D-Efficient Mixed-Level Foldover Designs for Screening Experiments

16

Nam-Ky Nguyen, Ron S. Kenett, Tung-Dinh Pham, and Mai Phuong Vuong

Contents

Abstract

Definitive screening design (DSD) is a new class of three-level screening designs proposed by Jones and Nachtsheim [3] which only requires $2m+1$ runs for experiments with m three-level quantitative factors. The design matrices for DSDs are of the form $(\mathbf{C}', \ -\mathbf{C}', \ \mathbf{0})'$ where $\mathbf{C}$ is a $(0, \pm 1)$ submatrix with zero diagonal and $\mathbf{0}$ is a column vector of 0's. This paper reviews recent development on D-efficient mixed-level foldover designs for screening experiments. It then discusses a fast coordinate-exchange algorithm for constructing D-efficient DSD-augmented designs (ADSDs). This algorithm provides a new class of conference matrix-based mixed-level foldover designs (MLFODs) for screening experiments as introduced by Jones and Nachtsheim [4]. In addition, the paper also provides an alternative class of D-efficient MLFODs and an exhaustive algorithm for constructing the new designs. A case study comparing two candidate MLFODs for a large mixed-level screening experiment with 17 factors used is used to demonstrate the properties of the new designs.

Keywords

Conference matrix · Coordinate-exchange algorithm · Foldover design · Hadamard matrix · Plackett-Burman design

N.-K. Nguyen (✉)
Vietnam Institute for Advanced Study in Mathematics, Hanoi, Vietnam
e-mail: nknam@viasm.edu.vn

R. S. Kenett
KPA Ltd., Samuel Neaman Institute, Technion, Israel
e-mail: ron@kpa-group.com

T.-D. Pham
VNU University of Science, Hanoi, Vietnam
e-mail: tungpd@vnu.edu.vn

M. P. Vuong
Hanoi University of Science & Technology, Hanoi, Vietnam
e-mail: phuong.vuongmai@hust.edu.vn

16.1 Introduction

Screening experiments are designed to sort a typically long list of factors that can potentially affect the response variables of a product or process. The sorting highlights active factors. This experimentation strategy is widely applied in science and engineering. Another approach, pioneered by Genichi Taguchi, is to follow an experimental path of system design, parameter design, and tolerance design [6]. In this paper we consider the screening-optimizing continuum with the objective of improving the knowledge acquisition effort by increasing its quality and reducing the required effort. Most screening experiments in engineering and science involve both two-level and three-level factors. Yet, the most popular screening designs are two-level designs such as resolution III and IV fractional factorial designs (FFDs). Jones and Nachtsheim [3] pointed out the following disadvantages of using two-level FFDs to study quantitative factors:

(i) Quadratic effects are not estimable if they are included in the model;

H. Pham (ed.), *Springer Handbook of Engineering Statistics*, Springer Handbooks,
https://doi.org/10.1007/978-1-4471-7503-2_16

(ii) Main effects are not completely orthogonal to two-factor interactions as in the case of resolution III FFDs;
(iii) Certain two-factor interactions are fully aliased with one another as in the case of resolution IV FFDs;

A new class of three-level screening designs called DSDs introduced by Jones and Nachtsheim [3] eliminates these shortcomings. In addition, all quadratic effects of DSDs are orthogonal to main effects and not fully aliased with two-factor interactions. The design matrix for a DSD can be written as:

$$\begin{pmatrix} \mathbf{C} \\ -\mathbf{C} \\ \mathbf{0}' \end{pmatrix}, \tag{16.1}$$

where $\mathbf{C}$ is an $m \times m$ $(0, \pm 1)$ submatrix with zero diagonal and $\mathbf{0}$ is a column vector of 0's. Xiao et al. [14] pointed out that if we use a *conference* matrix of order m for $\mathbf{C}$, i.e., if $\mathbf{C}'\mathbf{C} = (m-1)\,\mathbf{I}_{m\times m}$, then the DSD is also orthogonal for main effects, i.e., all main effects are orthogonal to one another. For even $m \leq 50$ except for $m = 22$ and $m = 34$, the $\mathbf{C}$ matrices which are also conference matrices are given by Xiao et al. [14] and Nguyen and Stylianou [11]. All the $\mathbf{C}$ matrices we use in this paper are conference matrices with the exception of the one of order 22. Figure 16.1 shows the $\mathbf{C}$ matrices of order $m = 4$, 6, 8, and 10 generated by the cyclic generators given by Nguyen and Stylianou [11].

The limitation of a DSD is that all factors should be quantitative. Jones and Nachtsheim [4] (hereafter abbreviated as JN) introduced two types of conference matrix-based mixed-level screening designs. They called the more D-efficient, more economic one, DSD-augmented designs (ADSDs). ADSDs are in fact belonging to a class of mixed-level foldover designs (MLFODs), and, as such, they retain two advantages of the original DSD, namely, (i) all quadratic effects are orthogonal to main effects and (ii) all main effects are orthogonal to two-factor interactions. The latter feature is extremely useful when the experimenter wishes to include a specific two-factor interaction in the model. The limitation of this class of designs is that it has a high correlation among the quadratic effects. JN showed that this correlation was $\frac{1}{2} - \frac{2}{n-4}$ where n is the number of runs.

m=4	m=6	m=8	m=10
0+++	0+++++	0+++++++	0+−−+−++++
+0−+	+0−++−	+0++−+−−	+0+−−+−+++
++0−	+−0−++	+−0++−+−	−+0+−++−++
+−+0	++−0−+	+−−0++−+	−−+0++++−+
	+++−0−	++−−0++−	+−−+0++++−
	+−++−0	+−+−−0++	−++++−0−++−
		++−+−−0+	+−+++−−0−++
		+++−+−−0	++−+++−−0−+
			+++−+++−−0−
			++++−−++−−0

