Optimality and Efficiency of Circular Neighbor Balanced Design

Jeevitha M.¹, C. Santharam²

¹Department of Statistics, Loyola College, Chennai – 600034, Tamil Nadu, INDIA.

Corresponding Addresses:

¹jeevitha7@gmail.com, ²santharamsdeep@yahoo.com

Research Article

Abstract: This paper deals with the study of optimality of circular neighbor balanced designs for total effects when the one-sided or two sided neighbor effects are present in the models and the observation errors are correlated according to first order circular stationary autoregressive process. Some optimality results under some specified conditions are provided and the efficiencies of circular neighbor balanced designs relative to the optimal continuous block designs are also investigated. In order to discuss the efficiency of circular neighbor balanced designs among all possible block designs are characterized and the efficiencies of circular neighbor balanced with block of small size $k \le 16$ are illustrated.

Key Words: Auto regressive process, Block Design, Circular, Correlated observation, Neighbor balanced, Total effect, Universal optimality.

Introduction

Blocking of experimental units can efficiently eliminate heterogeneity in the experimental material and increase the sensitivity of data analysis. This technique has been popularly adopted in various scientific investigations and product quality improvement. It motivates many researchers to deeply study the optimality of block designs and their construction approaches. The detailed discussion can be found in most recent texts on the design of experiments, e.g., in Dey (1986), Pukelsheim (1993), Wu and Hamada (2000) and Box, Hunter and Hunter (2005). In many experiments, the response on one subject in a given period may be affected by the neighbor (or residual) effects of the treatments applied to that subject in the neighboring periods as well as by the direct effect of the current treatment. Under the linear models with the neighbor effects, many optimality results of block designs are established for treatment and neighbor effects separately. Hedayat and Afsarinejad (1978), Cheng and Wu (1980), Kunert (1984b) and Kushner (1997) for cross-over designs, Kunert (1984a) and AzaÄ³s, Bailey and Monod(1993), Druilhet (1999) and Filipialk and Markiewicz(2005) were dealt with circular neighbor- balanced designs. Bailey and Druilhet (2004) pointed out that the effect of most importance is the sum of the direct effect of the treatment and the neighbor effects of the same treatment that is the total

effect. Furthermore, they also showed that a circular neighbor-balanced design is universally optimal [in Kiefer's (1975) sense] for total effects under linear models containing the neighbor effects at distance one among the class of all designs with no treatment preceded by it. Optimality of circular neighbor - balanced designs for total effects with Autoregressive correlated observations was introduced by Yun long Yu, MingYao Ai, and Shayuan He. In this paper we study the universal optimality of circular neighbor-balanced designs for total effects, but when the observation errors are correlated according to a first-order circular autoregressive process. In this paper, Section 2 deals with some definitions and preliminaries. Section 3 presents the main results that circular neighbor- balanced designs are universally optimal under some conditions for the total effects in linear models which incorporate one-sided or two-sided neighbor effects when the observation errors are correlated according first-order to a circular autoregressive process. In order to discuss the efficiency of circular neighbor-balanced designs among all possible block designs with the same parameters, the optimal continuous block designs are characterized in Section 4. Section 5 presents the efficiency of circular neighborbalanced designs with blocks of small size $k \leq 16$ based on the previous structure of optimal equivalence classes of sequences.

2. Model and Definition

Let *ln denote* an *n*-dimensional vector of ones and the symbol \bigotimes denote the Kro- Necker product. Consider a set of circular block designs $\Omega_{(t,b,k)}$. For a design $d \in \Omega_{(t,b,k)}$, the left-Neighbor and two-sided Neighbor linear effect additive model can be written in vector form as (M_i)

$$Y = \mathbf{1}_{bk} \mu + T_d \tau + L_d \lambda + (I_b \otimes I_k) \beta + \epsilon$$
(1)
(M₂)

$$Y = \mathbf{1}_{bk}\mu + T_d \tau + L_d \lambda + R_d \rho + (I_b \otimes I_k)\beta + \varepsilon$$
Where $Y = (Y_{11}, \dots, Y_{1k}, \dots, Y_{b1}, \dots, Y_{bk})', Y_{ij}$ is the

