On the optimality of orthogonal arrays in case of correlated errors

Miwako Mishima[1,*] , Masakazu Jimbo[2] and Teruhiro Shirakura[3]

[1] *Department of Information Science, Gifu University, Gifu, 501-1193, Japan*
[2] *Department of Mathematics, Keio University, Yokohama, 223-8522, Japan*
[3] *Department of Human Development, Kobe University, Kobe, 657-8501, Japan*

Abstract

It is well-known that a two-level orthogonal array of strength 2 is universally optimum for the estimation of main effects for uncorrelated errors. In this paper, the property of orthogonal arrays which are also optimum even for correlated errors is discussed and a construction for such optimal designs is presented. Furthermore, in case when there are correlations between observations which are caused by the closeness of the assemblies (treatment combinations) of experiments, it is shown that if the design matrix is a linear orthogonal array, then the OLSE and the GLSE of main effects are uncorrelated.

AMS classifications: primary 62K05, 62K15; secondary 05B15

Keywords: design of experiments; orthogonal array; correlated error; nearest neighbor correlation structure; universally optimum; weakly universally optimum; Hamming distance.

1. Introduction

Optimal designs of experiments in the case of correlated observations have been studied by many authors. Kiefer (1975) introduced a general notion of optimality. Kiefer and Wynn (1981) discussed the optimality of balanced incomplete block designs and Latin square (or Latin hypercube) designs for correlated observations, and gave some constructions for the optimal designs (see also Cheng (1983)). Here, we consider the optimality of orthogonal arrays of strength 2 for the case when errors are correlated depending on the "closeness" between experiments.

Let $A_1, \ldots, A_m$ be the same kind of factors (treatments) with two levels and let $\Gamma = (\gamma_{ij})$ be an array of assemblies (treatment combinations), where $\gamma_{ij} \in \{0, 1\}$ is the level of the j-th factor for the i-th experiment. We assume that there are no interaction effects between these factors. For example, consider an experiment for estimating the respective effect of a certain drug at each of seven periods. Suppose that there are sufficient number of newborn rats to get a reliable result and that they are randomly assigned to N groups. The experiment is done for seven consecutive weeks after their birth. At each period (the first, the second, ... and the seventh week after birth, respectively), individual rats in the same group are received an identical treatment, i.e., 'dosing' or 'not dosing' with the drug. After the last week, we measure the mean of rats' weight or some constituent of their blood, etc. for each group. This experiment has $m = 7$ factors $A_1, \ldots, A_7$ (treatments of seven periods) with two levels (dosing and not dosing). Similar situation may occur in agricultural

* Corresponding author. E-mail: miwako@info.gifu-u.ac.jp. Supported in part by the Inamori Foundation.

field experiments, the factors could be treatments of some periods after seeding, each of which may be equally, say, two weeks (or maybe more), with two levels (say, irrigation and no irrigation).

	Treatments of seven periods								
	A_1	A_2	A_3	A_4	A_5	A_6	A_7		
Group i	0	1	0	0	1	1	1	0:	not dosing
Group j	0	1	1	0	1	1	1	1:	dosing

In the rat dosing experiment, as seen in the table, the groups i and j are received the same treatments except for the treatment A_3 of the third week after birth. The experiments for these two groups are said to be close and expected to give correlated observations. So it is rather natural to think that there exists some correlation $\mathrm{cov}(\varepsilon_i, \varepsilon_j)$ between errors of the i-th and the j-th experiments depending on their closeness, that is, the closer they are, the larger their correlation becomes. In this paper, we will consider the following model of observations under such a kind of correlated errors.

$$\boldsymbol{y} = \mu \mathbf{1}_N + X\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \qquad \mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma, \tag{1.1}$$

where $\boldsymbol{y} = (y_1, \ldots, y_N)^t$ is an N-dimensional vector of observations, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)^t$ is an N-dimensional vector of errors, $\mathbf{1}_N$ is the N-dimensional all-one column vector, $\boldsymbol{\alpha} = (\alpha_0^1, \ldots, \alpha_0^m)^t$ is an m-dimensional vector of main effects, and $X = (x_{ij})$ is an $N \times m$ matrix such that

$$x_{ij} = \begin{cases} -1 & \text{if } \gamma_{ij} = 0, \\ 1 & \text{if } \gamma_{ij} = 1. \end{cases}$$

That is, X is the matrix obtained by rewriting the elements 0 and 1 of Γ to -1 and 1 respectively. X is called the *design matrix* corresponding to Γ.

Then the *ordinary least squares estimator* (OLSE) and the *generalized least squares estimator* (GLSE) of $\boldsymbol{\alpha}$ are given by solutions of normal equations such as

$$\left(X^t X - \frac{1}{N} X^t J_N X\right) \hat{\boldsymbol{\alpha}} = \left(X^t - \frac{1}{N} X^t J_N\right) \boldsymbol{y} \tag{1.2}$$

and

$$\begin{aligned} &\left(X^t \Sigma^{-1} X - (\mathbf{1}_N^t \Sigma^{-1} \mathbf{1}_N)^{-1} X^t \Sigma^{-1} J_N \Sigma^{-1} X\right) \hat{\boldsymbol{\alpha}} \\ &\qquad = \left(X^t \Sigma^{-1} - (\mathbf{1}_N^t \Sigma^{-1} \mathbf{1}_N)^{-1} X^t \Sigma^{-1} J_N \Sigma^{-1}\right) \boldsymbol{y}, \end{aligned} \tag{1.3}$$

respectively, where J_N is the $N \times N$ all-one matrix. If the main effect $\boldsymbol{\alpha}$ of (1.1) is estimable, then the normal equation (1.2) (or (1.3)) has a unique solution.

There are many types of "optimality" criteria to evaluate the efficiency of $\hat{\boldsymbol{\alpha}}$. Now we review some optimality criteria for an array of assemblies Γ. Let

$$D(\Gamma) = \mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \begin{cases} (X^t X - \frac{1}{N} X^t J_N X)^{-1} (X^t - \frac{1}{N} X^t J_N) \Sigma \\ \qquad (X - \frac{1}{N} J_N X)(X^t X - \frac{1}{N} X^t J_N X)^{-1} \\ \qquad\qquad \text{for the OLSE } \hat{\boldsymbol{\alpha}}, \\ (X^t \Sigma^{-1} X - (\mathbf{1}_N^t \Sigma^{-1} \mathbf{1}_N)^{-1} X^t \Sigma^{-1} J_N \Sigma^{-1} X)^{-1} \\ \qquad\qquad \text{for the GLSE } \hat{\boldsymbol{\alpha}} \end{cases} \tag{1.4}$$

for the design matrix X corresponding to Γ, and let $C(\Gamma) = D(\Gamma)^{-1}$. Since $D(\Gamma)$ of (1.4) is the covariance matrix for $\hat{\boldsymbol{\alpha}}$, $C(\Gamma)$ for the OLSE $\hat{\boldsymbol{\alpha}}$ coincides with the so-called C-matrix (information matrix) only when $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2 I_N$.

Suppose that Ξ is a set of $N \times m$ arrays Γ of assemblies by which the main effect $\boldsymbol{\alpha}$ is estimable with respect to the OLSE (or the GLSE) and that $\boldsymbol{M}$ is the set of $C(\Gamma)$ for $\Gamma \in \Xi$. Let $\mathcal{M}$ be the convex hull of $\boldsymbol{M}$ and Φ be the set of functions ϕ on $\mathcal{M}$ such that

(i) ϕ is convex,

(ii) $\phi(bC) \leq \phi(C)$ for $C \in \mathcal{M}$ and for any $b \geq 1$, and

(iii) for any orthogonal matrix $\mathcal{P}$ and for $C \in \mathcal{M}$, $\phi(\mathcal{P}^t C \mathcal{P}) = \phi(C)$ holds.