Fig. 16.1 Conference $\mathbf{C}$ matrices of order m (+ denotes +1 and − denotes −1)

Nguyen et al. [12] introduced a new class of Hadamard matrix-based mixed-level foldover designs (MLFODs) and an algorithm which produces these MLFODs. These MLFODs were constructed by converting some two-level columns of a Hadamard matrix to three-level ones (see [2] for the information on Hadamard matrices and their use in design construction). Like the two-level foldover designs (FODs), each new MLFOD was constructed by a half fraction and its foldover. These Hadamard matrix-based MLFODs require fewer runs and compare favorably with the conference matrix-based MLFODs of [4] in terms of the D-efficiencies and r_{max}, the maximum of the absolute values of the correlation coefficients among the columns of the model matrix. Like the ADSDs, these MLFODs are also *definitive* in the sense that the estimates of all main effects are unbiased with respect to any active second-order effects. The limitation of this class of Hadamard matrix-based MLFODs is that it is not very efficient when the number of three-level factors is greater than the number of two-level ones.

The design matrix of a conference matrix-based MLFOD for m_3 three-level factors and m_2 two-level factors has a foldover structure and can be written as:

$$\begin{pmatrix} \mathbf{D} \\ -\mathbf{D} \end{pmatrix}, \tag{16.2}$$

where the submatrix $\mathbf{D}_{m^*\times(m_3+m_2)}$ can be constructed from a matrix of order m $(m_3 + m_2 \leq m \leq m^*)$. In the following sections, we describe (i) a fast coordinate-exchange algorithm (CEA) [8] for constructing ADSDs which are conference matrix based and (ii) an exhaustive search algorithm or ESA for constructing an alternative class of MLFODs. Both algorithms attempt to transform a base matrix to the submatrix $\mathbf{D}$ in (16.2) from which a D-efficient MLFOD can be obtained. For ADSDs, the base matrix is a conference matrix. For the new MLFODs, the base matrix is a two-level orthogonal matrix such as Hadamard matrix or a Plackett-Burman design [13].

16.2 A CEA for Constructing D-efficient ADSDs

The Appendix shows that the first-order and second-order D-efficiencies d_1 and d_2 of an ADSD are functions of $|\mathbf{B}|$ $(= |\mathbf{D}'\mathbf{D}|)$ where $|\mathbf{B}|$ is the determinant of matrix $\mathbf{B}$. Our algorithm minimizes the sum of squares of the off-diagonal elements of $\mathbf{B}$ as an indirect attempt to maximize $|\mathbf{B}|$

(see, e.g., [9]). The steps of our CEA for obtaining a submatrix $\mathbf{D}$ in (2) from which a D-efficient ADSD for m_3 three-level factors and m_2 two-level factors are:

1. Form a starting design $\mathbf{D}_{(m+1)\times(m_3+m_2)} = (d_{ij})$ by first picking $m_3 + m_2$ columns from a conference matrix of order m at random and add an extra zero row to the bottom of these columns. Mark the positions of the $2m_2$ zero entries in the last m_2 columns of $\mathbf{D}$, and replace these 0's by ± 1 in a random manner. Calculate the vector $\mathbf{J}_i$ $(i = 1, \ldots, m+1)$ of length $m_3 m_2 + \binom{m_2}{2}$ for each row of $\mathbf{D}$ where $\mathbf{J}_i$ is defined as $(d_{i1}d_{i(m_3+1)}, \ldots, d_{im_3}d_{i(m_3+m_2)}, d_{i(m_3+1)}d_{i(m_3+2)}, \ldots, d_{i(m_3+m_2-1)}d_{i(m_3+m_2)})$. Let $\mathbf{J} = \sum_{i=1}^{m+1} \mathbf{J}_i$ and f equal the sum of squares of the elements of $\mathbf{J}$.
2. Among the $2m_2$ marked positions in Step 1, search for a position such that the sign switch in this position results in the biggest reduction in f. If the search is successful, update f, $\mathbf{J}$ and $\mathbf{D}$. Repeat this step until f equals the length of $\mathbf{J}$ (i.e., all elements of $\mathbf{J}$ equal ± 1) or this value cannot be reduced further by any sign switch.

Remarks

(i) The above steps correspond to one "try" of the CEA and each try produces a matrix $\mathbf{D}$. Among a large number of tries whose f value reaches its lower bound, i.e., f equals $\mathbf{J}$'s length (or f cannot be reduced further), the one with the largest value of $|\mathbf{B}|$ is selected.

(ii) The purpose of picking at random $m_3 + m_2$ columns from a conference matrix of order m $(m \geq m_3 + m_2)$ to form $\mathbf{D}$ in a random manner is to avoid being trapped in the local optima.

(iii) When the input matrix in Step 1 is a conference matrix, the first $m_3 m_2$ elements of $\mathbf{J}$ always take values ± 1. This is not the case when the input matrix is not a conference matrix such as the one of order 22.

(iv) Our CEA is less prone to the curse of dimensionality than JN's exhaustive algorithm for ADSD construction which attempts to maximize $|\mathbf{B}|$ from among the 2^{2m_2} arrangements for $2m_2$ entries in $\mathbf{D}$.

