Noname manuscript No.
(will be inserted by the editor)

Connected row-column L-designs for symmetrical parallel line assays with two preparations

Feng-Shun Chai · Rakhi Singh+ · John Stufken

Received: date / Accepted: date

Abstract For parallel line assays with two preparations, an L-design is an equireplicated design that accommodates the estimation of three specific contrasts of primary interest with full efficiency. We provide necessary and sufficient conditions for the existence of connected L-designs for symmetrical parallel line assays with two preparations that are conducted in a row-column design.

Keywords Bioassays · L-designs · parallel line assays · row-column designs

1 Introduction

Parallel line assays have long been an important tool in bioassay studies (cf. Finney, 1978). There are nonetheless still some intriguing mysteries in this area of research, and we will solve one of these. The parallel line assays with two preparations, one a standard preparation and the other a test preparation, are called *symmetrical* parallel line assays with two preparations if both preparations have the same number of doses. Such assays can be run in a variety of experimental designs, including completely randomized designs, block designs, and row-column designs. We will focus on the use of row-column designs. Finney (1978), page 177, already describes the use of Latin square designs and generalizations in the presence of two blocking systems formed by different

Feng-Shun Chai
Academia Sinica, Taipei, Taiwan

Rakhi Singh
University of North Carolina at Greensboro, Greensboro, NC, USA
E-mail: r_singh5@uncg.edu
+Corresponding author

John Stufken
University of North Carolina at Greensboro, Greensboro, NC, USA

sites and animals. Regulatory bodies also provide guidelines for running these experiments in row-column designs (cf. Council of Europe, 2020, Section 5.3), along with a statistical analysis that treats row and column effects as fixed effects. There are also publications that treat row and column effects as random effects, and others that assume a polynomial trend in row or column effects (cf. Schlain et al., 2001).

There are three specific contrasts that are normally of primary interest in these studies, and it has been proposed to use equireplicated designs in which these three contrasts can be estimated with full efficiency. Designs that accomplish this have been called L-designs (Gupta and Mukerjee, 1990, 1996). Our focus is on the existence of row-column L-designs. However, other contrasts can also be of interest. For example, Finney (1978) extensively discusses the use of balanced incomplete block designs for when all pairwise comparisons are of equal interest. Treatment contrasts within preparations can also be useful to study deviations from the parallel line assumption. For that reason, our focus is on the existence of connected row-column L-designs. More specifically, we will answer the following fundamental question: With $k \geqslant 2$ rows, $b \geqslant 2$ columns and $m \geqslant 2$ doses for each of the two preparations, what are the values of k, b and m for which a connected row-column L-design exists?

The problem of establishing the existence of symmetrical and asymmetrical L-designs has received considerable attention for block designs. References include Das and Kulkarni (1966); Kulshreshtha (1969); Win and Dey (1980); Nigam and Boopathy (1985); Das and Saha (1986); Gupta (1988, 1989); Gupta and Mukerjee (1990); Chai (2002). The latter established a result that, when applied to symmetrical parallel line assays with $m \geqslant 2$ doses for each of two preparations, implies that an L-design in b blocks of size k exists if and only if $k \geqslant 4$ and both k and $k(m+1)/2$ are even. A few methods of constructions of block designs for both symmetric and asymmetric parallel line assays have been studied in Bhar (2016) and Shekhar and Bhar (2016).

Establishing necessary and sufficient conditions for the existence of row-column L-designs is a much more difficult problem. By using results on magic rectangles and nearly magic rectangles, we are however able to obtain such conditions. More precisely, we show that a connected L-design in k rows, b columns and with $m \geqslant 2$ doses for each of the two preparations exists if and only if $k \geqslant 4$, $b \geqslant 4$, k, b, $k(m+1)/2$ and $b(m+1)/2$ are all even, with the exception that the design does not exist for $m = k = b = 4$. The sufficiency will be established by construction of connected L-designs for every m, k and b that satisfy these conditions. Thus, for each such combination of m, k and b, we are satisfied with establishing the existence of a single connected L-design. In general, there could be different designs of this type with different properties, and a user who is interested in a connected L-design with additional properties may not want to use the designs that we construct. Addressing such additional concerns is beyond the scope of this paper.

In Section 2, basic concepts and notation are introduced, followed by a formal definition and characterization of row-column L-designs for symmetrical parallel line assays. The construction of row-column L-designs is explored in Section 3. Section 4 concludes with some final remarks.

2 Symmetrical parallel line assays and row-column L-designs

Let m denote the common number of doses for the standard and test preparations in a bioassay, so that there are a total of $v = 2m$ treatments. Writing $s_1 < \cdots < s_m$ and $t_1 < \cdots < t_m$ to denote these doses, respectively, we write $x_i = \log(s_i),\ z_i = \log(t_i)$, and follow the common assumption that the x_i's and z_i's are equally spaced. The responses under the standard and test preparations can now be modeled using polynomial models in the x_i's and z_i's, respectively. In a line assay, these would be polynomials of degree 1, and in a parallel line assay the slopes of the two lines would be equal. Writing the treatment effects as $\tau_i = \eta_1(x_i),\ \tau_{m+i} = \eta_2(z_i),\ i = 1, \ldots, m$, for polynomials η_1 and η_2, it is easily seen that if η_1 and η_2 are first-order orthogonal polynomials, say $\eta_j(u) = \beta_{j0} + \beta_{j1}\phi_1(u)$, $j = 1, 2$, then the relationship between the β's and τ's is given by

$$\beta_{10} = \frac{1}{m}\sum_{i=1}^{m}\tau_i = \frac{1}{m}1_m'\tau_s, \qquad \beta_{20} = \frac{1}{m}\sum_{i=1}^{m}\tau_{m+i} = \frac{1}{m}1_m'\tau_t,$$
$$\beta_{11} = c_1 e_m'\tau_s, \qquad \beta_{21} = c_1 e_m'\tau_t,$$

where c_1 is an appropriate constant, $\tau = (\tau_1, \ldots, \tau_m, \tau_{m+1}, \ldots, \tau_{2m})' = (\tau_s', \tau_t')'$ is the vector of treatment effects, $e_m = (1, 2, \ldots, m)' - \frac{1}{2}(m+1)1_m$, and 1_m is the $m \times 1$ vector of ones.