An array Γ^* is said to be *universally optimum relative to* Ξ if

$$\phi(C(\Gamma^*)) = \min_{\Gamma \in \Xi} \phi(C(\Gamma))$$

holds for all functions $\phi \in \Phi$ satisfying (i), (ii) and (iii). The concept of universally optimum was introduced by Kiefer (1975). It is known that the criterion includes A-, D- and E-optimality as its special cases.

Proposition 1.1 (Kiefer (1975)) *An array* $\Gamma^* \in \Xi$ *is universally optimum relative to* Ξ *if*

(a) $C(\Gamma^*) = aI_m$ *and*

(b) $\mathrm{tr}(C(\Gamma^*)) = \max_{\Gamma \in \Xi} \mathrm{tr}(C(\Gamma))$

hold, where a *is a constant,* I_m *is the* $m \times m$ *identity matrix and* $\mathrm{tr}(C)$ *denotes the trace of a matrix* C.

Suppose that $\boldsymbol{M}'$ is the set of the covariance matrices $D(\Gamma)$ for any $\Gamma \in \Xi$, $\mathcal{M}'$ is the convex hull of $\boldsymbol{M}'$ and Ψ is the set of functions ψ on $\mathcal{M}'$ such that

(i)′ ψ is convex,

(ii)′ $\psi(bD) \geq \psi(D)$ for $D \in \mathcal{M}'$ and for any $b \geq 1$, and

(iii)′ for any orthogonal matrix $\mathcal{P}$ and for $D \in \mathcal{M}'$, $\psi(\mathcal{P}^t D \mathcal{P}) = \psi(D)$ holds.

An array Γ^* is said to be *weakly universally optimum relative to* Ξ if

$$\psi(D(\Gamma^*)) = \min_{\Gamma \in \Xi} \psi(D(\Gamma))$$

holds for all functions $\psi \in \Psi$ satisfying (i)′, (ii)′ and (iii)′. The concept of weakly universally optimum was introduced by Kiefer and Wynn (1981). It is known that the criterion includes A- and E-optimality and that Γ^* is weakly universally optimum relative to Ξ if it is universally optimum relative to Ξ.

Proposition 1.2 (Kiefer and Wynn (1981)) *An array $\Gamma^* \in \Xi$ is weakly universally optimum relative to Ξ if*

(a)$'$ $D(\Gamma^*) = aI_m$ *and*

(b)$'$ $\mathrm{tr}(D(\Gamma^*)) = \min_{\Gamma\in\Xi} \mathrm{tr}(D(\Gamma))$

hold, where a is a constant.

In any two columns of an $N \times m$ array Γ, if the ordered pairs $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$ occur equally often, then Γ is called an *orthogonal array of size N, m constraints, two levels and strength* 2, denoted by $OA(N, m, 2, 2)$ (for more general definitions, see, for example, Raghavarao (1971) and Beth, Jungnickel and Lenz (1985)). Without loss of generality, we can assume that the first row of Γ is $(0, \dots, 0)$. Noting that $X^t \mathbf{1}_N = 0$ and $X^t X = NI_m$ hold for the design matrix X corresponding to an orthogonal array Γ, the covariance matrix (1.4) for the OLSE of $\boldsymbol{\alpha}$ is reduced to

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{1}{N^2} X^t \Sigma X.$$

In the case when $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2 I_N$, it is well-known that an $OA(N, m, 2, 2)$ is universally optimum among arrays by which the main effect $\boldsymbol{\alpha}$ is estimable (see Kiefer (1975)). Our aim is to find an array Γ which is optimum not only for the case when $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2 I_N$ but also for a more general class of covariance structures given in the next sections.

In Section 2, we shall show the property of orthogonal arrays when the true covariance structure is correlated. In Section 3, when a linear orthogonal array is utilized, the covariance matrices for the OLSE and the GLSE of $\boldsymbol{\alpha}$ are given under a covariance structure which depends only on the Hamming distance. In Sections 4 and 5, under certain covariance structures the optimum designs are provided by using the result of Section 3. In Section 6, as a by-product, in the case of complete factorial designs the covariance matrices for the OLSE and the GLSE of main effect $\boldsymbol{\alpha}$ are shown under some covariance structures.

2. Optimality of orthogonal designs for correlated errors

For the model (1.1), we usually assume $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2 I_N$, but the true covariance structure may be different from this. Here, we consider the possibility of correlated errors. That is, if two rows of Γ resemble (close) each other, then there may exist correlation between errors of the experiments corresponding to these rows.

Let $\boldsymbol{\gamma}_i = (\gamma_{i1}, \dots, \gamma_{im})$ and $\boldsymbol{\gamma}_k = (\gamma_{k1}, \dots, \gamma_{km})$ be the i-th and the k-th rows of Γ. The number of j such that $\gamma_{ij} \neq \gamma_{kj}$ is called the *Hamming distance* between $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_k$, denoted by $d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k)$. And $d(\boldsymbol{\gamma}_i, (0, \dots, 0))$ is called the *Hamming weight* of $\boldsymbol{\gamma}_i$, denoted by $w(\boldsymbol{\gamma}_i)$. In our settings, we measure the closeness of the i-th and the k-th experiments by the Hamming distance $d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k)$. For an array Γ, if

$$\delta = \min\{d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k) |\ \boldsymbol{\gamma}_i,\ \boldsymbol{\gamma}_k \text{ are distinct two rows of } \Gamma\},$$

then Γ is called an array with *minimum distance* δ.

First, we assume the following covariance structure:

Covariance structure I:

$$\mathrm{cov}(\varepsilon_i, \varepsilon_k) \begin{cases} = \sigma^2 & \text{if } i = k, \\ \geq 0 & \text{if } d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k) \leq \delta, \\ = 0 & \text{if } d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k) > \delta, \end{cases}$$

where $\delta \leq \lfloor m/2 \rfloor$ is a given constant ($\lfloor a \rfloor$ implies the greatest integer not exceeding a).

If we do not know a covariance structure, we will use the OLSE to estimate the main effect $\boldsymbol{\alpha}$. Now, we will present an array of assemblies which is weakly universally optimum for such a case. Let Ξ be a set of $OA(N, m, 2, 2)$.

Theorem 2.1 *Under Covariance structure I, if an orthogonal array* Γ^* *satisfies* $d(\boldsymbol{\gamma}_i^*, \boldsymbol{\gamma}_k^*) > \delta$ *for any two rows* $\boldsymbol{\gamma}_i^*$ *and* $\boldsymbol{\gamma}_k^*$ *of* Γ^**, then* Γ^* *is weakly universally optimum relative to* Ξ*. In this case,* $\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = (\sigma^2/N)I_m$ *holds, which does not depend on the correlations of errors.*

Remark. A special case of Theorem 2.1 was given without proof by Jimbo (1986).

Proof of Theorem 2.1. Let X be the design matrix corresponding to an orthogonal array $\Gamma \in \Xi$. In this case, $X^t X = N I_m$ and $X^t J_N = 0$. Then it is easy to show that

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = (X^t X)^{-1} X^t \Sigma X (X^t X)^{-1} = \frac{1}{N^2} \sum_{i,k} \mathrm{cov}(\varepsilon_i, \varepsilon_k) \boldsymbol{x}_i^t \boldsymbol{x}_k, \tag{2.1}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{x}_k$ are the i-th and the k-th rows of X. Thus we obtain

$$\begin{aligned} \mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) = \mathrm{tr}(D(\Gamma)) &= \frac{1}{N^2} \sum_{i,k} \mathrm{cov}(\varepsilon_i, \varepsilon_k) \mathrm{tr}(\boldsymbol{x}_i^t \boldsymbol{x}_k) \\ &= \frac{m}{N}\sigma^2 + \frac{1}{N^2} \sum_{i \neq k} \mathrm{cov}(\varepsilon_i, \varepsilon_k)(m - 2d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k)) \\ &\geq \frac{m}{N}\sigma^2, \end{aligned}$$

since $\mathrm{tr}(\boldsymbol{x}_i^t \boldsymbol{x}_k) = \mathrm{tr}(\boldsymbol{x}_k \boldsymbol{x}_i^t) = m - 2d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k)$. We also obtain $\mathrm{tr}(D(\Gamma^*)) = (m/N)\sigma^2$, since $d(\boldsymbol{\gamma}_i^*, \boldsymbol{\gamma}_k^*) > \delta$ for any i and k. Hence the theorem is proved by Proposition 1.2. $\square$

Here, we consider a construction for orthogonal arrays which satisfy the condition of Theorem 2.1. If all the row vectors of Γ form an n-dimensional linear subspace of $GF(2)^m$, Γ is called a *linear orthogonal array*. In this case, $N = 2^n$ for some integer $n \leq m$. Usually, the following method is used to construct linear orthogonal arrays.

