Optimal Variance- and Efficiency-Balanced Designs for One- and Two-Way Elimination of Heterogeneity

By A. Das 1 and A. Dey 2

Summary: In this paper, a series of E-optimal non-binary variance balanced (block or row-column) designs and a series of E-optimal non-binary efficiency balanced (block or row-column) designs are provided in certain broad classes of competing designs. Furthermore, their high efficiencies by the usual A- and D-optimality criteria are shown.

Key words and phrases: E-optimality, A-efficiency, D-efficiency, block designs, row-column designs, variance balance, efficiency balance.

1 Introduction

Variance- and efficiency-balanced designs in one-way and two-way elimination of heterogeneity settings have been studied quite extensively in the literature. Though such balanced designs lead to considerable simplicity in the analysis, with the availability of high speed computers, simplicity in analysis alone does not justify the attractiveness of these designs, and, further statistical justification, in terms of optimality considerations, is necessary. In the literature, optimality results on balanced designs are available only for the equireplicate case (see e.g. Kiefer (1958, 1975)), and, not much is known about the optimality of variance- and efficiency-balanced designs when the treatments are not equally replicated, except for a recent paper by Mukerjee and Saha (1990), in which the optimality of efficiency — balanced block designs has been studied in some restricted classes of competing designs with unequal replications and unequal block sizes. The assumption of equal replication often puts a severe restriction on the other parameters (v, b, k)

Dr. Ashish Das, Stat-Math. Division, Indian Statistical Institute, 203 B.T. Road, Calcutta 700035, India.

² Dr. Aloke Dey, Stat-Math. Division, Indian Statistical Institute, 7 S. J. S. Sansanwal Marg, New Delhi 110016, India.

of the design (e.g., the number of blocks in block designs). From a practical point of view in situations where bk/v is not an integer, efficient designs with unequal replications are desirable. This paper attempts to present some efficient non-binary variance- or efficiency-balanced block and row-column designs with unequal replicates. These designs are E-optimal in certain broad classes of competing designs and also have high efficiencies as per the A- and D-optimality criterion.

2 Preliminaries

In the usual setting of block designs, let v denote the number of treatments, b, the number of blocks and k, the number of units per block. Any allocation of v treatments to the bk experimental units is a block design. Under the usual fixed effects, additive model with homoscedasticity and independence, the coefficient matrix of the reduced normal equations for estimating linear functions of treatment effects, using a block design d with parameters v, b, k is given by

$$C_d = R_d - k^{-1} N_d N_d' , (2.1)$$

where $R_d = \text{diag}(r_{d1}, \dots, r_{dv})$, r_{di} is the replication of the *i*th treatment in *d* and $N_d = ((n_{dii}))$ is the $v \times b$ incidence matrix of the design *d*.

The row-column designs considered here have bk experimental units arranged in a rectangular array of b columns and k rows such that each unit receives only one of the v treatments being studied. For an arbitrary row-column design d, the "C-matrix", under an appropriate model is given by

$$C_d^{(RC)} = R_d - k^{-1} N_{1d} N'_{1d} - b^{-1} N_{2d} N'_{2d} + (bk)^{-1} r_d r'_d$$

$$= R_d - k^{-1} N_{1d} N'_{1d} - b^{-1} N_{2d} (I - k^{-1} 11') N'_{2d} , \qquad (2.2)$$

where R_d is as defined earlier, $r_d = (r_{d1}, \ldots, r_{dv})'$, N_{d1} and N_{2d} are the $v \times b$ treatment-column and $v \times k$ treatment-row incidence matrices, respectively, I is an identity matrix (of appropriate order) and 1, a column vector of unities.

identity matrix (of appropriate order) and 1, a column vector of unities. It is known that C_d as in (2.1) and $C_d^{(RC)}$ as in (2.2) are symmetric, nonnegative definite matrices, with zero row sums. A block (resply. row-column) design d is called connected if and only if Rank $(C_d) = v - 1$ (Rank $(C_d^{(RC)}) = v - 1$). Henceforth, only connected designs are considered.

For given positive integers v, b, k, $D_0(v, b, k)$ will denote the class of all connected block designs with v treatments, b blocks and block size k. Similarly,

D(v, b, k) will denote the class of all connected *row-column* designs with v treatments, k rows and b columns.

