Ann Inst Stat Math (2015) 67:673–685
DOI 10.1007/s10463-014-0470-0

Minimax design criterion for fractional factorial designs

Yue Yin · Julie Zhou

Received: 12 November 2013 / Revised: 15 March 2014 / Published online: 18 May 2014
© The Institute of Statistical Mathematics, Tokyo 2014

Abstract An A-optimal minimax design criterion is proposed to construct fractional factorial designs, which extends the study of the D-optimal minimax design criterion in Lin and Zhou (*Canadian Journal of Statistics* **41**, 325–340, 2013). The resulting A-optimal and D-optimal minimax designs minimize, respectively, the maximum trace and determinant of the mean squared error matrix of the least squares estimator (LSE) of the effects in the linear model. When there is a misspecification of the effects in the model, the LSE is biased and the minimax designs have some control over the bias. Various design properties are investigated for two-level and mixed-level fractional factorial designs. In addition, the relationships among A-optimal, D-optimal, E-optimal, A-optimal minimax and D-optimal minimax designs are explored.

Keywords A-optimal design · D-optimal design · Factorial design · Model misspecification · Requirement set · Robust design

1 Introduction

Factorial designs are very useful in industrial experiments to investigate possible influential factors. A full factorial design allows us to examine all the main effects and interactions among the factors. However, the total number of runs in a full factorial design can be huge for a large number of factors, and often there may not be enough resources to run it. In this situation, a fractional factorial design (FFD) can be applied, although it does not allow us to estimate all the main effects and interactions. There are a couple of ways to deal with this. One way is to assume that higher-order interactions are negligible, and we only estimate lower-order effects. Another way is to specify and

Y. Yin · J. Zhou (✉)
Department of Mathematics and Statistics, University of Victoria, Victoria, BC V8W 2Y2, Canada
e-mail: jzhou@uvic.ca

analyze a subset of possible significant effects. The subset is referred as a requirement set that usually contains all the main effects and some interactions. In this paper, we construct optimal/robust FFDs for given requirement sets.

Suppose N is the run size of a full factorial design, and n is the affordable run size with $n < N$. There are $\binom{N}{n}$ possible choices to choose a FFD. Which one should we choose? Various optimal criteria have been explored in the literature. For example, see Mukerjee and Wu (2006) for regular FFD criteria based on the component hierarchy principle, and developments for nonregular FFDs are reviewed in Xu et al. (2009).

The commonly used A-, D- and E-optimal design criteria are model based. They have been investigated extensively in the literature, and optimal designs have been constructed for various regression models. For example, see Fedorov (1972) and Pukelsheim (1993). These criteria can also be applied to select optimal FFDs. In Tang and Zhou (2009, 2013), D-optimal designs for two-level FFDs are studied for various requirement sets and run sizes.

If the requirement set includes all the significant effects of the factors in the experiment, then the bias of the least squares estimator (LSE) is negligible. However, the requirement set may not be correctly specified, and in particular it may miss some significant effects. When this happens, the bias of the LSE is not negligible and needs to be considered in the construction of optimal FFDs. This motivated the research in Wilmut and Zhou (2011), where a D-optimal minimax criterion was proposed for two-level factorial designs. A D-optimal minimax design (DOMD) minimizes the maximum determinant of the mean squared error (MSE) of the LSE, and the maximum is taken over small possible departures of the requirement set. Lin and Zhou (2013) extended this criterion to mixed-level factorial designs and derived several interesting properties of DOMDs. These designs are robust against misspecification of the requirement set which is similar to the misspecification in the regression response function. The bias of the LSE is also the focus in minimum aberration designs, which have been studied by many authors including Cheng and Tang (2005).

In this paper, the minimax approach is further extended to the A-optimal minimax criterion, and properties of A-optimal minimax FFDs are explored. In addition, the relationships among A-optimal, D-optimal, E-optimal, A-optimal minimax and D-optimal minimax designs are investigated and several theoretical results are obtained.

2 Notation

Consider a full factorial design for k factors, $F_1, F_2, \ldots, F_k$, with $a_1, a_2, \ldots, a_k$ levels, respectively. The total number of runs is $N = a_1 a_2 \cdots a_k$. In this paper, all the main effects and interactions are coded to be orthogonal when we fit a full linear model to estimate those effects. There are $N - 1$ coded variables in the model and denoted by $x_1, x_2, \ldots, x_{N-1}$. Let $\mathbf{U}$ be the $N \times N$ model matrix of the full model, and its first column is a vector of ones $\mathbf{1}_N$ for the grand mean term.