Parallelism is studied with the contrast $\psi_1 = \beta_{11} - \beta_{21}$. If the assumption of parallelism is not violated, then interest focuses on estimating the potency, ρ, defined as the ratio of doses with equal effects for the two preparations, i.e., for $x = log(s)$ and $z = log(t)$ with $\eta_1(x) = \eta_2(z)$,

$$\rho = s/t = e^{(x-z)} = c_2 \exp^{\{-(\beta_{10}-\beta_{20})/\beta_1\}},$$

where $\beta_1 = (\beta_{11} + \beta_{21})/2$. This stresses the importance of the contrasts $\psi_2 = \beta_{11} + \beta_{21}$ and $\psi_3 = \beta_{10} - \beta_{20}$, which we would like to estimate with full efficiency. In terms of the τ_i's, these three contrasts correspond, except for a multiplicative constant, to

$$e_m'\tau_s - e_m'\tau_t, \quad e_m'\tau_s + e_m'\tau_t, \quad \text{and } 1_m'\tau_s - 1_m'\tau_t, \tag{1}$$

respectively.

We will consider the usual fixed-effects additive Gauss-Markov model for a row-column design with one observation per cell. Let X_1 be a $kb \times k$ row-incidence matrix, X_2 be a $kb \times b$ column-incidence matrix, and X_3 be the $kb \times v$ treatment-incidence matrix corresponding to a design d. Such a design d is a $k \times b$ matrix with elements from the set $\{1, 2, \ldots, 2m\}$ such that an entry in a cell corresponds to the application of that treatment in the corresponding row and column. We write the model as

$$y = X_1\alpha + X_2\gamma + X_3\tau + \epsilon, \tag{2}$$

where α is the $k \times 1$ vector of row effects, γ is the $b \times 1$ vector of column effects, τ is the $v \times 1$ vector of treatment effects with $v = 2m$ as defined above, and ϵ is the random error vector with mean 0_{kb} and variance $\sigma^2 I_{kb}$. We will write $M = X_3'X_1$ and $N = X_3'X_2$ for the treatment-row and treatment-column incidence matrices. If the design is equireplicated with common treatment replication r, then $1_v'M = b1_k'$, $M1_k = r1_v$, $1_v'N = k1_b'$, and $N1_b = r1_v$. The information matrix for τ may then be written as

$$C = X_3'(I_{kb} - P_{[X_1 \quad X_2]})X_3 = rI_v - \frac{1}{k}NN' - \frac{1}{b}MM' + \frac{r^2}{kb}1_v1_v',$$

where $P_A = A(A'A)^-A'$ denotes the orthogonal projection matrix onto the column space of A.

A design is connected if and only if $\text{rank}(C) = v-1$. An equireplicated row-column design can only be connected if r is at least equal to 2. We will use the notation $\mathcal{D}(v,k,b,r)$ for the class of all equireplicated, connected row-column designs with v treatments, each replicated $r = kb/v$ times, in k rows and b columns. Connected designs facilitate unbiased estimation of every treatment contrast, and the best linear unbiased estimator of $c'\tau$, where $c'1_v = 0$, is given by $c'\hat{\tau} = c'C^-Q$ where C^- is an arbitrary generalized inverse of C and $Q = X_3'(I_v - P_{[X_1 \quad X_2]})y$. It follows that the variance of this estimator is given by $\sigma^2 c'C^-c$. Since $C \leqslant X_3'(I_v - P_{1_{kb}})X_3 = r(I_v - \frac{1}{v}1_v1_v')$, it also follows that for any design in $\mathcal{D}(v,k,b,r)$ we have for any contrast $c'\tau$ that

$$\text{Var}(c'\hat{\tau}) \geqslant \sigma^2 c'c/r. \tag{3}$$

If equality holds in (3) for a contrast $c'\tau$, then we say that the design allows the estimation of this contrast with full efficiency.

Definition 1 An equireplicated design d is called a row-column L-design if it allows the estimation of the three contrasts in (1) with full efficiency under Model (2).

The next lemma will help with the characterization of row-column L-designs.

Lemma 1 *Under Model (2), a design $d \in \mathcal{D}(v,k,b,r)$ allows the estimation of a treatment contrast $c'\tau$ with full efficiency if and only if $c'M = 0_k'$ and $c'N = 0_b'$.*

Proof The contrast $c'\tau$ is estimated with full efficiency if and only if

$$c'X_3'(I_v - P_{[X_1 \quad X_2]})X_3c = c'X_3'(I_v - P_{1_{kb}})X_3c,$$

or equivalently

$$c'X_3'(P_{[X_1 \quad X_2]} - P_{1_{kb}})X_3c = 0. \tag{4}$$

But $1_{kb}'X_3c = r1_v'c = 0$, so that (4) is equivalent to $P_{[X_1 \quad X_2]}X_3c = 0_{kb}$. This in turn is equivalent to $X_1'X_3c = 0_k$ and $X_2'X_3c = 0_b$, or $M'c = 0_k$ and $N'c = 0_b$.

In order to emphasize the dependence of N and M on design $d \in \mathcal{D}(v,k,b,r)$, we will change the notation to N_d and M_d. Moreover, we will partition these matrices into two parts, with one corresponding to the first m treatments (doses for the standard treatment) and the other to the last m treatments: $N_d = (N_{d1}', N_{d2}')'$ and $M_d = (M_{d1}', M_{d2}')'$. Equations (5) and (6) in the following corollary are an immediate consequence of Definition 1 and Lemma 1.

Corollary 1 *A design $d \in \mathcal{D}(v,k,b,r)$ with $m \geqslant 2$ is a row-column L-design if and only if $k \geqslant 4$, $b \geqslant 4$ and*

$$1_m'N_{d1} = \frac{k}{2}1_b'; \quad e_m'N_{d1} = 0_b'; \quad 1_m'M_{d1} = \frac{b}{2}1_k'; \quad e_m'M_{d1} = 0_k', \tag{5}$$

$$1_m'N_{d2} = \frac{k}{2}1_b'; \quad e_m'N_{d2} = 0_b'; \quad 1_m'M_{d2} = \frac{b}{2}1_k'; \quad e_m'M_{d2} = 0_k'. \tag{6}$$

It is clear from (5) and (6) that k and b must be even, but neither can be equal to 2. To see the latter, let, without loss of generality, k be equal to 2. For both $e_m'N_{d1}$ and $e_m'N_{d2}$ to be $0_k'$, it must be that each column has either both treatments from $\{1,2,\ldots,m\}$ or from $\{m+1, m+2, \ldots, 2m\}$. But since $1_m'N_{d1} = 1_m'N_{d2} = \frac{k}{2}1_b'$, each column must have exactly one element from each of $\{1,2,\ldots,m\}$ and $\{m+1, m+2, \ldots, 2m\}$, which is a contradiction.

