DEPARTMENT OF ENGINEERING MANAGEMENT

Two-level orthogonal designs in 24 and 28 runs

Eric D. Schoen, Nha Vo-Thanh & Peter Goos

UNIVERSITY OF ANTWERP
Faculty of Applied Economics
City Campus
Prinsstraat 13, B.226
B-2000 Antwerp
Tel. +32 (0)3 265 40 32
Fax +32 (0)3 265 47 99
www.uantwerpen.be

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF ENGINEERING MANAGEMENT

Two-level orthogonal designs in 24 and 28 runs

Eric D. Schoen, Nha Vo-Thanh & Peter Goos

RESEARCH PAPER 2015-016
AUGUST 2015

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.226
phone: (32) 3 265 40 32
fax: (32) 3 265 47 99
e-mail: joeri.nys@uantwerpen.be

D/2015/1169/016

Two-level orthogonal designs in 24 and 28 runs

Eric D. Schoen
Faculty of Applied Economics, University of Antwerp, Belgium
TNO, Zeist, Netherlands

Nha Vo-Thanh
Faculty of Applied Economics, University of Antwerp, Belgium

Peter Goos
Faculty of Applied Economics and StatUa Center for Statistics, University of Antwerp, Belgium
Faculty of Bioscience Engineering and Leuven Statistics Research Centre (LSTAT), KULeuven, Belgium

August 5, 2015

Abstract

Due to incomplete previous approaches it is largely unclear which are the best 24-run and 28-run two-level designs in terms of any of the criteria currently in use to evaluate designs. In the present paper, we address this issue. First, we study the complete catalog of non-isomorphic orthogonal two-level 24-run designs involving 3–23 factors and we identify the best few designs in terms of aliasing between main effects and two-factor interaction effects and, subject to this, in terms of aliasing among two-factor interaction effects. Next, by modifying an existing enumeration algorithm, we identify the best few 28-run designs involving 3–14 factors. Based on a complete catalog of 7570 designs with 28 runs and 27 factors, we also seek good designs with more than 14 factors. To demonstrate the added value of our work, we provide a detailed comparison of our designs to the alternatives available in the literature.

KEY WORDS: G-Aberration; G_2-Aberration; Hadamard Matrix; Orthogonal Array; Plackett-Burman Design.

1 Introduction

In the early stages of process optimization, experimenters seek controllable factors that really affect the process. A usual approach to this end is to create a list of candidate factors, vary their settings according to a well-chosen experimental plan, and relate the changes observed in the process to the changes in the factor settings. This procedure is called factor screening. Orthogonal experimental designs in which all factors have two levels are commonly used plans for screening experiments. This is due to the fact that each of the four level combinations of any pair of factors is then tested equally often. Therefore, in a statistical model that includes only main effects, each of these effects is estimated with maximum precision and independently from other main effects. Two-level orthogonal experimental designs are therefore ideal when it comes to detecting main effects of factors, especially when these are the only active effects.

Some two-level orthogonal designs are regular, while others are nonregular. Regular two-level designs (see, e.g., Mee, 2009; Montgomery, 2009; Wu and Hamada, 2009) exist for run sizes that are powers of 2. The smallest regular designs have 4, 8, 16 or 32 runs and can accommodate up to 3, 7, 15 and 31 factors, respectively. One of the weaknesses of these designs is that main effects can be completely aliased with two-factor interactions and two-factor interactions can be completely aliased with other two-factor interactions. Completely aliased effects cannot be included in a model simultaneously. Therefore, the number of models that can be fitted to data from regular experimental designs is limited and there exists ambiguity when interpreting the experimental results.

Many nonregular designs do not involve completely aliased effects. For this reason, they permit more statistical models to be fitted to the data so that they offer more information on the factorial effects. In addition, in case certain interactions are active and a main-effects model is estimated, the estimates of the main effects usually have a smaller bias. The smallest nonregular designs have 12, 16 or 20 runs and involve up to 11, 15 and 19 factors, respectively. These designs have been studied extensively by Lin and Draper (1993), Deng and Tang (2002), and Sun et al. (2008).

In this paper, we study designs with 24 or 28 runs. Compared to designs with 12, 16 or 20 runs, these designs allow more factors to be included in the experiment and offer more precise estimates of the factor effects. Also, for any given number of factors, the aliasing between main effects and two-factor interactions and the aliasing among two-factor interactions are less severe for designs in 24 or 28 runs than for designs in 12, 16 or 20 runs. For example, for the best 13-factor 20-run designs listed in Deng and Tang (2002), the generalized word count of length 3, which measures the correlation between main effect contrast vectors and two-factor interaction contrast vectors, equals 15.92. For the best 24-run designs, this generalized word count is only 6, indicating a substantially smaller amount of aliasing.

Most of the applications of 24-run and 28-run orthogonal designs we have been able to trace in the applied literature are based on projections of the 23-factor 24-run design or the 27-factor 28-run design proposed by Plackett and Burman (1946). For example, Dürig and Fassihi (1993) obtained a 13-factor design in 24 runs by taking the first 13 columns of the Plackett-Burman design, while El Ati-Hellal et al. (2007) use the first 25 columns of the 27-factor Plackett-Burman design. Without performing a complete search for the best possible 13-factor 24-run design, Mee (2009, p.224) points out that the 13-factor design derived from the Plackett-Burman design can definitely be improved upon, and makes clear that starting from Plackett-Burman designs will generally not lead to an optimal design.

The main purpose of the present paper is to identify 24-run and 28-run designs with a minimum or a near-minimum degree of aliasing between main effects and two-factor interactions and, subject to this, a minimum or near-minimum degree of aliasing among two-factor interactions. The paper therefore reports designs that perform well in terms of various statistically meaningful criteria. In Section 2, we discuss the criteria used to quantify the degrees of aliasing as well as some other criteria that are relevant for comparing the performance of designs in terms of model fitting. In Section 3, we review the literature on 24-run and 28-run designs and identify several shortcomings. In Section 4, we discuss the results of a complete search over all non-isomorphic 24-run designs listed by Schoen et al. (2010), identify designs with top-3 values of the design evaluation criteria for each number of factors, compare the designs to those in the literature and provide recommendations on what designs to use. In Section 5, we modify the enumeration procedure of Schoen et al. (2010) to identify 28-run designs with top-3 values of the design evaluation criteria and up to 14 factors. We also obtain a complete list of all 7570 non-isomorphic designs with 28 runs and 27 factors. In order to find good designs involving 20–26 factors, we studied all possible projections of the best twelve 27-factor designs. Finally, in order to find good designs involving 15–19 factors, we studied all projections of the best six 20-factor designs obtained from the best twelve 27-factor designs. For each number of factors, we compare the 28-run designs we found with those in the literature and provide recommendations on which ones to use. Finally, in Section 6, we discuss the strengths and weaknesses of our approach and we summarize our findings.

2 Evaluation criteria

In this paper, we identify designs in 24 or 28 runs with a limited amount of aliasing between main effects and two-factor interactions and, subject to this, a limited amount of aliasing among two-factor interactions. To evaluate designs, we consider their strength (Hedayat et al., 1999), their confounding frequency vector (CFV; Deng and Tang, 1999, 2002) and their generalized word-length pattern (GWLP; Tang and Deng, 1999). We also evaluate the designs in terms of their ability to estimate models involving interactions. To this end, we calculate the rank of the model matrix corresponding to a model with an intercept, all main effects and all two-factor interactions. We call such a model an interaction model. Finally, we calculate the fraction of 5-factor subsets with an estimable interaction model, and we compute the average D-efficiency over all these models.

2.1 Strength

All designs considered here are two-symbol orthogonal arrays. These arrays are rectangular arrangements of the symbols, where the rows correspond to the experimental runs, the columns correspond to the factors, and the symbols correspond to the factor levels. An orthogonal array with two symbols and a strength of t is such that, in every set of t columns, each of the 2^t combinations of symbols occurs equally often (Rao, 1947). A strength of 3 implies that main effects can be estimated independently from two-factor interaction effects, while a strength of 2 implies that main effects can be estimated independently from other main effects, but not independently from two-factor interaction effects. The 24-run designs discussed in this paper are either of strength 3 or of strength 2. The 28-run designs are all of strength 2. They cannot have strength 3, because the run size of 28 is not divisible by $2^3 = 8$. Therefore, it is impossible for the eight level combinations of any three factors to occur equally often.

2.2 Confounding frequency vector

The CFV is based on so-called J_s-characteristics of s-factor interaction contrast vectors. For a two-level design, any contrast vector has elements -1 or $+1$. The J_s-characteristic of an s-factor interaction contrast vector is the absolute value of the sum of the vector's elements. Deng and Tang (1999) showed that the J_3-characteristics in two-level strength-2 designs are of the form $N - 8k$, where k is a nonnegative integer, N is the run size, and $k \leq N/8$. So, the only possible J_3-characteristics for 24-run designs are 24, 16, 8 and 0, while 28-run designs have J_3-characteristics of 28, 20, 12 or 4. A J_3-characteristic of N in an N-run design means that the main effect contrast vector of each factor involved and the two-factor interaction contrast vector of the remaining two factors are perfectly positively or negatively correlated. A J_3-characteristic of 16 or 8 in a 24-run design implies a correlation of $\pm 2/3$ or $\pm 1/3$ between any of the three main-effect contrast vectors, on the one hand, and the two-factor interaction contrast vector of the remaining factors, on the other hand. A J_3-characteristic of 20, 12 or 4 in a 28-run design implies correlations of $\pm 5/7$, $\pm 3/7$ or $\pm 1/7$. Ideally, a J_3-characteristic is as close to zero as possible. A zero J_3-characteristic means that the main effect contrast vectors of the factors involved are orthogonal to the two-factor interaction contrast vectors. For 28-run designs, a zero J_3-characteristic for a three-factor interaction contrast vector is not possible, as a result of which at least some aliasing between main effects and two-factor interaction effects is unavoidable. This is not the case for 24-run designs.

When J_3-characteristics are calculated for all possible three-factor interaction contrast vectors, the frequencies of the different outcomes can be collected in a vector F_3. For example, there is a 14-factor 24-run design with an F_3 vector equal to $F_3(24, 16, 8, 0) = (0, 4, 92, 268)$. So, among all $14!/(3!11!) = 364$ three-factor interaction contrast vectors, none have a J_3-characteristic of 24, four have a J_3-characteristic of 16, 92 have a J_3-characteristic of 8 and 268 have a J_3-characteristic of 0. It is customary to drop the last element of the F_3 vector in case it is zero, and to use $F_3(24, 16, 8) = (0, 4, 92)$ instead of $F_3(24, 16, 8, 0) = (0, 4, 92, 268)$.