A FFD of n runs is selected from the N rows of $\mathbf{U}$ without replacement. The FFD should allow us to estimate all the effects in a requirement set $\mathcal{R}$. Since the columns of $\mathbf{U}$ can be permuted, without loss of generality, we assume that the first q variables $x_1, \ldots x_q$ represent all the effects in $\mathcal{R}$. For the ith run in the FFD, those variables take values as $x_{i1}, \dots x_{iq}$, which comes from a selected row of $\mathbf{U}$. Then the linear model for $\mathcal{R}$ is, $y_i = \theta_0 + \theta_1 x_{i1} + \ldots + \theta_q x_{iq} + \epsilon_i, \quad i = 1, \ldots, n$, where y_i is the observed response at the ith run, and the errors ϵ_i's are assumed to be i.i.d. with mean 0 and variance σ^2. Let $\mathbf{X}_1$ be the model matrix, $\mathbf{y} = (y_1, \ldots, y_n)^\top$, $\boldsymbol{\theta}_1 = (\theta_0, \theta_1, \ldots, \theta_q)^\top$, and $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^\top$. Then the model can be written in a matrix form as

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\theta}_1 + \boldsymbol{\epsilon}. \tag{1}$$

The LSE of $\boldsymbol{\theta}_1$ is $\hat{\boldsymbol{\theta}}_1 = \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} \mathbf{X}_1^\top \mathbf{y}$, and $\mathrm{Cov}(\hat{\boldsymbol{\theta}}_1) = \sigma^2 \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}$.

Partition matrix $\mathbf{U}$ into $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$, where $\mathbf{U}_1$ contains the column $\mathbf{1}_N$ and the columns for variables $x_1, \ldots, x_q$, and $\mathbf{U}_2$ contains the others. Since the columns of $\mathbf{U}$ are orthogonal, we have

$$\mathbf{U}^\top \mathbf{U} = \begin{pmatrix} \mathbf{U}_1^\top \mathbf{U}_1 & \mathbf{U}_1^\top \mathbf{U}_2 \\ \mathbf{U}_2^\top \mathbf{U}_1 & \mathbf{U}_2^\top \mathbf{U}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \end{pmatrix}, \tag{2}$$

where both $\mathbf{V}_1 = \mathbf{U}_1^\top \mathbf{U}_1$ and $\mathbf{V}_2 = \mathbf{U}_2^\top \mathbf{U}_2$ are diagonal matrices.

If $\mathcal{R}$ misses some significant effects, then model (1) is misspecified. A possible model with small departures from (1) is derived in Lin and Zhou (2013), which can be written as $\mathbf{y} = \mathbf{X}_1 \boldsymbol{\theta}_1 + \mathbf{X}_2 \boldsymbol{\theta}_2 + \boldsymbol{\epsilon}$, where $\mathbf{X}_2$ is the model matrix for variables $x_{q+1}, \ldots, x_{N-1}$, and $\boldsymbol{\theta}_2 = (\theta_{q+1}, \ldots, \theta_{N-1})^\top$ is an unknown parameter vector satisfying $\frac{1}{N} \boldsymbol{\theta}_2^\top \mathbf{V}_2 \boldsymbol{\theta}_2 \le \alpha^2$. Here $\alpha \ge 0$ controls the size of departures. When $\alpha = 0$, model (1) is assumed to be correct. It is clear that $\mathbf{X}_1$ and $\mathbf{X}_2$ are submatrices of $\mathbf{U}_1$ and $\mathbf{U}_2$, respectively.

3 Design criteria

Commonly used A-optimal, D-optimal, and E-optimal criteria are based on measures of the covariance matrix of the LSE. In particular, the A-optimal design (AOD), D-optimal design (DOD), and E-optimal design (EOD) minimize the trace, the determinant and the largest eigenvalue of $\mathrm{Cov}(\hat{\boldsymbol{\theta}}_1)$, respectively. However, if model (1) is misspecified, then the LSE $\hat{\boldsymbol{\theta}}_1$ is biased with bias $(\hat{\boldsymbol{\theta}}_1) = E(\hat{\boldsymbol{\theta}}_1) - \boldsymbol{\theta}_1 = \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} \mathbf{X}_1^\top \mathbf{X}_2 \boldsymbol{\theta}_2$. The MSE is

$$\mathrm{MSE}\,(\hat{\boldsymbol{\theta}}_1, \mathbf{X}_1, \boldsymbol{\theta}_2) = \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} \mathbf{X}_1^\top \mathbf{X}_2 \boldsymbol{\theta}_2 \boldsymbol{\theta}_2^\top \mathbf{X}_2^\top \mathbf{X}_1 \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} + \sigma^2 \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}. \tag{3}$$

Consider the loss function $\mathcal{L}_{DM}(\mathbf{X}_1) = \max_{\boldsymbol{\theta}_2 \in \Theta} \det \left(\mathrm{MSE}(\hat{\boldsymbol{\theta}}_1, \mathbf{X}_1, \boldsymbol{\theta}_2)\right)$, where $\Theta = \left\{\boldsymbol{\theta}_2 \mid \frac{1}{N} \boldsymbol{\theta}_2^\top \mathbf{V}_2 \boldsymbol{\theta}_2 \le \alpha^2\right\}$. Then the design minimizing $\mathcal{L}_{\mathrm{DM}}(\mathbf{X}_1)$ over all possible designs $\mathbf{X}_1$ is called a DOMD, which has been investigated in Wilmut and Zhou (2011) and Lin and Zhou (2013).

Similar to the D-optimal minimax criterion, we now propose the A-optimal minimax criterion based on the trace of the MSE as follows. Define a loss function

$$\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1) = \max_{\boldsymbol{\theta}_2 \in \Theta} \mathrm{trace}\left(\mathrm{MSE}(\hat{\boldsymbol{\theta}}_1, \mathbf{X}_1, \boldsymbol{\theta}_2)\right), \tag{4}$$

and an A-optimal minimax design (AOMD) minimizes $\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1)$ over $\mathbf{X}_1$.