Figure 16.2 shows the steps of constructing an ADSD for four three-level factors and four two-level factors. Figure 16.2a displays a starting design in Step 1. Figure 16.2b shows that the eight 0's in the last four columns of the design in Fig. 16.2a are being replaced by ± 1 in a random manner. At this point, the vector $\mathbf{J}$ is $(1, 1, 1, 1, 1, -1, 1, 1, 1, 1, 1, -1, -1, -1, 1, 1, -3, 1, -1, 1, -1, 3)$ and f is 38. Figure 16.2c,d correspond to Step 2. In Fig. 16.2c, the value 1 in the position $(1, 7)$ of the design in Fig. 16.2b is replaced by -1. At this point,

a)	b)	c)	d)
++++++0+	++++++++	++++++−+	++++++−+
−−−+0+++	−−−+−+++	−−−+−+++	−−−+++++
−++0−++−	−++0−++−	−++0−++−	−++0−++−
+−+−−0++	+−+−−+++	+−+−−+++	+−+−−+++
−+0−+−++	−+0−+−++	−+0−+−++	−+0−+−++
++−+−−+0	++−+−−++	++−+−−++	++−+−−++
+0−−+++−	+0−−+++−	+0−−+++−	+0−−+++−
0−+++−+−	0−+++−+−	0−+++−+−	0−+++−+−
00000000	0000+−++	0000+−++	0000+−++

Fig. 16.2 Steps of constructing an ADSD for four three-level factors and four two-level factors

the vector $\mathbf{J}$ is $(1, 1, -1, 1, 1, -1, -1, 1, 1, 1, -1, -1, -1, -1, -1, 1, -3, -1, -1, -1, -1, 1)$ and f is reduced to 30. In Fig. 16.2d, the value -1 in the position $(2, 5)$ of the design in Fig. 16.2c is replaced by 1. At this point, the vector $\mathbf{J}$ is $(-1, 1, -1, 1, -1, -1, -1, 1, -1, 1, -1, -1, 1, -1, -1, 1, -1, 1, 1, -1, -1, 1)$ and f is reduced to 22 which is its lower bound.

16

16.3 An ESA for Constructing D-efficient MLFODs

While the CEA attempts to convert some three-level columns of a conference matrix into two-level columns, the ESA attempts to convert some two-level columns of a two-level orthogonal matrix into three-level columns. The ESA requires three simple steps:

1. From each base matrix of order m, generate $m - 1$ additional matrices by shifting the columns of this matrix to the left cyclically. From each matrix, use the first $m_3 + m_2$ $(\leq m)$ columns to form a starting design $\mathbf{D}_{m\times(m_3+m_2)} = (d_{ij})$.
2. For each matrix obtained from Step 1, generate $\binom{m}{k}$ new matrices by replacing k elements in each of its first m_3 columns by 0's. Here $k = 2, \ldots, x$ where x is an integer chosen to be $\lceil \frac{m}{5} \rceil$ where $\lceil . \rceil$ denotes the ceiling function.

   The replacement is performed so that if $(i_1, j), (i_2, j), \ldots, (i_k, j)$ are entries in column j being replaced by 0's, and then the entries being replaced by 0's in the next column are $((i_1 + 1) \mod m, j+1), ((i_2 + 1) \mod m, j+1), \ldots, ((i_k + 1) \mod m, j + 1)$.
3. For each matrix in Step 2, calculate r_{max}, the maximum in terms of absolute value of the correlation coefficients among the columns of the model matrix. Then among the designs with smallest r_{max}, pick the one with the highest d_2, the D-efficiency for the pure quadratic model (see Eq. (16.8) in the Appendix).

(1)	(2)	(3)	(4)
0+++++++	0+++++++	0+++++++	0+++++++
00++−+−−	+0++−+−−	+0++−+−−	+0++−+−−
+00++−+−	0−0++−+−	+−0++−+−	+−0++−+−
+−00++−+	+0−0++−+	0−−0++−+	+−−0++−+
++−0+++−	++0−+++−	+0−−+++−	0+−−+++−
+−+−−+++	+−+0−+++	+−0−−+++	+0+−−+++
++−+−−++	++−+−−++	++−0−−++	++0+−−++
+++−+−−+	+++−+−−+	+++−+−−+	+++0+−−+

Fig. 16.3 Some candidate designs generated in Step 2 for an MLFOD for four three-level factors and four two-level factors

Remarks

(i) In Step 1, the base matrix is a two-level orthogonal matrix. This is a Hadamard matrix or a Plackett-Burman design if m is divisible by 4 or a conference matrix with the 0's on the diagonal being replaced by 1's if m is not divisible by 4 but is divisible by 2. The base matrix slightly affects the goodness of the resulting design.

(ii) For small m, say $m \leq 12$, the starting designs in Step 1 can also be constructed by randomly selecting a subset of $m_3 + m_2$ from m columns of the base matrix.

(iii) For each pair (m_3, m_2) and a given x, the number of zeros in each of the m_3 three-level columns, the number of candidate designs we have to consider is $m \sum_{k=2}^{x} \binom{m}{k}$.

(iv) r_{max} in Step 3 is calculated from the vector $\mathbf{J} = \sum_{i=1}^{m} \mathbf{J}_i$ of length $2\binom{m_3}{2} + m_3 m_2$ where $\mathbf{J}_i$ is defined as $(d_{i1}^2 d_{i2}^2, \ldots, d_{i(m_3-1)}^2 d_{im_3}^2, d_{i1} d_{i2}, \ldots, d_{i(m_3-1)} d_{im_3}, d_{i1} d_{i(m_3+1)}, \ldots, d_{im_3} d_{i(m_3+m_2)})$.

Figure 16.3 shows some candidate designs generated in Step 2 for an MLFOD for four three-level factors and four two-level factors.

16.4 Results and Discussion

Table 16.1 provides d_1, d_2, and $r_{\max}$ of 69 D-efficient ADSDs and two sets of corresponding new MLFODs with $m_3 = 4, \ldots, 12$, $m_2 = 1, \ldots, m_3$, and $n \geq 16$. The goodness statistics are the first-order D-efficiency and the second-order D-efficiency, the maximum in terms of the absolute value of the correlation coefficients among $2m_3 + m_2$ columns of the model matrix $\mathbf{X}$ for the pure quadratic model, respectively. The D-efficiencies d_1, d_2 of ADSDs are calculated according to Eqs. (16.5) and (16.8) in the Appendix. The first set of MLFODs labeled MLFOD_1 was obtained by selecting the first $m_3 + m_2$ columns of a conference matrix and then change the 0's to 1's in the last m_2 columns. The second set of MLFODs labeled MLFOD_2 was constructed by the ESA in Sect. 16.3 using $x = \lceil \frac{m}{5} \rceil$.