In a similar fashion, frequency vectors of J-characteristics for higher-order interactions can be calculated. The concatenation of all these vectors, $(F_3, F_4, F_5, \ldots, F_n)$ (where n is the total number of factors), is the CFV. A minimum G-aberration design sequentially minimizes the entries of the CFV from left to right. In this paper, to limit the computational burden, we restrict ourselves to the sub-vectors (F_3, F_4), (F_3, F_4, F_5) and (F_3, F_4, F_5, F_6). We refer to the corresponding design evaluation criteria as the GA4, GA5 and GA6 classifiers. Note that, if there is a unique best design according to any of these criteria, then that design has minimum G-aberration.

2.3 Generalized word-length pattern

The GWLP contains the sums of the squared correlations of the third-order and the higher-order interaction contrast vectors with the intercept. These sums are called the generalized word counts of length $3, 4, \ldots, n$, depending on the order of the interaction contrast vectors considered. The generalized word counts are denoted by $A_3, A_4, \ldots, A_n$. For example, the 14-factor 24-run design with $F_3(24, 16, 8, 0) = (0, 4, 92, 268)$ has a generalized word count of length 3 equal to $0 \times 1^2 + 4 \times (\pm 2/3)^2 + 92 \times (\pm 1/3)^2 + 268 \times 0^2 = 12$. The GWLP is the vector of all generalized word counts: $(A_3, A_4, \ldots, A_n)$. A design that sequentially minimizes the elements of this vector, from left to right, is a minimum G_2-aberration design.

There may be many designs with the same generalized word counts. To distinguish between designs with the same A_3 value, we use the F_3 vector, and to distinguish between designs with the same A_4 value, we utilize the F_4 vector. This leads us to seek designs that sequentially minimize the vector (A_3, F_3, A_4, F_4), which we refer to as the mixed A4 classifier. The second half of this vector, quantifying length-4 word counts and counting J_4 characteristics, was used by Schoen and Mee (2012) to classify strength-3 designs.

Using the mixed A4 classifier instead of the pure GWLP has two advantages. First, it emphasizes the desire to avoid strong correlations between main-effect contrast vectors and two-factor interaction contrast vectors more strongly. This is because sequentially minimizing the A_3 value and the elements of the F_3 vector results in the smallest frequencies possible for large J_3-characteristics among all designs that minimize the A_3 value. A second advantage of using the mixed A4 classifier is that it reduces the number of designs to consider substantially. As a matter of fact, there may be large numbers of designs which minimize the A_3 value, but, generally, only a few of them possess the most attractive F_3 vectors.

To illustrate the usefulness of the mixed A4 classifier, consider two 16-factor 24-run designs. The best design in terms of the mixed A4 classifier has an A_3 value of 24.44 and $F_3(24, 16, 8) = (0, 4, 204)$. Another design, ranked second in terms of the mixed A4 classifier, has the same A_3 value, but $F_3(24, 16, 8) = (0, 6, 196)$. Therefore, it involves two more correlations of $\pm 2/3$ than the first design. This feature makes the second design less suitable when the intention is to fit a main-effects model.

2.4 Rank of interaction model matrix

For a design with n factors, the interaction model matrix is an $N \times q$ matrix with a column for the intercept, n main-effect contrast vector columns and $n(n-1)/2$ two-factor interaction contrast vector columns. Hence, $q = 1 + n(n+1)/2$. The difference between the rank r of the interaction model matrix and $1+n$ (the number of main effects plus one for the intercept) indicates the number of estimable two-factor interaction effects. Therefore, designs with a large rank r of the interaction model matrix are attractive.

As an illustration of the usefulness of the rank criterion, the three best 7-factor 24-run designs in terms of the GA4 classifier have ranks of 19, 24 and 22, respectively, for the interaction model matrix. Therefore, the design ranked second in terms of the GA4 classifier might be preferred if the detection of two-factor interaction effects is deemed very important.

2.5 Estimability and efficiency in subsets of five factors

The designs we consider are intended for use in screening experiments. Oftentimes, only a limited number of factors turn out to have significant effects. Therefore, it makes sense to study the projections of these designs obtained by dropping columns. This result in designs with fewer factors. Ideally, these designs allow estimation of an interaction model in the remaining factors.

To quantify the potential of a design in this context, Loeppky et al. (2007) defined the projection estimation capacity (PEC) vector $(PEC_1, PEC_2, \ldots, PEC_n)$, where PEC_x is the fraction of x-factor subsets (out of a total of $\binom{n}{x}$ subsets) for which the interaction model can be estimated. For 24-run and 28-run designs, interaction models with any seven or more factors are not estimable, so that $PEC_x = 0$ for $x \geq 7$. In this paper, we use the PEC_5 value along with the average D-efficiency with which all five-factor interaction models can be estimated. We label the average D-efficiency with which all five-factor interaction models can be estimated PIC_5, where PIC is an abbreviation of projection information capacity. We work with the PEC_5 and PIC_5 values rather than the complete PEC and PIC vectors for three reasons. First, scalar criteria are simpler to handle. Second, $PEC_3 \geq PEC_4 \geq PEC_5$ (Loeppky et al., 2007). As a result, designs with a large PEC_5 value also have large PEC_4 and PEC_3 values. Similarly, designs with large PIC_5 values also possess large PIC_4 and PIC_3 values. Finally, the results of Loeppky et al. (2007) suggest that the PEC_6 values for 24-run designs are generally rather poor, while both the PEC_5 and PEC_6 values are quite high for the best 28-run designs. So, considering the entire PEC or PIC vector rather than just the PEC_5 and PIC_5 values would not help much to discriminate among designs.

3 Literature review

The 24-run and 28-run designs in the literature have been constructed by means of two different techniques. They are either based on projections from Hadamard designs with 23 or 27 factors, or on a bottom-up design enumeration.

3.1 Projection-based approach

The papers that discuss projections from Hadamard designs start from a complete collection of non-isomorphic Hadamard matrices such as the one given by Sloane (2015a). A Hadamard matrix of order N, denoted by H_N, is an $N \times N$ matrix satisfying the equality $H_N' H_N = NI_N$. As a result, the columns of a Hadamard matrix are orthogonal to each other. This explains why Hadamard matrices can be used as orthogonal designs. Note that the rows of a Hadamard matrix are also orthogonal to each other.

The Hadamard matrices relevant for this article have dimensions $N = 24$ and $N = 28$, respectively. The dimension N of a Hadamard matrix is often called its order. There are 60 non-isomorphic Hadamard matrices of order 24 and 487 non-isomorphic Hadamard matrices of order 28. In normalized form, the elements of the first column and the first row of a Hadamard matrix are all $+1$. A Hadamard design is obtained from a normalized Hadamard matrix by dropping the first column. In this way, Evangelaras et al. (2004), Deng and Tang (2002), Ingram and Tang (2005), Belcher-Novosad and Ingram (2003) and Loeppky et al. (2007) obtain 60 Hadamard designs involving 24 runs and 23 factors and 487 Hadamard designs involving 28 runs and 27 factors. The designs recommended by these authors are all based on selections of columns of these starting designs.

One major problem with this approach is that it overlooks many alternative 24-run designs involving 23 factors and 28-run designs with 27 factors, and therefore also many potentially interesting projections of these designs obtained by dropping columns. This is because any given Hadamard matrix may give rise to several non-isomorphic designs, possibly with different statistical properties. This is explained in detail in the appendix. It is known that there are 130 non-isomorphic 23-factor designs with 24 runs (Hedayat et al., 1999) and 7570 non-isomorphic 27-factor designs in 28 runs (Sloane, 2015b, A048885). Any search involving only the 60 23-factor Hadamard matrices and the 487 27-factor Hadamard matrices is therefore incomplete and may lead to suboptimal designs.

We now summarize the literature based on projections of 23-factor 24-run designs and 27-factor 28-run designs. Yamamoto et al. (1995) and Yumiba et al. (1997) enumerate all non-isomorphic projections of all 130 non-isomorphic designs with 23 factors and 24 runs into five and six factors, respectively. The focus of these authors was on enumeration rather than on recommending good experimental designs. A positive aspect is that these authors start their search from the complete set of 130 non-isomorphic orthogonal designs. Evangelaras et al. (2004) discuss non-isomorphic projections of the 60 Hadamard designs involving 24 runs in up to five factors, and report their projection properties, generalized resolution and GWLP. Deng and Tang (2002) list top-3 designs involving 24 runs and up to eight factors according to the GA5 classifier, obtained from projections of the 60 Hadamard designs for 23 factors. Ingram and Tang (2005) specify a stepwise algorithm to search for minimum G-aberration designs and minimum G_2-aberration designs based on projections from Hadamard designs. They provide a list of recommended designs with 3–23 factors obtained from the 60 Hadamard designs for 23 factors. Belcher-Novosad and Ingram (2003) discuss a modification of Ingram and Tang's algorithm to overcome the computational difficulties when addressing 28-run designs and report a near-minimum G-aberration design with 28 runs and 17 factors. Bulutoglu and Ryan (2015) used a heuristic approach to find 28-run designs with 15–26 factors with near minimum G_2-aberration based on all 7570 27-factor designs. Finally, starting from the 60 Hadamard designs with 24 runs and the 487 Hadamard designs with 28 factors, Loeppky et al. (2007) propose a systematic approach to search for designs with good PEC values, and present extensive tables including 24-run designs with 6–23 factors and 28-run designs with 6–27 factors.

3.2 Enumeration-based approach

Generally, any search based on projections from designs with many factors is incomplete, since it is possible that orthogonal designs exist that cannot be obtained from projections. This is indeed the case for designs with 24 and 28 runs. For example, Yumiba et al. (1997) report 1317 non-isomorphic six-factor projections

Table 1: Numbers of non-isomorphic 24-run two-level designs

Factors	Designs	Factors	Designs	Factors	Designs
3	4	10	38,592,861	17	4,385,567
4	10	11	52,912,678	18	1,502,242
5	63	12	51,154,497	19	409,478
6	1,350	13	43,092,737	20	86,725
7	57,389	14	31,833,387	21	13,833
8	1,470,157	15	19,960,039	22	1,604
9	12,952,435	16	10,351,396	23	130

from all 130 23-factor designs in 24 runs. Evangelaras et al. (2007) and Schoen et al. (2010), however, report 1350 non-isomorphic six-factor designs. As a result, there exist 33 six-factor designs that cannot be found by projection from 23-run designs.

