

An algorithm for finding exact D - and A - optimal designs with n observations and k two-level factors in the presence of autocorrelated errors

E. Bora-Senta and C. Moyssiadis*

Department of Mathematics
Aristotle University of Thessaloniki
54006 Thessaloniki. Greece

ABSTRACT. Exact designs with n observations and k two level factors in the presence of autocorrelated errors are considered. The problem of finding D - and A - optimal designs, is discussed. An algorithm for constructing such designs, using exhaustive search for different values of n and k , is developed. The application of this algorithm showed, that in the case of positive autocorrelation, the maximum possible number of interchanges of the factor levels provides almost optimal designs. On the contrary, in the case of negative autocorrelation, the minimum such number provides almost optimal designs. A list of the exact D - and A - optimal designs is given.

1 Introduction

In optimal design theory we are interested in constructing designs which are the best, in some sense, among all possible ones. There are many criteria of optimality which usually assume that the error terms involved, are uncorrelated with zero mean and common variance. In this paper we consider exact designs, having k two-level factors, when a given number of n ($n \geq k + 1$) observations must be taken and the errors are autocorrelated.

If we denote by \mathbf{Y} the vector of observed responses, we assume the following linear main effects model

$$\mathbf{Y} = X \boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (1.1)$$

*This research was partially supported by the grant No 1669/ΠΕΝΕΔ'95 of the secretariat for research and technology of Greece.

where $X = (x_{ij})$ is the $n \times (k + 1)$ design matrix ($i = 1, 2, \dots, n$ and $j = 0, 1, \dots, k$) with $x_{i0} = 1$ for every i and $x_{ij} = +1$ or -1 for every i and $j \neq 0$, $\beta = (\beta_0, \beta_1, \dots, \beta_k)'$ the vector of the unknown parameters (overall mean and main effects) and $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)'$ the vector of random errors. By assuming that the error terms are uncorrelated with zero mean and constant variance, the best linear unbiased estimators (BLUE's) for the elements of β are given by

$$\hat{\beta} = (X' X)^{-1} X' Y \tag{1.2}$$

where $C = X' X$ is the information matrix.

Under these assumptions a lot of research has been done on the problem of finding optimal designs, and miscellaneous results are available. See for example Ehlich ([8],[9]), Galil and Kiefer ([10], [11]), Cheng ([4]), Moysiadias and Kounias ([14]) etc. An extensive review can be found in Shah and Sinha ([15]).

However, in many cases where the observations are taken sequentially, it may be reasonable to suppose that the successive elements of vector ϵ have some kind of autocorrelation. Let the ϵ_i 's follow a first-order autoregressive model (AR(1)), i.e.

$$\epsilon_i = \rho \epsilon_{i-1} + \eta_i, \quad i = 1, 2, \dots, n, \quad \epsilon_0 = 0$$

where $E(\eta) = 0$, $V(\eta) = \sigma_\eta^2 I_n$ and $-1 < \rho < 1$. Therefore the vector ϵ must satisfy the following relations:

$$E(\epsilon) = 0 \text{ and } V(\epsilon) = \sigma^2 V \tag{1.3}$$

where $V = [v_{ij}]$, $v_{ij} = \rho^{|i-j|}$ for $i, j = 1, 2, \dots, n$.

The BLUE's for the elements of β are then given by

$$\hat{\beta} = (X' V^{-1} X)^{-1} X' V^{-1} Y \tag{1.4}$$

and the information matrix is $C_V = X' V^{-1} X$ where

$$V^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1 + \rho^2 & -\rho & \dots & 0 & 0 \\ 0 & -\rho & 1 + \rho^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 + \rho^2 & -\rho \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{pmatrix}$$

Jenkins and Chanmugam ([12]) studied various kinds of correlation between errors in the case of a single two-level factor ($k = 1$). Banerjee ([2]) considered the general case of k factors for some classes of weighing designs

under AR(1) errors. Constantine ([7]) using an approximation of the information matrix C developed methods of constructing efficient weighing and factorial designs starting out from Hadamard matrices. An extensive research has been done for the case of the complete 2^k factorial designs and the 2^{k-p} fractional factorial designs (see, for example, [5], [16], [13], [6] and [1]).

In this paper, we deal with two main questions concerning optimality, under the assumption that the errors follow a first order autoregressive model.

1. Given the number of observations n , the number of factors k and an estimation of the correlation coefficient ρ , what is the exact D - or A -optimal design?
2. What is the relation of an optimal design, under the assumption that successive observations are autocorrelated, to an optimal design with uncorrelated observations and with the same n and k ?

For the sake of brevity, we shall denote the D - (or A -) optimal design for the case where the observations are uncorrelated, as D_0 - (or A_0 -) optimal design.

2 Theoretical results

It is known that under the assumption of uncorrelated errors, the functions defining the optimality criteria ($Det(C)$ and $Tr(C^{-1})$) are invariant with respect to some transformations of the design matrix X . Such transformations are the interchange of two rows and/or columns of X , or the multiplication by -1 of a row and/or column of X . So, if the design matrix \tilde{X} is obtained from X by a number of successive transformations of the above type, it is considered to be *equivalent* to X . The corresponding information matrices are considered as *equivalent* too.