The advantage of the MLFODs over the orthogonal arrays (for the same number of three-level and two-level factors) is that the former require much less runs. At the same time, the former, unlike the latter, could guarantee that (i) all quadratic effects are orthogonal to main effects and (ii) all main effects are orthogonal to two-factor interactions. While orthogonality does not help in simplifying the data analysis which is now done entirely by computers, they help in the interpretations of the results which is the aim of the experimenters.

ADSDs, due to their method of construction, always have two runs more than the corresponding new MLFODs. It can be seen in Table 16.1 that all ADSDs have higher d_1's than the corresponding MLFOD_2's (but smaller d_1's than the corresponding MLFOD_1's). At the same time, nearly all MLFOD_2's have higher d_2's than the corresponding ADSDs. The $r_{\max}$'s of the ADSDs are always higher than the ones of MLFODs. This is due to the fact that the correlation between any two quadratic effect columns is $\frac{1}{2} - \frac{2}{n-4}$ (see JN p. 129). This value approaches 1/2 as n becomes large. In table 16.1, d_1 (or d_2) values of MLFODs printed in bold are higher than the ones of ADSDs for the same set of (m_3, m_2). $r_{\max}$ values of MLFODs printed in bold are smaller than the ones of ADSDs for the same set of (m_3, m_2).

With the exception of the ADSD for $m_3 = m_2 = 9$ and four ADSDs constructed from a $\mathbf{C}$ matrix of order 22 (see Table 16.1), all ADSDs in Table 16.1 have the f values reaching their lower bound. This fact guarantees that the constructed ADSDs will have minimum correlations between a two-level factor and a three-level factor and between any two two-level factors.

Unlike ADSDs, our MLFOD_2's do not guarantee zero correlation among three-level factors. At the same time, unlike our MLFOD_2's with $n = 16, 24, 32, 40$, and 48 (whose $\mathbf{D}$ matrices were constructed from the Hadamard matrices or Plackett-Burman designs), ADSDs do not have zero correlation among the two-level factors.

16.5 An Industrial Case Study

Lin and Kacker [7] describe an experiment aiming to improve the quality and productivity of wave soldering of circuit pack assemblies (CPA). This experiment, summarized in a study by Kenett et al. [6] (p. 474), has four yield variables and 17 factors (controllable variables). The four yield variables are: (i) Insulation resistance, (ii) Cleaning characterization, (iii) Soldering efficiency, and (iv) Solder mask cracking. The 17 factors are: (A) Type of activator, (B) Amount of activator, (C) Type of surfactant, (D) Amount of surfactant, (E) Amount of antioxidant, (F) Type of solvent, (G) Amount of solvent, (H) Amount of flux, (I) Preheat time, (J) Solder temperature, (K) Conveyor speed, (L) Conveyor angle, (M) Wave height setting, (N) Detergent concentration, (O) Detergent