Both AOMDs and DOMDs are robust against misspecification of the requirement set. It looks like we need to solve minimax problems to find AOMDs and DOMDs. Since the analytical expressions for $\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1)$ and $\mathcal{L}_{\mathrm{DM}}(\mathbf{X}_1)$ can be derived, we only need to solve minimization problems. Thus, it is not more difficult to find AOMDs and DOMDs than AODs and DODs. The expression for $\mathcal{L}_{\mathrm{DM}}(\mathbf{X}_1)$ is,

$$\mathcal{L}_{\mathrm{DM}}(\mathbf{X}_1) = \sigma^{2(q+1)} \frac{1 + \frac{N\alpha^2}{\sigma^2}\left(1 - \lambda_{\min}\left(\mathbf{V}_1^{-1/2}\mathbf{X}_1^\top \mathbf{X}_1 \mathbf{V}_1^{-1/2}\right)\right)}{\det\left(\mathbf{X}_1^\top \mathbf{X}_1\right)}, \tag{5}$$

where $\lambda_{\min}()$ is the smallest eigenvalue of a matrix, from Lin and Zhou (2013). The analytical expression for $\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1)$ is obtained in the next theorem.

Theorem 1 *The loss function in* (4) *equals to*

$$\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1) = \sigma^2 \left\{ trace\left(\left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}\right) + \frac{N\alpha^2}{\sigma^2}\lambda_{\max}\left(\left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} - \mathbf{V}_1^{-1}\right)\right\}, \tag{6}$$

where $\lambda_{\max}()$ *denotes the largest eigenvalue of a matrix.*

The proof of Theorem 1 is in the Appendix. This result is very useful to explore theoretical properties of AOMDs and find relationships with other optimal designs. From (5) and (6), it is clear that DOMDs and AOMDs depend on parameter α^2 only through a parameter $v = \alpha^2/\sigma^2$, which can be viewed as the bias to variance ratio. If the bias is not important, choose v close to zero. Some optimal designs do not depend on v, and detailed results and comments are given in Sects. 4 and 5.

Theorem 2 *For a given requirement set, the minimum loss for AOMDs,* $\min_{\mathbf{X}_1} \mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1)$*, is a nonincreasing function of* n*.*

The proof of Theorem 2 is in the Appendix. Similar results have been proved for other optimal designs, such as DOMDs in Lin and Zhou (2013).

4 Results for two-level designs

For a two-level factor, "+1" and "−1" are used to code the high and low levels. All the regression variables $x_1, \ldots, x_{N-1}$ take values ± 1 and are orthogonal for a full factorial design. Thus, from (2), matrix $\mathbf{V}_1 = N\,\mathbf{I}_{q+1}$ with $N = 2^k$, and the loss functions in (5) and (6) become,

$$\mathcal{L}_{\mathrm{DM}}(\mathbf{X}_1) = \sigma^{2(q+1)} \frac{1 + v \ \left(N - \lambda_{\min}\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)\right)}{\det\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)}, \tag{7}$$

$$\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1) = \sigma^2 \left\{ \mathrm{trace}\left(\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)^{-1}\right) + v \left(\frac{N}{\lambda_{\min}\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)} - 1\right)\right\}. \tag{8}$$

Also AOD, DOD and EOD minimize the following loss functions, respectively,

$$\mathcal{L}_A(\mathbf{X}_1) = \sigma^2 \mathrm{trace}\left(\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)^{-1}\right), \quad \mathcal{L}_D(\mathbf{X}_1) = \frac{\sigma^{2(q+1)}}{\det\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)},$$

$$\mathcal{L}_E(\mathbf{X}_1) = \frac{1}{\lambda_{\min}\left(\mathbf{X}_1^{\top}\mathbf{X}_1\right)}.$$

If $v = 0$, then there is no bias and the loss functions in (7) and (8) are the same as those for D-optimal and A-optimal criteria, respectively. In general, we have the following relationships,

$$\mathcal{L}_{\mathrm{DM}}(\mathbf{X}_1) = \mathcal{L}_D(\mathbf{X}_1)\left(1 + v\left(N - \frac{1}{\mathcal{L}_E(\mathbf{X}_1)}\right)\right),$$

$$\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1) = \mathcal{L}_A(\mathbf{X}_1) + \alpha^2\left(N\mathcal{L}_E(\mathbf{X}_1) - 1\right).$$

From these relationships, it is obvious that

(1) If a design is both DOD and EOD, then it is a DOMD for all values of v.
(2) If a design is both AOD and EOD, then it is an AOMD for all values of α^2.
(3) If $\alpha^2 = 0$, then a DOMD is a DOD and an AOMD is an AOD.
(4) For large α^2 and v, an EOD tends to be a DOMD and AOMD.

For a given requirement set $\mathcal{R}$, a design $\mathbf{X}_1$ is said to be orthogonal for $\mathcal{R}$ if it satisfies $\mathbf{X}_1^{\top}\mathbf{X}_1 = n\ \mathbf{I}_{q+1}$. This definition of orthogonal design is slightly different from the one in Tang and Zhou (2009) and other papers, where an orthogonal design implies that the main effects are orthogonal. Here orthogonal designs for $\mathcal{R}$ make all the effects in $\mathcal{R}$ orthogonal, so they are more restricted than the orthogonal designs for main effects only.

