, S., Styan, G.P., 1989. The equality of the ordinary least squares estimator and the best linear unbiased estimator (with discussion). Amer. Stat. 43, 153–164.
Santos, C., Nunes, C., Dias, C., Mexia, J.T., 2017. Joining models with commutative orthogonal block structure. Linear Algebra Appl. 517, 235–245.
Santos, C., Nunes, C., Dias, C., Mexia, J.T., 2020. Models with commutative orthogonal block structure: a general condition for commutativity. J. Appl. Stat. 47, 2421–2430.
Sprott, D.A., 1978. Gauss's contributions to statistics. Historia Math. 5, 183–203.
Vining, G.G., Kowalski, S.M., Montgomery, D.C., 2005. Response surface designs within a split-plot structure. J. Qual. Technol. 37, 115–129.
Wallis, W.D., 1973. On one-factorizations of complete graphs. J. Austr. Math. Soc. 16, 167–171.
Yates, F., 1947. Analysis of data from all possible reciprocal crosses between a set of parental lines. Heredity 1, 287–301.
Zmyślony, R., 1980. A characterization of best linear unbiased estimators in the general linear model. In: Mathematical Statistics and Probability Theory, Proceedings of the Sixth International Conference, Wisla, 1978. In: Lecture Notes in Statistics, Vol. 2, Springer, New York and Berlin, pp. 365–373.