A bottom-up enumeration of all possible designs would include both the latter 33 designs as well as the former 1317 designs. Various authors have used a bottom-up enumeration approach, in which all designs involving small numbers of factors are extended factor by factor to obtain all possible designs with large numbers of factors. Li et al. (2004) introduce a bottom-up enumeration method to generate designs for 3–6 factors in up to 28 runs. Starting from a complete set of three-factor designs, all possible extensions with one additional factor were generated. Each newly obtained set with an additional factor was then sorted according to the GA5 classifier, and up to ten of the designs in each class were extended further.

Evangelaras et al. (2007) express the design generation problem as a linear system of equations. For 24 runs, they generate all non-isomorphic designs with up to six factors, while, for 28 runs, they generate all non-isomorphic designs with up to five factors. They classify the designs in terms of G-aberration and D-efficiency for the interaction model involving all factors. Angelopoulos et al. (2007) generate all non-isomorphic 7-factor designs with 24 runs and all non-isomorphic 6-factor designs with 28 runs. Their primary classification criteria are the G_2-aberration criterion for the 24-run designs and the D-efficiency for the interaction model involving all factors for the 28-run designs. Finally, Bulutoglu and Margot (2008) generate all 7-factor 24-run designs and all strength-3 24-run designs with up to 11 factors.

As opposed to the authors previously named, Schoen et al. (2010) completely enumerate all non-isomorphic 24-run designs, for any number of factors. For 28 runs, they generate all non-isomorphic designs with up to seven factors. Bulutoglu and Ryan (2015) list the distance distributions of the minimum G_2-aberration 24-run designs. These distance distributions can be converted into GWLPs using Krawtchouk polynomials. For 28-run designs, they enumerate subsets of 3–14 factor designs that are guaranteed to include the minimum G_2-aberration designs, and they identify the GWLP of these minimum aberration designs.

4 Designs with 24 runs

The numbers of non-isomorphic 24-run designs generated by Schoen et al. (2010) for 3–23 factors are given in Table 1. The large numbers of designs necessitate the use of computationally cheap criteria to search for good designs.

We identified the top-3 designs according to the GA4 and mixed A4 classifiers. For 4–12 factors and 17–23 factors, the top-3 designs according to the GA4 classifier match the top-3 designs according to the mixed A4 classifier. For 13–16 factors, the classifiers give different top-3 designs. There can be multiple designs with the same ranking in terms of the GA4 classifier or the mixed A4 classifier. To provide a more detailed characterization of each individual design in that case, we also evaluated all top-3 designs using the GA6 classifier, the PIC_5 value and the PEC_5 value. The designs are available upon request. Detailed tables describing the features of all top-3 designs are given in Appendix B. Here, we only present the most important designs we identified and their key features. A top-3 design is presented in the tables in this article if it is best according to (i) the GA4 classifier (which is a shortened version of the CFV on which the G-aberration criterion is based), (ii) the mixed A4 classifier (which is related to the G_2-aberration criterion),

(iii) the rank of the full $N \times q$ interaction model matrix, (iv) the PIC_5 value, or (v) the PEC_5 value. In case there is a tie for one criterion, we use the other criteria as tie breakers. The selected designs with 4–12 factors and their main characteristics are given in Table 2, while the selected designs with 13–23 factors and their main characteristics are shown in Table 3.

The first columns in the tables identify individual designs and contain labels of the form $n.i$, where n is the number of factors and i is the ranking based on the GA6 classifier applied to the merged set of top-3 designs according to the GA4 and mixed A4 classifiers. For each number of factors n, the table entry with the smallest value of i corresponds to a design that is optimal in terms of this classifier. For cases with 15, 17, 18 and 19 factors, the smallest value of i is larger than one. This means that multiple designs exist which are equivalent in terms of the GA6 classifier, but not in terms of the PEC_5 and PIC_5 values. Similarly, there may be several designs with equal but suboptimal rank of the GA6 classifier that differ in PEC_5 and PIC_5 values.

The tables' next few columns contain details about the generalized word counts A_3 and A_4 and the frequency vectors F_3 and F_4. Note that J-characteristics not referred to in the tables have zero frequencies. The tables' last column presents PIC_5 values. Table 2 also shows the rank of the interaction model matrix. This rank is not shown in Table 3 because it equals 24 for all selected designs with 12 or more factors. Instead, Table 3 shows the PEC_5 values of the selected designs. We do not show the PEC_5 value in Table 2, because it equals one for all selected designs with up to 12 factors. Designs recommended by Deng and Tang (2002), Ingram and Tang (2005), Loeppky et al. (2007), Bulutoglu and Ryan (2015), Evangelaras et al. (2007), Angelopoulos et al. (2007) and Plackett and Burman (1946) are indicated in Tables 2 and 3 using superscripts.

4.1 4–12 factors

Table 2: Selected 24-run designs with 6–12 factors. Designs identified previously are indicated by superscripts. For all designs in the table, $F_3(24, 16) = (0, 0)$ and $F_4(24, 16) = (0, 0)$.

ID	A_3	$F_3(8)$	A_4	$F_4(8)$	Rank	PIC_5
4.1[1,3,4]	0.00	0	0.11	1	11	-
5.1[1,3,4]	0.00	0	0.56	5	16	0.86812
5.4	0.22	2	0.11	1	16	0.93901
6.1[1,3,4,5]	0.00	0	1.67	15	18	0.86812
6.3[4,5]	0.22	2	1.00	9	22	0.90049
7.1[1,3,4,6]	0.00	0	3.89	35	19	0.86812
7.2[4,6]	0.44	4	2.33	21	24	0.89668
8.1[1,3,4]	0.00	0	7.78	70	20	0.86812
8.2[4]	0.67	6	5.78	52	23	0.87690
8.3[4]	0.78	7	5.89	53	24	0.86945
9.1[1,2,3]	0.00	0	14.00	126	21	0.86812
9.2	1.11	10	10.89	98	24	0.86782
10.1[1,2,3]	0.00	0	23.33	210	22	0.86812
10.2	1.56	14	18.67	168	24	0.86493
11.1[1,2,3]	0.00	0	36.67	330	23	0.86812
11.2	2.00	18	30.00	270	24	0.86428
12.1[1,2,3]	0.00	0	55.00	495	24	0.86812

[1]Ingram and Tang (2005); [2]Loeppky et al. (2007); [3]Bulutoglu and Ryan (2015); [4]Deng and Tang (2002); [5]Evangelaras et al. (2007); [6]Angelopoulos et al. (2007).

The 24-run minimum G-aberration and G_2-aberration designs for 4–12 factors have a strength of 3. Therefore, these designs have a zero entry in the column labeled $F_3(8)$ in Table 2, as a result of which the entire $F_3(24, 16, 8)$ vector is zero. As a consequence, the main effects can be estimated independently from the two-factor interactions when these designs are used. For four as well as six factors, two different strength-3 designs exist. Table 2 only shows the better of the two designs in each case. For other numbers of factors, there is only one strength-3 design.

The recommended four-factor design, labeled 4.1, has a strength of 3 because it consists of a full 2^4 design and a regular half fraction of it. Among the top-3 designs, design 4.1 has the best D-efficiency for the interaction model: its D-efficiency value equals 0.9684.

For 5–12 factors, the strength of 3 for the minimum G-aberration and G_2-aberration designs comes at a price. For example, the minimum G-aberration and G_2-aberration design for 5 factors has a D-efficiency of 0.86812 for the full interaction model. The design ranked 4th in terms of the G-aberration criterion and labeled 5.4, which is a strength-2 design, has a D-efficiency of 0.93901. Our results also show that strength-3 designs in 6–12 factors allow fewer two-factor interaction effects to be estimated than alternative strength-2 designs. For instance, design 6.3 has an interaction model matrix with a rank of 22, while design 6.1 has a rank of 18 only for its interaction model matrix. As a result, with design 6.3, it is possible to estimate all 15 two-factor interaction effects. With design 6.1, only 11 two-factor interaction effects can be estimated. The PIC_5 value of design 6.3 is also better than that of design 6.1. For other numbers of factors, we observe similar patterns in the results. So, if the interest is in estimating many interaction effects, maximizing the strength of the orthogonal design is not a good idea.

4.2 13–23 factors

Table 3: Selected 24-run designs with 13–23 factors. Designs identified previously are indicated by superscripts. For design 23.3, $F_3(24) = 1$. All other designs have $F_3(24) = 0$.

ID	A_3	$F_3(16)$	$F_3(8)$	A_4	$F_4(24)$	$F_4(16)$	$F_4(8)$	PEC_5	PIC_5
13.1	10.00	0	90	41.67	1	0	366	0.96503	0.82092
13.7	8.00	2	64	48.33	0	0	435	0.99456	0.84857
13.8[1,2]	6.00	6	30	55.00	0	0	495	0.98834	0.84833
14.1	15.11	0	136	53.22	1	4	454	0.96054	0.80565
14.4	12.00	4	92	61.00	1	0	540	0.98701	0.83853
15.2	20.00	0	180	67.67	2	21	507	0.95738	0.80178
15.5[1]	18.11	3	151	73.00	3	0	630	0.97269	0.81988
15.10	18.11	8	131	73.00	1	4	632	0.97303	0.81937
16.1[1]	24.89	0	224	90.22	4	32	648	0.94872	0.79405
16.5	24.44	4	204	91.11	6	0	766	0.96062	0.80489
16.10	24.44	6	196	91.11	3	3	781	0.96268	0.80614
17.2	31.11	0	280	115.11	2	83	686	0.95669	0.79808
18.2[1]	37.33	0	336	148.00	9	81	927	0.94993	0.79346
19.3[1]	45.33	0	408	185.33	4	38	1480	0.93455	0.77885
19.5	45.33	0	408	185.33	4	40	1472	0.93627	0.78006
20.1[1]	53.33	0	480	231.67	5	48	1848	0.93421	0.77857
20.3	53.33	0	480	231.67	5	160	1400	0.95227	0.79290
21.1[1]	63.33	0	570	285.00	0	0	2565	0.90226	0.75259
21.3	63.33	0	570	285.00	5	160	1880	0.44346	0.77852
22.1[1,2]	73.33	0	660	348.33	0	0	3135	0.90226	0.75259
22.3	73.33	0	660	348.33	10	99	2649	0.94087	0.78134
23.1[1,2,P]	84.33	0	759	421.67	0	0	3795	0.90226	0.75259
23.3	83.33	12	702	421.67	5	60	3510	0.92781	0.77243

[1]Ingram and Tang (2005); [2]Loeppky et al. (2007); [P]Plackett and Burman (1946).