Lemma 1 *An orthogonal design for $\mathcal{R}$ is an AOD, DOD, EOD, AOMD and DOMD.*

It is well anticipated from the results on universal optimality for full rank models (Sinha and Mukerjee 1982). This result shows that optimal designs can be constructed by finding orthogonal designs for $\mathcal{R}$ if they exist. However, for many run sizes of n and requirement sets $\mathcal{R}$, orthogonal designs for $\mathcal{R}$ do not exist. For example, if n is not a multiple of 4, orthogonal designs do not exist for any $\mathcal{R}$ with $q \geq 2$. Next theorem shows another result that orthogonal designs do not exist.

Theorem 3 *Orthogonal designs for $\mathcal{R}$ do not exist for run size $n \in [N - q, N - 1]$, where q is the number of effects in $\mathcal{R}$.*

The proof of Theorem 3 is in the Appendix. This result indicates that orthogonal designs for $\mathcal{R}$ do not exist if the run size is close to $N-1$. Orthogonal designs for $\mathcal{R}$ may not exist even when n is a multiple of 4 and is small. For example, if n is not a multiple of 8, e.g.,$n = 12$ or 20, and $\mathcal{R}$ includes some two-factor interactions and/or three-factor interactions, then orthogonal designs for $\mathcal{R}$ do not exist from Deng and Tang (2002, Proposition 1).

If the regressors $x_1, \cdots, x_{N-1}$ are multiplied by constant nonzero numbers $b_1, \cdots,$ b_{N-1}, respectively, do optimal designs stay the same? This is a scale invariant issue. DODs are known to be scale invariant. Lin and Zhou (2013) also proved that DOMDs are scale invariant. However, AODs are not scale invariant. Since AODs are AOMDs when $v = 0$, AOMDs are not scale invariant either. Nevertheless, this lack of scale invariance is not a matter of concern because AODs and AOMDs are used when interest lies in the parameters as they stand and not in their multiples or linear transformations.

Another property is the level-permutation invariance. DODs and DOMDs are level-permutation invariant (Lin and Zhou 2013) for a special class of permutations. For two-level designs, there is only one level permutation: switching the two levels (high level $\leftrightarrow$ low level). It turns out that AODs and AOMDs are also level-permutation invariant, which is a result of the next theorem.

Theorem 4 *Let* $\tilde{\mathbf{X}}_1 = \mathbf{X}_1\mathbf{Q}$, *where* $\mathbf{Q}$ *is a diagonal matrix with diagonal elements being* ± 1. *Then we have* $\mathcal{L}_{\text{AM}}(\tilde{\mathbf{X}}_1) = \mathcal{L}_{\text{AM}}(\mathbf{X}_1)$, *where function* $\mathcal{L}_{\text{AM}}$ *is given in* (8).

The proof of Theorem 4 is straight forward and is omitted. If a factor's two levels are permuted, then the corresponding main effect variable and interaction variables involving this factor change a sign. Therefore, we have $\tilde{\mathbf{X}}_1 = \mathbf{X}_1\mathbf{Q}$, where $\mathbf{X}_1$ and $\tilde{\mathbf{X}}_1$ are the model matrices before and after the permutation, and the diagonal elements of $\mathbf{Q}$ are ± 1. The result in Theorem 4 implies that AOMDs are level-permutation invariant. This result can have the following two interpretations:

(1) It does not matter how we label the two levels as high and low levels. It is very helpful in practice, since the high and low levels are not clear for categorical factors, such as tool type or material type.
(2) AOMDs are not unique. Switching the levels of one or more factors in an AOMD yields another AOMD.

The following two examples will show the properties for two-level designs. Since optimal designs do not depend on the value of σ^2, it is set to be 1 in all the examples.

Example 1 Consider an experiment to investigate 4 factors and a requirement set $\mathcal{R} = \{F_1, F_2, F_3, F_4, F_1F_2, F_3F_4\}$. For this $\mathcal{R}$, we have $N = 16$ and $q = 6$. We construct optimal designs for $n = 8, 9, \cdots, 15$ to illustrate design properties discussed in this section. For each run size n, a complete search is done to find AODs, DODs, EODs, AOMDs and DOMDs. The following is the summary of the results.

(1) Orthogonal designs for $\mathcal{R}$ do not exist for any $n \in [8, 15]$. For $9 \le n \le 15$, this result is obvious from Theorem 3 and the comments after Lemma 1. For $n = 8$, it is due to the specific structure of the effects in $\mathcal{R}$.
(2) Optimal designs are not unique, which is consistent with the comments after Theorem 4.