Unfortunately, the above is not true if the errors in the model (1.1) are correlated. In fact the interchange of any two observations (rows of the design matrix), may give a completely different design with respect to optimality. In this case we have the following.

Proposition 2.1 *Let X be the design matrix of the main effects of the model with autocorrelation and V the covariance matrix of the errors. Then:*

- a) *The interchanging of any two columns of X and/or the multiplication of any column(s) by -1 , gives a matrix \tilde{X} , which is equivalent to X , with respect to A - and D -optimality criteria.*

- b) *The interchanging of any two rows of X and/or the multiplication of any row(s) by -1 , gives a matrix \check{X} , which, in general, is not equivalent to X .*

Proof: a) Let T be a matrix obtained from the identity matrix I_{k+1} following the same transformations as the ones giving matrix \check{X} from X . Then it is easy to see that $\check{X} = X \cdot T$, so, we have:

$$\tilde{C}_V = \check{X}' V^{-1} \check{X} = (X \cdot T)' V^{-1} (X \cdot T) = T' X' V^{-1} X T = T' C_V T.$$

Hence

$$\text{Det}(\tilde{C}_V) = \text{Det}(T' C_V T) = \text{Det}(C_V)$$

and

$$\text{Tr}(\tilde{C}_V^{-1}) = \text{Tr}(T^{-1} C_V^{-1} (T')^{-1}) = \text{Tr}(C_V^{-1})$$

i.e. the information matrix remains invariant with respect to D - and A -optimality criteria.

- b) If T is defined analogously as in (a), we have $\check{X} = T \cdot X$, so:

$$\text{Det}(\check{C}_V) = \text{Det}(X' T' V^{-1} T X)$$

which is, in general, different from $\text{Det}(C_V)$.

Similarly $\text{Tr}(\check{C}_V^{-1}) \neq \text{Tr}(C_V^{-1})$, in general, which completes the proof. \square

As a consequence of the above proposition, we need to modify, in the presence of autocorrelated errors, the definition of the equivalence of designs or of information matrices. We give the following definition:

The design matrix \check{X} which is obtained from X by a number of successive interchanges of columns or by multiplying some columns by -1 , will be called *c-equivalent* to X . The corresponding information matrices, will be called *c-equivalent* too.

From the first part of the proposition 2.1, we obviously have:

Corollary 2.1 *Every c-equivalent design of a D - (or A -) optimal design, is D - (or A -) optimal too. \square*

There is a special case, where the row transformations produce D -, or A -optimal designs. This is the case of saturated designs where the following proposition holds.

Proposition 2.2 *Let $n = k + 1$ i.e. the design is saturated. Then:*

- a) *If X is D_0 -optimal design, it is also D -optimal for every $\rho \neq 0$.*

b) If X is A_0 -optimal design and, in addition, X is a Hadamard matrix, then it is A -optimal for every $\rho \neq 0$ too.

Proof: Using the previous notation and since X is a square matrix it holds:

$$\text{Det}(\check{C}_V) = \text{Det}(C_V) = \text{Det}(X)^2 / \text{Det}(V),$$

which proves (a).

It is easy to see that if X is a Hadamard matrix, i.e. $X' \cdot X = n \cdot I_n$, then

$$\text{Tr}(\check{C}_V^{-1}) = \text{Tr}(C_V^{-1}) = n \cdot \text{Tr}(V)$$

which completes the proof. \square

Proposition 2.3 Let $\mathcal{D}(X)$ be the class of the design matrices consisted of the c -equivalent to X design matrices and of the matrices obtained by interchanging the rows of the matrices in the class¹. Then, the classes $\mathcal{D}(X_1)$ and $\mathcal{D}(X_2)$ corresponding to different design matrices X_1 and X_2 , are either identical or they do not have any common element.

Proof: It is obvious that the class $\mathcal{D}(X)$ is “closed” under the interchanging of the rows and/or columns of its elements, or by multiplying the columns of its elements by -1 . In fact, if $A \in \mathcal{D}(X)$ then A is obtained from X by interchanging the rows of X or of some c -equivalent to X . So, X can be obtained from A by using the same transformations in an inverse order. Hence the class $\mathcal{D}(A)$ is identical to the class $\mathcal{D}(X)$.

So, if A is a common element of the two classes $\mathcal{D}(X_1)$, $\mathcal{D}(X_2)$, then $\mathcal{D}(A)$ is identical to each one of the classes $\mathcal{D}(X_1)$ and $\mathcal{D}(X_2)$, which proves the proposition. \square

The above proposition means that in order to compare the two classes $\mathcal{D}(X_1)$, $\mathcal{D}(X_2)$ it is sufficient to check whether the matrix X_1 (or X_2) belongs to the class $\mathcal{D}(X_2)$ (or $\mathcal{D}(X_1)$).

Proposition 2.4 Let $X = (R'_1, R'_2, \dots, R'_n)'$, where R_i , $i = 1, 2, \dots, n$ denotes the i -th row of X , be the design matrix under the presence of autocorrelated errors. Let also $Z = (R'_n, R'_{n-1}, \dots, R'_1)'$ be the design matrix obtained by reversing the order of the rows of X . Then the information matrices corresponding to X and Z , are the same, i.e.

$$X' V^{-1} X = Z' V^{-1} Z.$$

¹Note that if the matrix X_1 is obtained from X by multiplying some row(s) by -1 , it does not necessarily belong to $\mathcal{D}(X)$.