Corollary 2 *If $d \in \mathcal{D}(2m,k,b,r)$ is a row-column L-design with $m \geqslant 2$, then $k \geqslant 4$, $b \geqslant 4$ and $k, b, k(m+1)/2$, and $b(m+1)/2$ must be all even.*

Proof That $k \geqslant 4$ and $b \geqslant 4$ must be even is obvious from Corollary 1. Moreover, the condition $e_m'N_{d1} = 0_b'$ in Corollary 1 is equivalent to $(1,2,\ldots,m)N_{d1} = (k(m+1)/4)1_b'$, so that $k(m+1)/2$ must be even. Similarly, it follows from $e_m'M_{d1} = 0_k'$ that $b(m+1)/2$ must be even.

In the next section, we will show that the necessary conditions for the existence of a connected equireplicated row-column L-design in Corollary 2 are also sufficient provided that $kb/2m$ is an integer and $r \geqslant 2$, with the single exception that no such design exists for $m = k = b = 4$, as established in the next lemma.

Lemma 2 *A connected equireplicated row-column L-design does not exist for $m = k = b = 4$.*

Proof Suppose the design does exist. From Corollary 1, it follows that (a) every row and every column must have exactly two treatments from $\{1,2,3,4\}$ and, thus, also exactly two from $\{5,6,7,8\}$, and (b) if 1 appears in a row or column, then so must 4. Similarly for 2 and 3, for 5 and 8, and for 6 and 7.

Since each treatment appears twice, after a permutation of rows and columns (a connected row-column L-design is invariant under such operations), we can assume without loss of generality, that the first row and column start as 1, 4. But that means that the second replication of treatment 4 must occur in row 2 and column 2. In the remainder of the first row, we must have 5 and 8 or 6 and 7. Without loss of generality, say, it is 5 and 8. Continuing like this, it is seen that the design must look like

$$\begin{bmatrix} 1 & 4 & 5 & 8 \\ 4 & 1 & 8 & 5 \\ 6 & 7 & 2 & 3 \\ 7 & 6 & 3 & 2 \end{bmatrix}$$

It is an L-design, but the information matrix C has only rank 5. Hence the design is not connected.

3 Construction of row-column L-designs

Lemma 3 shows the main structure of the desired row-column L-design d and Lemma 4 constructs a connected design d^C with each element in the set $\{1, 2, \dots, 2m\}$ occurring twice. Designs constructed in Theorem 1 are connected because they contain d^C as a subdesign. Occasionally we assume that $b \geqslant k$. This can be done without loss of generality, for otherwise a design could be obtained by simply transposing the original design.

From Corollary 2, k and b should both be even, say $k = 2k_1$ and $b = 2b_1$. The basic structure of $k \times b$ row-column designs d that we will construct is

$$d = \left[\begin{array}{c|c} S_1 & T_1 \\ \hline T_2 & S_2 \end{array} \right], \tag{7}$$

where the matrices S_j and T_j, $j = 1, 2$, are each of the order $k_1 \times b_1$; the elements of $S_j, j = 1, 2$, belong to the set $\{1, 2, \dots, m\}$ and the elements of $T_j, j = 1, 2$, belong to the set $\{m + 1, m + 2, \dots, 2m\}$.

Lemma 3 *Let d be a $k \times b$ row-column design with elements from the set $\{1, 2, \dots, 2m\}$ as in (7) and with the properties that*

(i) together S_1 and S_2 contain each treatment from the set $\{1, 2, \dots, m\}$ r times;
(ii) together T_1 and T_2 contain each treatment from the set $\{m+1, m+2, \dots, 2m\}$ r times;
(iii) each of the S_j and T_j, $j = 1, 2$, are constructed by juxtaposing multiple submatrices so that if one of these submatrices of S_j is of size $p_g \times q_g$, it has constant row sums equal to $q_g(m + 1)/2$ and constant column sums equal to $p_g(m + 1)/2$ and a $p_g \times q_g$ submatrix of T_j has constant row sums equal to $q_g(3m + 1)/2$ and constant column sums equal to $p_g(3m + 1)/2$.

Then, d is an equireplicated row-column L-design.

Let $\Pi_R(H)$ denote a matrix obtained by cyclically shifting each row of H to the next row and $\Pi_C(H)$ denote a matrix obtained by cyclically shifting each column of H to the next column. Also, let $\Pi_{CR}(H) = \Pi_C(\Pi_R(H))$.

Lemma 4 *Let $m = m_1 m_2$. Let S_1^C be a $m_1 \times m_2$ matrix containing each element from the set $\{1, 2, \ldots, m\}$ exactly once, $S_2^C = \Pi_{CR}(S_1^C)$, and $T_1^C = m1_{m_1}1'_{m_2} + S_1^C$. Then, $d^C = \begin{bmatrix} S_1^C & T_1^C \\ T_1^C & S_2^C \end{bmatrix} \in \mathcal{D}(v = 2m, k = 2m_1, b = 2m_2, 2)$ is a connected equireplicated row-column design.*

Proofs of Lemmas 3 and 4 are provided in the appendix. Designs d^C are used in the construction of connected L-designs for general values of m, b and k. It turns out that to construct a connected L-design for $m = 8$, $k = 4u$, and $b = 4l, l \geqslant 3,\ \ u \geqslant 1$, we need a different design to play the role of d^C. This design, also named d^C, is presented in the following remark.