Table 1 Optimal designs and minimum loss functions in Example 1

n	Optimal designs	$\mathcal{L}_A$	$\mathcal{L}_{\mathrm{AM}}$	$\mathcal{L}_D^{1/(q+1)}$	$\mathcal{L}_{\mathrm{DM}}^{1/(q+1)}$	$\mathcal{L}_E$
8	1, 2, 5, 8, 10, 11, 15, 16	1.3750	7.2034	0.1524	0.2236	0.4268
9	1, 2, 3, 5, 8, 10, 12, 15, 16	1.0417	4.0417	0.1281	0.1848	0.2500
10	1, 2, 4, 5, 6, 9, 11, 14, 15, 16	0.9072	3.9072	0.1127	0.1626	0.2500
11	AOD, DOD and DOMD: 1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15	0.7750		0.0993	0.1429	
	EOD and AOMD: 1, 2, 3, 5, 6, 8, 9, 11, 12, 13, 16		3.4237			0.2266
12	1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 16	0.6458	1.6458	0.0876	0.1200	0.1250
13	1, 2, 3, 4, 5, 6, 7, 9, 11, 12, 13, 14, 16	0.5909	1.5909	0.0804	0.1100	0.1250
14	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 16	0.5375	1.5375	0.0738	0.1010	0.1250
15	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	0.4861	1.2639	0.0679	0.0913	0.1111

(3) All the optimal designs are equivalent when $n = 9, 12, 13, 14, 15$. As a result, AOMDs and DOMDs do not depend on the value of v. When $n = 8$ and 10, there are more EODs than the other optimal designs. AODs, DODs, AOMDs and DOMDs are equivalent, and they are all EODs. Thus, AOMDs and DOMDs do not depend on the value of v either. Some EODs are not AODs, DODs, AOMDs and DOMDs. When $n = 11$, AODs and DODs are different from EODs, so AOMDs and DOMDs depend on the value of v. For $v = 1$, AOMDs and EODs are equivalent, while DOMDs and DODs are the same.

(4) Table 1 presents the minimum loss functions and optimal designs. For $\mathcal{L}_{\mathrm{AM}}$ and $\mathcal{L}_{\mathrm{DM}}$, the results are for $v = 1$. Since optimal designs are not unique, one design is given for each case. The numbers in the optimal designs are the run numbers from a full factorial design and generated as follows (Wilmut and Zhou 2011): the first run is for all the factors at level -1, then factor F_i alters between -1 and $+1$ for every 2^{i-1} runs, $i = 1, 2, 3, 4$. The results show that the loss functions decrease as n increases for all the criteria, which is consistent with Theorem 2.

Example 2 Consider a requirement set $\mathcal{R} = \{F_1, F_2, F_3, F_4, F_5, F_1F_2, F_1F_3\}$ with $N = 32$ and $q = 7$. We construct optimal designs for $n = 8, 12, 15, 16, 19, 20$. Since $\binom{N}{n}$ can be huge, we apply the simulated annealing algorithm in Wilmut and Zhou (2011) to search for the optimal designs. A simulated annealing algorithm is known to be effective finding optimal designs and has been applied by many authors; see other references and the detailed algorithm in Wilmut and Zhou (2011). Table 2 shows some optimal designs and the minimum loss functions, over a wide class of designs as covered by the simulated annealing algorithm. Here is the summary of the numerical results, where AOMDs and DOMDs are computed for $v = 1$.

(1) Orthogonal designs for $\mathcal{R}$ exist for $n = 8$ and 16. In fact, many regular fractional factorial designs (2^{5-2} and 2^{5-1}) are orthogonal designs for this $\mathcal{R}$.
(2) For $n = 8, 12, 16$ and 20, all the optimal designs are equivalent.
(3) For $n = 15$ and 19, the optimal designs are different. For $n = 15$, some AODs are DODs, some EODs are AOMDs, and some DOMDs are AODs. For $n = 19$,

Table 2 Optimal designs and minimum loss functions in Example 2

n	optimal designs	$\mathcal{L}_A$	$\mathcal{L}_{\text{AM}}$	$\mathcal{L}_D^{1/(q+1)}$	$\mathcal{L}_{\text{DM}}^{1/(q+1)}$	$\mathcal{L}_E$
8	1, 7, 12, 14, 18, 24, 27, 29	1.0000	4.0000	0.1250	0.1869	0.1250
12	2, 4, 6, 11, 13, 16, 17, 23, 24, 26, 28, 30	0.7292	3.7292	0.0871	0.1302	0.1250
15	AOD and DOD: 2, 3, 5, 8, 9, 12, 14, 15, 17, 20, 22, 23, 26, 27, 32	0.5625	3.5625	0.0682	0.1019	0.1250
	EOD and AOMD: 3, 5, 6, 10, 12, 13, 15, 16, 18, 20, 24, 25, 27, 30, 31	0.5728	2.9314	0.0690	0.1024	0.1050
	DOMD and AOD: 1, 2, 3, 4, 7, 13, 14, 16, 21, 22, 24, 26, 27, 28, 31	0.5625	3.2371	0.0683	0.1018	0.1148
16	3, 4, 5, 8, 9, 10, 14, 15, 18, 19, 21, 22, 25, 28, 31, 32	0.5000	1.5000	0.0625	0.0891	0.0625
19	AOD and DOD: 3, 4, 5, 6, 9, 10, 13, 15, 16, 17, 18, 21, 23, 24, 25, 27, 28, 30, 31	0.4362	1.6377	0.0536	0.0771	0.0688
	EOD, AOMD and DOMD: 1, 3, 4, 5, 9, 10, 14, 15, 16, 18, 21, 22, 23, 24, 25, 26, 27, 28, 29	0.4375	1.4375	0.0537	0.0765	0.0625
20	1, 3, 5, 6, 7, 10, 11, 12, 13, 16, 17, 18, 20, 23, 24, 25, 27, 29, 30, 31	0.4125	1.4125	0.0508	0.0723	0.0625

Table 3 Orthogonal codes for a three-level factor

Factor level	x_L	x_Q
0	-1	$+1$
1	0	-2
2	$+1$	$+1$

some EODs are both AOMDs and DOMDs. AOMDs and DOMDs depend on the value of v.