Proof: By denoting

$$S = (s_{ij}), \quad i, j = 1, 2, \dots, n, \quad \text{with } s_{ij} = \begin{cases} 1, & \text{if } j = n + 1 - i \\ 0, & \text{otherwise} \end{cases}$$

it is easy to see that

$$Z = SX, \quad \text{and } S'V^{-1}S = V^{-1},$$

which proves the statement. \square

Hence, if a matrix having a definite order of rows has been examined for optimality, the matrix having the reverse order of rows does not have to be examined. This reduces the number of "candidates" for optimality matrices to about a half of their total number.

The case where $k = 1$ has been studied theoretically in a paper by Angelis, Bora-Senta and Moyssiadis ([1]).

Using the $n \times 1$ vectors:

$\mathbf{g}_n = (+1, -1, +1, \dots, (-1)^{n+1})'$, $\mathbf{q}_{n;\nu} = (1'_\nu, -1'_{n-\nu})'$, $\nu = 1, 2, \dots, n - 1$, where $\mathbf{1}_n = (1, 1, 1, \dots, 1)'$ and denoting by $[a]$ the largest integer not exceeding a , the main result of the above-mentioned paper can be summarized as follows :

Theorem 2.1 *The design matrix of the D- optimal design with one factor and n observations can be written, except of equivalence, as:*

$$X = (\mathbf{1}_n, \mathbf{q}_{n;z}), \quad \text{if } \rho < 0 \quad \text{and} \quad X = (\mathbf{1}_n, \mathbf{g}_n), \quad \text{if } \rho > 0$$

where $z = \left[\frac{n+1}{2} \right]$. The same design is also A -optimal, except for the cases with odd n ($n > 3$) and $\rho > r$, where:

$$r = \frac{(n^2 - 2n - 1) - 2\sqrt{(n^2 - 3n + 1)(n - 2)}}{(n - 3)^2}. \quad (2.1)$$

In these cases the A-optimal design has a design matrix

$$X = \begin{pmatrix} 1 & 1 \\ \mathbf{1}_{n-1} & \mathbf{g}_{n-1} \end{pmatrix}.$$

\square

We define as:

$$NLC(X) = \frac{1}{2} \sum_{j=1}^k \sum_{i=1}^{n-1} (1 - x_{ij}x_{i+1,j}), \quad (2.2)$$

the total number of level changes over all columns of the design matrix X . From the Theorem 2.1, it is observed that the A - and D - optimal designs for $\rho < 0$ have $NLC(X) = 1$, which is the minimum number of level changes over all possible design matrices X . Similarly, the A - and D - optimal designs for $\rho > 0$ have $NLC(X) = n - 1$, which is the maximum number of level changes over all possible design matrices X . There is only one exception for the A -optimality and for the cases $\rho > r$, (r as in 2.1), where $NLC(X)$ is less than maximum, namely $NLC(X) = n - 2$. However, by comparing the trace of the inverse of the information matrix of the A -optimal design with the one of the design having the maximum number of level changes, it is found that there is not a significant difference. In fact, if we define as A -efficiency, the ratio of the above two traces, expressed as a percentage, we see that the A -efficiency of the last design, is greater than 99.3%. In figure 1 we can see that this efficiency is an increasing function of n for a given value of ρ .

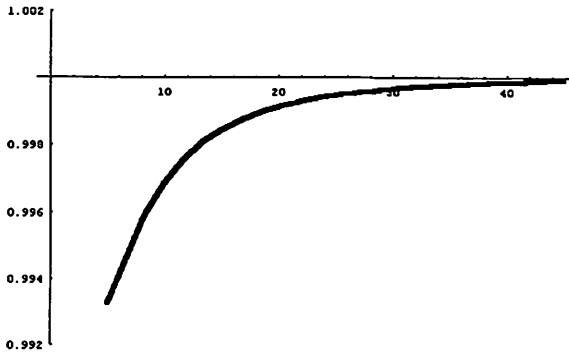


Figure 1. A -efficiency for the almost A -optimal designs for $\rho = 0.8$

Designs with great A -efficiency (greater than, say, 0.85) will be referred to, as “almost A -optimal designs”. Note that there are designs, which are A_0 -optimal, but they have low A -efficiency. For example if $\rho = 0.4$, $n = 9$ the design with $\mathbf{x} = (-1'_4, 1'_5)'$ has A -efficiency only 0.72.

According to the above observations, the proposition (2.4) can be restated as follows:

Corollary 2.2 *The design with a single factor and n observations, is A - and D - optimal, if the corresponding design matrix X attains the minimum value of level changes of the factor for $\rho < 0$, or the maximum value of level changes for $\rho > 0$, except for the case of A -optimality and $\rho > r$, as r is defined in (2.1). In this case the design is “almost A -optimal” with A -efficiency very close to 1. \square*

Unfortunately, the general case where $k \geq 2$ cannot be studied in a similar way. In fact, the expressions giving the determinant of the information matrix (or the trace of the inverse of the information matrix) are so complicated, that it is extremely difficult to maximize (or minimize) them, using algebraic operations.