For each number of factors from 13 to 23, the first design listed in Table 3 is a minimum G-aberration design, even if the i value in the design's label is larger than 1. For example, we selected design 15.2 rather than design 15.1 because it has a better PEC_5 value (0.95738 versus 0.95538) and a better PIC_5 value (0.80178 versus 0.80025). For similar reasons, designs 17.2, 18.2 and 19.3 were selected instead of designs 17.1, 18.1, 19.1 and 19.2.

None of the designs with 13 or more factors has a strength of 3. As a result, when 13 or more factors are studied, the main effects are aliased to some extent with the two-factor interactions. The maximum correlation between a main effect and any two-factor interaction is $\pm 1/3$ for most of the selected designs, as the designs have a zero entry in the column labeled $F_3(16)$. For the designs 13.7, 13.8, 14.4, 15.5, 15.10, 16.5, and 16.10, the maximum correlation between a main effect and any two-factor interaction is $\pm 2/3$, due to their nonzero entry in the column labeled $F_3(16)$. Nevertheless, these designs have better PEC_5 and PIC_5 values than the corresponding minimum G-aberration designs. Similarly, design 23.3 has $F_3(24) = 1$, but its PEC_5 and PIC_5 values are better than those of the minimum G-aberration design.

4.3 Comparison with designs from the literature

4.3.1 Comparison in terms of CFV and GWLP

Evangelaras et al. (2007) focus on six-factor designs. They list the seven best designs in terms of the G-aberration criterion and the eight best designs in terms of D-efficiency for the interaction model. Their first list includes all six-factor 24-run designs discussed in this paper. None of the designs in their second list is included here, because we focus on aberration criteria, which assess the suitability of designs in case not all of the main effects or interactions are active.

Angelopoulos et al. (2007) list the ten best seven-factor designs in terms of the G_2-aberration criterion. They additionally report the generalized resolution of these ten designs. Our seven-factor designs are included in their list.

Our 4-factor, 7-factor and 8-factor designs are also reported by Deng and Tang (2002). For five factors, we report one design in Table 2 that is not listed by Deng and Tang (2002), namely design 5.4. This design can be constructed by concatenating a regular 2^{5-1} design and a regular 2^{5-2} design. It performs equally well as design 5.3, which was reported by Deng and Tang (2002) in terms of the GA4 classifier. However, we prefer design 5.4 over design 5.3 because it has a better D-efficiency for the full interaction model (0.93901 versus 0.92601).

For six factors, Deng and Tang (2002) list three designs. The stated value of the GA5 classifier of their design 6.2 appears to be incorrect. It turns out that this design corresponds to our design 6.1. It permits the estimation of one more two-factor interaction effect than our design 6.2, which corresponds to design 6.1 of Deng and Tang (2002). Additionally, our design 6.1 has a better F_6 vector. However, for practical purposes, we recommend design 6.3, because it permits estimation of the interaction model in all six factors. This is why the design has a rank of 22 for the interaction model matrix. This design is also reported by Deng and Tang (2002).

Bulutoglu and Ryan (2015) identify the GWLP of the minimum G_2-aberration designs. Additional material provided by these authors via personal communication also shows the numbers of minimum G_2-aberration designs. For up to 12 factors, the minimum G_2-aberration designs are unique. For 13–20 factors, several non-isomorphic minimum G_2-aberration designs exist, and, for 21–23 factors, all orthogonal designs are equivalent in terms of G_2-aberration, as a result of which they are all minimum G_2-aberration designs. The sets of tied designs with 14 or more factors include at least one instance with either $F_3(24) > 0$ or $F_4(24) > 0$, implying completely aliased effects. Within the sets of tied designs, there are also alternative minimum G_2-aberration designs that do not have completely aliased effects. Therefore, it remains worthwhile to rank the entire catalog with the mixed A4 classifier, as done in the present paper. Our list of selected designs includes minimum G_2-aberration designs for each number of factors, but we report the best of the minimum G_2-aberration designs in terms of the mixed A4 classifier. Our minimum G-aberration designs with 13, 14, 15 and 16 factors in Table 3 differ from the minimum G_2-aberration designs in Bulutoglu and Ryan (2015).

For numbers of factors ranging from 4 to 23, Ingram and Tang (2005) present a single good design according to the G-aberration criterion and a single good design according to the G_2-aberration criterion for each number of factors. In the cases listed below, we obtained designs that outperform Ingram and Tang's designs in terms of G-aberration or in terms of the mixed A4 classifier. The preferred designs also have better PEC_5 or PIC_5 values than those of the earlier authors.

1. Our minimum G-aberration design involving 13 factors (labeled 13.1) has $F_6(16, 8) = (12, 456)$. The near-minimum G-aberration design reported by Ingram and Tang is second best, with $F_6(16, 8) = (48, 312)$.

2. Our minimum G-aberration design involving 14 factors, labeled 14.1, has $F_4(24, 16, 8) = (1, 4, 454)$. The near-minimum G-aberration design reported by Ingram and Tang is second best, with $F_4(24, 16, 8) = (1, 7, 438)$.
3. Our best 14-factor design according to the mixed A4 classifier, labeled 14.4, has $F_3(24, 16, 8) = (0, 4, 92)$. Ingram and Tang's best design in terms of the G_2-aberration criterion is second best, with $F_3(24, 16, 8) = (0, 12, 60)$, although it has the same A_3 and A_4 values as design 14.4.
4. Ingram and Tang's best 15-factor design in terms of the G-aberration criterion does not belong to the set of top-3 designs according to the GA4 classifier. Our design 15.4, with $F_4(24, 16, 8) = (3, 13, 530)$, is third best, while their design has $F_4(24, 16, 8) = (4, 8, 541)$.
5. The best 16-factor design according to the mixed A4 classifier, labeled 16.5, has $F_3(24, 16, 8) = (0, 4, 204)$. Ingram and Tang's best design in terms of G_2-aberration is second best, with $F_3(24, 16, 8) = (0, 6, 196)$, although it has the same A_3 and A_4 values as design 16.5.

Finally, one of the most frequently adopted approaches to construct a 24-run design in n factors involves taking the first n columns of the 23-factor Plackett-Burman (PB) design (Plackett and Burman, 1946). The full 23-factor PB design corresponds to design 23.1 in Table 3. We compared the $F_3(8)$ frequencies of the minimum G-aberration designs in 4–23 factors with the $F_3(8)$ frequencies of designs involving the first 4–23 columns of the PB design. For 4–12 factors, the $F_3(8)$ frequencies of the PB-derived designs increase from 1 to 94, while, for the minimum G-aberration designs, $F_3(8) = 0$. Indeed, the PB-derived designs have a strength of 2 only, while the minimum G-aberration designs for 4–12 factors have a strength of 3.

For 13–23 factors, the PB-derived designs and the minimum G-aberration designs all have a strength of 2. For 13–20 factors, the $F_3(8)$ frequencies of the PB-derived designs are 9 to 33 units larger than those of the minimum G-aberration designs. As a result, the PB-derived designs are inferior to the minimum G-aberration designs reported here, for 4–20 factors. For 21–23 factors, however, taking the first 21–23 columns of the PB design does result in the minimum G-aberration design.

4.3.2 Comparison in terms of PIC value

In this section, we compare the PIC_5 values of the designs from our search with those of the designs reported by Loeppky et al. (2007) as well as with those of the designs in Deng and Tang (2002), Ingram and Tang (2005), Evangelaras et al. (2007) and Angelopoulos et al. (2007). Loeppky et al. (2007) focused on finding designs that perform well in terms of PEC. Although we also computed the PEC, we prefer the PIC criterion because this criterion does not only record whether interaction models are estimable, but also how precisely they can be estimated.

Figure 1 compares the best PIC_5 values for the designs in Tables 2 and 3 with the best ones of the designs obtained by Loeppky et al. (2007), Deng and Tang (2002), Ingram and Tang (2005), Evangelaras et al. (2007) and Angelopoulos et al. (2007). For six factors, the best design found by Loeppky et al. (2007) matches the best design of Evangelaras et al. (2007) in terms of PIC_5 value, and outperforms ours. The best 7-factor design in terms of PIC_5 value is given by Angelopoulos et al. (2007, design 24.7.107). For 8–12 factors, our designs match the best designs of Loeppky et al. (2007), Deng and Tang (2002) and Ingram and Tang (2005). For 13, 14, 16, 17, 20 and 22 factors, our best designs in terms of PIC_5 value outperform those in the literature, although the differences are very small in a few of these cases. For 15 and 18 factors, the best designs of Ingram and Tang (2005) coincide with our best designs. Finally, for 19, 21 and 23 factors, the designs of Loeppky et al. (2007) have the best PIC_5 value. For 19 or more factors, these designs as well as our best ones outperform those proposed by Ingram and Tang (2005).

The PIC_5 values of the 23-factor PB design and the designs obtained by taking the first 6–22 columns of this designs are shown in Figure 1 by means of asterisks. The figure shows that PB-derived designs have substantially lower PIC_5 values than the best designs considered here.

Finally, Bulutoglu and Ryan (2015) do not list individual designs. For 4–12 factors, the minimum G_2-aberration designs whose GWLP they list are unique, and included in our own list. Regarding PIC_5, their designs in 5–9 factors are not the best of the competing designs, but for 10–12 factors, their designs have an optimal PIC_5. For 13–23 factors, the minimum G_2-aberration designs are not unique. For this reason, we do not consider their 13–23 factor designs in our comparison of PIC_5 values.