5 Results for mixed-level designs

In this section, we consider designs with some factors taking more than two levels. The following two cases are discussed in detail, but the methodology is quite general and can be applied to other cases.

Case 1: Factors have mixed two levels and three levels.

Case 2: All the factors have three levels.

A two-level factor is coded the same as in Sect. 4. A three-level factor needs two variables to represent the main effect, and they are called the linear x_L and quadratic x_Q components and coded orthogonally as in Table 3.

For two-level designs, we have $\mathbf{V}_1 = N\mathbf{I}_{q+1}$, so the loss functions of AOMDs and DOMDs can be simplified and several nice properties about the optimal designs have been derived in Sect. 4. However, mixed-level designs usually do not have $\mathbf{V}_1 = N\mathbf{I}_{q+1}$, which makes it harder to obtain minimax design properties.

Consider a level permutation π, and let $\mathbf{X}_1$ and $\mathbf{X}_1^{\pi}$ be the model matrices before and after the permutation, respectively. The permutation can involve level changes in one or more factors. Define a class of permutations,

$$\Pi = \left\{\pi \mid \mathbf{X}_1^{\pi} = \mathbf{X}_1\mathbf{Q}_{\pi}, \text{ where } \mathbf{Q}_{\pi} \text{ is a diagonal matrix with elements } \pm 1\right\}.$$

Next theorem shows that AOMDs are invariant under the level permutations in Π.

Theorem 5 *Let* $\mathbf{X}_1^{\pi} = \mathbf{X}_1\mathbf{Q}_{\pi}$, *where* $\mathbf{Q}_{\pi}$ *is a diagonal matrix with diagonal elements being* ± 1. *Then, we have* $\mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1^{\pi}) = \mathcal{L}_{\mathrm{AM}}(\mathbf{X}_1)$, *where function* $\mathcal{L}_{\mathrm{AM}}$ *is given in* (6).

The proof of Theorem 5 is in the Appendix. This result is true for factors with any levels. This Π includes the permutation for two-level factors discussed in Sect. 4 and the permutation for three-level factors by switching levels 0 and 2. Next two examples present some AOMDs for mixed-level designs and their properties.

Example 3 Suppose there are three factors, F_1, F_2 and F_3, and each has three levels. A requirement set includes all the main effects and the interaction between F_1 and F_2, i.e., $\mathcal{R} = \{F_1, F_2, F_3, F_1F_2\}$ with $N = 27$ and $q = 10$. We construct AOMDs for $n = 21$ and 24 using a complete search, and the results are presented in Tables 4 and 5. For $n = 21$, there is only one AOD, and there are eight AOMDs. The AOD is different from the AOMDs. All the eight AOMDs can be generated from the AOMD

Table 4 Optimal designs for $n = 21$ in Example 3

Factor	Factor levels																				
	AOD: $\mathcal{L}_A = 0.5394$																				
F_1	0	1	2	0	2	0	1	2	0	2	1	0	2	0	1	2	0	2	0	1	2
F_2	0	0	0	1	1	2	2	2	0	0	1	2	2	0	0	0	1	1	2	2	2
F_3	0	0	0	0	0	0	0	0	1	1	1	1	1	2	2	2	2	2	2	2	2
	AOMD: $\mathcal{L}_{AM} = 1.5574$																				
F_1	0	2	0	1	2	0	2	1	2	0	1	1	2	0	1	2	0	2	0	1	2
F_2	0	0	1	1	1	2	2	0	0	1	1	2	2	0	0	0	1	1	2	2	2
F_3	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2	2	2	2	2	2	2

Table 5 Optimal designs for $n = 24$ in Example 3

factor	factor levels																							
	AOD and AOMD: $\mathcal{L}_A = 0.4595$, $\mathcal{L}_{\text{AM}} = 0.9595$																							
F_1	0	1	2	0	1	2	0	1	2	0	1	2	1	0	1	2	0	1	2	0	2	0	1	2
F_2	0	0	0	1	1	1	2	2	2	0	0	0	1	2	2	2	0	0	0	1	1	2	2	2
F_3	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2

in Table 4 by permuting the levels 0 and 2 of F_1 and F_3 and by switching two factors F_1 and F_2. Permuting the levels 0 and 2 of F_2 does not generate more designs. The construction through switching factor levels is from the result in Theorem 5, while the construction through switching two factors works because of the symmetry of the requirement set $\mathcal{R}$ in F_1 and F_2. For $n = 24$, all the AODs and AOMDs are the same. There are four AODs and AOMDs. Table 5 shows one of them, and the other three can be generated by permuting the levels 0 and 2 of F_3 and by switching two factors F_1 and F_2.

Example 4 Consider an experiment with four factors: F_1 and F_2 with three levels, and F_3 and F_4 with two levels. We want to estimate all the main effects, the interaction between F_1 and F_3, and the interaction between F_3 and F_4, so requirement set $\mathcal{R} = \{F_1, F_2, F_3, F_4, F_1F_3, F_3F_4\}$, $N = 36$ and $q = 9$. For each run size $n = 12, 15, 18, 20, 24$ and 30, a simulated annealing algorithm is run 10 times and the design with the smallest loss function is taken as an AOMD. Table 6 presents one AOMD for $n = 15$.