observation response on plot *j* of block *i*, μ is the general mean, τ , λ and ρ are, respectively, the *t*-dimensional vectors of the direct effects, left-Neighbor effects and right-Neighbor effects of the *t* treatments, T_d , L_d and R_d are the corresponding incidence matrices, β is the *b*-dimensional vector of the block effects, and ε is the vector of random errors. Suppose that the errors in each block are correlated according to a first-order circular auto regressive process, denoted by AR(1,C). Details given in Kunert and Martin (1987). The AR(1,C) process can be represented in the recursive form $\varepsilon_i = v\varepsilon_{i-1} + \eta_i$ with |v| < 1, where the ηi 's are uncorrelated noises with $E(\eta i)=0$ and Var $(\eta_i)=\sigma 2$, and $E(\varepsilon_0)=0$. Then $E(\varepsilon) = 0$ Cov $(\varepsilon) = \sigma 2I_b \otimes S$ and

$$S^{-1} = (1 + v)^{2} I_{k} - v(H + H^{2})$$

$$= \begin{bmatrix} 1+v^{2} & -v & 0 & \dots & 0 & -v \\ -v & 1+v^{2} & -v & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1+v^{2} & -v \\ -v & 0 & 0 & \dots & -v & 1+v^{2} \end{bmatrix}$$

Where *H* denotes the $k \times k$ matrix with $h_{1k}=1$ and the $(i,j)^{\text{th}}$ element $h_{ij}=1$ if i-j=1 and 0 otherwise. Note that when v = 0, the structure of errors is reduced to the popular i.i.d. case. Let φ and ψ denote the total effects of the *t* treatments in the models (M_1) and (M_2) , respectively, that is $\varphi=\tau+\lambda$ and $\psi=\tau+\lambda+\rho$. Thus, we can obtain the following universal optimality results of CNBD's for the total effects.

Result 1

For $3 \le k \le t$, a CNBD (2) in $\Omega_{(t,b,k)}$ is universally optimal for the total effects in the model (M₁) among all the designs with no treatment Neighbor of itself when $0 \le v < 1$, and among all the designs with no treatment Neighbor fit self at distance 1 or 2 when -1 < v < 0.

Result 2

For $4 \le k \le t$, a CNBD (3)in $\Omega_{(t,b,k)}$ is universally optimal for the total effects in the model (M₂) among all the designs with no treatment Neighbor of itself at distances upto 2 when $0 \le v < 1$, and among all the designs with no treatment Neighbor of itself at distances upto 3 when -1 < v < 0.

3. Characterization of optimal continuous block designs

In the following sections, we are going to discuss the efficiency of CNBD (2) for the total effect in the model (M_1) , even if all the procedure can similarly be adapted to the case of CNBD (3) for the total effect in the $model(M_2)$. The optimal designs among all possible designs with the same parameters are characterized according to the method introduced by Kushner (1997). For details also refer to Kunert and Martin (2000) and Bailey and Druilhet (2004). For $u=1, 2, \ldots, b$, let T_{du} be the incidence matrix of the direct effects of the treatment in block u, $1 \le u \le b$. Then $T_d = (T_{d1}, T_{d2}, \dots, T_{db})$ ' is just the incidence matrix of the direct effects. For each u, define $L_{du}=H T_{du}$, $R_{du}=H'T_{du}$. Thus, it is obvious that $L_d=$ $(I_b \otimes H)T_d$ and $R_d = (I_b \otimes H') T_d$ are exactly the incidence matrices of the left-Neighbor effects and of the right-Neighbor effects. Define that two sequences of treatments on a block are equivalent if one sequence can be obtained from the other by relabeling the treatments and denote by s the equivalence class of the sequence l on the block u. Because $tr(C_{du})$ are in variant under permutations of treatment labels, so the value $tr(C_{du})$ remains the same for any sequence in the same equivalence class. Thus, we can define,

$$c(s) = tr(C_{du}) = \frac{1}{2} \left[(1 + v^2 - v)k + (1 - v)^2 \sum_{i=1}^{t} m_i - v \sum_{i=1}^{t} p_i - \frac{2(1 - v)^2}{k} \sum_{i=1}^{t} n_i^2 \right]$$
(3)

where n_i is the number of occurrences of treatment *i* in the sequence l,m_i is the number of times treatment *i* is on the left-hand side of itself in the sequence l and p_i is the number of plots having treatment *i* both on the left-hand side and on the right-hand side.