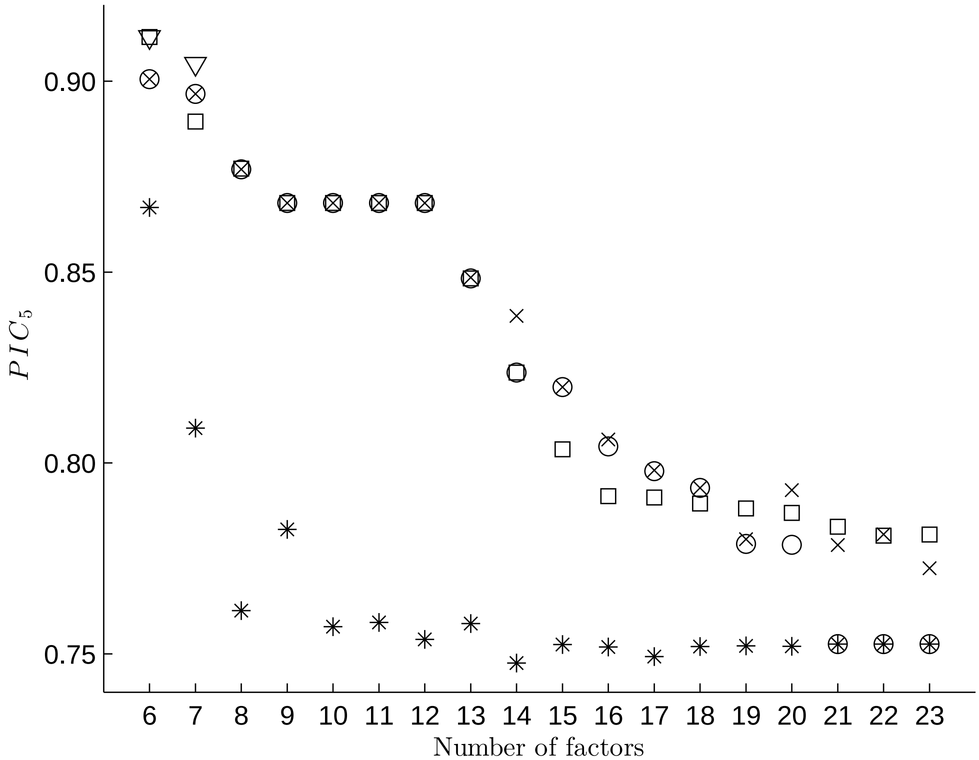

Figure 1: PIC_5 values of our 24-run designs with 6–23 factors (crosses), those obtained by Loeppky et al. (2007) (boxes), Deng and Tang (2002) or Ingram and Tang (2005) (circles), Evangelaras et al. (2007) or Angelopoulos et al. (2007) (triangles), and the first 6–23 columns of the Plackett-Burman design (asterisks).

4.4 Recommendations

It should be clear that all 24-run designs listed in Tables 2 and 3 perform well in terms of various design selection criteria. The same goes for the designs displayed in Figure 1. In this section, we attempt to provide practical guidelines regarding which design to use for each number of factors. We distinguish between the cases of 4 or 5 factors, 6 factors and 7–23 factors.

For 4 or 5 factors, the number of parameters in the full interaction model is substantially smaller than the number of runs, 24. Therefore, for 4 and 5 factors, we recommend the most D-efficient designs, which are listed in Table 2 as designs 4.1 and 5.4, respectively.

For 7–23 factors, the number of parameters in the full interaction model exceeds the run size. In these cases, a 24-run design is intended as a screening design and the primary interest is in detecting the active factors. One possible approach to factor screening, in line with the effect hierarchy principle, focuses on main effects and involves fitting a main-effects model first to identify active main effects. Minimum G-aberration designs facilitate the detection of main effects in that case because they minimize the aliasing of the main effects and the two-factor interactions. Therefore, we recommend the minimum G-aberration designs involving 7–23 factors if the factor screening starts with a main-effects model.

An alternative approach pays immediate attention to interaction effects as well and involves fitting interaction models in small numbers of factors. If that approach is utilized, it makes sense to pick designs that have a large PIC_5 value. Such designs also guarantee a precise estimation of models with fewer than five factors. Therefore, if only a few factors are expected to be active and an interaction model in these factors is of interest, we recommend the designs with the largest PIC_5 values shown in Figure 1. Most of these designs were identified by means of our evaluation of the top-3 designs we obtained, namely the designs 8.2, 9.1, 10.1, 11.1, 12.1, 13.9, 14.4, 15.10, 16.10, 17.2, 18.2, 20.3, and 22.3. Four of the designs with maximum PIC_5 value were identified by other authors, namely the 7-factor design labeled 24.7.107 by Angelopoulos et al. (2007) and the 19-, 21- and 23-factor designs found by Loeppky et al. (2007) by means of their top-down search and labeled 19.1, 21.3 and 23.2 in their paper. Note that, in some cases, the minimum G-aberration designs have the best PIC_5 values. Note also that we recommend the minimum G-aberration designs with 7–23 factors in case more than five factors are expected to be active.

For 6 factors, the number of parameters in the full interaction model equals 22. That model is estimable with an appropriate 24-run design. In case there is an immediate interest in quantifying all two-factor interaction effects, we recommend design 24.6.232 given by Evangelaras et al. (2007), which is one of the eight 6-factor designs with a D-efficiency of 0.784 for the full interaction model. When estimating the full interaction model, only two degrees of freedom remain for estimating the error variance. In case this is a concern and the intention is to start with a main-effects model, we recommend the minimum G-aberration design 6.1.

5 Designs with 28 runs

In this section, we discuss 28-run designs involving 3–27 factors. For 3–7 factors, catalogs with all non-isomorphic designs are available. The numbers of non-isomorphic designs for these numbers of factors are shown in the right column of Table 4; see Evangelaras et al. (2007) for the number of five-factor designs, Angelopoulos et al. (2007) for the number of six-factor designs and Schoen et al. (2010) for the number of designs with 3–7 factors. At present, there is no complete catalog containing all non-isomorphic 28-runs designs for numbers of factors ranging from 8 to 26. For 27 factors, there are 7570 non-isomorphic designs; see (Sloane, 2015b, sequence A048885).

In Section 5.1, we outline a step-by-step enumeration procedure to identify top-3 designs with 3–14 factors according to the GA4 and mixed A4 classifiers, and we characterize the resulting designs. In Section 5.2, we adopt a projection approach to find good designs involving 15–26 factors, starting from the 7570 non-isomorphic 27-factor designs.

5.1 3–14 factors

5.1.1 Enumeration approach

We obtained the designs for 3–14 factors using a procedure similar to the one adopted by Bulutoglu and Ryan (2015). There are four non-isomorphic designs involving 28 runs and three factors. These four designs have $F_3(28, 20, 12, 4)$ vectors equal to $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$. The best of the four three-factor designs has $(0, 0, 0, 1)$ as F_3 vector. That design involves three copies of the 2^3 factorial design plus a regular half fraction. We extended this design by adding one extra factor column at a time, using the enumeration procedure of Schoen et al. (2010). Each time a factor column has been added, we discard the designs with J_3-characteristic values greater than 4 and designs that are not in lexicographically minimal form. In other words, we construct sets of non-isomorphic designs with an $F_3(28, 20, 12, 4)$ vector of the form $(0, 0, 0, a)$, where $a = \binom{n}{3}$. The resulting sets of designs only involve the smallest possible J_3-characteristic. Therefore, these sets contain the designs that are optimal in terms of the GA4 classifier.

Using our step-by-step procedure, we were able to obtain the complete set of 28-run designs with $F_3(28, 20, 12, 4) = (0, 0, 0, a)$, involving 4–14 factors. No such designs exist with 15 or more factors. As a result, for 15 or more factors, we need to use a different procedure to find designs that perform well in terms of the GA4 classifier. In Table 4, we show the numbers of non-isomorphic designs we found with $F_3(28, 20, 12, 4) = (0, 0, 0, a)$, which match the results of Bulutoglu and Ryan (2015). These numbers are very small compared to the total number of non-isomorphic designs for 3–7 factors. The new enumeration approach thus saves substantial computing time, as fewer designs need to be extended and evaluated.

For any given number of factors, the 28-run designs with $F_3(28, 20, 12, 4) = (0, 0, 0, a)$ have the smallest possible A_3 values. This claim can be proven as follows. As shown by Deng and Tang (1999), any 28-run design has J_3-characteristics of 28, 20, 12 and 4, which correspond to correlations of ± 1, $\pm 5/7$, $\pm 3/7$ and $\pm 1/7$, respectively, and an F_3 vector of the form $F_3(28, 20, 12, 4) = (d, c, b, a)$, where $a = \binom{n}{3} - (b + c + d)$. As a result, the A_3 value of a 28-run design equals $d + c(\pm 5/7)^2 + b(\pm 3/7)^2 + [\binom{n}{3} - (b + c + d)](\pm 1/7)^2$, or

Table 4: Numbers of non-isomorphic 28-run designs

n	$F_3(28, 20, 12, 4)$	
	(0,0,0,a)	General
3	1	4
4	3	7
5	15	127
6	320	17,826
7	12,194	5,882,186
8	63,606	
9	20,552	
10	841	
11	45	
12	10	
13	2	
14	1	

$[48d + 24c + 8b + \binom{n}{3}]/49$. For a design with $b = c = d = 0$, $A_3 = \binom{n}{3}/49$. Now, suppose that a design would exist for which $A_3 \leq \binom{n}{3}/49$. For such a design, it would be required that $48d + 24c + 8b \leq 0$. As b, c and d are nonnegative integers, this is, however, impossible. As a result, no better designs exist in terms of A_3 value than those with $F_3(28, 20, 12, 4) = (0, 0, 0, a)$. To find 4–14 factor designs that are optimal according to the mixed A4 classifier, we therefore only have to rank the designs with $F_3(28, 20, 12, 4) = (0, 0, 0, a)$ according to their A_4 values, followed by the F_4 vector.

5.1.2 Results

Table 5: Selected 28-run designs with 4–14 factors. For all designs in the table, $F_3(28, 20, 12) = (0, 0, 0)$ and $F_4(28) = 0$.