Let $g_1, \ldots, g_m$ be distinct n-dimensional non-zero column vectors in the vector space $GF(2)^n$ and let $G = [g_1, \ldots, g_m]$. By arranging N $(= 2^n)$ m-dimensional row vectors θG $(\theta \in GF(2)^n)$ into an $N \times m$ matrix we obtain an $OA(N, m, 2, 2)$. In this case, the $OA(N, m, 2, 2)$ is called a linear orthogonal array *generated from* G.

Theorem 2.2 *Let $G = [I_n \vdots K]$ be an $n \times m$ matrix over $GF(2)$ and let Γ be a linear $OA(2^n, m, 2, 2)$ generated from G. If every δ column vectors of the $(m-n)\times m$ matrix $H = [I_{m-n} \vdots K^t]$ are linearly independent, then Γ has minimum distance at least $\delta + 1$ and under Covariance structure I, Γ is weakly universally optimum with respect to the set Ξ of orthogonal arrays.*

Proof. This is a direct consequence of the well-known result of coding theory that a linear code with a parity check matrix H has minimum distance at least $\delta+1$ if any δ distinct column vectors of H are linearly independent (see, for example, MacWilliams and Sloane (1977)). $\square$

3. Covariance structure depending only on the Hamming distance

In this section, we assume that the correlation between errors depends only on the Hamming distance and discuss some properties of covariance matrices for the OLSE and the GLSE of $\boldsymbol{\alpha}$ when a linear orthogonal array is utilized. Furthermore, a sufficient condition for the covariance matrices for the OLSE and the GLSE to coincide is provided. The results obtained in this section will be used in the next two sections to show the optimality of a certain type of linear orthogonal arrays.

$$\mathrm{cov}(\varepsilon_i, \varepsilon_k) = \begin{cases} \sigma^2 & \text{if } i = k, \\ \sigma^2 \rho_{d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k)} & \text{if } i \neq k. \end{cases} \tag{3.1}$$

This type of covariance structure was first introduced by Kiefer and Wynn (1981) just in the case when $\rho_1 \neq 0$ and $\rho_2 = \cdots = \rho_m = 0$ as a "nearest neighbor"(NN) correlation structure of an "m-way layout with m factors (treatments)".

For an array Γ of assemblies, define $N \times N$ matrices $D_\ell = (d_{ik}^{(\ell)})$ $(\ell = 0, 1, \ldots, m)$ as follows:

$$d_{ik}^{(\ell)} = \begin{cases} 1 & \text{if } d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k) = \ell, \\ 0 & \text{otherwise.} \end{cases}$$

D_ℓ is called the *ℓ-th adjacency matrix* of Γ. Then by using the adjacency matrices D_ℓ, the covariance matrix (3.1) of errors can be written as

$$\mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma = \sigma^2 (I_N + \rho_0 (D_0 - I_N) + \sum_{\ell=1}^{m} \rho_\ell D_\ell), \tag{3.2}$$

where ρ_ℓ $(\ell = 0, 1, \ldots, m)$ is the correlation coefficient for two experiments whose Hamming distance is ℓ. In usual cases, it is natural to be assumed that $\rho_0 \geq \rho_1 \geq \cdots \geq \rho_m \geq 0$.

If the number m of constraints (factors) is large, constructions for non-linear orthogonal arrays are so complicated that we will focus our attention on linear orthogonal arrays in the sequel of this paper. Since there are no repeated rows in a linear orthogonal array Γ, the 0-th adjacency matrix of Γ is represented as $D_0 = I_N$, that is, if $d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_k) = 0$, then $i = k$. For convenience, let $\rho_0 = 1$. Then the covariance matrix (3.2) is rewritten as

$$\mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma = \sigma^2 \left(I_N + \sum_{\ell=1}^{m} \rho_\ell D_\ell \right) = \sigma^2 \sum_{\ell=0}^{m} \rho_\ell D_\ell. \tag{3.3}$$

For a linear orthogonal array Γ, the number of rows $\boldsymbol{\gamma}_j$ which is at distance ℓ from a given row $\boldsymbol{\gamma}_i$ is constant not depending on the choice of $\boldsymbol{\gamma}_i$ because $d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_j) = w(\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_j)$. Thus $D_\ell \mathbf{1}_N = c_\ell \mathbf{1}_N$, where c_ℓ is the number of rows in Γ with Hamming weight ℓ. Hence $\Sigma \mathbf{1}_N = c\mathbf{1}_N$ for a constant c, which implies that $\Sigma^{-1}\mathbf{1}_N = c^{-1}\mathbf{1}_N$ since $\mathbf{1}_N = \Sigma^{-1}\Sigma\mathbf{1}_N = c\Sigma^{-1}\mathbf{1}_N$. Noting this fact and $X^t\mathbf{1}_N = 0$, the covariance matrix (1.4) for the GLSE of $\boldsymbol{\alpha}$ is reduced to

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = (X^t\Sigma^{-1}X)^{-1}.$$

Before stating the next theorem, we will provide two lemmas. For any given $m \times m$ permutation matrix Q if there exists an $N \times N$ permutation matrix P such that

$$\Gamma Q = P\Gamma,$$

then Γ is said to be *invariant* with respect to any column permutations.

Lemma 3.1 *Let Γ be an $N \times m$ linear orthogonal array which is invariant with respect to any column permutations and let A be an $N \times N$ matrix whose (i,j)-th element depends only on the Hamming distance between the i-th and the j-th row vectors of Γ. If A is not singular, the elements of A^{-1} also depend only on the Hamming distance between the row vectors of Γ.*

Proof. Let $\boldsymbol{\gamma}_i$ be the i-th row of Γ. Transform every row vector $\boldsymbol{\gamma}_i$ by adding a given vector $\boldsymbol{\gamma}_h$ of Γ. Since the set of all the row vectors of Γ forms a linear subspace of $GF(2)^m$, there exists exactly one k such that $\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_h = \boldsymbol{\gamma}_k$ for any i. Thus the transformation induces a permutation on the subspace. Let P_1 be a permutation matrix corresponding to the transformation. Then P_1 exchanges every two rows $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_k$ of Γ such that $\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_h = \boldsymbol{\gamma}_k$. It is easy to show that

$$P_1 A P_1^t = A, \tag{3.4}$$

since $d(\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_h, \boldsymbol{\gamma}_j + \boldsymbol{\gamma}_h) = d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_j)$ and the (i,j)-th element a_{ij} of A depends only on the Hamming distance between the i-th and the j-th rows of Γ. On the other hand, let P_2 be an $N \times N$ permutation matrix induced by a column permutation of Γ. Then this type of permutation also satisfies (3.4).