With each design $d \in D(v, b, k)$ are associated the block designs d_{N_1} and d_{N_2} with incidence matrices N_{1d} and N_{2d} respectively, i.e., $d_{N_1}(d_{N_2})$ is the block design obtained by treating the {columns} ({rows}) of d as blocks. Then, from (2.2), it follows that

$$C_d^{(RC)} = C_d^N - b^{-1} N_{2d} (I - k^{-1} 11') N_{2d}^{\prime} , \qquad (2.3)$$

where

$$C_d^N = R_d - k^{-1} N_{1d} N_{1d}' (2.4)$$

is the C-matrix of d_{N_1} . We denote by $D_N(v, b, k)$, the class of designs d_{N_1} corresponding to $d \in D(v, b, k)$ and consisting of all connected block designs having v treatments, b blocks and block size k.

We now have the following definitions.

Definition 2.1: A connected block (resply. row-column) design d is said to be variance-balanced if and only if it permits the estimation of all normalized treatment contrasts with the same variance.

A connected block (resply, row-column) design d is variance-balanced if and only if

$$C_d(C_d^{(RC)}) = \theta(I - v^{-1}11')$$
, (2.5)

where $\theta(>0)$ is the unique non-zero eigenvalue of $C_d(C_d^{(RC)})$.

Let d denote a connected block (resply, row-column design). The positive eigenvalues of the matrix $R_d^{-1/2}C_dR_d^{-1/2}$ (resply, $R_d^{-1/2}C_d^{(RC)}R_d^{-1/2}$) are called the canonical efficiency-factors of the design d.

Definition 2.2: A connected block (resply. row-column) design d is said to be efficiency-balanced if and only if the canonical efficiency-factors of d are all equal.

A connected block (resply. row-column) design d is efficiency-balanced if and only if

$$C_d(C_d^{(RC)}) = \alpha (R_d - n^{-1} r_d r_d')$$
, (2.6)

where $0 < \alpha \le 1$ is a scalar and n = bk.

It is known that a variance-balanced (block or row-column) design with v>2 is efficiency-balanced, and conversely, if and only if the design is equireplicate. Also, it may be noted that in the class of proper (equal block sized) designs, any binary variance- or efficiency-balanced block design is necessary equireplicate.

For a block design $d \in D_0(v, b, k)$, let $0 = z_{d0} < z_{d1} \le z_{d2} \le \ldots \le z_{d, v-1}$ denote the eigenvalues of C_d . Similarly, let $0 = z_{d0}^* < z_{d1}^* \le z_{d2}^* \le \ldots \le z_{d, v-1}^*$ denote the eigenvalues of $C_d^{(RC)}$ for $d \in D(v, b, k)$.

Definition 2.3: Let d^* be a block (resply. row-column) design belonging to $D_0(v,b,k)(D(v,b,k))$. If $z_{d^*1} \ge z_{d1}(z_{d^*1}^* \ge z_{d1}^*)$ for any other design $d \in D_0(v,b,k)$ $(d \in D(v,b,k))$, then d^* is E-optimal in $D_0(v,b,k)$ (resply. in D(v,b,k)).

It is well-known that a design is *E*-optimal if and only if it minimizes the maximum variance of the best linear unbiased estimator of normalized treatment contrasts.

Finally, we quote some results and definitions from Das and Dey (1989).

Definition 2.4: A $k \times b$ array containing entries from a finite set $\Omega = \{1, 2, ..., v\}$ of v treatment symbols is called a Youden Type (YT) row-column design if the ith treatment symbol occurs in each row of the array m_i times, for i = 1, 2, ..., v, where $m_i = r_i/k$ and r_i is the replication of the ith treatment symbol in the array.

Theorem 2.1: A necessary and sufficient condition for the existence of a YT design is that r_i/k is an integer, for i = 1, 2, ..., v.

Theorem 2.2: A necessary and sufficient condition for $C_d^{(RC)} = C_d^N$ is that $d \in D(v, b, k)$ is a YT design.

Remark 1: In view of Theorem 2.2, it is clear that if the block design d_{N_1} corresponding to a row-column design $d \in D(v, b, k)$ is ϕ -optimal according to some non-increasing optimality criterion ϕ , then d is also ϕ -optimal, provided d is a YT design (An optimality criterion ϕ is non-increasing if $\phi(A) \le \phi(B)$ whenever A-B is non-negative definite). Thus, in the case of YT designs, the search for optimal designs in a three-way setting reduces to that in a two-way setting.