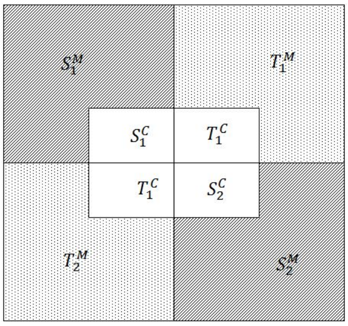
Remark 1 For $m = 8, k = 4, b = 12$, define design d^C as in (8). It is not too hard to see that d^C satisfies Lemma 3 and that it is connected. Therefore, d^C is a connected equireplicated row-column L-design in $\mathcal{D}(16, 4, 12, 3)$.

$$d^C = \left[\begin{array}{cccccc|cccccc} 1 & 5 & 6 & 7 & 3 & 5 & 9 & 13 & 14 & 15 & 11 & 13 \\ 8 & 4 & 3 & 2 & 6 & 4 & 16 & 12 & 11 & 10 & 14 & 12 \\ \hline 16 & 16 & 12 & 11 & 10 & 10 & 8 & 8 & 4 & 3 & 2 & 2 \\ 9 & 9 & 13 & 14 & 15 & 15 & 1 & 1 & 5 & 6 & 7 & 7 \end{array}\right] \tag{8}$$

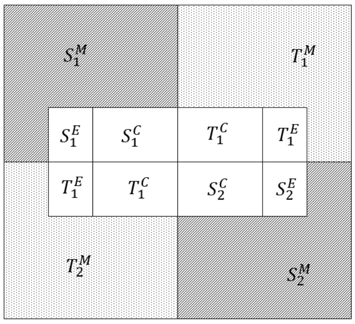
We now state the sufficiency result as already announced at the end of Section 2.

Theorem 1 *Sufficient conditions for the existence of a connected equireplicated row-column L-design with $m \geqslant 2$ are that each of k, b, $k(m+1)/2$ and $b(m+1)/2$ are even provided that $r = kb/2m$ ($\geqslant 2$) is an integer, $k \geqslant 4$ and $b \geqslant 4$, except that such a design does not exist for $m = k = b = 4$.*

In what follows, we first present a general structure for construction of a row-column L-design $d \in \mathcal{D}(v = 2m, k, b, r)$ and then construct the individual components. The construction makes use of the two types of designs shown in Figure 1, both of which follow the basic structure in (7). The difference is that, depending on the values of m, k, and b, in some cases we are able to construct each of the parts S_j and T_j , $j = 1, 2$, by juxtaposing equireplicated submatrices as in Lemma 3(iii), in which case we use designs of type I. In other cases where this is not possible, the construction becomes a bit more tricky and we need to revert to type II designs. Thus, in the first cases, we have stronger properties than those in Lemma 3(i) and (ii) because each S_j and T_j are equireplicated on their respective sets of treatments. The second cases are trickier because the individual quadrants in (7) are not equireplicated. In those cases we need to make use of the submatrices S_j^E and T_j^E in designs of type II. Take for example the case $m = 16$, $k = 8$, and $b = 12$. The array S_1 is then of size 4×6. Since S_1^C must contain each of $1, 2, \ldots, m$ once, it means



(a) Design type I



(b) Design type II

Fig. 1: Structure of design d. The parts with the superscript C represent the connected component d^C as in Lemma 4. Other parts need to be completed as described in the proof of Theorem 1.

that S_1^M in designs of type I cannot be equireplicated on $1, 2, \ldots, m$. But we show how to construct the corresponding design by using type II designs.

Before we show the constructions, we note the following properties about the individual components of Figure 1. If we can construct a design d in Figure 1 satisfying (a)-(c) along with satisfying Lemma 3, then d is a desired connected row-column L-design. Our proof then follows by showing the method by which these components can be constructed.

(a) The elements of S_1 and S_2 (NW and SE quarters in Figure 1) together contain r replications of each treatment from the set $\{1, 2, \ldots, m\}$, and the elements of T_1 and T_2 (NE and SW quarters in Figure 1) together contain r replications of each treatment from the set $\{m+1, m+2, \ldots, 2m\}$;
(b) T_1^E, T_1^C, T_1^M, and T_2^M are obtained from S_1^E, S_1^C, S_1^M, and S_2^M, respectively, by replacing i with $m+i$, $i = 1, \ldots, m$.
(c) The middle component of d, that is $\begin{bmatrix} S_1^C & T_1^C \\ T_1^C & S_2^C \end{bmatrix}$, corresponds to d^C in Lemma 4 or in Remark 1. The adopted construction method ensures that d^C has each treatment from the set $\{1, 2, \ldots, 2m\}$ equally often and is a connected subdesign of d.

Proof (Proof of Theorem 1) It is sufficient to show the construction of S_j, $j = 1, 2$, since T_j are obtained from S_j. We will distinguish between two cases based on whether m is even or odd.

In what follows, $P_{a\times b}$ denotes a magic rectangle of order $a \times b$ (Kudrle and Menard, 2007) and $Q_{a\times b}$ a nearly magic rectangle of order $a \times b$ (Chai et al., 2019). The notation $[i : j]$ denotes a row vector $[i\ i+1\ i+2\ \ldots\ j]$ when $i \leqslant j$ and $[i\ i-1\ i-2\ \ldots\ j]$ when $i > j$. Recall that $S_2^C = \Pi_{CR}(S_1^C)$, as in Lemma 4, for both design types I and II.

Case (A): m is even. When m is even, from Corollary 2, both k_1 and b_1 must be even. For the bulleted items $S_1^M \cup S_2^M$ in what follows, x copies indicates that the dark hatched part of Figure 1 (corresponding to S_1^M and S_2^M) jointly contains x copies of each of the matrices $H_1, \ldots, H_{m/2}$, where H_i's are 2×2 matrices with diagonal values equal to i and off-diagonal values $m - i + 1$. We further segment the cases as below:

Case (A1): $m = 4m_0 + 2 = 2(2m_0 + 1)$. For $r = kb/2m = 4\frac{(k_1/2)(b_1/2)}{(2m_0+1)}$ to be an integer, $2m_0+1$ must divide $(k_1/2)(b_1/2)$, and, so, there are odd integers m_1 and m_2 so that $2m_0 + 1 = m_1 m_2$ and $k_1 = 2um_1$, $b_1 = 2lm_2$.
When $m = 2$, $m_1 = m_2 = 1$, so that $k_1 = 2u$ and $b_1 = 2l$, we use

- Design type: II,
- $S_1^C = [1\ 2]'$, $S_1^E = [2\ 1]'$ and $S_2^E = [1\ 2]'$, and
- $S_1^M \cup S_2^M$: $2(ul - 1)$ copies.

For $m > 2$, if $m_1 = 1$ or $m_2 = 1$, without loss of generality let $m_2 = 1$. Then $m_1 = 2m_0 + 1$, $k_1 = 2um_1$, $b_1 = 2l$, and we use

- Design type: II,
- $S_1^C = [1 : m]'$, $S_1^E = [m : 1]'$ and $S_2^E = \Pi_R([m : 1]')$, and
- $S_1^M \cup S_2^M$: $2(ul - 1)$ copies.