However, for the case $k = 2$, we give the following conjecture, based on the experience of several exhaustive searches.

Proposition 2.5 (Conjecture) *The A- and D- optimal designs for the case $k = 2$ and for different values of n , have design matrices given by:*

i) $n = 4\nu$,

$$X = \left(\mathbf{1}_n, \mathbf{q}_{n;2\nu}, \begin{pmatrix} \mathbf{q}_{2\nu;\nu} \\ -\mathbf{q}_{2\nu;\nu} \end{pmatrix} \right), \quad \text{for } \rho < 0$$

$$X = \left(\mathbf{1}_n, \mathbf{g}_n, \begin{pmatrix} \mathbf{g}_{2\nu} \\ -\mathbf{g}_{2\nu} \end{pmatrix} \right), \quad \text{for } \rho > 0$$

ii) $n = 4\nu + 1$,

$$X = \left(\mathbf{1}_n, \mathbf{q}_{n;2\nu+1}, \begin{pmatrix} \mathbf{q}_{2\nu+1;\nu+1} \\ -\mathbf{q}_{2\nu;\nu} \end{pmatrix} \right), \quad \text{for } \rho < 0$$

$$X = \left(\mathbf{1}_n, \mathbf{g}_n, \begin{pmatrix} \mathbf{g}_{2\nu+1} \\ \mathbf{g}_{2\nu} \end{pmatrix} \right), \quad \text{for } \rho > 0$$

iii) $n = 4\nu + 2$,

$$X = \left(\mathbf{1}_n, \mathbf{q}_{n;2\nu+1}, \begin{pmatrix} \mathbf{q}_{2\nu+1;\nu+1} \\ -\mathbf{q}_{2\nu+1;\nu} \end{pmatrix} \right), \quad \text{for } \rho < 0$$

$$X = \left(\mathbf{1}_n, \mathbf{g}_n, \begin{pmatrix} \mathbf{g}_{2\nu+1} \\ \mathbf{g}_{2\nu+1} \end{pmatrix} \right), \quad \text{for } \rho > 0$$

iv) $n = 4\nu + 3$,

$$X = \left(\mathbf{1}_n, \mathbf{q}_{n;2\nu+2}, \begin{pmatrix} \mathbf{q}_{2\nu+2;\nu+1} \\ -\mathbf{q}_{2\nu+1;\nu} \end{pmatrix} \right), \quad \text{for } \rho < 0$$

$$X = \left(\mathbf{1}_n, \mathbf{g}_n, \begin{pmatrix} \mathbf{g}_{2\nu+1} \\ \mathbf{g}_{2\nu+2} \end{pmatrix} \right), \quad \text{for } \rho > 0$$

□

This conjecture was found to be true for all the values of $n \leq 14$ and for a wide variety of values of ρ in $(-1, 1)$.

A construction of optimal designs, or at least highly efficient designs, for D - and A -optimality criteria, can be achieved with an exhaustive search, by developing a suitable computer algorithm.

In the following section we develop such an algorithm which gives a number of optimal designs for relatively small values of n and k . We tried to find out a relation between these designs and the corresponding ones for the case without autocorrelation, and to state a proposition extending the corollary (2.2).

3 The algorithm

For an exhaustive search over all possible information matrices we first need to find all the possible design matrices, then to form the corresponding information matrices, to calculate their determinants and the trace of their inverses and finally to point out the D - and A -optimal among them. For given values of n and k we can form 2^{nk} such matrices by taking $x_{ij} = +1$, or -1 , for $i = 1, 2, \dots, n$, and $j = 1, 2, \dots, k$. Obviously, many of them are equivalent to each other, so it is need to reduce their number, using the previous propositions. For this purpose, we have developed an algorithm which, by using a known class of matrices for some k and constant n , produces sequentially a class for $k+1$. The way of construction determines the order of matrices in this class, which in turn is used for the next step.

Let us denote this class by $C_{n,k}$. Its elements are $n \times (k+1)$ matrices of the form:

$$C_{n,k} = \{X : X = (\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)\}$$

where \mathbf{x}_i are $n \times 1$ vectors with the following properties:

(P1) Every \mathbf{x}_i has at most $\lfloor \frac{n}{2} \rfloor$ negative 1's.

In fact, by multiplying (if it is necessary) \mathbf{x}_i by -1 (proposition 2.1), we can always have $\mathbf{x}'_i \cdot \mathbf{1}_n \geq 0$ for n even or ≥ 1 for n odd.

(P2) If \mathbf{x}_i has t negative 1's, then \mathbf{x}_{i+1} has at most t negative 1's, i.e. if $\mathbf{x}'_i \cdot \mathbf{1}_n = n - 2t$, then $\mathbf{x}'_{i+1} \cdot \mathbf{1}_n \geq n - 2t$.

In fact, in any other case we can interchange \mathbf{x}_i with \mathbf{x}_{i+1} , (proposition 2.1).