3 E-Optimal Variance- and Efficiency-Balanced Block Designs

3.1 Variance-Balanced Designs

Consider a Balanced Incomplete Block (BIB) design with parameters v', b' = v'(v'-1)/3, r' = v'-1, k = 3, $\lambda' = 2$, and let N be the incidence matrix of such a BIB design. Such a BIB design is also called a two-fold triple system and general solutions to these designs are well-known (see e.g., Bose (1939)). Let d^* be a block design with incidence matrix

$$N_{d*} = \begin{bmatrix} N & I_{v'} \\ 0' & 21' \end{bmatrix} . {3.1}$$

Then, it is easy to see that the C-matrix of d^* is $C_{d^*} = (2v/3)(I-v^{-1}11')$ where v = v'+1. Thus, d^* is variance-balanced, with v = v'+1 treatments, $b = (v^2-1)/3$ blocks, block size k=3 and replications

$$r_{d*i} = v - 1 (= r_1, \text{ say})$$
 for $i = 1, 2, ..., v - 1$,
 $r_{d*v} = 2(v - 1) (= r_2, \text{ say})$. (3.2)

Jacroux (1980) proved that for any block design $d \in D_0(v, b, k)$,

$$z_{d1} \le \bar{r}(k-1)v/\{(v-1)k\} , \qquad (3.3)$$

where \bar{r} is the largest integer not exceeding bk/v. Since d^* is variance-balanced, $z_{d^*i} = 2v/3$ for i = 1, 2, ..., v - 1. It is now easy to see that $z_{d^*i} = 2v/3$ attains the upper bound specified in (3.3) and hence d^* is *E*-optimal over $D_0(v, b, 3)$ with v = v' + 1, $b = (v^2 - 1)/3$. We thus have

Theorem 3.1: The design d^* is variance-balanced with replications as in (3.2) and is E-optimal over $D_0(v, b, 3)$, with v = v' + 1, $b = (v^2 - 1)/3$, provided BIBD (v', b', r', 3, 2) exists.

The following are the only possible series of BIB designs which satisfy the conditions required for obtaining the design d^* of Theorem 3.1:

(i)
$$v' = 3t$$
, $b' = t(3t-1)$, $r' = 3t-1$, $k = 3$, $\lambda' = 2$, $t \ge 1$,
(ii) $v' = 3t+1$, $b' = t(3t+1)$, $r' = 3t$, $k = 3$, $\lambda' = 2$, $t \ge 1$.

We refer to Dey (1986) for their construction. Note that for t = 1 in series (i), the design reduces to a Randomized block design.

3.2 Efficiency-Balanced Designs

Let there exist a BIB design d' with parameters v', b, r', k, λ . Let these treatments be grouped into (p+1) disjoint groups, say $p_1, p_2, \ldots, p_{p+1}$, such that the first group p_1 contains (v'-2p) treatments and each of the remaining groups contain two treatments. Let $N_{d'}$ be the incidence matrix of d', such that the first (v'-2p) rows correspond to treatments in p_1 and the remaining rows correspond to treatments in p_i for $i=2,3,\ldots,p+1$. From $N_{d'}$, we get another matrix N_{d*} by adding the two rows corresponding to the two treatments in each of the groups $p_2, p_3, \ldots, p_{p+1}$; the first (v'-2p) rows are left unaltered. Then, N_{d*} is the incidence matrix of a block design with v=v'-2p+p=(v'-p) treatments, b blocks, block size k and replications

$$r_{d*i} = b k/v' = r' (= r_1, \text{ say})$$
 for $i = 1, 2, ..., v - p$,
= $2r' (= r_2, \text{ say})$ for $i = v - p + 1, ..., v$. (3.5)

