

# Screening by Conference Designs

**Clemens Elster**

*Physikalisch-Technische Bundesanstalt*

*Abbestr. 2-12*

*10587 Berlin*

*Germany*

*email: celster@chbrb.berlin.ptb.de*

**Arnold Neumaier**

*Institut für Mathematik*

*Universität Wien*

*Strudlhofgasse 4*

*A-1090 Wien*

*Austria*

*email: neum@cma.univie.ac.at*

**Abstract.** Screening experiments are addressed to the identification of the relevant variables within some process or experimental outcome potentially depending on a large number of variables. In this paper we introduce a new class of experimental designs called edge designs. These designs are very useful for screening experiments since they allow a model-independent estimate of the set of relevant variables, thus providing more robustness than traditional designs.

We give a bound on the determinant of the information matrix of certain edge designs, and show that a large class of edge designs meeting this bound can be constructed from conference matrices. We also show that the resulting conference designs have an optimal space exploration property which is important as a guard against unexpected nonlinearities. We survey the existence of and constructions for conference matrices, and give, for  $n < 50$  variables, explicit such matrices when  $n$  is a prime, and references to explicit constructions otherwise.

revised, February 1995

**Key words:** Conference design, conference matrix, edge design, fractional factorial design, linear model, model-independent estimate, screening experiments, skew Hadamard matrix

**1991 MSC Classification:** 62K05, secondary 05B20

# 1 Introduction

Screening experiments aim at the correct identification of the relevant variables out of a large number of variables that possibly influence some characteristic  $y$ . The aim of screening is to detect cheaply this subset of influential variables; then more accurate designs on the lower dimensional subspace are used to account for nonlinearities. Since the noninfluential variables are not explored in subsequent measurements, an important requirement for screening designs is that they place measurements such that the region of interest is well-explored. The number of trials  $N$  needed for the identification of the active factors should be kept as small as possible.

More specifically, screening is addressed to estimating a multivariate function  $y$  such that

$$y^i \approx y(x^i), \quad x^i = (x_1^i, \dots, x_n^i)^t \in [a_1, b_1] \times \dots \times [a_n, b_n],$$

but that within the measurement noise,  $y$  only depends on  $k$  variables. Here  $n$  is assumed to be large and  $k$  rather small;  $n > 15$ ,  $k < 5$  might be typical. The experimental region is chosen as a rectangular box, the most typical situation when screening experiments take place. Without loss of generality we may assume that the  $x_\alpha^i$  vary between  $-1$  and  $1$ .

As an example, consider an industrial process where a large number of variables is involved but the process actually depends only on a small but unknown subset of these variables. Since the experimental expense for empirical

model building depends drastically on the number of variables involved, e.g. the number of experiments  $N$  needed for quadratic modelling is  $N = O(n^2)$ , a correct reduction in complexity with  $N = O(n)$  experiments is of great importance when  $n$  is large.

For small  $n$ , screening is usually performed with 2-level fractional factorial designs suitable for optimal estimation of the coefficients  $\{\theta\}$  in a model such as

$$y = \theta_0 + \sum_{i=1}^n \theta_i x_i + \sum_{i < j}^n \theta_{ij} x_i x_j.$$

However in order to estimate all main effects  $\theta_i$  and interaction effects  $\theta_{ij}$ , again a large number of experiments  $N = O(n^2)$  is needed, e.g. 466 for  $n = 30$ . As a consequence, screening for a large number of variables is frequently done by the use of linear models and orthogonal designs like fractional factorial and Plackett-Burman designs (Box & Draper, 1987, Plackett & Burman, 1946).

Although linear models may serve only as a crude approximation, screening with linear models and  $O(n)$  experiments often performs successfully, especially when combined with appropriate data transformations. However, in Section 2 we demonstrate by an example the risk of screening with linear models and traditional designs.

A more robust screening method uses the “one factor at a time” design, where  $n$  experiments around a point  $x^0$  are performed in such a manner that in each case only one component differs from  $x^0$ . These designs allow a

model-independent check of the influence of each variable around  $x^0$ .

However, in the case where a linear approximation is adequate, possibly after some transformation of the data, such designs have a very low efficiency since only 2 of the  $n + 1$  data points of the design are used for the determination of the influence of a particular component, in contrast to the conventional designs where all experiments contribute to the estimates of all coefficients. Moreover, for large  $n$ , “one factor at a time” designs cluster the design points around  $x^0$ , and are therefore not necessarily representative for the region of interest.

In this paper we introduce some new designs for screening called edge designs. They combine the good efficiency properties of optimal designs for linear models with the robustness of “one factor at a time” designs. We discuss the properties and use of edge designs. Using conference matrices, we construct a large class of such designs having several optimality properties.

In the following,  $I$  denotes an identity matrix,  $j$  a vector with all entries equal 1, and  $J = jj^t$ . Further  $\|x\|$  denotes the Euclidean norm, i.e.  $\|x\|^2 = x_1^2 + \dots + x_n^2$ , and  $(x, y) = x_1y_1 + \dots + x_ny_n$  is the standard inner product in  $\mathbb{R}^n$ .