(P3) For every $\ell = 1, 2, \dots, (k-1)$ exists a partition of the rows of the matrix $X_{\ell+1}$, consisting of the first $\ell+1$ columns of X , in s parts, having rows with all the elements but the last equal, i.e.

$$X_{\ell+1} = \begin{pmatrix} M_1 & \delta_1 \\ M_2 & \delta_2 \\ \dots & \dots \\ M_s & \delta_s \end{pmatrix}, \text{ where } \begin{array}{l} M_i \text{ is a } \nu_i \times \ell \text{ matrix, } \nu_i \geq 1, \\ \text{with } \nu_1 + \nu_2 + \dots + \nu_s = n \\ \text{and with all rows identical,} \\ \delta_i \text{ is a } \nu_i \times 1 \text{ vector.} \end{array}$$

Then the vectors δ_i can always be written in the form $\delta_i = (\mathbf{1}'_{t_i}, -\mathbf{1}'_{\nu_i - t_i})'$, for $t_i \leq \nu_i$.

In fact, if we rearrange the ν_i rows of the i -th part of the partition of the matrix $X_{\ell+1}$, then the matrix M_i remains unchanged. So the vector δ_i can always take the above form.

This property of matrices X reduces the number of the representatives in the class $C_{n,k}$.

(P4) The matrices in class $C_{n,k}$ must be of full rank, i.e. $\text{rank}(X) = k+1$.

In fact, by using the known property that for any two matrices A and B the rank of the product AB cannot exceed the rank of either A or B , we find that if $\text{rank}(X) \leq k$, then $\text{rank}(X'V^{-1}X) \leq k$, which means that if $\text{rank}(X) < k+1$, then the information matrix is singular.

(P5) To every matrix $X = (\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ we correspond the $k \times 1$ vector $\mathbf{a}(X)$ defined by:

$$\mathbf{a}(X) = \frac{1}{2} \left(\mathbf{J}_{k \times n} + \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \dots & x_{kn} \end{pmatrix} \right) \cdot \begin{pmatrix} 2^{n-1} \\ 2^{n-2} \\ \vdots \\ 2^0 \end{pmatrix},$$

where $\mathbf{J}_{k \times n}$ is the $k \times n$ matrix consisting only of 1's. Note that the binary representation of the i -th element of $\mathbf{a}(X)$ is $\frac{1}{2}(\mathbf{1}' + \mathbf{x}'_i)$. Using proposition 2.1, we can rearrange the columns of the matrix X , so as the elements of the vector $\mathbf{a}(X)$ to be in decreasing order i.e. $\mathbf{a}(X) = (a_{(k)}, a_{(k-1)}, \dots, a_{(1)})'$.

So, the problem of finding if a matrix Z belongs to a class $\mathcal{D}(X)$ is reduced to the one of comparing the corresponding vector $\mathbf{a}(Z)$ with the vectors $\mathbf{a}(X_i)$ of the class $\mathcal{D}(\mathbf{a}(X))$ (proposition 2.3).

We proceed, now, with the description of the algorithm. First of all the elements of the class $C_{n,1}$ are determined and are ordered in decreasing order, according to their total number of negative 1's, i.e.

$$C_{n,1} = \left\{ \left(\begin{array}{cc} 1_{n-t} & 1_{n-t} \\ \mathbf{1}_t & -\mathbf{1}_t \end{array} \right), t = s, s-1, \dots, 1, \text{ where } s = \left\lfloor \frac{n}{2} \right\rfloor \right\}.$$

The optimum design in this class is known (see theorem 2.1). Suppose we have studied the class $C_{n,k-1}$ for $k = 2, 3, \dots$. Then we continue as follows:

Step 1. Begin with the first element, say Y , of the class $C_{n,k-1}$, having,

say, t_0 negative 1's in its last column. Let's suppose Y can be written as:

$$Y = \begin{pmatrix} M_1 \\ M_2 \\ \dots \\ M_s \end{pmatrix}, \text{ where } \begin{array}{l} M_i \text{ is a } \nu_i \times k \text{ matrix, } \nu_i \geq 1, \\ \text{with } \nu_1 + \nu_2 + \dots + \nu_s = n \\ \text{and with all rows identical.} \end{array} .$$

For each t , $t = t_0, t_0 - 1, \dots, 1$ (property P2), construct a vector δ consisting of s segments corresponding to M_i 's of the matrix Y , i.e.

$$\delta = (\delta'_1, \delta'_2, \dots, \delta'_s)' , \text{ with } \delta_i = (\mathbf{1}'_{\nu_i - t_i}, -\mathbf{1}'_{t_i})'$$

where $\sum t_i \leq t$ (property P3). Berry's algorithm AS179 ([3]), gives the complete set of all such vectors δ 's, by taking all the partitions of $\sum t_i$ into s non-ordered, non-empty subsets t_1, t_2, \dots, t_s . For every δ form the augmented matrix $Z = (Y, \delta)$ and continue to Step 2.

Step 2. Compute the rank of the matrix $Z = (Y, \delta)$. If this rank is equal to $k + 1$ (property P4), then go to Step 3. Else, ignore the matrix Z and continue with the next δ .

Step 3. For the matrix $Z = (Y, \delta)$, form the corresponding class of vectors $\mathcal{D}(\mathbf{a}(Z))$ (property P5). Compare each one of its vectors with the vectors $\mathbf{a}(X_i)$ of the matrices X_i that have already been recorded in the class $\mathcal{C}_{n,k}$. If all these comparisons fail, include matrix Z in the class $\mathcal{C}_{n,k}$. If δ is not the last one, take next δ and go to Step 2. Else, if $t > 1$ take $t \leftarrow t - 1$ and go to Step 1. Else continue to Step 4.