The C-matrix of d^* is

$$C_{d*} = (\lambda/k) \begin{bmatrix} v_I' - 11' & -211' \\ -21'1 & 2v'I - 411' \end{bmatrix},$$
(3.6)

where in (3.6), the first principal submatrix is of order v-p and the second, of order p; $\lambda = r'(k-1)/(v'-1)$. Simple calculations yield

$$C_{d^*} = \alpha (R_{d^*} - n^{-1} r_{d^*} r'_{d^*}) , \qquad (3.7)$$

where
$$\alpha = \lambda v'/kr'$$
, $R_{d^*} = \begin{bmatrix} r'I \\ 2r'I \end{bmatrix}$, $r'_{d^*} = (r'1', 2r'1')$ and $n = bk$. Thus, d^*

is efficiency-balanced. In fact, that $d^{\frac{1}{4}}$ is efficiency-balanced follows from a result of Puri and Nigam (1975), though the proof given here is some what different. Further, from Bagchi (1988), it follows that for p>0, d^* is E-optimal over $D_0(v,b,k)$ provided the following two conditions are satisfied:

(i)
$$v - pr' \ge 2$$
,
(ii) $v - v(v - pr')^{-1} \ge p\lambda$.

Thus, we have

Theorem 3.2: The design d^* (if it exists) is (i) efficiency-balanced and (ii) is E-optimal over $D_0(v, b, k)$ provided the conditions of (3.8) are met.

In particular, the following two series of BIB designs satisfy (3.8):

(i)
$$v' = s^2 + s + 1 = b$$
, $r' = s + 1 = k$, $\lambda = 1$,
 s a prime power and $p \le s - 1$,
(ii) $v' = 4t - 1 = b$, $r' = 2t - 1 = k$, $\lambda = t - 1$, $t > 1$, $p = 1$.

For construction methods, refer Dey (1986).

4 E-Optimal Variance- and Efficiency-Balanced Row-Column Designs

From Definition 2.4, Theorems 2.1 and 2.2, and Remark 1, the following results are obvious.

Theorem 4.1: The block contents of the block design d^* (if it exists) in Theorem 3.1 can be rearranged to yield a YT design provided (v-1) is divisible by 3. In such a case, the YT design is variance-balanced and E-optimal in D(v, b, 3).

In particular, the series (i) of (3.4) can be used to obtain E-optimal row-column designs.

Theorem 4.2: The block contents of the design d^* (if it exists) in Theorem 3.2 can be rearranged to yield a YT design, provided r' is divisible by k. Further, in such a case, the YT design is efficiency-balanced and E-optimal provided the conditions (3.8) are met.

Note that for the series of designs in (3.9), the conditions in Theorem 4.2 are satisfied and hence, these designs can be used to obtain *E*-optimal row-column designs.

5 Efficiency of Designs as per A- and D-Criterion

The designs constructed in Section 3 and 4 are shown to be *E*-optimal. It may be of further interest to see how these designs perform under a change of criterion. For a block (resply. row-column) design d belonging to $D_0(v, b, k)$ (resply. D(v, b, k)) let,

$$\phi_{A}(d) = \sum_{i=1}^{v-1} z_{di}^{-1} \quad \left(\text{resply.} \quad \sum_{i=1}^{v-1} z_{di}^{*-1}\right) \quad \text{and}$$

$$\phi_{D}(d) = \prod_{i=1}^{v-1} z_{di}^{-1} \quad \left(\text{resply.} \quad \prod_{i=1}^{v-1} z_{di}^{*-1}\right) . \tag{5.1}$$

A design is A-optimal (D-optimal) if it minimizes $\phi_A(d)(\phi_D(d))$ over all the designs in $D_0(v, b, k)$ or D(v, b, k). The A- and D-efficiency of a design d is defined as

$$e_A(d) = \phi_A(d_A^*)/\phi_A(d)$$
 and
$$e_D(d) = \{\phi_D(d_D^*)/\phi_D(d)\}^{1/(v-1)},$$
 (5.2)

where $d_A^*(d_D^*)$ is the A-optimal (D-optimal) design. One difficulty with these definitions of efficiency is that A-(or D-)optimal designs are known only for some specific values of v, b, k. Alternatively, one can obtain simple lower bounds of e_A and e_D as conservative measures of efficiency (see, e.g., Cheng and Wu (1981)). It has seen shown by Kiefer (1958, 1975) that for any design $d \in D_0(v, b, k)$ (or $d \in D(v, b, k)$),

$$\phi_A(d) \ge (v-1)^2/\{b(k-1)\}$$
 and
$$\phi_D(d) \ge [(v-1)/\{b(k-1)\}]^{v-1} .$$
 (5.3)

These lower bounds are the ϕ_A and ϕ_D values of a BIB (resply., Youden) design with parameters v, b, k. The efficiency lower-bounds are then,

$$e'_{A}(d) = (v-1)^{2}/\{b(k-1)\phi_{A}(d)\}\$$

$$e'_D(d) = (v-1)/[b(k-1)\{\phi_D(d)\}^{1/(v-1)}]$$
.