We denote by  $\mathbb{Q}$  the cube  $\{x \in \mathbb{R}^n : -1 \leq x_\alpha \leq 1, \alpha = 1, \dots, n\}$  and by  $X$  the  $N \times (n + 1)$  design matrix with  $X_{ij} = f_j(x^i)$ . Here  $x^i \in \mathbb{Q}$  is the  $i$ th setting of the variable vector  $x^i = (x_1^i, \dots, x_n^i)^t$  and  $f$  is given by the assumed

linear relation

$$y^i = \theta^t f(x^i) + \eta_i = \theta_0 + \sum_{\alpha=1}^n \theta_\alpha x_\alpha^i + \eta_i, \quad (1)$$

where  $y^i$  denotes the  $i$ th measurement and  $\eta_i$  the measurement error; the  $\eta_i$  ( $i = 1, \dots, N$ ) are assumed to be independent and identically distributed.

Thus

$$X = \begin{pmatrix} 1 & x_1^1 & \cdots & x_n^1 \\ \vdots & \vdots & & \vdots \\ 1 & x_1^N & \cdots & x_n^N \end{pmatrix}. \quad (2)$$

## 2 Edge designs

In this section we shall introduce a new class of screening designs which allow a model-independent test for active variables. This is achieved by arranging the measurements into a set  $E$  of pairs such that within these pairs the coordinates differ in one component only. We shall call such pairs edges since in the optimal case they are located at the edges of the cube and we refer to designs consisting of a collection of edges as edge designs.

Independent of any particular model, data collected with edge designs may be evaluated using the assumption that only a few, say  $p$ , of the  $n$  factors are active, i.e., contribute to the variability in the observations. This so-called factor sparsity assumption mentioned e.g. by Lenth (1989) is very natural in screening experiments, and implies that almost all differences

$$z_{ij} := y_i - y_j, \quad (i, j) \in E,$$

consist of noise only. If we assume that the noise in the data is additive, normally distributed with zero mean and variance  $\sigma^2$ ,  $n - p$  of the  $z_i$  are normally distributed with zero mean and variance  $2\sigma^2$ . Because of the unknown number of outliers, the variance must be estimated in a robust way. For example, we can use the median estimate

$$\tilde{\sigma} = \frac{\text{median}\{|z_{ij}| : (i, j) \in E\}}{2^{\frac{1}{2}} \cdot 0.675} \quad (3)$$

which is consistent when  $p = 0$  (Lenth, 1989), and hence can be expected to give reliable results when  $p \ll n$ . Outliers then determine active factors.

When each factor is varying in the same number  $r$  of edges one can improve on this by guessing  $p$  and discarding the  $rp$  largest  $|z_k|$  before using the median estimate. This gives an estimate  $\tilde{\sigma}(p)$  for each guess  $p$ ; the best value of  $p$  is found by matching  $rp$  with the number  $\omega(p)$  of distinct factors varying in edges  $(i, j)$  carrying an outlier with  $|z_{ij}| > \kappa \cdot 2^{\frac{1}{2}} \tilde{\sigma}(p)$ , using  $\kappa = 3$ . By sorting the  $|z_{ij}|$ , it is very easy to compute  $\tilde{\sigma}(p)$  and  $\omega(p)$  for all  $p$ . If  $q$  denotes the smallest positive number with  $\omega(q) > q$  then  $p = \omega(q - 1)$  is a reliable estimate for the number of active factors, and the procedure also reveals the active factors themselves. However, if  $\omega(q) - q$  has several sign changes, the decision is less straightforward, and is best done graphically.

The need for edge designs is illustrated by the following example where conventional screening analysis leads to completely unsatisfactory results.

To allow explicit calculations, the small value  $n = 7$  was chosen. As small conventional screening designs we used (i) a  $2_{III}^{7-4}$  fractional factorial Plackett-

Burman design and (ii) a  $2_{IV}^{7-3}$  design. We used the generators  $x_4 = x_1x_2$ ,  $x_5 = x_1x_3$ ,  $x_6 = x_2x_3$ ,  $x_7 = x_1x_2x_3$  in the first case and  $x_5 = x_1x_2x_3$ ,  $x_6 = x_2x_3x_4$  and  $x_7 = x_1x_3x_4$  in the latter case, the default suggestion of the BBN software RS1/Discover (1992) for screening with (i) a linear model and (ii) with a so-called linear plus model, i.e., a design for which no main effect is confounded with any two-factor interaction. Within these designs the third column is confounded with the three factor interaction of columns 5, 6 and 7. The resulting least squares estimates of the linear models for noiseless data given by the function  $y = x_5x_6x_7$  are shown in tables 1 and 2. We chose this function for demonstrating the difficulties since it assumes only two different values on the vertices of the cube  $\mathbb{Q}$ ; hence the results obtained are invariant under scaling.

Since  $x_3$  is confounded with the three factor interaction  $x_5x_6x_7$  in both classical screening designs, these designs ‘explain’ the function values at the design points by the model

$$\hat{y} = x_3,$$

suggesting a completely wrong set  $\{3\}$  of active factors. The estimated variance is zero, giving no indication of failure.

The least squares results for a  $2n = 14$  edge design, in fact the conference design used also in Table 7 below, are given in Table 3. These properly reflect the strong nonlinearity, indicated by the large value of  $\hat{\sigma}$ , and by analyzing

the variation along the 7 edges

$$z_k := y_{k+7} - y_k \quad (k = 1, \dots, 7),$$

(cf. Table 4), one recognizes  $x_5, x_6, x_7$  as relevant variables, and the others as probably irrelevant ones.