Assume that $d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_j) = d(\boldsymbol{\gamma}_k, \boldsymbol{\gamma}_\ell)$, then $a_{ij} = a_{k\ell}$. By combining two types of permutations P_1 and P_2 we can obtain a permutation P which exchanges the (i,j)-th and the (k,ℓ)-th elements of A each other. Needless to say, $PAP^t = A$ holds. Since $P^tP = PP^t = I_N$, we have

$$A(PA^{-1}P^t) = (PAP^t)(PA^{-1}P^t) = PAA^{-1}P^t = I_N.$$

Hence

$$PA^{-1}P^t = A^{-1}.$$

Thus the (i,j)-th element and the (k,ℓ)-th element of A^{-1} should coincide, which means that A^{-1} is also a matrix whose (i,j)-th element depends only on the Hamming distance between the i-th and the j-th rows of Γ. $\square$

Lemma 3.2 *Let Γ be an $N \times m$ linear orthogonal array and let $A = (a_{ij})$ be an $N \times N$ matrix, where a_{ij} depends only on the Hamming distance between the i-th and the j-th rows of Γ, defined by $a_{ij} = \upsilon_{d(\gamma_i, \gamma_j)}$. Suppose that X is the design matrix corresponding to Γ. Then*

$$AX = X \sum_{i=1}^{N} \upsilon_{w(\gamma_i)} Z_{\boldsymbol{x}_i}$$

holds, where $Z_{\boldsymbol{x}_i} = -\mathrm{diag}(x_{i1}, \ldots, x_{im})$ for the i-th row $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{im})$ of X.

Proof. Define $N \times N$ matrices $R_i = (r^i_{jk})$ as follows:

$$r^i_{jk} = \begin{cases} 1 & \text{if } \boldsymbol{\gamma}_j + \boldsymbol{\gamma}_k = \boldsymbol{\gamma}_i, \\ 0 & \text{otherwise.} \end{cases}$$

Since the set of row vectors of Γ is a linear subspace of $GF(2)^m$, for a given $\boldsymbol{\gamma}_i$, there exists a unique $\boldsymbol{\gamma}_k$ such that $\boldsymbol{\gamma}_j + \boldsymbol{\gamma}_k = \boldsymbol{\gamma}_i$ for any $\boldsymbol{\gamma}_j$. Thus R_i is a permutation matrix which exchanges every two rows $\boldsymbol{x}_j$ and $\boldsymbol{x}_k$ of X such that $\boldsymbol{\gamma}_j + \boldsymbol{\gamma}_k = \boldsymbol{\gamma}_i$. Accordingly the matrix A can be represented as

$$A = \sum_{i=1}^{N} \upsilon_{w(\boldsymbol{\gamma}_i)} R_i,$$

and then

$$AX = \sum_{i=1}^{N} \upsilon_{w(\boldsymbol{\gamma}_i)} R_i X.$$

Furthermore, by noting that $R_i X$ can be obtained by adding $\boldsymbol{\gamma}_i$ to every row of Γ and rewriting 0 and 1 to -1 and 1, respectively, $R_i X = X Z_{\boldsymbol{x}_i}$ is observed. Thus

$$AX = X \sum_{i=1}^{N} \upsilon_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i}.$$

We have just completed the proof of Lemma 3.2. □

As an immediate consequence, we have the following theorem.

Theorem 3.1 *Let Γ be an $N \times m$ linear orthogonal array and let X be the design matrix corresponding to Γ. Under the covariance structure* (3.3), *for the OLSE $\hat{\boldsymbol{\alpha}}_O$,*

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}_O) = \frac{1}{N^2} X^t \Sigma X = \frac{\sigma^2}{N} \sum_{i=1}^{N} \rho_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i} \tag{3.5}$$

holds.

Now we can show the following result for the GLSE of $\boldsymbol{\alpha}$.

Theorem 3.2 *Besides the assumption of Theorem 3.1, assume that the (i, j)-th element of Σ^{-1} depends only on the Hamming distance between the i-th and the j-th row vectors of Γ. Then the covariance matrix for the GLSE $\hat{\boldsymbol{\alpha}}_G$ coincides with that for the OLSE $\hat{\boldsymbol{\alpha}}_O$ of the form* (3.5).

Remark. For example, if Γ is a linear orthogonal array which is invariant with respect to any column permutations, then Σ^{-1} satisfies the condition of Theorem 3.2 by virtue of Lemma 3.1.

Proof of Theorem 3.2. Since $\Sigma = \sigma^2 \sum_{\ell=0}^{m} \rho_\ell D_\ell$ for the adjacency matrices D_ℓ ($\ell = 0, 1, \ldots, m$) of Γ, under the covariance structure (3.3)

$$\Sigma X = \sigma^2 X \sum_{i=1}^{N} \rho_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i}$$

follows from Lemma 3.2. Similarly, by using Lemma 3.2 with the assumption on Σ^{-1},

$$\Sigma^{-1} X = \frac{1}{\sigma^2} X \sum_{i=1}^{N} \nu_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i}$$

is obtained, where ν_ℓ ($\ell = 1, 2, \ldots, m$) is a certain series of constants satisfying $\Sigma^{-1} = (1/\sigma^2) \sum_{\ell=0}^{m} \nu_\ell D_\ell$. Since $X^t X = N I_m$, we have

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}_G) = (X^t \Sigma^{-1} X)^{-1} = \left(\frac{N}{\sigma^2} \sum_{i=1}^{N} \nu_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i} \right)^{-1}. \qquad (3.6)$$

Furthermore, by noting $(\Sigma X)^t (\Sigma^{-1} X) = X^t \Sigma \Sigma^{-1} X = N I_m$,

$$\sum_{i=1}^{N} \nu_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i} = \left(\sum_{i=1}^{N} \rho_{w(\boldsymbol{\gamma}_i)} Z_{\boldsymbol{x}_i} \right)^{-1} \qquad (3.7)$$

holds. Hence the theorem is proved by (3.5), (3.6) and (3.7). □

4. Optimum factorial designs for the nearest neighbor covariance structure

In this section, we consider the same covariance structure that Kiefer and Wynn (1981) treated as a "nearest neighbor" (NN) correlation structure, i.e., $\rho_1 = \rho$ (≥ 0) and $\rho_2 = \cdots = \rho_m = 0$ in (3.3), which is a kind of the "moving-average" model of order 1.

Covariance structure II:

$$\mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma = \sigma^2 (I_N + \rho D_1).$$

Let Γ be a $2^n \times m$ linear orthogonal array and s be the number of row vectors with Hamming weight 1 in Γ. Without loss of generality, we may assume that in Γ there exists exactly one row vector with Hamming weight 1 such that the i-th coordinate is 1 for $i = 1, \ldots, s$. Then the subspace W consisting of the rows of Γ is expressed as a direct sum of the s-dimensional linear space $W_1 = GF(2)^s$ and an $(n-s)$-dimensional linear subspace W_2 of the linear space $GF(2)^{m-s}$, that is,

$$W = W_1 \oplus W_2. \qquad (4.1)$$

Obviously the Hamming weight of each row vector in W_2 is 0 or greater than 1. Let X_s be the $2^s \times s$ matrix obtained by arranging the row vectors of W_1 and by rewriting the elements 0 and 1 to -1 and 1, respectively. Similarly, let U be the $2^{n-s} \times (m-s)$ matrix obtained by arranging the vectors of W_2 and by rewriting the elements 0 and 1 to -1 and 1, respectively. For the design matrix X corresponding to Γ, without loss of generality, we assume that the first row of X is $(-1, \ldots, -1)$ and write

$$X = \left[\mathbf{1}_{2^{n-s}} \otimes X_s \vdots U \otimes \mathbf{1}_{2^s} \right], \tag{4.2}$$

where $\mathbf{1}_\ell$ is the ℓ-dimensional all-one column vector and $\otimes$ indicates a direct product.