Table 16.1 Comparison of D-efficiencies and r_{max} of MLFODs and ADSDs

			MLFOD$_1$			MLFOD$_2$			ADSD		
m_3	m_2	n^a	d_1	d_2	r_{max}	d_1	d_2	r_{max}	d_1	d_2	r_{max}
4	3	16	**0.909**	0.443	**0.143**	0.831	**0.484**	**0.333**	0.858	0.478	0.357
5	2	16	**0.899**	0.390	**0.143**	0.795	0.412	**0.333**	0.836	0.426	0.357
6	1	16	**0.892**	0.348	**0.143**	0.764	0.352	**0.333**	0.818	0.386	0.357
4	4	16	**0.911**	0.469	**0.143**	0.839	**0.508**	**0.333**	0.862	0.502	0.357
5	3	16	**0.900**	0.415	**0.143**	0.797	0.432	**0.333**	0.843	0.450	0.357
6	2	16	**0.892**	0.371	**0.143**	0.744	0.364	**0.333**	0.826	0.408	0.357
7	1	16	**0.888**	0.331	**0.143**				0.812	0.376	0.357
5	4	20	**0.902**	0.414	**0.200**	0.822	0.452	**0.375**	0.882	0.455	0.389
6	3	20	**0.907**	0.375	**0.200**	0.811	0.411	**0.375**	0.869	0.411	0.389
7	2	20	**0.910**	0.342	**0.200**	0.803	0.359	**0.375**	0.856	0.376	0.389
8	1	20	**0.911**	0.312	**0.111**	0.784	0.314	**0.375**	0.845	0.347	0.389
5	5	20	**0.888**	0.430	**0.200**	0.814	0.466	**0.375**	0.884	0.475	0.389
6	4	20	**0.896**	0.392	**0.200**	0.804	0.425	**0.375**	0.873	0.431	0.389
7	3	20	**0.903**	0.359	**0.200**	0.795	0.373	**0.375**	0.861	0.395	0.389
8	2	20	**0.907**	0.329	**0.200**	0.785	0.330	**0.375**	0.850	0.365	0.389
9	1	20	**0.909**	0.300	**0.111**	0.768	0.293	**0.375**	0.841	0.340	0.389
6	5	24	**0.939**	0.403	**0.091**	0.806	**0.481**	**0.333**	0.900	0.434	0.409
7	4	24	**0.933**	0.366	**0.091**	0.793	**0.443**	**0.333**	0.890	0.396	0.409
8	3	24	**0.929**	0.335	**0.091**	0.764	**0.406**	**0.333**	0.881	0.365	0.409
9	2	24	**0.926**	0.309	**0.091**	0.738	**0.370**	**0.333**	0.872	0.339	0.409
10	1	24	**0.924**	0.285	**0.091**	0.714	**0.339**	**0.333**	0.865	0.316	0.409
6	6	24	**0.940**	0.422	**0.091**	0.811	**0.496**	**0.333**	0.903	0.452	0.409
7	5	24	**0.934**	0.384	**0.091**	0.791	**0.456**	**0.333**	0.893	0.413	0.409
8	4	24	**0.930**	0.352	**0.091**	0.767	**0.419**	**0.333**	0.884	0.381	0.409
9	3	24	**0.926**	0.325	**0.091**	0.743	**0.383**	**0.333**	0.876	0.354	0.409
10	2	24	**0.924**	0.300	**0.091**	0.722	**0.352**	**0.333**	0.869	0.331	0.409
11	1	24	**0.923**	0.276	**0.091**	0.696	**0.322**	**0.333**	0.863	0.312	0.409
7	6	28	**0.923**	0.381	**0.143**	0.816	**0.463**	**0.273**	0.912	0.416	0.423
8	5	28	**0.927**	0.351	**0.143**	0.803	**0.432**	**0.273**	0.905	0.383	0.423
9	4	28	**0.930**	0.326	**0.143**	0.786	**0.406**	**0.273**	0.898	0.354	0.423
10	3	28	**0.933**	0.303	**0.143**	0.771	**0.382**	**0.273**	0.892	0.331	0.423
11	2	28	**0.934**	0.283	**0.143**	0.757	**0.361**	**0.273**	0.886	0.310	0.423
12	1	28	**0.934**	0.265	**0.077**	0.744	**0.340**	**0.273**	0.881	0.292	0.423
7	7	28	**0.917**	0.395	**0.143**	0.808	**0.472**	**0.273**	0.913	0.431	0.423
8	6	28	**0.921**	0.365	**0.143**	0.798	**0.442**	**0.273**	0.907	0.398	0.423
9	5	28	**0.925**	0.339	**0.143**	0.780	**0.416**	**0.273**	0.901	0.369	0.423
10	4	28	**0.929**	0.317	**0.143**	0.766	**0.391**	**0.273**	0.895	0.345	0.423
11	3	28	**0.931**	0.296	**0.143**	0.753	**0.370**	**0.273**	0.889	0.323	0.423
12	2	28	**0.932**	0.277	**0.143**	0.743	**0.350**	**0.273**	0.884	0.305	0.423
8	7	32	**0.954**	0.374	**0.067**	0.857	**0.467**	**0.231**	0.923	0.400	0.433
9	6	32	**0.951**	0.345	**0.067**	0.835	**0.432**	**0.231**	0.917	0.371	0.433
10	5	32	**0.948**	0.321	**0.067**	0.815	**0.402**	**0.231**	0.912	0.345	0.433
11	4	32	**0.945**	0.299	**0.067**	0.795	**0.375**	**0.231**	0.906	0.323	0.433
12	3	32	**0.944**	0.281	**0.067**	0.778	**0.351**	**0.231**	0.902	0.304	0.433
8	8	32	**0.955**	0.388	**0.067**	0.860	**0.479**	**0.231**	0.925	0.414	0.433
9	7	32	**0.951**	0.359	**0.067**	0.834	**0.443**	**0.231**	0.919	0.384	0.433
10	6	32	**0.948**	0.334	**0.067**	0.814	**0.412**	**0.231**	0.914	0.358	0.433
11	5	32	**0.946**	0.312	**0.067**	0.793	**0.385**	**0.231**	0.909	0.336	0.433
12	4	32	**0.944**	0.293	**0.067**	0.778	**0.361**	**0.231**	0.904	0.316	0.433
9	8	36	**0.939**	0.357	**0.111**	0.845	**0.447**	**0.200**	0.929	0.386	0.441

(continued)

Table 16.1 (continued)

			MLFOD$_1$			MLFOD$_2$			ADSD		
m_3	m_2	n[a]	d_1	d_2	r_{max}	d_1	d_2	r_{max}	d_1	d_2	r_{max}
10	7	36	**0.941**	0.333	**0.111**	0.840	**0.422**	**0.200**	0.925	0.359	0.441
11	6	36	**0.943**	0.312	**0.111**	0.832	**0.397**	**0.200**	0.921	0.336	0.441
12	5	36	**0.945**	0.293	**0.111**	0.827	**0.373**	**0.200**	0.917	0.316	0.441
9	9	36	**0.935**	0.368	**0.111**	0.844	**0.456**	**0.200**	0.929	0.398	0.441
10	8	36	**0.938**	0.344	**0.111**	0.838	**0.431**	**0.200**	0.926	0.371	0.441
11	7	36	**0.940**	0.323	**0.111**	0.829	**0.406**	**0.200**	0.922	0.348	0.441
12	6	36	**0.942**	0.304	**0.111**	0.821	**0.381**	**0.200**	0.919	0.328	0.441
10	9	40	**0.963**	0.351	**0.053**	0.867	**0.440**	**0.216**	0.937	0.374	0.447
11	8	40	**0.961**	0.328	**0.053**	0.857	**0.414**	**0.216**	0.933	0.350	0.447
12	7	40	**0.959**	0.308	**0.053**	0.844	**0.390**	**0.216**	0.930	0.329	0.447
10	10	40	**0.963**	0.363	**0.053**	0.867	**0.450**	**0.216**	0.938	0.385	0.447
11	9	40	**0.961**	0.339	**0.053**	0.858	**0.424**	**0.216**	0.935	0.361	0.447
12	8	40	**0.959**	0.319	**0.053**	0.845	**0.400**	**0.216**	0.931	0.339	0.447
11	10	44	**0.946**	0.337	**0.095**	0.841	**0.460**	**0.222**	0.935	0.361	0.452
12	9	44	**0.946**	0.317	**0.095**	0.832	**0.438**	**0.222**	0.931	0.339	0.452
11	11	44	**0.944**	0.347	**0.095**	0.842	**0.469**	**0.222**	0.933	0.370	0.452
12	10	44	**0.943**	0.326	**0.140**	0.830	**0.446**	**0.222**	0.932	0.349	0.452
12	11	48	**0.969**	0.333	**0.043**	0.874	**0.457**	**0.200**	0.947	0.353	0.457
12	12	48	**0.969**	0.343	**0.043**	0.877	**0.466**	**0.200**	0.948	0.362	0.457

[a]Run size of MLFODs. For the same set of (m_3, m_2) ADSD requires two extra runs

temperature, (P) Cleaning conveyor speed (Q) Rinse water temperature. Out of these 17 factors, 7 factors (A), (C), (F), (M), (N), (O), and (Q) are two-level factors and the rest are three-level factors. The aim of this experiment is to single out the active factors and then apply a full quadratic model in these factors.