**Table 1:** Least squares estimates for the  $2_{III}^{7-4}$  design.

$N$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$\hat{\theta}_7$	$\hat{\sigma}$
8	0	0	0	1	0	0	0	0	-

**Table 2:** Least squares estimates for the  $2_{IV}^{7-3}$  design.

$N$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$\hat{\theta}_7$	$\hat{\sigma}$
16	0	0	0	1	0	0	0	0	0

**Table 3:** Least squares estimates for the conference design of Table 7.

$N$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$\hat{\theta}_7$	$\hat{\sigma}$
14	0	0	0.5	0.5	0	0	0	0	1.16

**Table 4:** Model-independent checks for the conference design

$N$	$z_1$	$z_2$	$z_3$	$z_4$	$z_5$	$z_6$	$z_7$
	$y^8 - y^1$	$y^9 - y^2$	$y^{10} - y^3$	$y^{11} - y^4$	$y^{12} - y^5$	$y^{13} - y^6$	$y^{14} - y^7$
14	0.0	0.0	0.0	0.0	-2.0	-2.0	-2.0

While this example is particularly extreme, it is very easy to run into troubles



of the same kind with more realistic, strongly nonlinear functional relationships. We conclude that for screening with linear models it is important to have model-independent checks for the selection of active factors.

In general, the use of edge designs guarantees that irrelevant variables are never treated as relevant, in contrast to the above example with classical designs, and there is only a very small chance that a relevant variable is not correctly recognized, which occurs when the two function values nearly agree on the corresponding edges. Thus screening with edge designs is robust.

### 3 Efficiency and D-optimality

The design matrix  $X$  of any design for the linear model (1) on the cube  $\mathbb{Q}$  satisfies

$$\det X^t X \leq N^{n+1} \tag{4}$$

(Shah & Sinha, 1989), with equality iff

$$X^t X = NI, \quad \text{and hence all } X_{ik} \in \{-1, 1\}. \tag{5}$$

The efficiency  $D_{eff}$  (Galil & Kiefer, 1980) of a design is therefore defined by

$$D_{eff} = \left( \frac{D}{D^*} \right)^{\frac{1}{n+1}}, \tag{6}$$

where  $D$  is the determinant of  $X^t X$  and

$$D^* = N^{n+1} \tag{7}$$

denotes the upper bound in (4). A design achieving the bound (4) has efficiency 1 and is called D-optimal (Shah & Sinha, 1989).

For screening itself, the efficiency measure is immaterial. However, when – perhaps after suitable transformations of the data – a linear relation happens to be the correct model, then a high efficiency implies that good use was made of the given number of experiments.

The simplest edge designs, “one factor at a time” designs, with

$$x^0 = j, x^i = j - 2e^i, (i = 1, \dots, n), \quad (8)$$

where  $e^i = (0, \dots, 1, 0, \dots)^t$  with the 1 at the  $i$ th position, have very low efficiency; indeed,

$$X = \begin{pmatrix} 1 & j^t \\ j & J - 2I \end{pmatrix}, \quad \det X^t X = (\det X)^2 = 2^{2n}, \quad (9)$$

with marginal efficiency

$$D_{eff} = 2^{\frac{2n}{n+1}} / (n+1) \rightarrow 0 \quad \text{for } n \rightarrow \infty. \quad (10)$$

When there are enough edges, D-optimal edge designs do exist. For example, we may consider edge designs consisting of  $4n$  trials, where the corresponding  $4n \times (n+1)$  design matrix  $X$  has the special form

$$X = \begin{pmatrix} j & C + I \\ j & C - I \\ j & -C + I \\ j & -C - I \end{pmatrix} \quad (11)$$

with a  $n \times n$  matrix  $C$  satisfying  $C_{ii} = 0$ . Here the  $i$ th row together with the  $(n + i)$ th row as well as the  $(2n + i)$ th row together with the  $(3n + i)$ th row define edges. Optimality conditions are given by

**Theorem 1:** The design matrix (11) satisfies

$$X^t X = \begin{pmatrix} 4n & 0 \\ 0 & 4C^t C + 4I \end{pmatrix}. \quad (12)$$

In particular, the design is D-optimal whenever

$$C^t C = (n - 1)I. \quad (13)$$

Proof: (12) follows directly from (11), and D-optimality follows from (7) since (13) implies  $X^t X = 4nI$  and  $N = 4n$ .  $\square$

If  $C$  is a conference matrix, as defined in Section 4, (13) is satisfied. The resulting “double” conference designs give many examples of D-optimal edge designs with  $4n$  trials. We did not investigate the existence of D-optimal edge designs for other values of  $N$  though surely many exist.

When  $n$  is large and experiments are expensive it is important to keep the number of screening experiments as low as possible. We call an edge design minimal if it has just  $n$  edges, one for each variable. We conjecture that a minimal edge design cannot be D-optimal. We are able to prove this for the most natural class of minimal edge designs, namely those with  $N = 2n$  trials and  $n$  disjoint edges. In this case, the design matrix  $X$  may be written in

the form

$$X = \begin{pmatrix} j & S + I \\ j & S - I \end{pmatrix}. \quad (14)$$

$S$  is an  $n \times n$  matrix with

$$S_{ii} = 0 \text{ and } |S_{ij}| \leq 1, \quad (15)$$

so that the  $i$ th row and the  $(n + i)$ th row define an edge contained in  $\mathbb{Q}$ . If we denote by  $x^1, \dots, x^{2n}$  the points of the design with design matrix (14), then  $x^i$  and  $x^{i+n}$  form an edge, differing only in the  $i$ th coordinate, and the corresponding midpoints

$$v^i = \frac{1}{2}(x^i + x^{i+n}) \quad (16)$$

are the transposed rows of  $S$ . From (14) we obtain

$$X^t X = 2 \begin{pmatrix} n & j^t S \\ S^t j & I + S^t S \end{pmatrix}. \quad (17)$$

We first prove two auxiliary results.