For $m > 2$, if both $m_1 > 1$ and $m_2 > 1$, so that $k_1 = 2um_1$ and $b_1 = 2lm_2$, we use

- Design type: II,
- $S_1^C = Q_{2m_1 \times m_2}$, $S_1^E = \Pi_R(S_1^C)$ and $S_2^E = \Pi_R(\Pi_{CR}(S_1^C))$, and
- $S_1^M \cup S_2^M$: $2(ul - 1)$ copies.

To satisfy Lemma 3 (iii), the rows of $Q_{2m_1 \times m_2}$ should be arranged such that their sums alternate, that is, in the first row the elements sum up to s, in the second to $s + 1$, in the third to s, and so on, where $s = (m_2(m + 1) - 1)/2$. This can always be done because exactly half of the row sums are s, the other half being $s + 1$. By doing this, and by combining S_1^C and S_1^E according to the design type II, we get the constant row and column sums.

Case (A2): $m = 4m_0$, m_0 is odd, and r is even. For $r = kb/2m = 2\frac{(k_1/2)(b_1/2)}{m_0}$ to be an integer, m_0 must divide $(k_1/2)(b_1/2)$, and, so, there must be odd integers m_1 and m_2 so that $m_0 = m_1 m_2$, $k_1 = 2um_1$, $b_1 = 2lm_2$.
When $m = 4$, then $m_1 = m_2 = 1$, $k_1 = 2u$ and $b_1 = 2l$. Since $l = 1$ and $u = 1$ is not possible because the design does not exist for $m = k = b = 4$ (see Lemma 2), we take, without loss of generality, $u > 1$ and $l \geqslant 1$. We use

- Design type: II,
- $S_1^C = [1 : 4]'$, $S_1^E = [4 : 1]'$ and $S_2^E = \Pi_R([4 : 1]')$, and
- $S_1^M \cup S_2^M$: $ul - 2$ copies.

When $m > 4$, for $k_1 = 2um_1$ and $b_1 = 2lm_2$, we use

- Design type: I,
- $S_1^C = P_{2m_1 \times 2m_2}$, and
- $S_1^M \cup S_2^M$: $ul - 1$ copies.

Case (A3): $m = 8m_0$. For $m = 2ab$, where a and b are even integers not simultaneously equal to 2, we define matrices $F_{a\times b}$ and $G_{a\times b}$ of order $a \times b$ such that each of the row sums are equal to $b(1+m)/2$, each of the column sums are equal to $a(1+m)/2$, and elements of F and G together contain treatments from the set $\{1,\dots,m\}$ exactly once. One way of constructing F and G, through a decomposition of magic rectangles, is given in Lemma A1 in the appendix. For $r = kb/2m = \frac{(k_1/2)(b_1/2)}{m_0}$ to be an integer, m_0 must divide $(k_1/2)(b_1/2)$, and, so, there are two integers m_1 and m_2, so that $m_0 = m_1m_2$, $k_1 = 2um_1$, $b_1 = 2lm_2$. Then, $r = ul$. We first discuss the case when r is odd, implying both l and u are odd. If both $l = u = 1$, then $r = 1$, but we have that $r \geqslant 2$. So, without loss of generality we take $l > 1$.

When $m = 8$, and r is odd, for $k_1 = 2u$ and $b_1 = 2l, l \geqslant 3$, we use

- Design type: I,
- $d^C = \begin{bmatrix} S_1^C & T_1^C \\ T_1^C & S_2^C \end{bmatrix}$ is from Remark 1, and
- $S_1^M \cup S_2^M$: $(ul-3)/2$ copies.

When $m > 8$, and r is odd, for $k_1 = 2um_1$ and $b_1 = 2lm_2, l \geqslant 3$, we use

- Design type: II,
- $S_1^C = P_{2m_1\times 4m_2}$, $S_1^E = F_{2m_1\times 2m_2}$ and $S_2^E = G_{2m_1\times 2m_2}$, and
- $S_1^M \cup S_2^M$: $(ul-3)/2$ copies.

When $m = 8m_0 = 8m_1m_2$ and r is even, for $k_1 = 2um_1$ and $b_1 = 2lm_2$, without loss of generality, let l be even. We use

- Design type: I,
- $S_1^C = P_{2m_1\times 4m_2}$, and
- $S_1^M \cup S_2^M$: $(ul-2)/2$ copies.

To complete this proof, we need to show that each of the $p_g \times q_g$ decomposed submatrices of S_j and T_j, $j = 1,2$, have row sums equal to $q_g(m+1)/2$ and column sums equal to $p_g(m+1)/2$. The submatrices $H_1,\dots,H_{m/2}$ are 2×2 matrices with the row and column sums equal to $m+1$. It is also not too hard to see that S_1^C and S_1^E taken together according to design type have row sums equal to $p_g(m+1)/2$ and column sums equal to $q_g(m+1)/2$ and so is the case for S_2^C and S_2^E taken together.

Case (B): m is odd. For this case, $r = kb/v = (2k_1)(2b_1)/2m = 2k_1b_1/m$, and hence, r must be even. For odd p and q, $p \geqslant 3$, $q \geqslant 3$, let $A_{p\times q} = [a_1\ a_2\ \dots\ a_p]'$ denotes a $p\times q$ matrix constructed as in Lemma A2, and a_i' is the ith row of the matrix A, $i = 1,\dots,p$. For the bulleted items $S_1^M \cup S_2^M$ in what follows, $x \circ Z$ indicates that the dark hatched part of Figure 1 (corresponding to S_1 and S_2) jointly contains x copies of the matrix Z.

Since m is odd, k_1b_1 must be a multiple of m, so that there are odd integers m_1 and m_2 such that $m = m_1m_2$, $k_1 = um_1$ and $b_1 = lm_2$.

Case (B1): When $m_1 > 1$ and $m_2 > 1$, for $k_1 = um_1$ and $b_1 = lm_2$, we use

- Design type: I,

- $S_1^C = P_{m_1 \times m_2}$ and
- $S_1^M \cup S_2^M$: $2(ul-1) \circ P_{m_1 \times m_2}$.

Case (B2): When one of m_1 and m_2 is 1, take $m_2 = 1$ without loss of generality. Then $k_1 = um_1 = um$ and $b_1 = l$ with $l \geqslant 2$.
When $b_1 = l$ is odd, we use

- Design type: II,
- $S_1^C = a_l$, $S_1^E = [a_1\ a_2\ \ldots\ a_{l-1}]$, and $S_2^E = \Pi_R(S_1^E)$, and
- $S_1^M \cup S_2^M$: $2(u-1) \circ A'_{l \times m}$.