Optimal equivalence classes of sequences when k = 3 or 4

When k = 3 or 4, all the non-equivalent sequences are listed in the following two tables.

Table 1: All the non-equivalent sequences when k=3

No.	Sequence	v	v_I	C(S)
1	aaa	0	1	0
2	aab	1	1	$1/3(v^2+v+1)$
3	abc	3	0	$1/2(v^2+v+1)$

Гí	ab	le 1	2:	All	the	non-	equ	ival	ent	sequ	uence	es w	hen	<u>k</u> =4

No.	Sequence	v	v_1	C(S)
1	aaaa	0	1	0
2	aabb	0	2	$v^2 + 1$
3	aaab	1	1	$1/2(v^2+1)$
4	aabc	2	1	$v^2 + 1$
5	abcd	4	0	$v^2 + 1$

Proposition 1:

When k = 3 or 4, for any $v \in (-1,1)$, a CNBD (2) is universally optimal for the total effects in the model (M₁) among all possible designs with equal size.

Proposition 2:

When $k \ge 5$, $v \ge 2$ and v1 = 0 or 1 in any optimal sequence.

Proposition 3

$$v \in \left(\frac{v-v}{2}\right), 1$$

When for any positive integer k ≥ 5 , if k is odd, then the optimal sequence has the form of $a_1a_2a_2a_3a_3\cdots a_{[k/2]}a_{[k/2]}$, while if k is even, then the optimal sequence has the form of $a_1a_1a_2a_2\cdots a_{[k/2]}a_{[k/2]}$, where $a_1,...,a_{[k/2]}$ are distinct treatments.

Proposition 4

When any equivalence class of sequences c(s) $-1 < v \le \frac{3 - \sqrt{5}}{2}$

with $\frac{-1 < v > 2}{2}$ has the following upper bound: $c(s) \le f(v_1, v_2)$

$$= (1-v)^{2}(k-1) - \frac{1}{2}(v^{2} - 3v + 1)v_{1} - \frac{1}{2}(v^{2} - 4v + 1)v_{2} + \frac{(1-v)^{2}}{k}v_{2}\left[\frac{(K-v_{1})^{2}}{v_{2}} - \frac{(1-v)^{2}}{k}(2K - 2v_{1} - v_{2})\left[\frac{(K-v_{1})^{2}}{v_{2}}\right]^{2}$$
(4)

With equality if and only if for all $i \in N_2$

(*i*)
$$m_i = n_i - 1, \ p_i = n_i = 2;$$

(*ii*) $n_i = \left[\frac{(K - v_1)}{v_2}\right]_{or} \left[\frac{(K - v_1)}{v_2}\right]^2 + 1$

Proposition 5

When $-1 < v \le 0$ and for $k \ge 8$, no optimal sequence contains any treatment just once, i.e., $v_1 = 0$ for any optimal sequence.

4. Optimal equivalence classes of sequences when $k \ge 5$

Let *l* be sequence in an equivalence class. Denote by N_1 and N_2 , respectively, the sets of treatments appearing just once and at least twice in 1. Then $N=N_1U$ N_2 is the set of distinct treatments in *l*. Let $v_1 = |N_1|, v_2 =$ $|N_2|$ and v = |N|, where |N| denotes the cardinality of the set N. For illustration, under the condition of, $-1 < v < 3 - \sqrt{5}/2$

 $-1 < v < 3 - \sqrt{5}/2$ the optimal treatment sequences for the given parameters $\{v_1, v_2\}$ are listed together with the corresponding tr (C_{du}) for k = 5, 6, 7, 8, ... 16 in Tables 3 to 14, respectively. Note that the sequence for a CNBD (2) is also listed in the last row for the convenience of comparison.