Let $D'_0(= I_{2^s}), D'_1, \ldots, D'_s$ be the adjacency matrices of W_1 and let $V = D'_0 + \rho D'_1$. Since (4.2) implies that X can be divided into 2^{n-s} subblocks with 2^s rows each and the covariance of every two rows contained in distinct subblocks is 0, $\mathrm{cov}(\boldsymbol{\varepsilon})$ is a block diagonal matrix as follows:

$$\mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma = \sigma^2 \mathrm{diag}(\underbrace{V, \ldots, V}_{2^{n-s}}). \tag{4.3}$$

A preliminary result is needed for further discussion.

Lemma 4.1 *Let X be the design matrix of the form* (4.2) *corresponding to a linear orthogonal array Γ. Let $B = (b_{ij})$ be a $2^s \times 2^s$ matrix, where b_{ij} depends only on the Hamming distance between the i-th and the j-th row vectors $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_j$ of W_1, defined by $b_{ij} = \upsilon_{d(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_j)}$. Suppose that $A = \mathrm{diag}(\underbrace{B, \ldots, B}_{2^{n-s}})$ is an $N \times N$ block diagonal matrix $(N = 2^n)$. Then*

$$AX = X \cdot \mathrm{diag}(\underbrace{\lambda, \ldots, \lambda}_{s}, \underbrace{\kappa, \ldots, \kappa}_{m-s}),$$

holds, where

$$\kappa = \sum_{\ell=0}^{s} \upsilon_\ell \binom{s}{\ell} \qquad \text{and} \qquad \lambda = \kappa - 2 \sum_{\ell=1}^{s} \upsilon_\ell \binom{s-1}{\ell-1}.$$

Proof. Let $\boldsymbol{u}_j = (u_{j1}, \ldots, u_{j,m-s})$ be the j-th row of U in (4.2). Assume that X is divided into 2^{n-s} subblocks as follows:

$$X = \begin{bmatrix} X_{(1)} \\ X_{(2)} \\ \vdots \\ X_{(2^{n-s})} \end{bmatrix},$$

where $X_{(j)} = [X_s \vdots \boldsymbol{u}_j \otimes \mathbf{1}_{2^s}]$. Then,

$$AX = \mathrm{diag}(B, \ldots, B) \cdot X = \begin{bmatrix} BX_{(1)} \\ BX_{(2)} \\ \vdots \\ BX_{(2^{n-s})} \end{bmatrix}.$$

Now, let $U_j = \mathrm{diag}(\underbrace{1, \ldots, 1}_{s}, u_{j1}, \ldots, u_{j,m-s})$. In this case,

$$BX_{(j)} = B\ [X_s \vdots \boldsymbol{u}_j \otimes \mathbf{1}_{2^s}] = B\ [X_s \vdots -J]\ U_j,$$

where J is the $2^s \times (m-s)$ all-one matrix. Let $\bar{\boldsymbol{x}}_i$ be the i-th row of $[X_s \vdots -J]$ and $\bar{\boldsymbol{\gamma}}_i$ be the vector obtained by rewriting the elements -1 and 1 of $\bar{\boldsymbol{x}}_i$ to 0 and 1, respectively. By using Lemma 3.2 we obtain

$$\begin{aligned} BX_{(j)} = B\ [X_s \vdots -J]\ U_j &= [X_s \vdots -J] \sum_{i=1}^{2^s} \upsilon_{w(\bar{\gamma}_i)} Z_{\bar{\boldsymbol{x}}_i} U_j \\ &= [X_s \vdots -J]\ U_j \sum_{i=1}^{2^s} \upsilon_{w(\bar{\gamma}_i)} Z_{\bar{\boldsymbol{x}}_i} = X_{(j)} \sum_{i=1}^{2^s} \upsilon_{w(\bar{\gamma}_i)} Z_{\bar{\boldsymbol{x}}_i}, \end{aligned}$$

where $Z_{\bar{x}_i}$ is defined also in Lemma 3.2. Therefore, it follows that

$$AX = X \sum_{i=1}^{2^s} \upsilon_{w(\bar{\gamma}_i)} Z_{\bar{\boldsymbol{x}}_i}.$$

Since the rows of X_s consist of all the s-dimensional vectors with elements ± 1, in X_s there are $\binom{s}{\ell}$ row vectors with Hamming weight ℓ and among them there are $\binom{s-1}{\ell-1}$ row vectors such that each of them has 1 at the i-th coordinate. By taking account of this fact, we obtain

$$\sum_{i=0}^{2^s} \upsilon_{w(\bar{\gamma}_i)} Z_{\bar{\boldsymbol{x}}_i} = \mathrm{diag}(\underbrace{\lambda, \ldots, \lambda}_{s}, \underbrace{\kappa, \ldots, \kappa}_{m-s}),$$

where

$$\kappa = \sum_{\ell=0}^{s} \upsilon_\ell \binom{s}{\ell} \qquad \text{and} \qquad \lambda = \kappa - 2 \sum_{\ell=1}^{s} \upsilon_\ell \binom{s-1}{\ell-1}.$$

Thus the lemma is proved. □

The following is a direct consequence of Lemma 4.1.

Theorem 4.1 *Under Covariance structure II, if Γ is an $N \times m$ linear orthogonal array, then $C(\Gamma)$ for the GLSE $\hat{\boldsymbol{\alpha}}_G$ coincides with that for the OLSE $\hat{\boldsymbol{\alpha}}_O$. Further $C(\Gamma)$ is a diagonal matrix which consists of s and $m-s$ diagonal elements with values*

$$\frac{N}{\sigma^2} \cdot \frac{1}{1+(s-2)\rho} \qquad \textit{and} \qquad \frac{N}{\sigma^2} \cdot \frac{1}{1+s\rho},$$

respectively, where s is the number of vectors with Hamming weight 1 among the row vectors of Γ.

Proof. Let X be the design matrix of the form (4.2) corresponding to Γ. Then Σ is represented by (4.3) under Covariance structure II and

$$\Sigma^{-1} = \frac{1}{\sigma^2} \mathrm{diag}(V^{-1}, \ldots, V^{-1})$$

holds. Suppose that the row vectors in W_1 of (4.1) are numbered arbitrarily. It is obvious that the (i, j)-th element of V depends only on the Hamming distance between the i-th and the j-th row vectors of W_1. Since W_1 is the linear space $GF(2)^s$, it is invariant with respect to any column permutations of coordinates. Hence we can claim that the (i, j)-th element of V^{-1} also depends only on the Hamming distance between the i-th and the j-th row vectors of W_1 (see Lemma 3.1). Since

$$\Sigma X = \sigma^2 X \cdot \mathrm{diag}(\underbrace{\lambda, \ldots, \lambda}_{s}, \underbrace{\kappa, \ldots, \kappa}_{m-s})$$

follows from Lemma 4.1, where $\kappa = \binom{s}{0} + \rho\binom{s}{1}$ and $\lambda = \kappa - 2\rho\binom{s-1}{0}$, by using Lemma 4.1 again and by noting $(\Sigma X)^t(\Sigma^{-1}X) = NI_m$, we have

$$\Sigma^{-1}X = \frac{N}{\sigma^2} X \cdot \mathrm{diag}(\lambda^{-1}, \ldots, \lambda^{-1}, \kappa^{-1}, \ldots, \kappa^{-1}).$$

Thus it follows that

$$\begin{aligned} \mathrm{cov}(\hat{\boldsymbol{\alpha}}_G) &= (X^t\Sigma^{-1}X)^{-1} = \frac{\sigma^2}{N}\mathrm{diag}(\lambda, \ldots, \lambda, \kappa, \ldots, \kappa) \\ &= \frac{1}{N^2} X^t \Sigma X = \mathrm{cov}(\hat{\boldsymbol{\alpha}}_O), \end{aligned}$$

which proves the theorem. □

Theorem 4.1 allows us to state the following theorem. Let Ξ_L be the set of $N \times m$ linear orthogonal arrays by which the main effect $\boldsymbol{\alpha}$ is estimable.