ID	A_3	$F_3(4)$	A_4	$F_4(20, 12, 4)$			PIC_5
4.1	0.08	4	0.02	0	0	1	-
5.3	0.20	10	0.1	0	0	5	0.94091
6.3	0.41	20	0.31	0	0	15	0.93726
7.1	0.71	35	0.88	0	1	34	0.93102
8.1	1.14	56	2.9	0	9	61	0.91075
9.1	1.71	84	5.51	0	18	108	0.90738
10.1	2.45	120	11.47	0	44	166	0.89314
10.8	2.45	120	10.82	2	34	174	0.89629
10.10	2.45	120	10.49	6	20	184	0.89621
11.1	3.37	165	21.43	0	90	240	0.87942
11.4	3.37	165	18.82	3	65	262	0.88890
12.1	4.49	220	32.14	0	135	360	0.87934
12.3	4.49	220	28.22	6	93	396	0.88859
13.1	5.84	286	46.43	0	195	520	0.87933
14.1	7.43	364	65	0	273	728	0.87931

For 4–14 factors, our complete listing includes 77 groups of designs that are best, second best or third best according to the GA4 classifier or the mixed A4 classifier. All these designs permit estimation of the interaction models involving any subset of five factors. Hence, $PEC_5 = 1$ for each of the designs. The 77 groups differ according to the GA6 classifier, which was used for further differentiation. Each group can have multiple designs. In Appendix C, we show the values of the designs' GA6 classifier along with the word counts of lengths 3 up to 6 and their PIC_5 values. The generalized word counts of lengths 3 and 4, the F_3 and F_4 vectors and the PIC_5 values for a selection of the designs are shown in Table 5. For each number of

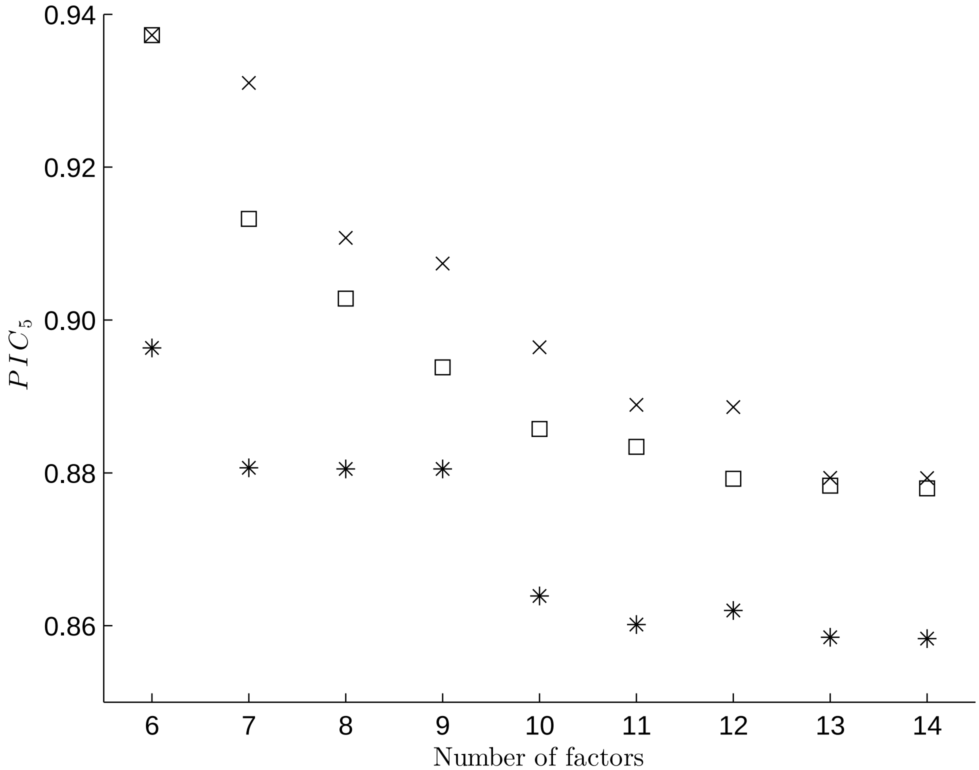

Figure 2: Best PIC_5 values of our 28-run designs with 6–14 factors (crosses), those obtained by Loeppky et al. (2007) (white boxes), and those derived from the first 6–14 columns of the Plackett-Burman design (asterisks).

factors, the table includes the designs with the best PIC_5 values we found in all of the 77 groups, the designs with the best PIC_5 values among those that are optimal in terms of the GA4 classifier and those with the best PIC_5 values among those that are optimal in terms of the mixed A4 classifier. All listed designs have the maximum possible rank of the interaction model matrix.

Because all designs in Table 5 have $F_3(28, 20, 12, 4) = (0, 0, 0, a)$, we only show the value of a in the table. That value always equals $\binom{n}{3}$. This result implies that the correlation between any main-effect contrast vector and any two-factor interaction contrast vector involving two other factors equals $\pm 1/7$. As a result, when estimating a main-effects model based on any of the designs in the table, the main effects estimates are biased only to a small extent by active two-factor interactions.

The table shows that the minimum G-aberration designs for 28 runs and 5–14 factors all have an F_4 vector of the form $F_4(28, 20, 12, 4) = (0, 0, b, a)$, with $a + b = \binom{n}{3}$ and $b \geq 0$. This implies that the designs generally perform well when a few two-factor interactions need to be estimated. This is confirmed by the high PIC_5 values: the smallest values in the table are all close to 0.88. Table 5 shows four designs that are not optimal in terms of the GA4 classifier, namely designs 10.8, 10.10, 11.4 and 12.3. The latter three designs are included because they are optimal in terms of the mixed A4 classifier. Their A_4 values are smaller than those of the designs which are optimal in terms of the GA4 classifier. Design 10.8 is second best according to the mixed A4 classifier. It is included in Table 5 because it has the best PIC_5 value among the top-3 10-factor designs.

In terms of the GA4 classifier, our 28-run designs with 7–14 factors are all better than those in Loeppky et al. (2007). For six factors, our best design and the best one of Loeppky et al. (2007) are equivalent in terms of the GA4 classifier. However, our design performs better in terms of the GA5 classifier. The PIC_5 values of our best six-factor designs and those in Loeppky et al. (2007), Angelopoulos et al. (2007), and

Li et al. (2004) coincide, while taking the first six columns from the Plackett-Burman design in 27 factors leads to inferior PIC_5 values. Figure 2 compares the PIC_5 values of the designs with 6–14 factors found by Loeppky et al. (2007) with those of our designs and the PB-derived designs. The latter group of designs is clearly inferior in terms of PIC_5. For 7–14 factors, our best designs outperform those of Loeppky et al. (2007).

Finally, Bulutoglu and Ryan (2015) present the GWLP of minimum G_2-aberration designs. Our list of selected designs in 4–14 factors includes minimum G_2-aberration designs, but we report the best of the minimum G_2-aberration designs in terms of the mixed A4 classifier.

5.2 15–27 factors

5.2.1 Approach

For 15 or more factors, no 28-run designs exist with an F_3 vector of the form $F_3(28, 20, 12, 4) = (0, 0, 0, a)$. The next best form of F_3 vector is $F_3(28, 20, 12, 4) = (0, 0, b, a)$ with $b > 0$. By exploring the complete set of non-isomorphic seven-factor designs, we identified 5,434,439 designs with such an F_3 vector. Extending these designs by adding one factor at a time would allow us to obtain a complete catalog of designs with the same kind of F_3 vector. If this factor-by-factor extension would reach designs with 15 factors or more, we would be able to identify the minimum G-aberration designs for these numbers of factors. Unfortunately, this approach is computationally infeasible due to the extremely high numbers of non-isomorphic designs involving eight or more factors. For example, extending the first 7,000 non-isomorphic seven-factor designs resulted in 1,792,801 non-isomorphic eight-factor designs. So, a rough estimate of the number of eight-factor designs of the type we are interested in is $1,792,801 \times 5,434,439/7,000 = 1,391,838,239$. Therefore, except in the special case of 27 factors, identifying the globally best 28-run designs for 15 or more factors according to the GA4 classifier is not feasible.

The infeasibility of the enumeration approach prompted us to adopt a projection approach for finding good designs involving 15–26 factors, starting from the complete catalog of non-isomorphic 27-factor designs. Before doing so, however, we first identify top-3 designs with 27 factors in terms of the GA4 classifier. To this end, we construct all 7570 non-isomorphic 27-factor designs starting from the complete catalog of all 487 non-isomorphic Hadamard matrices of order 28 listed by Sloane (2015a) and constructed by Kimura and Ohmori (1986), Kimura (1994a), and Kimura (1994b). All orthogonal 27-factor designs in 28 runs have the same GWLP, but may differ in CFV. Therefore, the top-3 designs in terms of the GA4 classifier match the top-3 designs in terms of the mixed A4 classifier. There were twelve 27-factor designs with $F_3(28, 20, 12, 4) = (0, 0, 351, 2574)$. These twelve designs are best in terms of the GA4 classifier and in terms of the mixed A4 classifier.

As the best F_3 vector for 27 factors is of the form $F_3(28, 20, 12, 4) = (0, 0, b, a)$, any minimum G-aberration design with fewer than 27 factors must also have an F_3 vector of the form $F_3(28, 20, 12, 4) = (0, 0, b, a)$. The smaller the number of factors, the smaller the anticipated b value of the corresponding minimum G-aberration design. For 20–26 factors, we therefore identified the top-3 designs in terms of the GA4 classifier among all possible projections from these twelve designs. In doing so, we considered more than 10,000,000 projections into 20-factor designs. As a consequence of this large number of relevant 20-factor designs, it was computationally infeasible to study all 15–19-factor projections of the best 12 designs with 27 factors. For this reason, for the cases of 15–19 factors, we studied all projections from the six best 20-factor designs we obtained in terms of the GA4 classifier. Using the projection approach to identify the best designs in terms of the mixed A4 classifier was computationally infeasible, because this would involve all 7570 27-factor design to start with.

Every time we encountered large numbers of designs with the same performance in terms of the GA4 classifier in the course of our projection approach, we restricted attention to the top-3 designs in terms of the GA5 classifier within the class of designs with a given performance in terms of the GA4 classifier. For example, the set of 16-factor designs that perform second best according to the GA4 classifier all have $F_3(12, 4) = (47, 513)$ and $F_4(20, 12, 4) = (21, 199, 1600)$. The three best F_5 vectors for these designs are $F_5(16, 8) = (54, 1686)$, $(56, 1678)$, and $(60, 1662)$.

Table 6: Selected designs with 28 runs and 20–27 factors. For all designs in the table, $F_3(28, 20) = (0, 0)$ and $F_4(28) = 0$.