Let us consider two candidate MLFODs for ten three-level factors and eight two-level factors (including a blocking factor): (a) a 36-run MLFOD constructed by the ESA and (b) a 38-run ADSD. Both designs were constructed from a conference matrix of size 18. We do not consider the 30-run mixed-level screening design of [10] and the orthogonal arrays (http://support.sas.com/techsup/technote/ts723_Designs.txt) for the same number of three- and two-level factors as they are not MLFODs and as such might not possess the advantages of an MLFOD, namely, (i) all quadratic effects are orthogonal to main effects and (ii) all main effects are orthogonal to two-factor interactions. Besides, a 72-run orthogonal array exceeds the available budget for this experiment. Figure 16.4 displays the **D** matrices of two mentioned candidate MLFODs.

Table 16.1 shows the goodness statistics of the two candidate designs. The d_1, d_2, and r_{max} of the 38-run ADSD are 0.926, 0.371, and 0.441, and the one of the corresponding 36-run MLFOD are 0.838, 0.410, and 0.2. While the 38-run ADSD is superior to the 36-run MLFOD in terms of d_1, it is inferior to the latter in terms of d_2 and r_{max}. This pattern can be observed in Table 16.1 for all MLFOD$_2$'s with $n \geq 24$.

The correlation cell plots of the two candidate designs are shown in Fig. 16.5. These plots, used in Jones and Nachtsheim [3], display the magnitude of the correlation between main effects, quadratic effects of three-level factors and two-factor interactions in screening designs. The color of each cell in these plots goes from white (no correlation)

a)

```
++++++++0+++++++++
+−−−+−++−0++−+−−−+
−−+++−+−+−+−−+−+++
+−−+++−++−−+−−+−++
0+−−+++−++−−+−−+−+
−0−++−−−−+−++−++−+
−+0+−−++++−+−−+−−+
−+−0−++−−−−+−++−++
0−+−0−++−−−−+−++−+
−0+−+0+−+−+++−+−−+
0+0−+−0−−++−−−+−++
+0+0−+−0−−++−−−+−+
−−0−0+−+0++−−+++−+
−++0+0−+−0−−++−−−+
−+−−0−0++−+++−−+++
+−+−−0−0++−+++−−++
++−+−−0−0−+−+++−−+
−−−+−++0−0+−+−−−++
```

b)

```
−−−−−−−−−−−−−−−−−+
+−−+−−−++++−−+++−+
++−−−−0−+++++−−−++
−−0−++−+−++−++−−++
−+−0−+−−+−−−++++++
++−−+++−−−++−++−−+
+++−+−−0−+−−+−++−+
+−−++++−0+−−−−−+++
−−+−−−+−−+−+−+++++
++++−−++−−−−++−−++
−++++0−−++−+−+−−−+
−+−++−−+−−++−−−+++
0−−−+−+++−−+++−+−+
−++−−++++0+−−−−+−+
−−+++−+−+−+−+−+−−+
+−++−+−−−−++++−+−+
−0−+−+++−+−++−+−−+
+−+−0+−++−−+−−+−++
0000000000+++−++++
```

Fig. 16.4 The **D** matrices of two candidate MLFODs for ten three-level factors and eight two-level factors (including a blocking factor) for the PCA experiment described by Kenett et al. [6]: (**a**) of a 36-run MLFOD constructed by the ESA and (**b**) of a 38-run ADSD