Theorem 4.2 *Let $m \geq 3$ and let Γ^* be an $N \times m$ linear orthogonal array with minimum distance at least 2. Then under Covariance structure II, Γ^* is*

(i) *universally optimum relative to Ξ_L for $0 \leq \rho < 1 - 2/m$, and*

(ii) *weakly universally optimum relative to Ξ_L for $0 \leq \rho$ (< 1)*

with respect to the OLSE and the GLSE of $\boldsymbol{\alpha}$.

Remark. Theorem 2.1 shows (ii) only for the OLSE $\hat{\boldsymbol{\alpha}}$ but under a more general covariance structure.

Proof of Theorem 4.2. Let $\Gamma \in \Xi_L$. Assume that $s \geq 1$ for Γ, where s is the number of vectors with Hamming weight 1 among the row vectors of Γ.

(i) Since the minimum distance of Γ^* is at least 2, the case when $s = 0$ in Theorem 4.1 gives $C(\Gamma^*) = (N/\sigma^2)I_m$. Thus if we can show $\mathrm{tr}(C(\Gamma)) < \mathrm{tr}(C(\Gamma^*))$ for $0 \leq \rho < 1 - 2/m$ and for $s \geq 1$, then (i) is proved by Proposition 1.1.

Let X be the design matrix corresponding to Γ. Theorem 4.1 yields

$$\mathrm{tr}(C(\Gamma)) = \frac{N}{\sigma^2}\left\{\frac{s}{1 + (s-2)\rho} + \frac{m - s}{1 + s\rho}\right\}$$

and especially,

$$\mathrm{tr}(C(\Gamma^*)) = \frac{N}{\sigma^2} m$$

for Γ^*. Since

$$\frac{s}{1+(s-2)\rho} + \frac{m-s}{1+s\rho} < m$$

for any $s \geq 1$ and for any $0 \leq \rho < 1 - 2/m$, $\mathrm{tr}(C(\Gamma)) < \mathrm{tr}(C(\Gamma^*))$. Thus the assertion of (i) follows.

(ii) In a manner similar to (i), we obtain

$$\mathrm{tr}(D(\Gamma)) = \frac{\sigma^2}{N}\{s(1+(s-2)\rho) + (m-s)(1+s\rho)\}$$

for any $\Gamma \in \Xi_L$. In case of Γ^*,

$$\mathrm{tr}(D(\Gamma^*)) = \mathrm{tr}\left(C(\Gamma^*)^{-1}\right) = \frac{\sigma^2}{N} m$$

holds. Then it is readily checked that $\mathrm{tr}(D(\Gamma)) - \mathrm{tr}(D(\Gamma^*)) > 0$ for $s \geq 1$, $0 \leq \rho$ (< 1) and $m \geq 3$. Hence (ii) is proved by Proposition 1.2. □

5. Further results in the case of large experiments for the OLSE

In this section, we assume two types of covariance structures of errors, like the "moving-average" model of order 2 and the "autoregressive" model of order 1, and consider the optimality only for the OLSE of $\boldsymbol{\alpha}$, simply denoted by $\hat{\boldsymbol{\alpha}}$, unless otherwise specified.

Covariance structure III:

$$\mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma = \sigma^2(I_N + \rho_1 D_1 + \rho_2 D_2).$$

Covariance structure IV:

$$\mathrm{cov}(\boldsymbol{\varepsilon}) = \Sigma = \sigma^2 \sum_{\ell=0}^{m} \rho^\ell D_\ell.$$

From Theorem 2.1 we know that under Covariance structure III, an orthogonal array with minimum distance 3 is weakly universally optimum for the OLSE, if it exists. Here we discuss the optimality of linear orthogonal arrays for the OLSE under Covariance structure III in the case when a linear orthogonal array with minimum distance 3 does not exist. Similar discussion will be developed under Covariance structure IV.

In order to show the following theorems, we need to present three lemmas beforehand, which are also closely related to coding theory. The first one follows from Theorem 3.1.

Lemma 5.1 *Let Γ be an $N \times m$ linear orthogonal array and let $\hat{\boldsymbol{\alpha}}$ be the OLSE of $\boldsymbol{\alpha}$. Under Covariance structure IV,* $\mathrm{cov}(\hat{\boldsymbol{\alpha}})$ *is a diagonal matrix and*

$$\frac{N}{\sigma^2}\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) = mf(\rho) - 2\rho f'(\rho)$$

holds, where $f(\rho) = \sum_{i=1}^{N} \rho^{w(\boldsymbol{\gamma}_i)}$ and $\boldsymbol{\gamma}_i$ is the i-th row vector of Γ.

Remark. $f(\rho)$ is called the *weight enumerator* of the linear subspace consisting of the rows of Γ.

The following lemmas on the weight enumerator are obtained by counting directly the number of 1 in each row of the linear code generated from G, or by applying the MacWilliams theorem concerned with a linear code and its dual code (see, for example, MacWilliams and Sloane (1977)).

Lemma 5.2 *Let $K = [\underbrace{0,\ldots,0}_{s},\underbrace{1,\ldots,1}_{m-s-1}]^t$ and $G = [I_{m-1} \vdots K]$, then the weight enumerator of the linear code generated from G is*

$$f(x) = \frac{(1+x)^m}{2}\left\{1 + \left(\frac{1-x}{1+x}\right)^{m-s}\right\}.$$

Lemma 5.3 *Let*

$$K = \begin{bmatrix} 0\cdots0 & 1\cdots1 & 1\cdots1 & 0\cdots0 \\ \underbrace{0\cdots0}_{s} & \underbrace{0\cdots0}_{a-1} & \underbrace{1\cdots1}_{b} & \underbrace{1\cdots1}_{c-1} \end{bmatrix}^t$$

and $G = [I_{m-2} \vdots K]$, then the weight enumerator of the linear code generated from G is given by

$$f(x) = \frac{(1+x)^m}{4}\{1 + (\delta^{a+b} + \delta^{b+c} + \delta^{c+a})\},$$

where $s + a + b + c = m$ and $\delta = (1-x)/(1+x)$.

Now, we can show the following two theorems.

Theorem 5.1 *Let $K^* = [\underbrace{1,\ldots,1}_{m-1}]^t$ and let Γ^* be the $2^{m-1} \times m$ linear orthogonal array generated from $G^* = [I_{m-1} \vdots K^*]$ for $m \geq 4$.*

(i) *Under Covariance structure III, Γ^* is weakly universally optimum relative to the set Ξ_L of $2^{m-1} \times m$ linear orthogonal arrays for $0 \leq \rho_2 < \frac{m-2}{(m-1)(m-4)}\rho_1$ (< 1), and then*

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{\sigma^2}{2^{m-1}}\left\{1 + \frac{(m-1)(m-4)}{2}\rho_2\right\} I_m. \tag{5.1}$$

(ii) *Under Covariance structure IV, Γ^* is weakly universally optimum relative to Ξ_L for $0 \leq \rho$ (< 1), and then*

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{\sigma^2}{2^m}(1-\rho^2)\left\{(1+\rho)^{m-2} + (1-\rho)^{m-2}\right\} I_m. \tag{5.2}$$

Proof. It is easy to see that the Hamming weight of every row of Γ^* is even. Therefore Γ^* is invariant for any column permutations. In this case, (3.5) in Theorem 3.1 is reduced to $\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = aI_{2^{m-1}}$ for some constant a under each of the covariance structures III and IV. Hence, in order to prove the theorem, we have only to show that $\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}}))$ attains the minimum at $s = 0$. Let $K = [\underbrace{0,\ldots,0}_{s},\underbrace{1,\ldots,1}_{m-s-1}]^t$ and let Γ be the $2^{m-1} \times m$ linear orthogonal array generated from $G = [I_{m-1} \vdots K]$.