We use the above lower bounds of A- and D-efficiencies for the designs constructed in Sections 3 and 4.

For the variance-balanced (block or row-column) design d^* , we easily have,

$$e'_A(d^*) = e'_D(d^*) = v/(v+1)$$
 (5.5)

We have computed and presented in Table 1 these lower bounds to A- and D-efficiency for variance-balanced designs d^* with $v < b \le 50$ obtained from the two series, as in (3.4), of BIB designs.

Table 1. Parametric values of *E*-optimal Variance-balanced block and row-column designs based on Theorems 3.1 and 4.1 and their *A*- and *D*-efficiency lower bounds

S. No.	υ	b	k	r ₁	<i>r</i> ₂	$e_A'(d^*)$	$e_D'(d^*)$
1	4	5		3	6	0.800	0.800
2*	5	8	3	4	8	0.833	0.833
3	7	16	3	6	12	0.875	0.875
4*	8	21	3	7	14	0.889	0.889
5	10	33	3	9	18	0.909	0.909
6*	11	40	3	10	20	0.917	0.917

For the efficiency-balanced block (resply. row-column) design d^* , using the expression (3.6) for the C-matrix, one can show, after some routine calculations, that the positive eigenvalues of C_{d^*} (resply., $C_{d^*}^{(RC)}$) are $2\lambda (v+p)/k$, $\lambda (v+p)/k$ and $2\lambda v/k$ with respective multiplicities (p-1), (v-p-1) and 1. Thus, for efficiency-balanced design d^* , we have,

$$e'_{A}(d^*) = 2v(v-1)/\{(2v-p)(v+p-1)\},$$

and (5.6)

$$e'_{D}(d^{*}) = \frac{v-1}{(v+p-1)} \{v2^{p}/(v+p)\}^{1/(v-1)}$$
.

These lower bounds to A- and D-efficiency for efficiency-balanced design d^* (in the parametric range $v < b \le 50$, $3 \le k \le 15$), which are derivable from existing

Table 2. Parametric values of E-optimal Efficiency-balanced block and row-column designs based on Theorems 3.2 and 4.2 and their A- and D-efficiency lower bounds

S. No.	p	υ	b	<u>k</u>	<i>r</i> ₁	<i>r</i> ₂	$e'_A(d^*)$	$e'_D(d^*)$
1	1	6	7	3	3	6	0.909	0.928
2*	1	8	12	3	4	8	0.933	0.950
3	1	12	26	3	6	12	0.957	0.959
4*	1	14	35	3	7	14	0.963	0.974
5	1	6	7	4	4	8	0.909	0.928
6*	1	9	15	4	6	12	0.941	0.957
7	2	11	13	4	4	8	0.917	0.941
8	1	12	13	4	4	8	0.957	0.969
9	1	12	26	4	8	16	0.957	0.969
10*	2	14	20	4	5	10	0.933	0.954
11*	1	15	20	4	5	10	0.966	0.976
12*	1	15	40	4	10	20	0.966	0.976
13	2	23	50	4	8	16	0.958	0.973
14	1	24	50	4	8	16	0.979	0.986
15	1	10	11	5	5	10	0.947	0.962
16	3	18	21	5	5	10	0.927	0.952
17	2	19	21	5	5	10	0.950	0.967
18	1	20	21	5	5	10	0.974	0.983
19	1	20	42	5	10	20	0.974	0.983
20*	3	22	30	5	6	12	0.939	0.960
21*	2	23	30	5	6	12	0.958	0.973
22*	1	24	30	5	6	12	0.979	0.986
23	1	10	11	6	6	12	0.947	0.962
24	2	14	16	6	6	12	0.933	0.954
25	1	15	16	6	6	12	0.966	0.976
26*	1	15	24	6	9	18	0.966	0.976
27	1	15	32	6	12	24	0.966	0.976
28	1	20	42	6	12	24	0.974	0.983
29	4	27	31	6	6	12	0.936	0.959
30	3	28	31	6	6	12	0.951	0.968
31	2	29	31	6	6	12	0.967	0.978
32	1	30	31	6	6	12	0.983	0.989
33	1	14	15	7	7	14	0.963	0.974
34*	1	20	30	7	10	20	0.974	0.983
35	1	21	44	7	14	28	0.976	0.984
36*	2	26	36	7	9	18	0.963	0.976
37*	1	27	36	7	9	18	0.981	0.988
38	1	14	15	8	8	16	0.963	0.974
39	1	12	13	9	9	18	0.957	0.969
40	1	18	19	9	9	18	0.971	0.981
41*	1	20	35	9	15	30	0.974	0.983
42	2	23	25	9	9	18	0.958	0.973
43	1	24	25	9	9	18	0.979	0.986
44	1	24	50	9	18	36	0.979	0.986
45*	1	26	39	9	13	26	0.980	0.987
46*	2	31	44	ģ	12	24	0.969	0.980
47*	1	32	44	ģ	12	24	0.984	0.990
48	3	34	37	ģ	9	18	0.959	0.974
49	2	35	37	ģ	ģ	18	0.972	0.982
50	1	36	37	ģ	ģ	18	0.986	0.991