ID	A_3	$F_3(12)$	$F_3(4)$	A_4	$F_4(20)$	$F_4(12)$	$F_4(4)$	PEC_5	PIC_5
20.1	44.00	127	1013	196.84	30	510	4305	0.99542	0.85923
20.3	44.16	128	1012	196.02	12	559	4274	0.99845	0.86129
20.x	44.82	132	1008	195.04	0	589	4256	1	0.86037
21.1	52.12	153	1177	240.67	34	624	5327	0.99568	0.85815
21.4	52.29	154	1176	240.67	8	702	5275	0.99833	0.86000
21.x	52.61	156	1174	240.18	0	723	5262	1	0.85980
22.1	60.98	181	1359	293.12	10	851	6454	0.99848	0.85949
22.4	60.98	181	1359	293.12	11	848	6456	0.99859	0.85957
22.x	61.31	183	1357	292.80	0	879	6436	1	0.85919
23.1	70.59	211	1560	354.43	0	1064	7791	1	0.85904
23.6	70.59	211	1560	354.43	8	1040	7807	0.99878	0.85916
24.1	80.82	242	1782	425.18	0	1276	9350	1	0.85885
24.3	80.82	242	1782	425.18	11	1243	9372	0.99885	0.85899
25.1	92.00	276	2024	506.00	0	1518	11132	1	0.85866
25.5	92.00	276	2024	506.00	15	1473	11162	0.99881	0.85872
26.1	104.00	312	2288	598.00	0	1794	13156	1	0.85866
27.1[P]	117.00	351	2574	702.00	0	2106	15444	1	0.85866

[P] Plackett and Burman (1946).

5.2.2 Results for 20–27 factors

Our list of top-3 designs according to the GA4 classifier for 20–27 factors includes 42 groups of designs in total. The groups differ according to the GA6 classifier. In Appendix C, we show the GA6 classifier of a representative of each of the groups, along with the PEC_5 and PIC_5 values. The generalized word counts of length 3 and 4, the F_3 and F_4 vectors and the PEC_5 and PIC_5 values for a selection of the designs are shown in Table 6. Designs were initially included in Table 6 only if they have the best PIC_5 value, or if they have the best PIC_5 value among the designs that perform best in terms of the GA4 classifier. Remarkably, this procedure did not result in designs involving 20–22 factors with a PEC_5 value of 1, even though design 23.1 involves a larger number of factors and yet has a PEC_5 value equal to 1. For this reason, we added the designs 20.x, 21.x and 22.x, which are projections from design 23.1, to the table. While these designs are neither optimal in terms of the GA4 classifier nor in terms of the PIC_5 value, they are good design options for 20–22 factors with $PEC_5 = 1$.

Each of the designs in the table has a PIC_5 value greater than 85% and a PEC_5 value greater than 99.5%. Designs with a given number of factors only show very small differences in PEC_5 and PIC_5 values. Therefore, whenever a practitioner anticipates that at most five factors will be active, we recommend a design with $PEC_5 = 1$ over designs with a slightly higher PIC_5 value and $PEC_5 < 1$. If more than five factors are expected to be active, we recommend the best tabulated designs in terms of the GA4 classifier.

Design 27.1 in our list corresponds to the 28-run PB design. In other words, in case 27 factors are studied, we do recommend the PB design. Note that this design permits estimation of interaction models involving any subset of five factors, so that $PEC_5 = 1$. The average D-efficiency with which all these models can be estimated is 0.85866.

A simple way of constructing 28-run designs with 20-26 factors is to take the first 20–26 columns of the 28-run PB design. For 25 and 26 factors, the designs obtained in this fashion correspond to design 25.1 and 26.1 in Table 6. However, our best 20–24 factor designs have $F_3(12)$ frequencies that are 11, 9, 5, 2 and 1 units smaller than those obtained by taking the first 20–24 columns of the PB design. As a result, the PB design is generally not a good starting point for finding the best 28-run design.

We compare our best designs to those of Loeppky et al. (2007) and those by taking the first 20–26 columns of the Plackett-Burman design for 27 factors in Figure 3. The figure's left panel shows the PIC_5 values. Our best designs involving 20–22 runs outperform those of Loeppky et al. (2007) in terms of PIC_5 value,

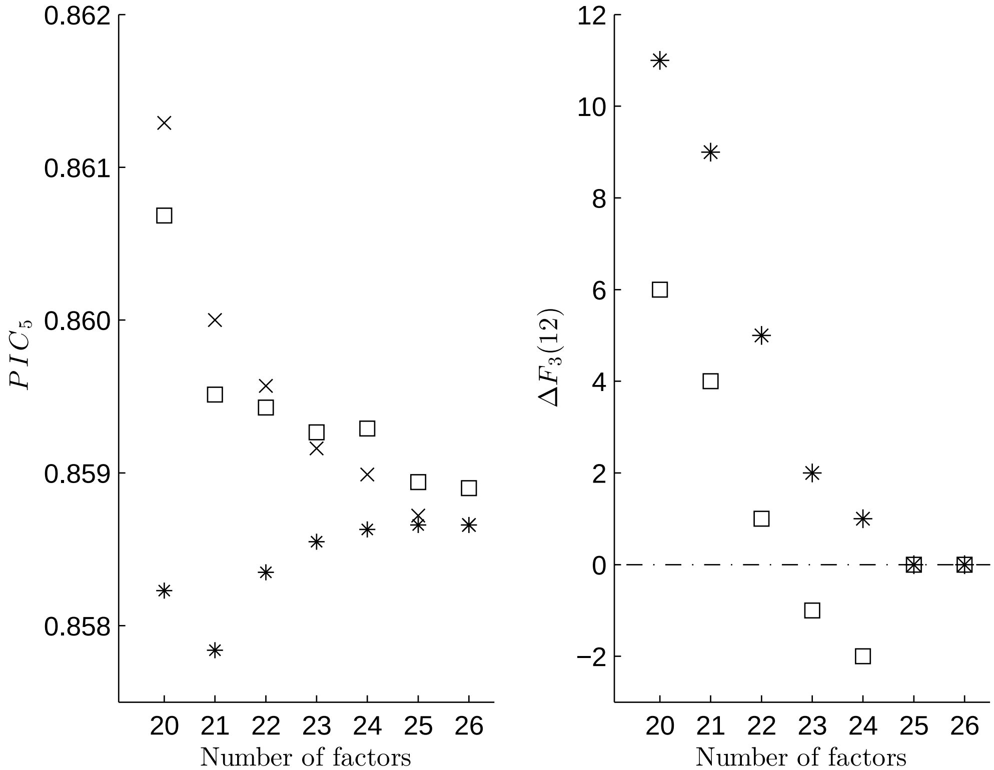

Figure 3: PIC_5 values and differences in $F_3(12)$ frequencies for the best 28-run designs 20–26 factors obtained by Loeppky et al. (2007) (white boxes), the designs derived from the first 20–26 columns of the Plackett-Burman design (asterisks) and the best designs we obtained from the 12 best 27-factor designs (crosses). Frequency differences are with respect to our designs.

while the best designs with 23–26 factors in Loeppky et al. (2007) are better than ours. The PB-derived designs all have lower PIC_5 values than the best of the other designs. However, the differences are minor. The right panel of Figure 3 shows the difference in the frequency of the J_3-characteristic of 12 between the best designs of Loeppky et al. (2007) and the PB-derived designs, on the one hand, and our best designs, on the other hand. If this difference is positive for a certain design, our design outperforms the alternative design. If the difference is negative, our design is inferior. The figure shows that our designs are better than those of Loeppky et al. (2007) for 20–22 factors, worse for 23 and 24 factors, and equally good for 25 and 26 factors. The PB-derived designs in 20–26 factors are worse for 20–24 factors and equally good for 25 and 26 factors.

Neither in terms of the F_3 vectors nor in terms of the PIC_5 values, our designs are consistently better or worse than those of Loeppky et al. (2007). For the performance in terms of the F_3 vectors, this is due to the fact that there exists no complete catalog of non-isomorphic designs with 20–26 factors which we can explore. For the performance in terms of the PIC_5 values, it is due to the fact that, for computational reasons, neither our approach nor that of Loeppky et al. (2007) evaluates all available designs in terms of their PIC_5 value. Our incomplete search for designs with 20–26 factors that perform well in terms of the GA4 classifier and the mixed A4 classifier sometimes results in a design with a better PIC_5 value. In other cases, it is the incomplete search of Loeppky et al. (2007) for designs with a large PEC_5 value that results in a design with a better PIC_5 value. In any case, Figure 3 shows that both approaches yield very similar results.

Bulutoglu and Ryan (2015) conducted a heuristic search for good 28-run designs in terms of G_2-aberration. Our designs in 20–24 factors turn out to have slightly higher A_3 values (difference at most

0.49) and slightly lower A_4 values (difference at most 1.15) than theirs. For 25–27 factors, our best designs have the same A_3 and A_4 values as those of Bulutoglu and Ryan (2015). Generally, our best designs should be more desirable in terms of the GA4 classifier, because this was the primary criterion in our heuristic search.

5.2.3 Results for 15–19 factors

Table 7: Selected 28-run designs with 15–19 factors. For all designs in the table, $F_3(28, 20) = (0, 0)$ and $F_4(28) = 0$.

ID	A_3	$F_3(12)$	$F_3(4)$	A_4	$F_4(20)$	$F_4(12)$	$F_4(4)$	PEC_5	PIC_5
15.1	14.02	29	426	65.57	21	168	1176	0.99767	0.87134
15.x	15.98	41	414	57.41	0	181	1184	1	0.87081
16.1	17.31	36	524	87.43	28	224	1568	0.99725	0.87075
16.x	20.73	57	503	74.69	0	230	1590	1	0.86631
17.1	23.18	57	623	105.06	28	262	2090	0.99628	0.86704
17.5	23.67	60	620	104.73	28	260	2092	0.99644	0.86503
17.x	25.96	74	606	96.57	0	294	2086	1	0.86342
18.3	29.71	80	736	128.73	28	322	2710	0.99568	0.86324
18.x	31.67	92	724	123.35	0	373	2687	1	0.86184
19.2	36.59	103	866	159.10	28	406	3442	0.99579	0.86107
19.7	36.76	104	865	158.78	28	404	3444	0.99596	0.86078
19.x	37.90	111	858	156.00	0	471	3405	1	0.86102

Our top list for 15–19 factors includes 32 designs. In Appendix C, we show the GA6 classifier of the designs along with their PEC_5 and PIC_5 values. The word counts of length 3 and 4, the F_3 and F_4 vectors as well as the PEC_5 and PIC_5 values of a selection of the designs are shown in Table 7. Initially, designs were included in Table 7 only if they have the best PIC_5 value, or if they have the best PIC_5 value among the designs that perform best in terms of the GA4 classifier. This procedure did not result in designs with a PEC_5 value of one. For this reason, we included the additional designs 15.x, 16.x, 17.x, 18.x and 19.x, which are projections from design 20.x, shown in Table 6. While these designs are neither optimal according to the GA4 classifier nor according to the PIC_5 value, they are good designs with 15–19 factor and with $PEC_5 = 1$.