(i) By counting the right-hand side of (3.5),

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{\sigma^2}{2^{m-1}}\mathrm{diag}(\underbrace{a,\ldots,a}_{s},\underbrace{b,\ldots,b}_{m-s})$$

can be obtained, where

$$a = 1 + \left\{\binom{s}{1} - 2\right\}\rho_1 + \left\{\binom{s}{2} - 2\binom{s-1}{1} + \binom{m-s}{2}\right\}\rho_2$$

and

$$b = 1 + \binom{s}{1}\rho_1 + \left\{\binom{s}{2} + \binom{m-s}{2} - 2\binom{m-s-1}{1}\right\}\rho_2.$$

Hence we have

$$\begin{aligned}\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) &= \frac{\sigma^2}{2^{m-1}}\{sa + (m-s)b\} \\ &= \frac{\sigma^2}{2^{m-1}}\rho_2(m-4)\left(s + \frac{(m-2)\rho_1 - m(m-4)\rho_2}{2(m-4)\rho_2}\right)^2 + \Delta,\end{aligned}$$

where Δ is a term which does not contain s. When

$$-\frac{(m-2)\rho_1 - m(m-4)\rho_2}{2(m-4)\rho_2} < \frac{1}{2},$$

that is, when $0 \leq \rho_2 < \frac{m-2}{(m-1)(m-4)}\rho_1$ (< 1), it is readily checked that the value of $\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}}))$ attains the minimum at $s = 0$. Thus Γ^* is weakly universally optimum relative to Ξ_L and we obtain (5.1). Hence (i) is proved.

(ii) By Lemmas 5.1 and 5.2 we have

$$\begin{aligned}\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) &= \frac{\sigma^2}{2^{m-1}}\{mf(\rho) - 2\rho f'(\rho)\} \\ &= \frac{\sigma^2}{2^{m-1}}\left\{\frac{m(1+\rho)^m}{2} - m\rho(1+\rho)^{m-1} + \frac{(1+\rho)^{m-1}}{2(1-\rho)}Q(s)\right\},\end{aligned}$$

where $Q(s) = \{(1-\rho)/(1+\rho)\}^{m-s}\{m(1+\rho)^2 - 4\rho s\}$. If $Q(s) \leq Q(s+1)$ for any $s \geq 0$, we can conclude that Γ^* is weakly universally optimum relative to Ξ_L by Proposition 1.2. Here, the inequality $Q(s) \leq Q(s+1)$ can be reduced to

$$s \leq \frac{(1+\rho)\{m(1+\rho)-2\}}{4\rho}. \tag{5.3}$$

Since the dimension of the linear subspace generated from G^* is $m-1$, the left-hand side of (5.3) is not less than $m-2$, that is, $0 \leq s \leq m-2$. Therefore, (5.3) holds for any $0 \leq \rho < 1$ and we have (5.2) for $s = 0$. Hence (ii) is proved. □

Remark. (a) Γ^* is a linear orthogonal array of strength $m-1$ and it is the even parity code if it is regarded as a linear code.
(b) Let $\hat{\boldsymbol{\alpha}}_O$ and $\hat{\boldsymbol{\alpha}}_G$ be the OLSE and the GLSE of the main effect $\boldsymbol{\alpha}$, respectively. Since Γ^* is invariant with respect to any column permutations, $\mathrm{cov}(\hat{\boldsymbol{\alpha}}_O) = \mathrm{cov}(\hat{\boldsymbol{\alpha}}_G)$ can be shown for Γ^* by using Lemma 3.1 and Theorem 3.2.

Theorem 5.2 *For $m = 3p$ and $p \geq 2$, let*

$$K^* = \left[\begin{array}{ccc} 1\cdots1 & 1\cdots1 & 0\cdots0 \\ \underbrace{0\cdots0}_{p-1} & \underbrace{1\cdots1}_{p} & \underbrace{1\cdots1}_{p-1} \end{array}\right]^t$$

and let Γ^* be the $2^{m-2} \times m$ linear orthogonal array generated from $G^* = [I_{m-2} \vdots K^*]$.

(i) *Under Covariance structure III, Γ^* is weakly universally optimum relative to the set Ξ_L of $2^{m-2} \times m$ linear orthogonal arrays for $0 \leq \rho_2 < \frac{3}{m-4}\rho_1$, and then*

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{\sigma^2}{2^{m-2}}\left\{1 + \frac{(m-3)(m-4)}{6}\rho_2\right\} I_m. \tag{5.4}$$

(ii) *Under Covariance structure IV, Γ^* is weakly universally optimum relative to Ξ_L for $0 \leq \rho$ (< 1), and then*

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{\sigma^2}{2^m} \cdot \frac{(1+\rho)^{m-1}}{1-\rho}\left\{(1-\rho)^2 + (3+2\rho+3\rho^2)\left(\frac{1-\rho}{1+\rho}\right)^{\frac{2}{3}m}\right\} I_m. \tag{5.5}$$

Remark. In this case, Γ^* is a linear orthogonal array of strength $2p-1$, which implies that Γ^* has the maximum strength among $2^{3p-2} \times 3p$ linear orthogonal arrays since the minimum distance of Γ^* is 2. In other words, if Γ^* is regarded as a linear code, the probability of detecting errors is maximum among all $(3p-2)$-dimensional linear codes with length $3p$.

Proof of Theorem 5.2. Let

$$H = [I_2 \vdots K^{*t}] = \left[\begin{array}{ccccc} 1 & 0 & 1\cdots1 & 1\cdots1 & 0\cdots0 \\ 0 & 1 & \underbrace{0\cdots0}_{p-1} & \underbrace{1\cdots1}_{p} & \underbrace{1\cdots1}_{p-1} \end{array}\right].$$

Then H is a parity check matrix of Γ^* when Γ^* is regarded as a linear code.

Let $E_1 = \{1, 3, 4, \ldots, p+1\}$, $E_2 = \{p+2, \ldots, 2p+1\}$ and $E_3 = \{2, 2p+2, \ldots, 3p\}$. Then $\boldsymbol{\gamma}_i = (\gamma_{i1}, \ldots, \gamma_{im})$ is a row (codeword) of Γ^* if and only if

$$\sum_{j \in E_1 \cup E_2} \gamma_{ij} = 0 \qquad \text{and} \qquad \sum_{j \in E_2 \cup E_3} \gamma_{ij} = 0 \tag{5.6}$$

over $GF(2)$. The condition (5.6) is equivalent to

$$\sum_{j \in E_1} \gamma_{ij} = \sum_{j \in E_2} \gamma_{ij} = \sum_{j \in E_3} \gamma_{ij}.$$

Therefore the Hamming weights for the respective sets E_1, E_2 and E_3 of coordinates in each row of Γ^* are odd or even simultaneously. By this symmetry, it is easy to show that

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = D(\Gamma^*) = aI_{2^{m-2}}$$

for some constant a under each of Covariance structures III and IV. In the rest of the proof, we shall show that $\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) = \mathrm{tr}(D(\Gamma^*))$ attains the minimum at $s = 0$ for each case of (i) and (ii). Let

$$K = \begin{bmatrix} 0 \cdots 0 & 1 \cdots 1 & 1 \cdots 1 & 0 \cdots 0 \\ \underbrace{0 \cdots 0}_{s} & \underbrace{0 \cdots 0}_{a-1} & \underbrace{1 \cdots 1}_{b} & \underbrace{1 \cdots 1}_{c-1} \end{bmatrix}^t,$$

where $s + a + b + c = m$, and let Γ be the $2^{m-2} \times m$ linear orthogonal array generated from $G = [I_{m-2} \vdots K]$.