When $b_1 = l$ is even but not a power of 2, that is, $b_1 = l = 2^w l_1$, where $l_1 > 1$ is odd, we use

- Design type: II,
- $S_1^C = a_{l_1}$, $S_1^E = [a_1\ a_2\ \ldots\ a_{l_1-1}]$, and $S_2^E = \Pi_R(S_1^E)$, and
- $S_1^M \cup S_2^M$: $2(u2^w-1) \circ A'_{l_1 \times m}$.

When $b_1 = l$ is a power of 2, that is, $b_1 = l = 2^w$, we use

- Design type: II,
- $S_1^C = [1:m]'$, $S_1^E = [m:1]'$, and $S_2^E = [1\ [m:2]]'$, and
- $S_1^M \cup S_2^M$: $(u2^w - 2) \circ \begin{bmatrix} [m:1] \\ [1:m] \end{bmatrix}'$.

It is again easy to verify that the constructed designs satisfy the conditions in Lemma 3 and properties (a)–(c) presented after Theorem 1. This completes the proof.

To illustrate the construction, we now provide some examples. For design type I, we provide examples from Case (B1) in the proof of Theorem 1.

Example 1 Consider $v = 18, k = b = 6$. This implies that $m = 9, m_1 = m_2 = 3$ and $u = l = 1$. With $P_{3\times 3} = \begin{bmatrix} 7 & 5 & 3 \\ 2 & 9 & 4 \\ 6 & 1 & 8 \end{bmatrix}$, an equireplicated row-column L-design d is

$$\left[\begin{array}{c|c} S_1^C & T_1^C \\ \hline T_1^C & S_2^C = \Pi_{CR}(S_1^C) \end{array}\right] = \left[\begin{array}{ccc|ccc} 7 & 5 & 3 & 16 & 14 & 12 \\ 2 & 9 & 4 & 11 & 18 & 13 \\ 6 & 1 & 8 & 15 & 10 & 17 \\ \hline 16 & 14 & 12 & 8 & 6 & 1 \\ 11 & 18 & 13 & 3 & 7 & 5 \\ 15 & 10 & 17 & 4 & 2 & 9 \end{array}\right].$$

Similarly, for $v = 18, k = 6$ and $b = 12$, a design is

$$\left[\begin{array}{c|c|c|c} S_1^M & S_1^C & T_1^C & T_1^M \\ \hline T_2^M & T_1^C & S_2^C & S_2^M \end{array}\right] = \left[\begin{array}{ccc|ccc|ccc|ccc} 7 & 5 & 3 & 7 & 5 & 3 & 16 & 14 & 12 & 16 & 14 & 12 \\ 2 & 9 & 4 & 2 & 9 & 4 & 11 & 18 & 13 & 11 & 18 & 13 \\ 6 & 1 & 8 & 6 & 1 & 8 & 15 & 10 & 17 & 15 & 10 & 17 \\ \hline 16 & 14 & 12 & 16 & 14 & 12 & 8 & 6 & 1 & 7 & 5 & 3 \\ 11 & 18 & 13 & 11 & 18 & 13 & 3 & 7 & 5 & 2 & 9 & 4 \\ 15 & 10 & 17 & 15 & 10 & 17 & 4 & 2 & 9 & 6 & 1 & 8 \end{array}\right].$$

For design type II, we provide an example from Case (A1) in the proof of Theorem 1.

Example 2 Consider $v = 4, k = 8$ and $b = 8$. This implies that $m = 2$, $u = 2$ and $l = 2$. An equireplicated row-column L-design d is

$$\left[\begin{array}{cccc|cccc} S_1^M & & & & & & & T_1^M \\ & S_1^E & S_1^C & & & T_1^C & T_1^E & \\ \hline & T_1^E & T_1^C & & & S_2^C & S_2^E & \\ T_2^M & & & & & & & S_2^M \end{array}\right] = \left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 2 & 3 & 4 & 3 & 4 \\ 2 & 1 & 2 & 1 & 4 & 3 & 4 & 3 \\ 1 & 2 & 2 & 1 & 3 & 4 & 3 & 4 \\ 2 & 1 & 1 & 2 & 4 & 3 & 4 & 3 \\ \hline 3 & 4 & 4 & 3 & 2 & 1 & 1 & 2 \\ 4 & 3 & 3 & 4 & 1 & 2 & 2 & 1 \\ 3 & 4 & 3 & 4 & 1 & 2 & 1 & 2 \\ 4 & 3 & 4 & 3 & 2 & 1 & 2 & 1 \end{array}\right].$$

4 Conclusions

The necessary conditions for the existence of a connected row-column L-design for symmetrical parallel line assay are provided. These conditions are also proved sufficient by providing a proof via constructing a design for each parameter set, except that the design for $v = 8$, $k = b = 4$ and $r = 2$ does not exist.

The L-design property assures that the three treatment contrasts of primary interest can be estimated with full efficiency. The connectedness of the designs guarantees that any other treatment contrast can be estimated unbiasedly. If one is interested in finding an equireplicated row-column L-design that, in some sense, is optimal for estimation of other treatment contrasts, one could search for the best row-column L-design in the class $\mathcal{D}(v, k, b, r)$ using a criterion of most interest. That is a problem that is beyond the scope of this paper.

Acknowledgements Part of this work was completed when both RS and JS were visiting Academia Sinica and we gratefully acknowledge hospitality and support from Academia Sinica. The research of JS is partially supported by NSF grant DMS–1935729.

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix

Lemma A1 *Let* $m = 2pq$, *where* p *and* q *are even integers that are not simultaneously equal to 2. Then there exist two matrices of order* $p \times q$, *say* $F_{p\times q}$ *and* $G_{p\times q}$, *such that they together contain numbers* $1, 2, \ldots, m$ *exactly once* *and each of them has row sums equal to* $q(1+m)/2$ *and column sums equal to* $p(1+m)/2$. *When combined, to form a* $2p \times q$ *matrix, the result is a magic rectangle on* $\{1, 2, \ldots, m\}$.