**Lemma 1:** For  $X$  of the form (14), we have

$$\det X^t X = n2^{n+1} \det(I + S^t S - \frac{1}{n} S^t J S) = n2^{n+1} \det(I + S^t K S),$$

where

$$K = I - \frac{1}{n} J. \quad (18)$$

*Proof:* Straightforward.  $\square$

**Lemma 2:** If  $S$  satisfies (15) then

$$\det(I + S^t K S) \leq (n + 1)^{n-1}.$$

*Proof:* Since (18) implies  $Kj = 0$  we conclude that  $S^t K S$  is singular. We denote by  $\lambda_1, \dots, \lambda_n$  the eigenvalues of  $S^t K S$  and assume without loss of generality  $\lambda_1 = 0$ . Since  $S^t K S$  is positive semidefinite, the  $\lambda_i$  are nonnegative and we obtain

$$\begin{aligned} \{\det(I + S^t K S)\}^{\frac{1}{n-1}} &= \left\{ \prod_{i=2}^n (1 + \lambda_i) \right\}^{\frac{1}{n-1}} \leq \frac{1}{n-1} \sum_{i=2}^n (1 + \lambda_i) \\ &= \frac{1}{n-1} (n-1 + \text{tr } S^t K S) \\ &= 1 + \frac{\text{tr } S^t S}{n-1} - \frac{1}{n(n-1)} \text{tr } (S^t j j^t S) \\ &\leq 1 + \frac{\text{tr } S^t S}{n-1} \leq 1 + \frac{n(n-1)}{n-1} = n + 1. \end{aligned}$$

□

The following bound shows that minimal edge designs of the form (14) cannot be D-optimal; but the characterization of the equality case opens the way to the construction of a large class of minimal edge designs which are at least asymptotically D-optimal.

**Theorem 2:** Let  $X$  have the form (14) with  $S$  restricted by (15). Then

$$\det X^t X \leq 2^{n+1} n (n + 1)^{n-1}, \quad (19)$$

with equality iff

$$S S^t = nI - J, \quad S^t j = 0, \quad (20)$$

and hence also

$$S_{ii} = 0, \quad S_{ij} \in \{1, -1\} \text{ for } i \neq j. \quad (21)$$

*Proof:* (19) follows directly from the two lemmata. If (20) holds then

$$\det(I + S^t K S) = \det(I + S^t S) = \det(I + S S^t)$$

since  $S^t S$  and  $S S^t$  have the same characteristic polynomial. Therefore,

$$\det(I + S^t K S) = \det \{(n+1)I - J\} = (n+1)^{n-1},$$

and Lemma 2 implies equality. Conversely, if equality holds in (19) then the estimates in the proof of Lemma 2 are all equalities, hence

$$\begin{aligned} \lambda_2 = \dots = \lambda_n &= n, \\ 0 &= \text{tr}(S^t j j^t S) = \|S^t j\|^2, \\ n(n-1) &= \text{tr } S^t S = \sum S_{ik}^2. \end{aligned}$$

The second relation forces  $S^t j = 0$ . Therefore  $S^t K S = S^t S$ , and since  $S^t S$  and  $S S^t$  have the same eigenvalues, the first relation shows that  $nI - S S^t$  has rank 1. Since this matrix is symmetric,  $nI - S S^t = \alpha u u^t$  for some suitable  $\alpha \in \mathbb{R}, u \in \mathbb{R}^n$ . Multiplication with  $j$  shows that  $nj = \alpha u u^t j$ . Hence  $u$  is a multiple of  $j$ , and without loss of generality,  $u = j$ ,  $u u^t = J$ . Then  $n = \alpha u^t j = \alpha j^t j = \alpha n$ , hence  $\alpha = 1$  and (20) holds. Finally, since  $S_{ii} = 0$ , the third relation forces  $S_{ij} = \pm 1$  for  $i \neq j$ , hence (21).  $\square$

We shall call a design with design matrix (14), where  $S$  satisfies (20), (21) and

$$S j = 0 \quad (22)$$

a conference design. Indeed, (20), (21) and (22) hold iff the matrix

$$C := \begin{pmatrix} 0 & j^t \\ j & S \end{pmatrix}$$

is a so-called conference matrix. Condition (22) is very convenient, since it relates our designs to an extensive body of knowledge about conference matrices, cf. Section 4.

If we multiply some columns of any  $S$  satisfying (20), (21) and (22) by  $-1$  we get another matrix satisfying (20) and (21), but usually not (22). It would be interesting to know whether every matrix  $S$  with (20) and (21) arises in this way.

**Theorem 3:** The efficiency of conference designs is given by

$$D_{eff} = \frac{n+1}{n} \left( \frac{n}{(n+1)^2} \right)^{\frac{1}{n+1}} = 1 - \frac{\log n}{n} + O\left(\frac{1}{n}\right),$$

and

$$\lim_{n \rightarrow \infty} D_{eff} = 1.$$

*Proof:* By Theorem 2, conference designs have maximal determinant among all designs of the form (14),

$$\det X^t X = 2^{n+1} n(n+1)^{n-1}. \quad (23)$$

Thus the result follows directly from equation (7).  $\square$

For conference designs,  $D_{eff}$  grows monotonically for  $n \geq 5$ , and for large  $n$ , conference designs are asymptotically D-optimal. This increase in efficiency

should be compared with the decay to zero of the efficiency (10) of "one factor at a time" designs:

$n$	5	15	25	35	45
$D_{eff}(\text{conference design})$	0.864	0.893	0.916	0.930	0.940
$D_{eff}(\text{one factor at a time})$	0.529	0.229	0.146	0.107	0.084

Note, however, that the number of design points is  $N = 2n$  in the case of conference designs and  $N = n + 1$  in the case of one factor at a time designs.