Step 4. Compute matrix V^{-1} for a given value of ρ . For each element X in the class $\mathcal{C}_{n,k}$, do the following. Form the class $\mathcal{D}(X)$ taking into account the proposition 2.4. (That means include in $\mathcal{D}(X)$, only the one of the two matrices having their rows identical in reverse order). Find the D - and A -optimal design among the elements of all the classes $\mathcal{D}(X)$. For the optimal design(s) X^* , store its determinant, the trace of its inverse, the number of level changes, etc.

Then Stop.

4 The results

The above algorithm was executed for the cases $k = 2$ and $n \leq 14$, $k = 3$ and $n \leq 10$, $k = 4$ and $n \leq 8$, $k = 5$ and $n = 7$, each for $\rho = -0.9 + 0.1 \cdot h$, $h = 0, 1, \dots, 18$. For greater values of the parameters k and n , the computer time needed for exhaustive search is huge. However, the cases we dealt with, reveal the general tendency.

The optimal designs are given in the appendix. The D -optimal designs for $k = 2$ and $n \leq 14$ are not listed explicitly, because they are the same as

the ones given in the conjecture 2.5. The A -optimal designs which are the same as the corresponding D -optimal ones, are not listed too.

By studying the structure and the properties of these designs, we tried to answer the following questions:

- (a) "In which cases the D - or A -optimal designs are also D_0 - or A_0 -optimal?"
- (b) "Can the D - or A -optimal designs be obtained by rearranging the rows of a D_0 - or A_0 -optimal design?"
- (c) "If we consider as D - or A -optimal the best design (with respect to D - or A -optimality criteria) generating by the corresponding D_0 - or A_0 -optimal one, in the case where these designs are different, what is the maximum error?"

The answer in question (a) is the case of the saturated designs ($n = k+1$) (proposition 2.2)

For the cases $n = 5, 7, 8, 9, 11, 12, 13$ and $k = 2$ there is only one, w.r.t. c -equivalence, D_0 - optimal design, which is A_0 -optimal too. It was found that the D -or A - optimal design is generated from the corresponding D_0 -or A_0 - optimal ones, i.e. the answer to question (b) is positive. In all these cases the D -optimal designs coincide with the one given in conjecture 2.5. The cases for which the A - optimal design is different from the corresponding D -optimal are listed in Appendix (section 6.1).

For the rest of the other cases examined, there are more than one D_0 - or A_0 -optimal designs. One of them i.e. $n = 6, k = 2$ will be discussed in details. For the rest cases we shall give a summary of the results.

4.1 The case $n = 6, k = 2$

There are two different design matrices, i.e.

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{pmatrix}, Y = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix},$$

which are both D_0 - and A_0 - optimal. These two matrices are equivalent, but not c-equivalent. In fact, $Y = T \cdot X \cdot S$ where

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}, S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

while Y can not be obtained from X by rearranging some columns or by multiplying some of them by -1 .

It was found that the D -optimal design is obtained by taking the permutation 123645 of the rows of the matrix Y , when $\rho < 0$ or by taking the permutation 162435, when $\rho > 0$, i.e.

$$X^* = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \text{ for } \rho < 0, \text{ and } X^* = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \text{ for } \rho > 0.$$

These designs coincide with the corresponding ones given in the conjecture 2.5. It was also found that the D -optimal design can not be obtained by rearranging the rows of X . So, in this case the answer to question (b) depends on the choice of the D_0 -optimal design.

If the matrix X was selected instead of Y and the best (w.r.t. the D -optimality criterion) design generated by X was considered as the D -optimal design, then the ratio of the determinants of the corresponding information matrices, expressed as a percentage, is a measure of the goodness of this choice. If X is D_0 -optimal too, as it happens here, we call this ratio D -efficiency of the best design w.r.t. the D -optimal one.

When the D -efficiency is large enough (for example greater than 85%), the error of taking as D -optimal design the best design obtained by some of the D_0 -optimal ones, is small. We computed the range of the D -efficiencies for different ranges of ρ and we distinguished the cases where the D -efficiencies were greater than 85%. The output was listed in the appendix (section 6.8).

The same is almost true for the A -optimality. In fact the D -optimal designs found above are also A -optimal for all values ρ , except for $\rho \geq 0.7$. In this case the A -optimal design is also obtained from the matrix Y by taking a different permutation 123645 of its rows (section 6.1 in the appendix).

The *A-efficiency* of the best design w.r.t. the *A*-optimal one is defined in an analogous way, by considering the traces of the inverses of the information matrices instead of the determinants.

If the *A*-efficiency is large enough (for example greater than 85%), the error of taking as *A*-optimal design the best (w.r.t. the *A*-optimality criterion) design obtained by some of the A_0 -optimal ones, is small. We computed the range of the *A*-efficiencies for different ranges of ρ and we discriminated the cases where the *A*-efficiencies were greater than 85%. The output was listed in the appendix (section 6.8).