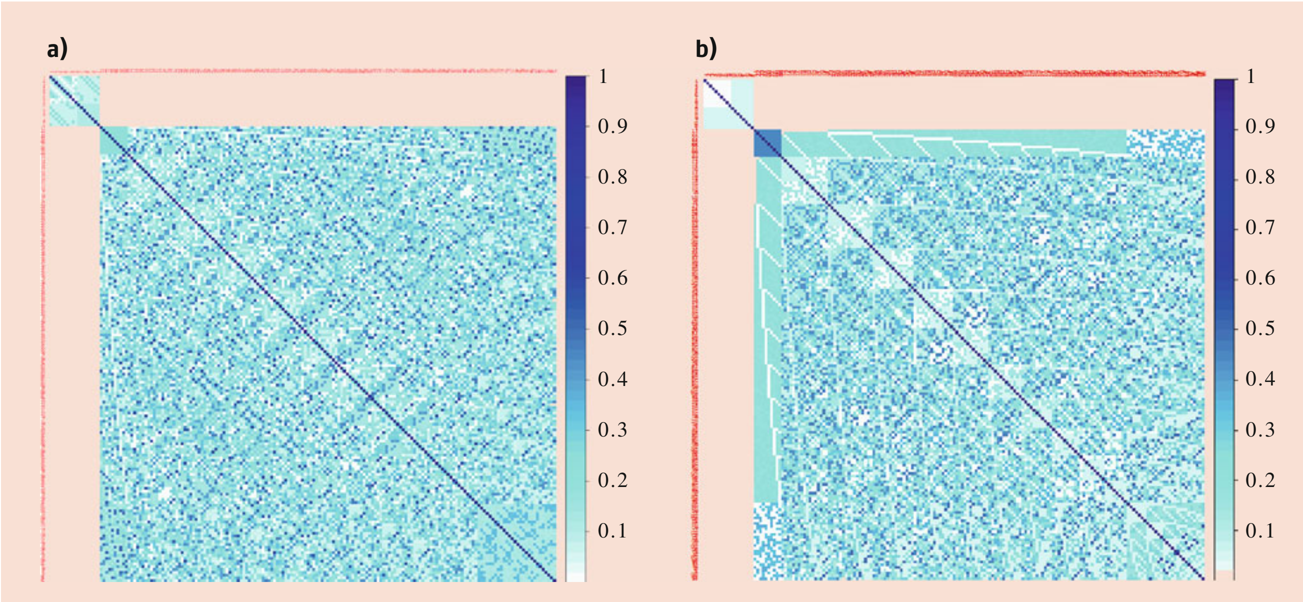

Fig. 16.5 Correlation cell plots of (**a**) 36-run MLFOD constructed by the ESA and (**b**) a 38-run ADSD for the experiment with ten three-level and eight two-level factors (including a blocking factor)

to dark (correlation of 1 or close to 1). Figure 16.5 confirms that for both designs, the main effects are orthogonal to the quadratic effects and two-factor interactions. It can be seen in Fig. 16.5b that while the main effects of three-level factors have zero correlation, the quadratic effects of these factors have fairly high correlation.

Both algorithms in Sects. 16.2 and 16.3 appear to be very fast and do not seem to be affected by the curse of dimensionality for design optimization. Both algorithms construct the abovementioned 38-run ADSD and 36-run MLFOD for ten three-level factors and eight two-level factors in less than 1 s on an HP EliteBook 8770w laptop. Note that the CEA uses 10,000 tries and out of 10,000 tries, 366 have the f values reaching the lower bound and all of them have the same values of $(d_1, d_2, r_{\max})$.

16.6 Conclusion

Screening designs precede efforts to optimize a product or process. Their goal is to reduce a long list of factors so that the optimization effort can focus on a shorter list of factors. The literature on screening experiments in engineering and science popularized experiments with two-level factors such as orthogonal FFDs [1]. This constraint slows down the knowledge acquisition process by producing information only on main effects and interactions. In mixed-level screening designs, one combines factors at two and three levels, thus also producing information on quadratic effects. This paper presents two new classes of MLFODs. Both classes are economical and efficient. These designs provide more choices for the experimenters and help them to "design for experiments instead of experiment for the design." A wider scope to this work is the augmentation of data collected in nonexperimental contexts with experimentally designed add-on observations. The analysis of observed data provides initial information on the X space characteristics. To achieve the optimized conditions, the designs discussed in this paper can be used to expand the initial data sets. This approach was mentioned by Kenett and Nguyen [5] with examples of tools used to assess the X space statistical properties. In summary, we aim here at expanding the screening-optimization continuum with more flexible and optimized designs.

The zipped file containing the **D** matrices (i.e., half fractions) for the designs in Table 16.1 and the input matrices we use to construct these **D** matrices as well as the Java programs implementing the algorithms in Sects. 16.2 and 16.3 is downloadable from https://drive.google.com/open?id=16thajCMzr4WLIb-GkyKrBv2HHnziAIMr.

Appendix: Calculating (d_1, d_2) Values of an MLFOD

Recall that $\mathbf{D}_{m^*\times(m_3+m_2)}(= (d_{ij}))$ in (16.2) is the submatrix from which an MLFOD for m_3 three-level factors and m_2 two-level factors can be constructed from a matrix of order m $(m_3 + m_2 \leq m \leq m^*)$. For ADSDs $m^* = m + 1$ and for our MLFODs $m^* = m$. The first-order D-efficiency d_1 and the pure quadratic D-efficiency d_2 of this MLFOD can be calculated as:

$$|\mathbf{X}'\mathbf{X}|^{1/p}/n \qquad (16.3)$$

where $\mathbf{X}$, p, and n are the model matrix, the number of parameters for the models, and the number of runs, respectively.

For the first-order model, $p = 1 + m_3 + m_2$ and the ith row of $\mathbf{X}$ can be written as $(1, d_{i1}, \ldots, d_{i(m_3+m_2)})$. Thus the (information matrix) $\mathbf{X}'\mathbf{X}$ will be of the form

$$2\begin{pmatrix} m^* & \mathbf{0}' \\ \mathbf{0} & \mathbf{B} \end{pmatrix}, \tag{16.4}$$

where $\mathbf{0}_{m_3+m_2}$ is a column vector of 0's and $\mathbf{B}_{(m_3+m_2)\times(m_3+m_2)} = \mathbf{D}'\mathbf{D}$. The determinant of $\mathbf{X}'\mathbf{X}$ for the first-order model can now be calculated as

$$|\mathbf{X}'\mathbf{X}| = 2^{1+m_3+m_2}\, m^*\, |\mathbf{B}|. \tag{16.5}$$

For the pure quadratic model, $p = 1 + m_3 + m_3 + m_2$ and the ith row of $\mathbf{X}$ can be written as $(1, d_{i1}^2, \ldots, d_{im_3}^2, d_{i1}, \ldots, d_{i(m_3+m_2)})$. Thus, the matrix $\mathbf{X}'\mathbf{X}$ will be of the form

$$2\begin{pmatrix} \mathbf{A} & \mathbf{0}' \\ \mathbf{0} & \mathbf{B} \end{pmatrix}, \tag{16.6}$$

where $\mathbf{0}_{(m_3+m_2)\times(1+m_3)}$ a matrix of 0's and $\mathbf{A}_{(1+m_3)\times(1+m_3)}$ is a matrix of the form

$$\begin{pmatrix} m^* & b\mathbf{1}' \\ b\mathbf{1} & \mathbf{A}^* \end{pmatrix} \tag{16.7}$$

assuming each of the m_3 three-level columns of $\mathbf{D}$ has a fixed number b of ± 1's. Here $\mathbf{1}_{m_3}$ is a column vector of 1's and $\mathbf{A}^*_{m_3\times m_3}$ the core of $\mathbf{A}$ in (16.7), i.e., the matrix $\mathbf{A}$ without its first row and first column. The determinant of $\mathbf{X}'\mathbf{X}$ for the pure quadratic model can now be calculated as

$$|\mathbf{X}'\mathbf{X}| = 2^{1+2m_3+m_2}\, m^*\, |\mathbf{A}^* - \frac{b^2}{m^*}\mathbf{J}|\, |\mathbf{B}|. \tag{16.8}$$