Table 2 (continued)

S. No.	p	υ	b	k	r_1	<i>r</i> ₂	$e'_A(d^*)$	$e'_D(d^*)$
 51	1	15	16	10	10	20	0.966	0.976
52	1	18	19	10	10	20	0.971	0.981
53*	1	24	40	10	16	32	0.979	0.986
54	2	29	31	10	10	20	0.967	0.978
55	1	30	31	10	10	20	0.983	0.989
56	1	22	23	11	11	22	0.977	0.984
57*	1	32	48	11	16	32	0.984	0.990
58	1	22	23	12	12	24	0.977	0.984
59	3	42	45	12	12	24	0.966	0.979
60	2	43	45	12	12	24	0.977	0.986
61	1	44	45	12	12	24	0.989	0.993
62	1	26	27	13	13	26	0.980	0.987
63	2	38	40	13	13	26	0.974	0.984
64	1	39	40	13	13	26	0.987	0.992
65	1	26	27	14	14	28	0.980	0.987
66	1	30	31	15	15	30	0.983	0.989
67	2	34	36	15	15	30	0.971	0.982
68	1	35	36	15	15	30	0.986	0.991
69*	1	35	48	15	20	40	0.986	0.991

BIB designs (or their complements) listed in Hall (1986), have been computed and presented in Table 2.

In these tables the designs marked with asterisk cannot be converted to a YT design. As such these parameters refer only to the block designs. It is apparent from these tables that the designs, apart from being *E*-optimal, have high *A*- and *D*-efficiencies as well.

Acknowledgements: The authors are thankful to the referee for his valuable comments on a previous draft.

References

Bagchi S (1988) A class of non binary unequally replicated *E*-optimal designs. Metrika 35:1-12 Bose RC (1939) On the construction of balanced incomplete block designs. Ann Eugen 9:353-399 Cheng CS, Wu CF (1981) Nearly balanced incomplete block designs. Biometrika 68:493-500 Das A, Dey A (1989) A generalization of systems of distinct representatives and its applications. Calcutta Statist Assoc Bull 38:57-63

Dey A (1986) Theory of block designs. Wiley, New York Hall M Jr (1986) Combinatorial theory, 2nd ed. Wiley, New York

Jacroux M (1980) On the determination and construction of E-optimal block designs with unequal number of replicates. Biometrika 67:661-667

Kiefer J (1958) On the nonrandomized optimality and randomized non-optimality of symmetrical designs. Ann Math Statist 29:675-699

Kiefer J (1975) Constructions and optimality of generalized Youden designs. In: Specialized JN (ed)
A Survey of Statistical Designs and Linear Models. North Holland, Amsterdam pp 333-352
Mukerjee R, Saha GM (1990) Some optimality results on efficiency-balanced designs frankhyā (to appear)

Puri PD, Nigam AK (1975) A note on efficiency balanced designs. Sankhyā B 37:45 - 460

Received 12 July 1989 Revised version 19 July 1990

.