Optimal equivalence classes of sequences when *k*=5

Table 3: Optimal sequences for all possible pairs of $\{v, v_1\}$ for k=5

S.No	Optimal Sequence	v	v ₁	Tr (C _{du})
1	Aabbb	2	0	$1/5(7v^2-4v+7)$
2	Abbbb	2	1	$1/5(7v^2-4v+7)$
3	Aabbc	3	0	1/10(17 v ² -9 v +17)
4	aabcc	3	0	1/10(17 v ² -9 v +17)
5	aabcc	3	0	$1/10(17 v^2-9 v + 17)$
6	abcde	5	0	$1/2 (3v^2 - v + 3)$

Among the above sequences, the sequence "abbcc" is the optimal sequence by Proposition 3.3 **Optimal equivalence classes of sequences when** k=6

Table 4: Optimal sequences for all possible pairs of $\{v, v_1\}$ for k=6

S.No	Optimal Sequence	V	v ₁	Tr (C _{du})
1	aaabbb	2	0	$2(v^2-v+1)$
2	abbbbb	2	1	$1/3(2v^2-v+2)$
3	aabbcc	3	0	$1/2(5v^2-4v+5)$
4	abbccc	3	0	$1/6(13v^2 - 11v + 13)$
5	abcdef	6	0	$2v^2-v+2$

Among the above sequences, the sequence "aabbcc" is the optimal sequence by Proposition 3.3 Similarly for k=7,...16, the optimal sequences can be derived and it can be concluded from the above tables that below are the optimal sequences for the corresponding block size "k"

|--|

	Optimal
Block Size	Sequence
5	abbcc
6	aabbcc
7	aabbccc
/	abbccdd
8	aabbccdd
	aaabbbccc
9	aabbccddd
	abbccddee
10	aabbccddd
10	abbccddee
	aabbbcccddd
11	aabbccddeee
	abbccddeeff
	aabbccddeeff
12	aaabbbcccddd
	aaaabbbbbcccc
	abbbbccccdddd
13	aaabbbcccdddd
	abbccddeeffgg
14	aabbccddeeffgghh
	aaaaabbbbbbccccc
15	aaabbbccccdddeee
	abbccddeeffgghh
	aaaaabbbbbbcccccc
16	aaaabbbbbccccdddd
	aabbccddeeffgghh

The below table represents all the optimal sequences for $5 \le k \le 16$. Also Note that the below table shows the optimal sequence and the last column lists the values Tr(C_{du}) of a CNBD (2) d.

Block Size	Optimal Sequence	c (s*)	tr(C _{du})
5	abbcc	$1/10(17v^2 - 9v + 17)$	$1/2 (3v^2 - v + 3)$
6	aabbcc	$1/2 (5v^2 - 4v + 5)$	$(2v^2 - v + 2)$
7	aabbccc abbccdd	$\frac{1/14}{(44v^2 - 44v + 43)} \frac{1/14}{1/14}$	$1/2 (5v^2 - 3v + 5)$
8	aabbccdd	$4(v^2 - v + 1)$	$(3v^2 - 2v + 3)$
9	aaabbbccc aabbccddd abbccddee	$\frac{1/2}{-8\nu+7} \frac{(9\nu^2 - 12\nu + 9)}{1/18} \frac{2/3}{(83\nu2)} \frac{(7\nu^2 - 8\nu + 7)}{-85\nu + 83)}$	$1/2 (7v^2 - 5v + 7)$
10	aabbccddd abbccddee	$\frac{1/2}{(11v^2 - 34v + 27)} \frac{1/2}{11v^2 - 12v + 11}$	$(4v^2 - 3v + 4)$
11	aabbbcccddd aabbccddeee abbccddeeff	$\frac{1/11}{1/22} (68v^2 - 62v + 68) \\ \frac{1}{22} (137v^2 - 164v + 137) \\ \frac{1}{22} (134v2 - 147v + 134)$	$1/2 (9v^2 - 7v + 9)$
12	aabbccddeeff aaabbbcccddd aaaabbbbcccc	(8v2 - 7v + 8) 1/4(27v2 - 36v + 27) 1/2(13v2 - 20v + 13)	$(5v^2 - 4v + 5)$
13	abbbbccccdddd aaabbbcccdddd abbccddeeffgg	1/26 (197v ² – <i>137v</i> + 197) 1/26 (221v ² – 296v +221) 1/26 (197v2 – 137v + 197)	$1/2 (11v^2 - 9v + 11)$
14	aabbccddeeffgghh	1/2 (17v2-20v+17)	$(6v^2 - 5v + 6)$
15	aaaaabbbbbbccccc aaabbbccccdddeee abbccddeeffgghh	$\frac{1/2 (17v^2 - 24v + 17)}{(141v^2 - 192v + 141)} \frac{1/2}{1/2}$ $(262v^2 - 229v + 262)$	$1/2 (13v^2 - 11v + 13)$
16	aaaaabbbbbbcccccc aaaabbbbbccccdddd aabbccddeeffgghh	$\frac{1}{1/2} (17v^2 - 24v + 17)$ 1/4(41v^2 - 62v +41) (10v^2 - 14v + 10)	$(7v^2 - 6v + 7)$