4.2 The other cases

The *D*- optimal designs for the other cases were found and were listed in the appendix (sections 6.2, 6.4). Similarly, the *A*- optimal designs for the other cases were found and were listed in the appendix (sections 6.1, 6.3, 6.5). We now state the following general remarks.

1. For most of the cases examined (except for $n = 9, k = 3$ and $n = 7, k = 5$), there are two different, with respect to the equivalence, D_0 -optimal designs, which are A_0 -optimal too. Only one of them generates the *D*- or the *A*-optimal design.
2. For the case $n = 9, k = 3$ there are three different, with respect to the equivalence, D_0 -optimal designs, which are A_0 -optimal too. Again, only one of them generates the *D*- or the *A*-optimal design.
3. For the case $n = 7, k = 5$ there are three D_0 -optimal designs and four A_0 -optimal ones, which are all different to each other.
4. There are cases (namely $n = 7, k = 3, \rho < -0.9$, $n = 7, k = 4, \rho > 0.5$, $n = 7, k = 5, \rho < -0.25$), where the *D*- optimal design is not D_0 - optimal design.
5. Similarly, there are cases (namely $n = 7, k = 3, \rho < -0.5$, $n = 6, k = 4, \rho < -0.85$, $n = 7, k = 4, \rho < -0.5$ or $\rho > 0.3$, $n = 7, k = 5, \rho < -0.65$ or $\rho > 0.75$), where the *A*- optimal design is not A_0 - optimal design.

5 Conclusion

Summarizing the above, we can say the following:

For absolutely small values of ρ (say $|\rho| < .4$) the *D*-optimal design or at least a design very close to it, can be constructed as follows: Begin from the known D_0 -optimal designs (*D*-optimal in the case without correlation) and by rearranging its rows, find the best, with respect to *D*-optimality

criterion. In most of the cases the total sum of the level changes (see 2.2) attains its maximum (or near the maximum) value when $\rho > 0$, or its minimum (or near the minimum) value when $\rho < 0$.

The same holds for the A -optimality.

For larger values of ρ the above process can not work. In these cases the optimal designs can be taken from the appendix.

6 Appendix

6.1 A -optimal Designs for $k = 2$

(Only those which differ from the corresponding D -optimal ones, which are given in general form in the conjecture 2.5)²

| $n=5, \rho < -.5$ | $n=5, \rho > .4$ | $n=6, \rho > .7$ | $n=7, \rho > .4$ | $n=8, \rho > .8$ |
|-------------------|------------------|------------------|------------------|------------------|
| + + + | + + + | + + + | + + - | + + + |
| + + - | + + - | + - + | + + + | + - - |
| + - - | + - + | + + - | + - - | + + + |
| + - + | + + + | + - + | + + + | + - + |
| + + + | + - - | + + + | + - + | + + - |
| | | + - - | + + - | + - + |
| | | | + - + | + + - |
| | | | | + - - |

| $n=9, \rho > .6$ | $n = 10, .7 < \rho < .8$ | $n=10, \rho > .9$ | $n=11, \rho > .6$ |
|------------------|--------------------------|-------------------|-------------------|
| + + + | + + + | + + - | + + - |
| + - - | + - - | + - + | + + + |
| + + + | + + + | + + + | + - - |
| + + - | + - - | + - - | + + + |
| + - + | + + + | + + + | + - - |
| + + - | + - + | + - - | + + + |
| + - + | + + - | + + + | + - + |
| + + + | + - + | + - + | + + - |
| + - - | + + - | + + - | + - + |
| | + - - | + - + | + + - |
| | | | + - + |

²+ stands for +1, - stands for -1

6.5 A-optimal Designs for $k = 4$.

(Only those which differ from the corresponding D -optimal ones)

| | | |
|--------------------|------------------------|-----------------------|
| $n=6, \rho < -.85$ | $n=6, -.85 < \rho < 0$ | $n=6, 0 < \rho < .75$ |
| + + + - + | + + + + - | + - - + + |
| + + - - + | + - - - - | + + + + - |
| + - - + + | + + + - + | + + + - - |
| + - + + + | + - - + + | + + + + + |
| + - - - - | + + + + + | + - - - - |
| + - - - - | + + + + + | + - - - - |

| |
|-------------------|
| $n=6, \rho > .75$ |
| + + + + - |
| + + + + - |
| + - - + + |
| + - - + + |
| + - - + + |
| + - - + + |

| |
|-------------------|
| $n=8, \rho > .85$ |
| + + + - + |
| + - - + + |
| + - - + + |
| + - - + + |
| + - - + + |
| + - - + + |
| + - - + + |

| | | |
|-------------------|--------------------------|------------------------|
| $n=7, \rho < -.5$ | $n=7, -.5 < \rho < -.25$ | $n=7, -.25 < \rho < 0$ |
| + + + + - | + + + + + | + + + - - |
| + + - - + | + + + - - | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |

| | | |
|-----------------------|-------------------------|-------------------|
| $n=7, 0 < \rho < .25$ | $n=7, .25 < \rho < .55$ | $n=7, \rho > .55$ |
| + + + + - | + + + + - | + + + + - |
| + + - - + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |
| + - - + + | + - - + + | + - - + + |