## 4 A survey of conference matrices

In this section we discuss properties and construction of the matrices  $S$  and  $C$  needed for the new designs.

A conference matrix of order  $n + 1$  is a  $(n + 1) \times (n + 1)$  matrix  $C$  satisfying

$$C^t C = C C^t = nI,$$

$$C_{ii} = 0, \quad i = 1, \dots, n + 1,$$

$$C_{ij} \in \{-1, 1\} \quad \text{for } i \neq j.$$

The name derives from an application to telephone conference networks (Belevitch, 1968). Multiplying any column or row of a conference matrix



by  $-1$ , and permuting rows and columns leads again to a conference matrix, and conference matrices related in this way are called equivalent. The following central result was proved by Delsarte, Goethals & Seidel (1971).

**Theorem 4:** Any conference matrix of order  $n + 1$  is equivalent to a matrix of the form

$$C = \begin{pmatrix} 0 & \epsilon j^t \\ j & S \end{pmatrix}, \quad \epsilon = (-1)^{\frac{n-1}{2}}, \quad (24)$$

where the  $n \times n$  matrix  $S$  satisfies the relations (20) and (21), and

$$S^t = \epsilon S, S j = 0. \quad (25)$$

Conversely, (24) defines a conference matrix for any matrix  $S$  satisfying (20) and (21) – but not necessarily (25).  $\square$

The relations  $S^t j = S j = 0$  state that each row and column of  $S$  contains the same number  $(n - 1)/2$  of ones, hence  $n$  must be odd.

**Case 1.**  $n \equiv 1 \pmod{4}$ . Then  $\epsilon = 1$  and the matrices (24) are symmetric. For symmetric conference matrices, another restriction was proved by Raghavarao (1960):

**Theorem 5:** If a symmetric conference matrix of order  $n \equiv 1 \pmod{4}$  exists then  $n$  is a sum of two squares.  $\square$

A simpler proof of this restriction was given by van Lint & Seidel (1966). For  $n \leq 50$ , this excludes  $n = 21$  and  $n = 33$ . They conjecture that symmetric

conference matrices of order  $n + 1$  exist whenever  $n \equiv 1 \pmod{4}$  and  $n$  is a sum of two squares; the constructions known until 1972 and the list of dimensions  $\leq 4000$  for which they apply are listed in Wallis, Street & Wallis (1972), Appendix C and Appendix I. A further construction was given by Mathon (1978), and generalized by Seberry & Whiteman (1983, 1988). This construction covers  $n = 45$ ; hence the first unsettled case is now  $n = 65$ . A report by Bussemaker, Mathon & Seidel (1979) contains comprehensive tables of the known symmetric conference matrices for  $n + 1 \leq 50$ , phrased in terms of the equivalent concept of regular two-graphs with  $n + 1$  vertices and eigenvalues  $\pm n^{\frac{1}{2}}$ . In particular, the number  $N(n)$  of inequivalent symmetric conference matrices of order  $n + 1$  is given by

$n$	1	5	9	13	17	21	25	29	33	37	41	45	49
$N(n)$	1	1	1	1	1	0	4	$\geq 6$	0	$\geq 11$	$\geq 18$	$\geq 80$	$\geq 18$

**Case 2.**  $n \equiv 3 \pmod{4}$ . Then  $\epsilon = -1$ , the matrices (24) are skew-symmetric. For a skew conference matrix  $C$ , the matrix  $H = I + C$  satisfies  $H^t H = I + C^t + C + C^t C = (n + 1)I$ . Thus  $H$  is a Hadamard matrix, and such Hadamard matrices are called skew, though they are not skew-symmetric. Wallis (1971) conjectured that skew Hadamard matrices exist for all dimensions divisible by 4; the known constructions and a list of dimensions  $\leq 4000$  for which they apply are listed in Seberry & Yamadi (1992), Chapter 7. The first unsettled case corresponds to  $n = 187$ . The number  $N(n)$  of inequivalent skew Hadamard matrices of order  $n + 1$  is given by

$n$	3	7	11	15	19	23	27	31	35	39	43	47
$N(n)$	1	1	1	2	1	16	$\geq 49$	$\geq 6$	$\geq 1$	$\geq 1$	$\geq 3$	$\geq 1$

Conference matrices of order  $n + 1$  are also discussed in Geramita & Seberry (1979) under the name of weighing matrices  $W(n + 1, n)$ .

Various conference matrices can be obtained by difference set methods: When  $n$  is an odd prime and  $R$  denotes the set of quadratic residues mod  $n$ , then the cyclic matrix  $S \in \mathbb{R}^{n \times n}$  with

$$S_{ik} = \begin{cases} 0 & \text{if } i = k, \\ 1 & \text{if } k - i \in R \text{ mod } n, \\ -1 & \text{otherwise} \end{cases} \quad (26)$$

satisfies (20) and (21) and thus defines a conference matrix (Paley, 1933). Results by Schur (1933) imply that, for  $n \equiv 1 \pmod{4}$ , any set  $R$  such that (24), (26) determine a conference matrix must be of the form given by Paley. For easy reference, we display in Table 5 the sets  $R$  for the odd primes below 50; an example for the use of the table is given in Section 6 below. For larger primes, consult the tables of indices in Jacobi (1956) or Nichol, Selfridge & McKee (1962);  $R$  consists of the set of numbers with even indices.