For some MLFODs such as ADSDs, MLFOD$_1$, and MLFOD$_2$ with $n = 16$ and 38, the matrix $\mathbf{A}^*$ in (16.7) will be of the form $c\mathbf{J} + d\mathbf{I}$ where $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is a matrix of 1's. Nguyen et al. [12] denoted this class of MLFODs as MLFOD*s. In these cases, $\mathbf{A}^* - \frac{b^2}{m^*}\mathbf{J}$ will also be of the form $c\mathbf{J} + d\mathbf{I}$. The determinants of a matrix of this form can be calculated as $d^{m_3}\ (1 + \frac{c\ m_3}{d})$. For ADSDs, $c = m - 2 - \frac{(m-1)^2}{m+1}$ and $d = 1$. For MLFOD$_1$, $c = m - 2 - \frac{(m-1)^2}{m}$ and $d = 1$.

References

1. Box, G.E., Hunter J.S., Hunter W.G.: Statistics for Experimenters: Design, Innovation, and Discovery, 2nd edn. John Wiley & Sons Ltd., New Jersey (2005)
2. Hedayat, A.S., Wallis, W.D.: Hadamard Matrices and Their Applications. Ann. Stat. **6**, 1184–1238 (1978)
3. Jones, B., Nachtsheim, C.J.: A class of three levels designs for definitive screening in the presence of second-order effects. J. Qual. Technol. **43**, 1–15 (2011)
4. Jones, B., Nachtsheim, C.J.: Definitive screening designs with added two-level categorical factors. J. Qual. Technol. **45**, 121–129 (2013)
5. Kenett R.S., Nguyen, N.K.: Experimental Learning: generate high information quality by comparing alternative experimental designs. Qual. Prog. **2017**, 40–47 (2017)
6. Kenett, R.S., Zacks, S., Amberti, D.: Modern Industrial Statistics: with applications in R, MINITAB and JMP, 3rd edn. John Wiley & Sons Ltd., West Sussex (2021)
7. Lin, K.M., Kacker R.N.: Optimization of the wave soldering process. In: Dehand, K. (eds.) Quality Control and Robust Design, and The Taguchi Method, pp. 143–157. Wadsworth BrooksCole, Pacific (1989)
8. Myer, R.K., Nachtsheim, C.J.: The coordinate-exchange algorithm for constructing exact optimal experimental designs. Technometrics **37**, 60–69 (1995)
9. Nguyen, N.-K.: An algorithmic approach to constructing supersaturated designs. Technometrics **38**, 69–73 (1996)
10. Nguyen, N.-K., Pham, D.-T.: Small mixed-level screening designs with orthogonal quadratic effects. J. Qual. Technol. **48**, 405–414 (2016)
11. Nguyen, N.-K., Stylianou, S.: Constructing definitive screening designs using cyclic generators. J. Stat. Theory Pract. **7**, 713–724 (2013)
12. Nguyen, N.-K., Pham, D.-T., Vuong, M.P.: Constructing D-efficient mixed-level foldover designs using Hadamard matrices. Technometrics **62**, 48–56 (2020)
13. Plackett, R.L., Burman, J.P.: The design of optimum multifactorial experiments. Biometrika **33**, 305–25 (1946)
14. Xiao, L.L, Lin, D.K.J., Bai, F.S.: Constructing definitive screening designs using conference matrices. J. Qual. Technol. **44**, 2–8 (2012)

Dr Nam-Ky Nguyen got his PhD. at the Indian Agricultural Research Institute, New Delhi, India in 1983. He has been a seasoned statistician with over 20 years of diversified experience in tertiary-level teaching, consulting and research at various universities and research organizations in Australia, USA, Vietnam, Thailand, Cambodia, Nigeria and Ivory Coast. Accredited by the Statistical Society of Australia (AStat) in 1998. Author of around 50+ publications in applied statistics/computing and collaborative research. Author of the Gendex DOE toolkit. This toolkit is currently in use at several universities and research organizations worldwide. Research interest includes design of experiments, statistical computing and statistical modelling.

Professor Ron Kenett is Chairman of the KPA Group, Israel and Senior Research Fellow at the Neaman Institute, Technion, Haifa, Israel. He is an applied statistician combining expertise in academic, consulting and business domains. He serves on the editorial board of several international journals and was awarded the 2013 Greenfield Medal by the Royal Statistical Society and, in 2018, the Box Medal by the European Network for Business and Industrial Statistics. Ron holds a BSc in Mathematics (with first class honors) from Imperial College, London University and a PhD in Mathematics from the Weizmann Institute of Science, Rehovot, Israel.

Mai Phuong Vuong PhD is a lecturer at the School of Applied Mathematics and Informatics, Hanoi University of Science and Technology. She holds a Bachelor in applied Mathematics and Informatics from Vietnam National University, Hanoi, and a PhD in GeoInformatics from TU Bergakademie Freiberg, Germany. Her current research areas are design of experiments and data science.

Dr Tung Dinh Pham is lecturer at the Vietnam National University – Hanoi University of Science, Vietnam. He is an applied statistician combining expertise in academic, consulting domains. Tung holds a BSc in Mathematics and PhD in Mathematical Statistics from Hanoi University of Science.