(i) By virtue of Theorem 3.1,

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) = \frac{\sigma^2}{2^{m-2}} \mathrm{diag}(\underbrace{t, \ldots, t}_{s}, \underbrace{u, \ldots, u}_{a}, \underbrace{v, \ldots, v}_{b}, \underbrace{w, \ldots, w}_{c})$$

is obtained, where

$$\begin{aligned}
t &= 1 + \left\{ \binom{s}{1} - 2 \right\} \rho_1 + \left\{ \binom{s}{2} - 2\binom{s-1}{1} + \binom{a}{2} + \binom{b}{2} + \binom{c}{2} \right\} \rho_2, \\
u &= 1 + \binom{s}{1} \rho_1 + \left\{ \binom{s}{2} + \binom{a}{2} - 2\binom{a-1}{1} + \binom{b}{2} + \binom{c}{2} \right\} \rho_2, \\
v &= 1 + \binom{s}{1} \rho_1 + \left\{ \binom{s}{2} + \binom{a}{2} + \binom{b}{2} - 2\binom{b-1}{1} + \binom{c}{2} \right\} \rho_2, \\
w &= 1 + \binom{s}{1} \rho_1 + \left\{ \binom{s}{2} + \binom{a}{2} + \binom{b}{2} + \binom{c}{2} - 2\binom{c-1}{1} \right\} \rho_2.
\end{aligned}$$

Immediately,

$$\begin{aligned}
\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) &= \frac{\sigma^2}{2^{m-2}} (st + au + bv + cw) \\
&= \frac{\sigma^2}{2^{m-2}} \left\{ m + s(m-2)\rho_1 + \frac{m-4}{2}(s^2 + a^2 + b^2 + c^2 - m)\rho_2 \right\}
\end{aligned}$$

is shown. By noting that $a^2+b^2+c^2 \geq (a+b+c)^2/3$ holds with the equality if and only if $a=b=c(=(m-s)/3)$, we have

$$\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) \geq \frac{\sigma^2}{2^{m-2}} \cdot \frac{2(m-4)}{3}\rho_2 \left(s + \frac{3(m-2)\rho_1 - m(m-4)\rho_2}{4(m-4)\rho_2} \right)^2 + \Delta,$$

where Δ is the term not containing s. If

$$-\frac{3(m-2)\rho_1 - m(m-4)\rho_2}{4(m-4)\rho_2} < \frac{1}{2},$$

then the value of $\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}}))$ attains the minimum at $s=0$, that is, Γ^* is weakly universally optimum relative to Ξ_L for $0 \leq \rho_2 < \frac{3}{m-4}\rho_1$, and then we obtain (5.4). Hence (i) is proved.

(ii) From Lemmas 5.1 and 5.3,

$$\begin{aligned}
\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) &= \frac{\sigma^2}{2^{m-2}}\{mf(\rho) - 2\rho f'(\rho)\} \\
&= \frac{\sigma^2}{2^{m-2}} \left\{ \frac{m(1+\rho)^{m-1}}{4}(1-\rho)(1+\delta^{a+b}+\delta^{b+c}+\delta^{c+a}) \right. \\
&\quad \left. + \frac{(1+\rho)^{m-1}}{1-\rho}\rho\left((a+b)\delta^{a+b} + (b+c)\delta^{b+c} + (c+a)\delta^{c+a}\right) \right\}
\end{aligned}$$

follows. It is easy to see that

$$\delta^{a+b} + \delta^{b+c} + \delta^{c+a} \geq 3(\delta^{a+b+c})^{\frac{2}{3}} (= 3\delta^{\frac{2}{3}(m-s)}) \tag{5.7}$$

and

$$\begin{aligned}
&(a+b)\delta^{a+b} + (b+c)\delta^{b+c} + (c+a)\delta^{c+a} \\
&\quad \geq 2\sqrt{ab}\cdot\delta^{a+b} + 2\sqrt{bc}\cdot\delta^{b+c} + 2\sqrt{ca}\cdot\delta^{c+a} \\
&\quad \geq 6(abc\delta^{2(m-s)})^{\frac{1}{3}}
\end{aligned} \tag{5.8}$$

for $0 \leq s \leq m-2$. The equalities in (5.7) and (5.8) hold if and only if $a=b=c=(m-s)/3$. Furthermore, the right-hand sides of (5.7) and (5.8) attain the minimum simultaneously at $s=0$. In the case when $s=0$,

$$\mathrm{tr}(\mathrm{cov}(\hat{\boldsymbol{\alpha}})) = \frac{\sigma^2}{2^m}\cdot\frac{m(1+\rho)^{m-1}}{1-\rho}\left\{(1-\rho)^2 + (3+2\rho+3\rho^2)\left(\frac{1-\rho}{1+\rho}\right)^{\frac{2}{3}m}\right\}$$

holds and then (5.5) is obtained immediately. Hence (ii) is proved. □

6. Complete factorial designs

In this section, as a by-product, we provide the covariance matrices for the OLSE and the GLSE of $\boldsymbol{\alpha}$ in the case of complete factorial designs under some covariance structures. In the case of complete factorial designs with m factors, the rows of Γ form the linear space $GF(2)^m$, thus Γ is invariant with respect to any column permutations. Hence by applying Lemma 3.1, Theorem 3.2 and Lemma 4.1 for $s=m=n$, we have the following theorem and its corollary.

Theorem 6.1 *For the complete factorial design with m factors, under the covariance structure* (3.1)*, the covariance matrices for the OLSE $\hat{\boldsymbol{\alpha}}_O$ and for the GLSE $\hat{\boldsymbol{\alpha}}_G$ coincide, and then*

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}_O) = \mathrm{cov}(\hat{\boldsymbol{\alpha}}_G) = \frac{\sigma^2}{N}\left\{\sum_{\ell=0}^{m}\rho_\ell\binom{m}{\ell} - 2\sum_{\ell=1}^{m}\rho_\ell\binom{m-1}{\ell-1}\right\} I_m$$

holds.

Corollary 6.1 *For the complete factorial design with m factors, under the covariance structures II, III and IV, the covariance matrices for the OLSE $\hat{\boldsymbol{\alpha}}_O$ and for the GLSE $\hat{\boldsymbol{\alpha}}_G$ are given as follows*:

$$\text{II: } \mathrm{cov}(\hat{\boldsymbol{\alpha}}_O) = \mathrm{cov}(\hat{\boldsymbol{\alpha}}_G) = \frac{\sigma^2}{2^m}\{1 + (m-2)\rho\} I_m,$$

$$\text{III: } \mathrm{cov}(\hat{\boldsymbol{\alpha}}_O) = \mathrm{cov}(\hat{\boldsymbol{\alpha}}_G) = \frac{\sigma^2}{2^m}\left\{1 + (m-2)\rho_1 + \frac{(m-1)(m-4)}{2}\rho_2\right\} I_m,$$

$$\text{IV: } \mathrm{cov}(\hat{\boldsymbol{\alpha}}_O) = \mathrm{cov}(\hat{\boldsymbol{\alpha}}_G) = \frac{\sigma^2}{2^m}(1+\rho)^{m-1}(1-\rho) I_m.$$

Acknowledgement

The authors are thankful to the referees for their scrutiny of the manuscript and their helpful comments.

References

Beth, T., D. Jungnickel and H. Lenz (1985). *Design Theory.* Bibliographisches Institute-Wissenschftsverlag.

Cheng, C. S. (1983). Construction of optimal balanced incomplete block designs for correlated observations. *Ann. Statisit.* **11**, 240–246.

Jimbo, M. (1986). Robustness of orthogonal arrays for correlated errors. *The 2-nd Japan-China Symposium on Statistics, Kyusyu Univ..*

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In: J. N. Srivastava, Ed., *A Survey of Statistical Designs and Linear Models.* North Holland, Amsterdam, 333–353.

Kiefer, J. and H. P. Wynn (1981). Optimum balanced incomplete block and Latin square designs for correlated observation. *Ann. Statist.* **9**, 737–757.

MacWilliams, F. J. and N. J. A. Sloane (1977). *The theory of error-correcting codes.* North-Holland, New York.

Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments.* Wiley, New York.