Table 6: Optimal sequences for $5 \le k \le 16$

5. Efficiency of CNBD (2) with blocks of size $5 \le k \le 16$

In this section we are going to discuss the Efficiency of CNBD (2) for the block size k=5, 6... 16.

For a fixed k, we can find an optimal equivalence class of sequence s*, which maximizes c(s) in (3). Any sequence in the optimal equivalence class is called optimal sequence. It was shown in Theorem 10 of Bailey and Druilhet (2004) that a designed* which has each sequence in s* equally often is universally optimal among all possible designs with the same size. Since the values t $r(C_{du})$ are invariant to any block *u* for a CNBD (2), so we can define the efficiency of a CNBD (2) *d* relative to the optimal continuous block design d* as

$$Eff(d) = \frac{tr(C_{d})}{tr(C_{d*})} = \frac{tr(C_{du})}{c(s^{*})}$$

The below tables show the calculations of Tr (C_{du}) and $c(s^*)$ when k=5, 6, 7,..., 16 Efficiency of CNBD (2) when the block size k=5

	Tabi	e /: Emci	lency of C	MDD(2)) when k=J	
S.No	v		c(S*)		tr(Cdu)	Eff(d)
1	-1	43.00	4.30	7.00	3.50	0.8140
2	-0.8	35.08	3.51	5.72	2.86	0.8153
3	-0.6	28.52	2.85	4.68	2.34	0.8205
4	-0.4	23.32	2.33	3.88	1.94	0.8319
5	-0.2	19.48	1.95	3.32	1.66	0.8522
6	0	17.00	1.70	3.00	1.50	0.8824
7	0.2	15.88	1.59	2.92	1.46	0.9194
8	0.4	16.12	1.61	3.08	1.54	0.9553
9	0.6	17.72	1.77	3.48	1.74	0.9819
10	0.8	20.68	2.07	4.12	2.06	0.9961
11	1	25.00	2.50	5.00	2.50	1.0000

Table 7: Efficiency of CNBD (2) when k=5

						Block	k Size					
v	5	6	7	8	9	10	11	12	13	14	15	16
-1	0.8140	0.7143	0.7165	0.6667	0.6813	0.6471	0.6627	0.6364	0.6511	0.6296	0.6431	0.6250
-0.8	0.8153	0.7158	0.7179	0.6680	0.6825	0.6483	0.6639	0.6376	0.6522	0.6308	0.6443	0.6262
-0.6	0.8205	0.7217	0.7232	0.6735	0.6876	0.6534	0.6687	0.6425	0.6569	0.6355	0.6488	0.6308
-0.4	0.8319	0.7351	0.7353	0.6859	0.6991	0.6651	0.6798	0.6537	0.6677	0.6465	0.6594	0.6415
-0.2	0.8522	0.7600	0.7580	0.7097	0.7213	0.6879	0.7014	0.6757	0.6888	0.6679	0.6802	0.6625
0	0.8824	0.8000	0.7955	0.7500	0.7590	0.7273	0.7388	0.7143	0.7259	0.7059	0.7169	0.7000
0.2	0.9194	0.8545	0.8483	0.8095	0.8153	0.7876	0.7963	0.7746	0.7838	0.7661	0.7751	0.7600
0.4	0.9553	0.9143	0.9086	0.8816	0.8844	0.8643	0.8696	0.8537	0.8597	0.8464	0.8525	0.8412
0.6	0.9819	0.9636	0.9605	0.9474	0.9483	0.9381	0.9405	0.9322	0.9350	0.9281	0.9310	0.9250
0.8	0.9961	0.9920	0.9912	0.9881	0.9883	0.9858	0.9863	0.9843	0.9849	0.9832	0.9839	0.9824
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