Proof Let $p = 2p_1$ and $q = 2q_1$. From the literature on magic rectangles (Chai et al., 2019), it is not hard to construct a $2 \times q$ magic rectangle $V_1 = P_{2\times q}$ based on $\{1, 2, \ldots, 2q\}$. Define $V_z = V_1 + (z-1)2q1_2'1_q$, $z = 2, \ldots, p$. For $z = 1, \ldots, p_1$, exchange the first q_1 columns of V_z with the first q_1 columns of V_{p+1-z} and then, using the resulting matrices, exchange the second row of the zth matrix with the second row of the $(p+1-z)$th matrix. Call the matrices obtained in this way R_z and R_{p+1-z}. Note that the row and column sums of each of these matrices are equal to $q(1+m)/2$ and $1+m$, respectively. Then, $F_{p\times q} = [R_1', \ldots, R_{p_1}']'$ and $G_{p\times q} = [R_{p_1+1}', \ldots, R_p']'$. Also, $P_{2p\times q} = [R_1' \ldots R_p']'$ is a magic rectangle based on $\{1, 2, \ldots, m\}$.

Lemma A2 (Lemma 2.2 of Chai et al. (2019)) *Let* $p \geqslant 3$ *and* $q \geqslant 3$ *be odd. Then, there exists a* $p \times q$ *matrix* A *with equal column sums such that each row is a permutation of* $[1:q]$.

Proof of Lemma 3. That d is equireplicated is satisfied because of (i) and (ii) in the statement of the lemma. Additionally, the necessary and sufficient conditions for a design $d \in \mathcal{D}(v, k, b, r)$ to be a row-column L-design are provided in Corollary 1. Conditions $1_m'N_{d1} = \frac{k}{2}1_b'$ and $1_m'M_{d1} = \frac{b}{2}1_k'$ of Corollary 1 translate into each of the columns and rows having $k/2$ and $b/2$ treatments from $\{1, \ldots, m\}$, respectively. This is true because of the structure in Figure 1 and (i) and (ii) in the statement of the lemma. For the same reasons, conditions $1_m'N_{d2} = \frac{k}{2}1_b'$ and $1_m'M_{d2} = \frac{b}{2}1_k'$ are true.

Let M_{d1j}, $j = 1, 2$, be the $m \times k_1$ treatment-row incidence matrix corresponding to a design with the same treatments as in the rows of S_j, and no treatments in the other rows. Similarly, let M_{d2j}, $j = 1, 2$, be the $m \times k_1$ treatment-row incidence matrix corresponding to a design with the same treatments as in the rows of T_j, and no treatments in the other rows. Then, for a design d, the treatment-row incidence matrix is

$$M_d = \begin{bmatrix} M_{d11} & M_{d12} \\ M_{d21} & M_{d22} \end{bmatrix}.$$

From (iii) in the statement of the lemma, each of the S_j and T_j, $j = 1, 2$, can be decomposed into $p_g \times q_g$ submatrices S_j^g and T_j^g. Let M_{d1j}^g, $j = 1, 2$, be the $m \times k_1$ treatment-row incidence matrix where rows correspond to the m treatments in S_j, and columns correspond to the k_1 rows of S_j. For rows that do not intersect with S_j^g, the entire column in M_{d1j}^g is 0, and for treatments that do not appear in S_j^g, the entire row in M_{d1j}^g is 0. Define M_{d2j}^g similarly using T_j^g. Then, $M_{d1j} = \sum_g M_{d1j}^g$, and $M_{d2j} = \sum_g M_{d2j}^g$, $j = 1, 2$.

If we can show that $e_m'M_{d1j}^g = 0'$ for all g, then from the above equations, $\sum_g e_m'M_{d1j}^g = e_m'M_{d1j} = 0'$, implying that, in the notation of Corollary 1, $e_m'M_{d1} = 0'$. Arguments that the other conditions in Corollary 1, $e_m'N_{d1} = 0$,

$e'_m M_{d2} = 0$, and $e'_m N_{d2} = 0$, hold would be similar. This would imply the desired result.

So, $e'_m M^g_{d1j} = 0'$ if and only if $\sum_{w=1}^m w m_{wh} = ((m+1)/2)(\sum_{w=1}^m m_{wh})$, $h = 1, ..., k_1$, where m_{wh} is the (w, h)th element of M^g_{d1j}. Additionally, $\sum_{w=1}^m w m_{wh} = \sum_{\ell=1+(j-1)b_1}^{jb_1} c_{h\ell}$, where $c_{h\ell}$ is the (h, ℓ)th element of S^g_j and is defined as 0 for columns ℓ that do not intersect with S^g_j. Note that h is now restricted to those rows that intersect with S^g_j and $\sum_{w=1}^m m_{wh} = q_g$. Therefore, $e'_m M^g_{d1j} = 0'$ if and only if $\sum_{\ell=1+(j-1)b_1}^{jb_1} c_{h\ell} = q_g(m + 1)/2$. □

Proof of Lemma 4. Let $m = m_1 m_2$. The proof is divided in two cases:
Case (i): $m_1 = 1$ *or* $m_2 = 1$. Without loss of generality, let $m_1 = 1$, $S^C_1 = [1 : m]$ and $S^C_2 = \Pi_{CR}(S^C_1) = \Pi_C(S^C_1) = [m \;\; 1 : (m - 1)]$. Then d^C is (a) a $2 \times 2m$ row-column design with each treatment from the set $\{1, \ldots, 2m\}$ replicated exactly twice, (b) row-treatment and column-treatment connected, and (c) $N'_{d^C} M_{d^C} = 2 1_{2m} 1'_2$. Hence, from Theorem 2.1 of Raghavarao and Federer (1975), $d^C \in \mathcal{D}(2m, 2, 2m, 2)$ is a connected row-column design.

Case (ii): $m_1 > 1$ *and* $m_2 > 1$. Without loss of generality, let the qth column of S^C_1 be $c^{S_1}_q = (q - 1)m_1 1_{m_1} + [1 : m_1]'$, $q = 1, \ldots, m_2$. Then, since $T^C_1 = m 1_{m_1} 1'_{m_2} + S^C_1$ and $S^C_2 = \Pi_{CR}(S^C_1)$, the qth column of T^C_1 is $c^{T_1}_q = m 1_{m_1} + c^{S_1}_q$ and the qth column of S^C_2 is $c^{S_2}_q$ such that for $q = 1$, $c^{S_2}_1 = [m_2 m_1, ((m_2 - 1)m_1 + 1) : ((m_2 - 1)m_1 + m_1 - 1)]'$ and for $q = 2, \ldots, m_2$, $c^{S_2}_q = [(q-2)m_1 + m_1, ((q-2)m_1+1) : ((q-2)m_1+m_1-1)]'$. Then, d^C is a $2m_1 \times 2m_2$ row-column design with each treatment from the set $\{1, \ldots, 2m\}$ replicated twice.