6 Discussion

All the results are derived for an orthogonal parameterization in this paper. In practice, there are situations for a non-orthogonal parameterization, such as in Mukerjee and Tang (2012), where two-level FFDs are constructed using the minimum aberration

Table 6 Optimal design for $n = 15$ in Example 4

Factor	Factor levels														
	AOMD: $\mathcal{L}_{\text{AM}} = 3.8237$														
F_1	0	1	2	0	1	1	2	0	0	1	2	2	0	1	2
F_2	1	0	2	2	1	2	0	0	2	1	0	1	0	1	2
F_3	−1	−1	−1	1	1	1	1	−1	−1	−1	−1	−1	1	1	1
F_4	−1	−1	−1	−1	−1	−1	−1	1	1	1	1	1	1	1	1

criterion. It needs further research to develop minimax design theories and construct minimax FFDs with non-orthogonal parameterizations. In addition, since both the minimax design criterion and the minimum aberration criterion concern the bias of the LSE, it will be interesting to investigate the connection between the two criteria. This could be a future research topic.

Appendix: Proofs

Proof of Theorem 1 From (2), we have $\left(\mathbf{V}_1^{-1/2} \oplus \mathbf{V}_2^{-1/2}\right) \mathbf{U}^\top \mathbf{U} \left(\mathbf{V}_1^{-1/2} \oplus \mathbf{V}_2^{-1/2}\right) = \mathbf{I}_N$, which implies that $\mathbf{U} \left(\mathbf{V}_1^{-1} \oplus \mathbf{V}_2^{-1}\right) \mathbf{U}^\top = \mathbf{I}_N$. From $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2)$, we get

$$\mathbf{U}_1 \mathbf{V}_1^{-1} \mathbf{U}_1^\top + \mathbf{U}_2 \mathbf{V}_2^{-1} \mathbf{U}_2^\top = \mathbf{I}_N. \tag{9}$$

Define an $N \times N$ diagonal matrix $\mathbf{W}$ with diagonal elements $w_i, i = 1, \ldots, N$, where $w_i = 1$ if the ith row of $\mathbf{U}_1$ is selected in design $\mathbf{X}_1$, and 0 otherwise. Then, it is easy to verify that

$$\mathbf{X}_1^\top \mathbf{X}_1 = \mathbf{U}_1^\top \mathbf{W} \mathbf{U}_1, \quad \mathbf{X}_1^\top \mathbf{X}_2 = \mathbf{U}_1^\top \mathbf{W} \mathbf{U}_2, \quad \mathbf{W}^2 = \mathbf{W}. \tag{10}$$

From (3), the loss function in (4) is,

$$\begin{aligned} \mathcal{L}_{\text{AM}}(\mathbf{X}_1) &= \sigma^2 \operatorname{trace}\left(\left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}\right) + \max_{\boldsymbol{\theta}_2 \in \Theta} \operatorname{trace}\left(\left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} \mathbf{X}_1^\top \mathbf{X}_2 \boldsymbol{\theta}_2 \boldsymbol{\theta}_2^\top \mathbf{X}_2^\top \mathbf{X}_1 \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}\right) \\ &= \sigma^2 \operatorname{trace}\left(\left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}\right) + N\alpha^2 \lambda_{\max}\left(\left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1} \mathbf{X}_1^\top \mathbf{X}_2 \mathbf{V}_2^{-1} \mathbf{X}_2^\top \mathbf{X}_1 \left(\mathbf{X}_1^\top \mathbf{X}_1\right)^{-1}\right). \end{aligned} \tag{11}$$

Now using (9) and (10), we get

$$\begin{aligned}
&\lambda_{\max}\left(\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\mathbf{X}_1^\top\mathbf{X}_2\mathbf{V}_2^{-1}\mathbf{X}_2^\top\mathbf{X}_1\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\right)\\
&=\lambda_{\max}\left(\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\mathbf{U}_1^\top\mathbf{W}(\mathbf{I}-\mathbf{U}_1\mathbf{V}_1^{-1}\mathbf{U}_1^\top)\mathbf{W}\mathbf{U}_1\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}\right)\\
&=\lambda_{\max}\left(\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}-\mathbf{V}_1^{-1}\right).
\end{aligned} \tag{12}$$