**Table 5:** Difference sets for conference matrices

$n$	$R$
3	1
5	1,4
7	1,2,4
11	1,3,4,5,9
13	1,3,4,9,10,12
17	1,2,4,8,9,13,15,16
19	1,4,5,6,7,9,11,16,17
23	1,2,3,4,6,8,9,12,13,16,18
29	1,4,5,6,7,9,13,16,20,22,23,24,25,28
31	1,2,4,5,7,8,9,10,14,16,18,19,20,25,28
37	1,3,4,7,9,10,11,12,16,21,25,26,27,28,30,33,34,36
41	1,2,4,5,8,9,10,16,18,20,21,23,25,31,32,33,36,37,39,40
43	1,4,6,9,10,11,13,14,15,16,17,21,23,24,25,31,35,36,38,40,41
47	1,2,3,4,6,7,8,9,12,14,16,17,18,21,24,25,27,28,32,34,36,37,42

When  $n$  is an odd prime power, a construction by Delsarte, Goethals & Seidel (1971) gives  $M \subseteq \{1, \dots, v\}$  ( $v = n + 1$ ) such that the “negacyclic” matrix

$C \in \mathbb{R}^{v \times v}$  with

$$C_{ik} = \begin{cases} 0 & \text{if } k = i, \\ 1 & \text{if } k \geq i \text{ and } k - i \in M, \\ 1 & \text{if } k < i \text{ and } k - i + v \notin M, \\ -1 & \text{otherwise} \end{cases}$$

is a conference matrix. Unfortunately, their construction requires calculations in prime field extensions, hence explicit sets  $M$  are not easy to compute. They conjecture that conference matrices of this negacyclic form exist only when  $n$  is a prime power, and they prove this for  $n \leq 225$ .

Among the other constructions known (Wallis, Street & Wallis (1972) and Mathon (1978)) we only mention the doubling construction: If  $C$  is a skew conference matrix of order  $n + 1$  then

$$\bar{C} := \begin{pmatrix} C & C + I \\ C - I & C \end{pmatrix} \quad (27)$$

is a skew conference matrix of order  $2n + 2$ . Assuming the validity of Wallis' conjecture, this yields conference matrices of all orders  $n + 1 \equiv 0 \pmod{8}$  from their halves.

**Table 6:** Conference matrices for odd nonprimes  $< 50$ .

$n$	
9	(Bussemaker, Mathon & Seidel, 1979)
15	double $n = 7$
21	nonexistent by Theorem 4
25	(Bussemaker, Mathon & Seidel, 1979)
27	(Wallis, Street & Wallis, 1972)
33	nonexistent by Theorem 4
35	double $n = 19$
45	(Bussemaker, Mathon & Seidel, 1979)
49	(Bussemaker, Mathon & Seidel, 1979)

## 5 Space exploration properties

In most situations where screening experiments are done, there is no prior information on the relationship between the variables. Since after screening, only the variables found active are varied in subsequent experiments, mistakes in the decision of which variables are active cannot be corrected later.

We showed already the risk involved in screening solely based on a linear model. On the other hand, for a nonlinear model, the shape of a response surface may be different in different regions of the box of interest. Since this

may affect the decision on the right set of variables, it is important that the design points are located such as to explore the whole design space  $\mathbb{Q}$  to get a good global view of the response surface and to guard against unexpected nonlinearities.

For example, when  $n$  is large, “one factor at a time” designs (8) cluster the design points around  $x^0$ , and are therefore not necessarily representative for the region of interest. If we want to avoid such clustering, a natural criterion to consider is the minimum distance between the midpoints (8) of different edges of the design. If this distance is large, edges will be far away and hence obtain information about different regions of the box.

It is therefore interesting to observe that in terms of this minimum distance criterion, conference designs have an optimal space exploration property:

**Theorem 6:** For arbitrary  $v^1, \dots, v^n \in \mathbb{Q}$  with  $v_i^i = 0$  ( $i = 1, \dots, n$ ) we have

$$\min_{i \neq j} \|v^i - v^j\| \leq (2n)^{\frac{1}{2}}, \quad (28)$$

and equality holds iff the matrix  $S$  whose rows are  $v^1, \dots, v^n$  satisfies (20) and (21).

*Proof:* Note that

$$0 \leq \left\| \sum_{i=1}^n v^i \right\|^2 = \sum_{i=1}^n \|v^i\|^2 + \sum_{i \neq j} (v^i, v^j) \quad (29)$$

implies

$$\max_{i \neq j} (v^i, v^j) \geq -\frac{\sum \|v^i\|^2}{n(n-1)} \geq -1. \quad (30)$$

Therefore, for at least one pair of edges,

$$\|v^i - v^j\|^2 = \|v^i\|^2 + \|v^j\|^2 - 2(v^i, v^j) \leq 2n - 2 - 2(v^i, v^j) \leq 2n,$$

and (28) follows. Equality in (28) holds iff  $\|v^i\|^2 = n - 1$  and equality holds throughout (29) and (30), hence iff  $S^t j = \sum v^i = 0$ ,  $(SS^t)_{ii} = \|v^i\|^2 = n - 1$  for all  $i$ , and  $(SS^t)_{ij} = (v^i, v^j) = -1$  for all  $i \neq j$ , which is equivalent to (20), and  $S_{ij} = v_j^i \in \{1, -1\}$  for  $i \neq j$  since  $\sum_{i \neq j} (v_j^i)^2 = \|v^i\|^2 = n - 1$  and  $|v_j^i| \leq 1$ .  $\square$