Let $y_{p,q}$ be the response variable for the $(p, q)th$ cell of d^C, $p = 1, \ldots, 2m_1$, and $q = 1, \ldots, 2m_2$. Consider the first and the pth response from each of the 4 columns $c^{S_1}_1$, $c^{T_1}_{m_2}$, $c^{T_1}_{m_2}$, and $c^{S_2}_2$. Under model (2), we obtain the following expected value identity:

$$E(y_{1,1} - y_{p,1} - y_{1,2m_2} + y_{p,2m_2} + y_{(m_1+1),m_2} - y_{(m_1+p),m_2} \\ -y_{(m_1+1),(m_2+2)} + y_{(m_1+p),(m_2+2)}) = \tau_1 - \tau_p - \tau_{m_1} + \tau_{(p-1)}.$$

Summing these expectations for $p = 2, \ldots, m_1$, we get

$$\sum_{p=2}^{m_1} E(y_{1,1} - y_{p,1} - y_{1,2m_2} + y_{p,2m_2} + y_{(m_1+1),m_2} - y_{(m_1+p),m_2} - \\ y_{(m_1+1),(m_2+2)} + y_{(m_1+p),(m_2+2)}) = m_1(\tau_1 - \tau_{m_1}),$$

implying that $\tau_1 - \tau_{m_1}$ is estimable, and hence, treatment 1 is connected to treatment m_1. Substituting this back into the above equations, we get that treatment 1 is connected to treatment p, $p = 2, \ldots, m_1 - 1$. Hence, $\{1, 2, \ldots, m_1\}$ is a connected treatment group.

Similarly, considering the first and the pth observation from each of the 4 columns $c^{S_1}_1$, $c^{T_1}_{m_2}$, $c^{T_1}_{m_2}$, and $c^{S_2}_{\ell+1}$ $\ell = 2, \ldots, m_2$, ($\ell + 1 = 1$ when $\ell = m_2$) and applying the above procedure, we get that $\{(\ell - 1)m_1 + 1, (\ell - 1)m_1 + 2, \ldots, (\ell - 1)m_1 + m_1\}$ is a connected treatment group. Similarly, it can be shown that $\{m+(\ell-1)m_1+1, m+(\ell-1)m_1+2, \ldots, m+(\ell-1)m_1+m_1\}$ is a connected treatment group, $\ell = 1, \ldots, m_2$.

By the connectedness checking algorithm in Park and Shah (1995), all treatments in a connected group can be replaced by any one of them to reduce the number of treatments in a design. We use the first treatment from each group as a representative of the group. Then, checking connectedness of d^C is equivalent to checking the connectedness of the following design with $2m_2$ treatments

$$\begin{bmatrix} 1 & m_1+1 & \cdots & m-m_1+1 & m+1 & m+m_1+1 & \cdots & 2m-m_1+1 \\ m+1 & m+m_1+1 & \cdots & 2m-m_1+1 & m_1+1 & 2m_1+1 & \cdots & 1 \end{bmatrix}.$$

Just as in Case (i), the above design is a connected row-column design. This means that treatment representatives $1, m_1+1, 2m_1+1, \ldots, m-m_1+1, m+1, m+m_1+1, \ldots, 2m-m_1+1$ are all connected and hence, d^C is a connected row-column design. □

References

Bhar, L. (2016). Efficient block designs for symmetric parallel line assays. *Cogent Mathematics*, 3:1238181.

Chai, F.-S. (2002). Block designs for asymmetric parallel line assays. *Sankhyā: The Indian Journal of Statistics, Series B*, 64(2):162–178.

Chai, F.-S., Singh, R., and Stufken, J. (2019). Nearly magic rectangles. *Journal of Combinatorial Designs*, 27(9): 562-567.

Council of Europe (2020). European Pharmacopoeia. 10th ed. Strasbourg: Council of Europe.

Das, A. and Saha, G. (1986). Incomplete block designs for asymmetrical parallel line assays. *Calcutta Statistical Association Bulletin*, 35:51–58.

Das, M. and Kulkarni, G. (1966). Incomplete block designs for bio-assays. *Biometrics*, 22(4):706–729.

Finney, D. J. (1978). *Statistical method in biological assay.* Charles Griffin: London, 3rd edition.

Gupta, S. (1988). Designs for symmetrical parallel line assays obtainable through group divisible designs. *Communications in Statistics-Theory and Methods*, 17(11):3865–3868.

Gupta, S. (1989). On block designs for symmetrical parallel line assays with even number of doses. *Journal of Statistical Planning and Inference*, 21(3):383–389.

Gupta, S. and Mukerjee, R. (1990). On incomplete block designs for symmetric parallel line assays. *Australian Journal of Statistics*, 32(3):337–344.

Gupta, S. and Mukerjee, R. (1996). Developments in incomplete block designs for parallel line bioassays. In *Design and analysis of experiments*, volume 13 of *Handbook of Statistics*, pages 875–901. North-Holland, Amsterdam.

Kudrle, J. M. and Menard, S. B. (2007). Magic squares. In *The CRC handbook of combinatorial designs, C.J. Colbourn, J.H. Dinitz, editor,* , pages 524–527. CRC press, Boca Raton, FL.

Kulshreshtha, A. (1969). Modified incomplete block bioassay designs. In *Proc. Indian Sci. Congress, 56th Session*, volume 33.

Nigam, A. and Boopathy, G. (1985). Incomplete block designs for symmetrical parallel line assays. *Journal of Statistical Planning and Inference*, 11(1):111–117.

Park, D. and Shah, K. (1995). On connectedness of row-column designs. *Communications in Statistics-Theory and Methods*, 24(1):87–96.

Raghavarao, D. and Federer, W. (1975). On connectedness in two-way elimination of heterogeneity designs. *The Annals of Statistics*, 3(3):730–735.

Schlain, B., Jethwa, H.S., Subramanyam, M., Moulder, K., Bhatt, B., and Molloy, M. (2001). Designs for bioassays with plate location effects BioPharm International, November 2001, 40–44.

Shekhar, S. and Bhar, L. (2016). Incomplete block designs for asymmetric parallel line assays. *Communications in Statistics—Theory and Methods*, 45: 2413–2425.

Win, K. and Dey, A. (1980). Incomplete block designs for parallel-line assays. *Biometrics*, 33(3):487–492.