Putting (12) into (11) gives the result in (6). □

Proof of Theorem 2 For a given requirement set and run size n, define $l(n) = \min_{\mathbf{X}_1}\mathcal{L}_{\text{AM}}(\mathbf{X}_1)$, and we want to show that $l(n+1) \le l(n)$. Suppose $\mathbf{X}_1^*(n)$ minimizes $\mathcal{L}_{\text{AM}}(\mathbf{X}_1)$ for run size n, and let $\mathbf{A}(n) = \left(\mathbf{X}_1^*(n)\right)^\top\mathbf{X}_1^*(n)$. Matrix $\mathbf{A}(n)$ must be positive definite. When the run size is $n+1$, let design $\mathbf{X}_1^c(n+1)$ contain all the n runs in $\mathbf{X}_1^*(n)$ and one more run, say $\mathbf{u}^\top$ (a row vector), that is not in $\mathbf{X}_1^*(n)$, i.e., $\mathbf{X}_1^c(n+1) = \begin{pmatrix}\mathbf{X}_1^*(n)\\ \mathbf{u}^\top\end{pmatrix}$. Then we have $\mathbf{B}(n+1) := \left(\mathbf{X}_1^c(n+1)\right)^\top\mathbf{X}_1^c(n+1) = \mathbf{A}(n)+\mathbf{u}\mathbf{u}^\top$. It is easy to see that $\mathbf{B}(n+1)-\mathbf{A}(n)$ and $\mathbf{A}^{-1}(n)-\mathbf{B}^{-1}(n+1)$ are positive semidefinite matrices. Thus, trace $\left(\mathbf{B}^{-1}(n+1)\right) \le$ trace $\left(\mathbf{A}^{-1}(n)\right)$ and $\lambda_{\max}\left(\mathbf{B}^{-1}(n+1)-\mathbf{V}_1^{-1}\right) \le \lambda_{\max}\left(\mathbf{A}^{-1}(n)-\mathbf{V}_1^{-1}\right)$, which implies that $\mathcal{L}_{\text{AM}}(\mathbf{X}_1^c(n+1)) \le l(n)$. Since $l(n+1) = \min_{\mathbf{X}_1}\mathcal{L}_{\text{AM}}(\mathbf{X}_1)$, we have $l(n+1) \le \mathcal{L}_{\text{AM}}(\mathbf{X}_1^c(n+1)) \le l(n)$. □

Proof of Theorem 3 Let $\mathbf{H}_1$ be the $(N-n)\times(q+1)$ matrix consisting of those rows of $\mathbf{U}_1$ which are not rows of $\mathbf{X}_1$. Then, $\mathbf{X}_1^\top\mathbf{X}_1+\mathbf{H}_1^\top\mathbf{H}_1 = \mathbf{U}_1^\top\mathbf{U}_1 = N\mathbf{I}_{q+1}$. So, if $\mathbf{X}_1^\top\mathbf{X}_1 = n\mathbf{I}_{q+1}$, then $\mathbf{H}_1^\top\mathbf{H}_1 = (N-n)\mathbf{I}_{q+1}$, i.e., $\mathbf{H}_1$ has rank $q+1$. Since $\mathbf{H}_1$ has $N-n$ rows, this yields $N-n \ge q+1$, which contradicts $n \in [N-q, N-1]$. □

Proof of Theorem 5 Since $\mathbf{V}_1$ and $\mathbf{Q}_\pi$ are diagonal matrices and $\mathbf{Q}_\pi^2 = \mathbf{I}$, we have $\mathbf{Q}_\pi\mathbf{V}_1^{-1}\mathbf{Q}_\pi = \mathbf{V}_1^{-1}\mathbf{Q}_\pi^2 = \mathbf{V}_1^{-1}$ and $\left((\mathbf{X}_1^\pi)^\top\mathbf{X}_1^\pi\right)^{-1}-\mathbf{V}_1^{-1} = \mathbf{Q}_\pi\left(\left(\mathbf{X}_1^\top\mathbf{X}_1\right)^{-1}-\mathbf{V}_1^{-1}\right)\mathbf{Q}_\pi$, and the result follows. □

Acknowledgments The authors would like to thank an associate editor and two referees for helpful comments and suggestions which have led to improvements in the presentation of this paper. This research work is supported by Discovery Grants from the Natural Science and Engineering Research Council of Canada.

References

Cheng, C. S., Tang, B. (2005). A general theory of minimum aberration and its applications. *Annals of Statistics*, *33*, 944–958.

Deng, L., Tang, B. (2002). Design selection and classification for Hadamard matrices using generalized minimum aberration criteria. *Technometrics*, *44*, 173–184.

Fedorov, V.V. (1972). *Theory of Optimal Experiments*. New York: Academic Press.

Lin, D. K. J., Zhou, J. (2013). D-optimal minimax fractional factorial designs. *Canadian Journal of Statistics*, *41*, 325–340.

Mukerjee, R., Tang, B. (2012). Optimal fractions of two-level factorials under a baseline parameterization. *Biometrika*, *99*, 71–84.

Mukerjee, R., Wu, C. F. J. (2006). *A Modern Theory of Factorial Designs*. New York: Springer.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. New York: Wiley.

Sinha, B. K., Mukerjee, R. (1982). A note on the universal optimality criterion for full rank models. *Journal of Statistical Planning and Inference*, *7*, 97–100.

Tang, B., Zhou, J. (2009). Existence and construction of two-level orthogonal arrays for estimating main effects and some specified two-factor interactions. *Statistica Sinica*, *19*, 1193–1201.

Tang, B., Zhou, J. (2013). D-optimal two-level orthogonal arrays for estimating main effects and some specified two-factor interactions. *Metrika*, *76*, 325–337.

Wilmut, M., Zhou, J. (2011). D-optimal minimax design criterion for two-level fractional factorial designs. *Journal of Statistical Planning and Inference*, *141*, 576–587.

Xu, H., Phoa, F. K. H., Wong, W. K. (2009). Recent developments in nonregular fractional factorial designs. *Statistics Surveys*, *3*, 18–46.