In particular, the minimal distance  $(2n)^{\frac{1}{2}}$  of the midpoints of edges of a conference design is more than one third of the maximal distance  $2n^{\frac{1}{2}}$  realizable within  $\mathbb{Q}$ . For comparison, note that for “one factor at a time” designs, we have  $\|v^i - v^j\|_2 = 2^{\frac{1}{2}}$  and  $\|x^i - x^j\|_2 \leq 8^{\frac{1}{2}}$ , which means that these designs explore the design space only near  $x^0$ .

It is interesting to note that Hadamard designs, i.e., minimal D-optimal designs where  $X \in \mathbb{R}^{N \times (n+1)}$  satisfies (5) for  $N = n + 1$  have a similar optimality property; but lacking a means to test for the presence of nonlinearities, this property is here not very helpful.

**Theorem 7:** For arbitrary  $x^1, \dots, x^{n+1} \in \mathbb{Q}$ , we have

$$\min_{i \neq j} \|x^i - x^j\| \leq (2n + 2)^{\frac{1}{2}}, \quad (31)$$

and equality holds iff the matrix  $X$  defined by (2) is a Hadamard matrix.



*Proof:* Now we have

$$0 \leq \left\| \sum_{i=1}^{n+1} x^i \right\|^2 = \sum_{i=1}^{n+1} \|x^i\|^2 + \sum_{i \neq j} (x^i, x^j), \quad (32)$$

giving

$$\max_{i \neq j} (x^i, x^j) \geq -\frac{\sum \|x^i\|^2}{n(n+1)} \geq -1. \quad (33)$$

Since

$$\|x^i - x^j\|^2 = \|x^i\|^2 + \|x^j\|^2 - 2(x^i, x^j) \leq n + n - 2(x^i, x^j),$$

we conclude that

$$\min_{i \neq j} \|x^i - x^j\|^2 \leq n + n - 2 \max_{i \neq j} (x^i, x^j) \leq 2n + 2,$$

giving (31). Equality holds in (31) iff  $\|x^i\|^2 = n$  for all  $i$  and  $\max_{i \neq j} (x^i, x^j) = -1$ .

The first condition says that the  $x^i$  must be  $(1, -1)$ -vectors, and the second condition together with (32) and (33) forces  $(x^i, x^j) = -1$  for all  $i \neq j$ , giving  $XX^t = (n+1)I$ . This is equivalent with  $X^t = (n+1)X^{-1}$  and hence with  $X^tX = (n+1)I$ . Thus  $X$  must be a Hadamard matrix.  $\square$

## 6 Practical aspects

Tables 5 and 6 of Section 4 give explicitly a large number of conference matrices and hence of conference designs and double conference designs; further conference designs of the same order can be obtained by permuting the variables, and further double conference designs by permuting variables and/or

by switching the signs of some columns of  $C$ . This allows for randomization in the use of these designs.

For cases where no conference matrix exists or is available, in particular for all even  $n$ , one can use the next higher “good” value of  $n$  by introducing one or more extra variables without influence on the data. Thus edge designs which are optimal, or nearly optimal when  $n$  was increased, are available for all dimensions likely to be encountered in practice.

Edge designs may be evaluated in two independent ways: One can use the median estimates of Section 2 to decide on the correct set of active factors and get an estimate  $\tilde{\sigma}$  of the variance. This part of the analysis is model-independent. On the other hand, one can fit a linear model to the data or some nonlinear transformation of it and use the fitted model to identify the active factors by the corresponding significant main effects. If the two subsets of estimated active factors coincide one may conclude that the active factors were found. If the two subsets do not coincide, however, further measurements are needed, based on a design for a higher order model in the space determined by all factors found to be active by one of the methods of analysis.

When there are outliers in the data or the noise comes close in magnitude to the influence of some variable, the assumptions underlying our analysis are no longer justified, and further research is needed to analyse these cases properly.

**Table 7:** Conference design and rounded function values

Run	1	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$y = 0.9(x_1 - x_2)^2 + 0.7x_1 + 0.4x_2x_3 + \epsilon$
1	+	+	+	+	-	+	-	-	0.97
2	+	-	+	+	+	-	+	-	3.10
3	+	-	-	+	+	+	-	+	-1.24
4	+	+	-	-	+	+	+	-	4.80
5	+	-	+	-	-	+	+	+	2.46
6	+	+	-	+	-	-	+	+	3.79
7	+	+	+	-	+	-	-	+	0.24
8	+	-	+	+	-	+	-	-	3.42
9	+	-	-	+	+	-	+	-	-1.13
10	+	-	-	-	+	+	-	+	-0.38
11	+	+	-	-	-	+	+	-	4.61
12	+	-	+	-	-	-	+	+	2.42
13	+	+	-	+	-	-	-	+	3.98
14	+	+	+	-	+	-	-	-	0.24

We conclude our work with the application of a conference design to a simulated screening scenario. The data are simulated with the noisy quadratic function

$$f(x_1, \dots, x_7) = 0.9(x_1 - x_2)^2 + 0.7x_1 + 0.4x_2x_3 + \epsilon,$$

where

$$x_i \in [-1, 1], i = 1, \dots, 7, \quad \epsilon \sim N(0, \sigma^2)$$

and  $\sigma = 0.1$ . A conference design for dimension  $n = 7$  was generated from the difference set  $R = \{1, 2, 4\}$  given in Table 5, using the construction (14) with  $S$  defined by (26). Table 7 shows the resulting conference design together with the obtained function values. Note that the two sets of 7 rows differ only on the diagonal, showing that rows  $i$  and  $i + 7$  form an edge.