6.8 D- and A-efficiencies for different values of n, k, ρ

| k | n | ρ | D-efficiency (%) | A-efficiency (%) |
|--------------|-------------------|-------------------|------------------|------------------|
| 2 | 6 | $\rho < -0.4$ | 67 – 80 | 83 – 97 |
| | | $-0.4 < \rho < 0$ | 80 – 97 | |
| | 10 | $0 < \rho < 0.5$ | 85 – 96 | 98 – 99 |
| | | $\rho > 0.5$ | 80 – 85 | |
| 14 | $-1 < \rho < 1$ | 93 – 98 | 95 – 99 | |
| | $-1 < \rho < 1$ | 90 – 97 | 94 – 98 | |
| 3 | 5 | $\rho < -0.3$ | 50 – 85 | 85 – 98 |
| | | $-0.3 < \rho < 0$ | 85 – 95 | |
| | | $\rho > 0$ | 80 – 95 | |
| | 6 | $\rho < -0.5$ | 78 – 95 | 75 – 85 |
| | | $-0.5 < \rho < 0$ | 78 – 95 | 85 – 97 |
| | | $\rho > 0$ | 57 – 87 | 96 – 99 |
| | 7 | $\rho < 0$ | 88 – 98 | 90 – 98 |
| | | $\rho > 0$ | 48 – 83 | |
| | 8 | $\rho < 0$ | 78 – 95 | 92 – 98 |
| | | $\rho > 0$ | 40 – 83 | |
| | 9 | $\rho < 0$ | 85 – 96 | 96 – 99 |
| | | $\rho > 0$ | 45 – 80 | |
| 10 | $\rho < 0$ | 85 – 92 | 97 – 99 | |
| | $0 < \rho < 0.3$ | 85 – 94 | | |
| | $\rho > 0.3$ | 76 – 85 | | |
| 4 | 6 | $\rho < -0.4$ | 69 – 85 | 96 – 99 |
| | | $-0.4 < \rho < 0$ | 85 – 98 | |
| | | $\rho > 0$ | 90 – 98 | |
| | 7 | $\rho < -0.2$ | 67 – 83 | 85 – 98 |
| | | $-0.2 < \rho < 0$ | 83 – 91 | 96 – 98 |
| | | $\rho > 0$ | 85 – 99 | |
| 8 | $\rho < -0.5$ | 34 – 83 | 78 – 85 | |
| | $-0.5 < \rho < 0$ | 34 – 83 | 85 – 96 | |
| | $0 < \rho < 0.4$ | 85 – 95 | 98 – 99 | |
| $\rho > 0.4$ | 78 – 85 | | | |
| 5 | 7 | $\rho < -0.2$ | 82 – 99 | 96 – 98 |
| | | $-0.2 < \rho < 0$ | 100 | |
| | | $\rho > 0$ | 100 | |

References

- [1] E. Angelis, L. Bora-Senta and C. Moyssiadis, Optimal designs with a single two-level factor and n autocorrelated observations, To appear in *Utilitas Mathematica*.
- [2] K.S. Banerjee, On Hotteling's weighing designs under autocorrelation of errors, *Annals of Mathematical Statistics* **36** (1965), 1829–1834.
- [3] K.J. Berry, Enumeration of all permutations of multi-sets with fixed repetition numbers, *Applied Statistics, Series C* **31** (1982), 169–173.
- [4] C.S. Cheng, Optimality of some weighing and 2^n fractional factorial designs, *Ann. Statist.* **8** (1980), 436–446.
- [5] C-S. Cheng and M. Jacroux, The construction of trend-free run orders of two-level factorial designs, *Journal of the American Statistical Association* **83** (1988), 1152–1158.
- [6] C-S. Cheng and D.M. Steinberg, Trend robust two-level factorial designs, *Biometrika* **78** (1991), 325–336.
- [7] G.M. Constantine, Robust designs for serially correlated observations, *Biometrika* **76** (1989), 245–251.
- [8] H. Ehlich, Determinantenabschätzungen für binäre Matrizen, *Math. Zeitschr.* **83** (1964), 123–132.
- [9] H. Ehlich, Determinantenabschätzungen für Matrizen mit $n \equiv 3 \pmod{4}$, *Math. Zeitschr.* **84** (1964), 438–447.
- [10] Z. Galil and J. Kiefer, D-optimum weighing designs, *Ann. Statist.* **8** (1980), 1293–1306.
- [11] Z. Galil and J. Kiefer, Construction methods for D-optimum weighing designs when $n \equiv 3 \pmod{4}$, *Ann. Statist.* **10** (1982), 502–510.
- [12] G.M. Jenkins and J. Chanmugam, The estimation of slope when the errors are autocorrelated, *Journal of Royal Statistical Society* **24** (1962), 199–214.
- [13] P.W.M. John, Time trends and factorial experiments, *Technometrics* **32** (1990), 275–282.
- [14] C. Moyssiadis and S. Kounias, The exact D-optimal first order saturated design with 17 observations, *J.S.P.I.* **7** (1982), 13–27.
- [15] K.R. Shah and B.K. Sinha, *Theory of Optimal Designs*, Springer-Verlag, Berlin, 1989.
- [16] D.M. Steinberg, Factorial experiments with time trends, *Technometrics* **30** (1988), 259–269.