From the above table it is evident that the efficiency of a CNBD (2) approaches to 1 as v tends to 1 for k = 5. In the same manner, it can verified that the efficiency of a CNBD (2) approaches to 1 as v tends to 1 for k=6,7,8,...16. The Efficiency of CNBD (2) d for v belongs to (-1, 1) are given in the above table for k=5, 6...16.

The following figure shows the relationship between the efficiency Eff(d) of a CNBD (2) d and v for $5 \le k \le 16$. It can be seen that the efficiency of a CNBD (2) approaches to 1 as v tends to 1 for any k.



Figure 1: Efficiency of CNBD (2) when $5 \le k \le 16$

Summary and Conclusion

In this research paper, we have investigated the optimality and efficiency of circular Neighbor balanced block design. We have constructed the efficiency of circular neighbor balanced designs among all possible block designs with the same parameters the continuous block designs are characterized and the efficiencies of circular neighbor balanced designs with blocks of small size $k \le$ 16 are illustrated. From Fig 1, we could see that the efficiency of CNBD (2) approaches 1 as *v* tends to 1 for block sizes k=5, 6,...,16. So we can conclude that the Circular neighbor balanced design is an efficient design. Thus we can conclude that CNBD (2) is always a good choice when the adjacent observation errors have strong positive correlation.

6. Reference

- AZA IS, J.M., BAILEY, R.A. and MONOD, H., (1993). Acatalogue of efficient Neighbor- designs with border plots. Biometrics 49.
- BAILEY, R.A. and DRUILHET, P., (2004). Optimality of Neighbor-balanced designs for total effects. Ann. Statist. 32(4), 1650-1661.
- BOX, G.E.P., HUNTER, J.S. and HUNTER, W.G., (2005). Statistics for Experimenters. John Wiley & Sons, New Jersey.
- 4. CHENG, C.S. and WU, C.F.J.,(1980). Balanced repeated measurements designs. Ann. Statist. 8, 1272-1283.
- DRUILHET,P., (1999). Optimality of Neighbor balanced designs. J. Statist. Plann. Inference 81,141-152.
- FILIPIAK, K. and MARKIEWICZ, A., (2005). Optimality and efficiency of circular Neighbor balanced designs for correlated observations.Metrika61,17-27.
- HEDAYAT,A. and AFSARINEJAD, K., (1978). Repeated measurements design II. Ann. Statist. 6, 619 -628.
- KUNERT, J. and MARTIN, R.J., (2000). On the determination of optimal designs for an interference model. Ann. Statist.28,1728-1742.
- KUNERT, J., (1984a). Design balanced for circular residual effects. Comm.Statist.A-TheoryMethods13,2665-2671.
- KUNERT, J., (1984b). Optimality of balanced uniform repeated measurements de-signs. Ann. Statist.12,1006-1017.
- 11. KUSHNER, H.B., (1997). Optimal repeated measurements designs: The linear optimality equations. Ann. Statist.25,2328-2344.
- MAGDA, C.D., (1980). Circular balanced repeated measurements design. Comm. Statist. A- Theory Methods 9,1901-1918.
- SANTHARAM, C, PONNUSAMY, K.N., AND CHANDRASEKAR B., 1996: Universal Optimality of nearest Neighbor balanced block designs using ARMA models. Biometrical J. 38, 725 – 730.
- SANTHARAM, C, and PONNUSAMY, K.N.,1997: On the efficiencies of nearest Neighbor balanced block design with correlated error structure. Biometrical J. 39, 85 – 98.
- YUNLONG YU, MINGYAO AI, SHUYUAN He (2007) Optimality of Circular Neighbor – balanced Designs for Total Effects with Autoregressive Correlated Observations.