An analysis of the data with the software package RS1/Discover (1992) revealed no significant main effect, and thus gave an estimated linear model  $f(x) = 1.94 + \epsilon$ , with  $\epsilon$  of mean zero and standard deviation  $\sigma = 1.99$ .

On the other hand, an analysis of the edges, cf. Table 8, shows that the first three of the 7 differences of function values along the 7 edges of the design are significant, with an absolute value of more than four times the robust estimation  $\tilde{\sigma} = 0.20$  of  $\sigma$  defined in (3).

Thus both the need for a nonlinear model, and the set of active factors were discovered correctly, and further experiments are needed to estimate an appropriate 3-dimensional nonlinear model.

**Table 8:** Model-independent checks for the conference design

$z_1$	$z_2$	$z_3$	$z_4$	$z_5$	$z_6$	$z_7$	
$y^1 - y^8$	$y^2 - y^9$	$y^3 - y^{10}$	$y^4 - y^{11}$	$y^5 - y^{12}$	$y^6 - y^{13}$	$y^7 - y^{14}$	$\tilde{\sigma}$
-2.45	4.23	-0.86	0.19	0.04	-0.19	0.00	0.20

## References

- [1] Belevitch, V. (1968). Conference networks and Hadamard matrices. *Ann. Soc. Sci. Bruxelles (Ser.1)* **82**, 13-32.
- [2] RS1/Discover (1992). Bolt, Beranek and Newman Inc. Software Products Division, 150 Cambridge Park Drive, Cambridge, MA 02140.
- [3] Box, G.E.P., and N.R. Draper (1987). *Empirical Model-Building and Response Surfaces*. New York, Wiley 1987.
- [4] Bussemaker, F.C., R.A. Mathon and J.J. Seidel (1979). Tables of two-graphs, *TH-Report 79-WSK-05*. Dept. of Mathematics, Technological Univ. Eindhoven, Netherlands. Also published in: pp. 70-112 of: *Combinatorics and Graph theory* (S.B. Rao, ed.), *Lecture Notes in Mathematics* **885**, Springer, Berlin 1981.
- [5] Delsarte, P., J.M. Goethals and J.J. Seidel (1971). Orthogonal matrices with zero diagonal. II, *Canad. J. Math.* **23**, 816-832.
- [6] Galil, Z., J. Kiefer (1980). Time- and space-saving computer methods, related to Mitchell's DETMAX, for finding D-optimum designs. *Technometrics* **22**, 301-313.
- [7] Geramita, A.V. and S. Seberry (1979). *Orthogonal designs*. Dekker, New York 1979.

- [8] Jacobi, C. G. (1956). *Canon arithmeticus*, revised ed., Akademie-Verlag, Berlin 1956.
- [9] Lenth, R.S. (1989). Quick and easy analysis of unreplicated factorials. *Technometrics*, **31**, 469-473.
- [10] van Lint, S.H. and J.J. Seidel (1966). Equilateral point sets in elliptic geometry, *Indag. Math.* **28**, 335-348.
- [11] Mathon, R.A. (1978). Symmetric conference matrices of order  $pq^2 + 1$ , *Cand. J. Math.* **30**, 321-331.
- [12] Nichol, C. A., Selfridge, J. L., and McKee, L. (1962). A table of indices and power residues for all primes and prime powers below 2000, Norton, New York 1962.
- [13] Paley, R.E.A.C. (1933). On orthogonal matrices. *J. Math. and Phys.* **12**, 311-320.
- [14] Plackett, R.L. and J.P. Burman (1946). The design of optimum multifactorial experiments. *Biometrika* **33**, 305-325.
- [15] Raghavarao D. (1960). Some aspects of weighing designs. *Ann. Math. Stat.* **31**, 878-884.
- [16] Schur, I. (1933). Zur Theorie der einfach transitiven Permutationsgruppen. *Sitz. Ber. Preuss. Akad. Wiss., Phys.-Math. Kl.* 598-623.

- [17] Seberry, J. and A.L. Whiteman (1983). A new construction for conference matrices, *Ars. Combinatoria*, **16**, 119-127.
- [18] Seberry, J. and A.L. Whiteman (1988). New Hadamard matrices and conference matrices obtained via Mathon's construction, *Graphs and Combinatorics*, **4**, 355-377.
- [19] Seberry, J. and M. Yamada (1992). Hadamard matrices, sequences and block designs, pp. 431-560 in: *Contemporary design theory: a collection of surveys* (J.H. Dinitz and D.R. Stinton, eds.), Wiley, New York 1992.
- [20] Shah, K.R. and B.K. Sinha (1989). Theory of optimal design, *Lecture Notes in Statistics* **54**, Springer, New York 1989.
- [21] Wallis, J. (1971). Some (1,-1) matrices. *J. Combinatorial Theory*, Ser.B, **10**, 1-11.
- [22] Wallis, J., A.P. Street and J. Seberry Wallis (1972). Combinatorics: Room squares, sum-free sets, Hadamard matrices, *Lecture Notes in Mathematics* **292**, Springer, Berlin 1972.