Each of the designs in the table has a PIC_5 value greater than 86% and a PEC_5 value greater than 99.5%. Designs with a given number of factors only show very small differences in PEC_5 and PIC_5 values. Therefore, whenever a practitioner anticipates that at most five factors will be active, we recommend the designs with $PEC_5 = 1$ over designs with a slightly higher PIC_5 value and $PEC_5 < 1$. If more than five factors are expected to be active, we recommend designs that are best in terms of the GA4 classifier.

The best GWLPs identified by Bulutoglu and Ryan (2015) for 15- and 18-factor designs differ slightly from ours. The A_3 values for our best 15- and 18-factor designs are 1.31 and 0.32 units larger than those of Bulutoglu and Ryan (2015), while our A_4 values are 6.86 and 0.33 units smaller. For 16, 17, and 19 factors, our designs and the designs provided by Bulutoglu and Ryan (2015) have equal A_3 and A_4 values.

We compare our best designs to those of Loeppky et al. (2007), to the single 17-factor design from Belcher-Novosad and Ingram (2003) and to designs consisting of the first 15–19 columns of the 27-factor Plackett-Burman design in Figure 4. The figure's left panel shows PIC_5 values, while its right panel shows frequencies of J_3-characteristics of 12. Clearly, our best designs in terms of the GA4 classifier are better than the competing designs in terms of the $F_3(12)$ frequencies. We therefore recommend these designs when more than five factors are expected to be active and a main-effects model is used to identify active factors. Note that the design of Belcher-Novosad and Ingram (2003) is a near-minimum G-aberration 28-run design with 17 factors and $F_3(28, 20, 12, 4) = (0, 0, 59, 621)$. Our design 17.1 has a slightly better F_3 vector of $(0, 0, 57, 623)$.

The best PIC_5 values of our designs are up to 0.44% better than those of the best competing designs. For the case of 19 factors, our best design is slightly worse than the best design in Loeppky et al. (2007).

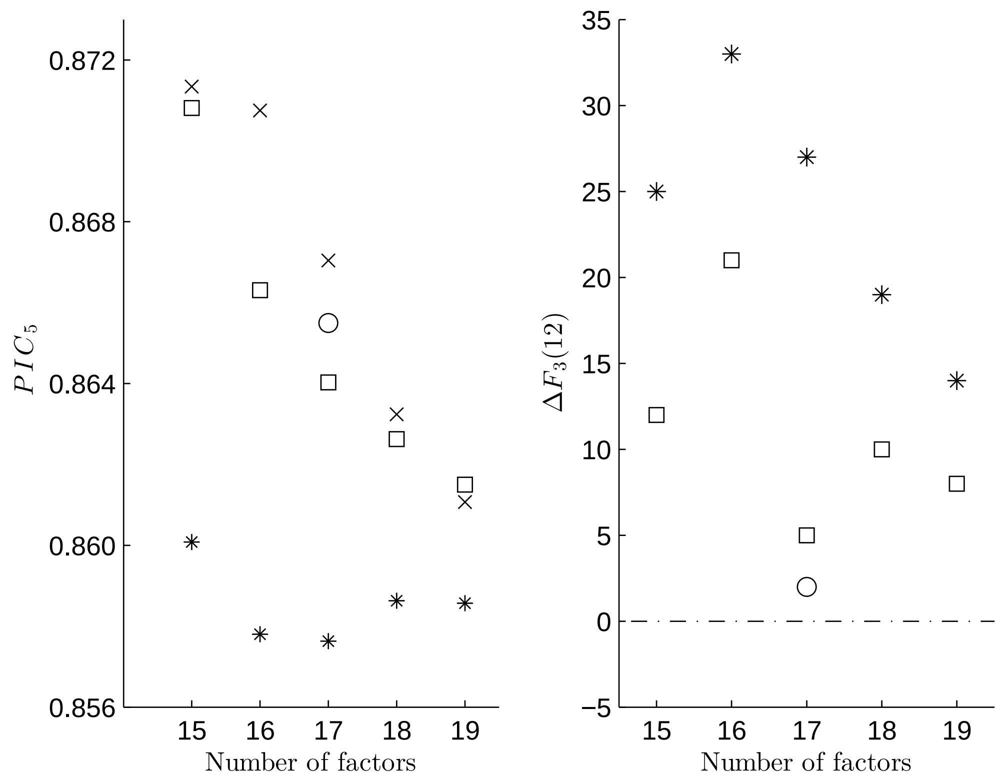

Figure 4: PIC_5 values and differences in $F_3(12)$ frequencies of the best 28-run designs 15–19 factors obtained by Loeppky et al. (2007) (white boxes), the designs derived from the first 15–19 columns of the Plackett-Burman design (asterisks), the 17-factor design from Belcher-Novosad and Ingram (2003) (circles) and the best designs we obtained from the six best 20-factor designs (crosses). Frequency differences are with respect to our designs.

Therefore, the best designs in terms of the PIC_5 value, both in Loeppky et al. (2007) and in our paper, are all good design choices in case at most five factors are expected to be active.

As before, the PB-derived designs are neither optimal in terms of G-aberration nor in terms of PIC_5 value.

6 Discussion

This paper describes the results of a complete search of the best 24-run designs involving up to 23 factors and of the best 28-run designs in terms of the GA4 and mixed A4 classifiers with up to 14 factors. For 28-run designs involving 15–26 factors, we searched for the best projections from 12 attractive 27-factor designs. All designs we obtained were compared to those from the literature. Our work, which supersedes all published work on designs with 24 and 28 runs in terms of thoroughness, resulted in a substantial number of designs that outperform those from the literature.

The paper features a partial enumeration procedure for orthogonal designs with run sizes that are not a multiple of eight, like the 28-run designs we studied. We extended the three-factor design whose only J_3-characteristic equals 4 by adding one extra factor column at a time. Each time a factor column had been added, we discarded the designs with J_3-characteristic values greater than 4 and designs that are not in lexicographically minimal form. We showed that this procedure allows us to find top 3 designs according to the GA4 and mixed A4 classifiers with up to a certain maximum number of factors. That maximum number is not known a priori. For the 28-run case, the partial enumeration procedure enabled us to find optimal designs with up to 14 factors.

Much has been said in this paper about incomplete searches for best designs. We showed that a substantial portion of the literature on designs based on Hadamard matrices ignores the fact that several non-isomorphic orthogonal designs can be constructed from a given Hadamard matrix. At the same time, there exist N-run designs with fewer than $N-1$ factors than cannot be obtained by projections from $(N-1)$-factor designs. However, our own search among all possible 24-run designs was also incomplete because, for computational reasons, we only evaluated all 24-run designs in terms of the GA4 and mixed A4 classifiers, and not in terms of the PEC_5 and PIC_5 values, for instance. Also, our search for good 28-run designs in 15–26 factors involved a projection approach starting with 12 good 27-factor designs. The only published other study of the complete catalog of 24-run designs, Bulutoglu and Ryan (2015), only considers the G_2-aberration criterion, and, unlike us, does not evaluate the designs in terms of model-based criteria that might be more appealing to statisticians. For 28-run designs in 15–26 factors, their approach, like ours, was heuristic.

While we think that the criteria we used are sensible, other criteria may lead to different designs being best. We encourage future research concerning designs' capacity to fit models with all main effects and subsets of g two-factor interactions. For $g = 1, 2, \ldots, 0.5n(n-1)$, an information capacity (IC) vector can be constructed with average D-efficiencies, and designs can be compared based on this vector (Li and Nachtsheim, 2000). However, this will certainly be computationally very demanding.

Finally, it would be of practical interest to be able to group the runs of the designs studied here in small blocks such that the main effects are not confounded with the blocks and the two-factor interactions are confounded with the blocks as little as possible.

Acknowledgements

The research of the first author was supported by the Flemish Foundation for Scientific Research FWO. The authors are grateful to Dursun Bulutoglu and Kenneth Ryan for providing more details of their search results.

A Non-isomorphic designs from a given Hadamard matrix

We illustrate the fact that one Hadamard matrix can give rise to multiple non-isomorphic designs using the 55th Hadamard matrix of order 24 from Sloane's website. The normalized Hadamard matrix is shown in Table 8.

Dropping the first column of the Hadamard matrix results in a 23-factor design with $F_3(24, 16, 8) = (1, 22, 662)$. We now transform the original Hadamard matrix into an isomorphic one which yields a non-isomorphic orthogonal design. To this end, we denote the columns of the original Hadamard matrix by $c_0, c_1, \ldots, c_{23}$. First, we swap columns c_0 and c_6. Next, we switch all the levels in the rows of the resulting matrix that start with a -1. In other words, we normalize the resulting matrix. The normalization is achieved by multiplying all columns with column c_6 in an element-wise fashion. This results in the matrix $[c_6.c_6, c_1.c_6, \ldots\ c_5.c_6, c_0.c_6, c_6.c_7\ \ldots, c_6.c_{23}]$, or $[c_0, c_1.c_6, \ldots, c_5.c_6, c_6, c_6.c_7, \ldots, c_6.c_{23}]$, since $c_6.c_6 = c_0$ and $c_0.c_6 = c_6$. Dropping the first column of the newly obtained matrix results in a new 23-factor orthogonal design with $F_3(24, 16, 8) = (1, 18, 678)$. This is a better F_3 vector than the original. Swapping the columns c_0 and c_2 in the original matrix and normalizing the result yields another orthogonal design with $F_3(24, 16, 8) = (1, 24, 654)$. This is a worse F_3 vector than the original.

At first sight, it seems contradictory that we can obtain several non-isomorphic orthogonal designs from one Hadamard matrix. The reason for this apparent contradiction is a difference between the definitions of isomorphism for Hadamard matrices and for orthogonal designs. Two Hadamard matrices are considered isomorphic when they can be obtained from each other using column permutations, row permutations, level switches in entire columns and level switches in entire rows. Therefore, the above operations using column c_6 or column c_2 result in isomorphic Hadamard matrices. Two orthogonal designs are isomorphic when they can be obtained from each other using column permutations, row permutations and level switches in entire columns, but not level switches in entire rows. Hence, as soon as the levels in certain rows of a Hadamard matrix are switched, it is possible that a new, non-isomorphic orthogonal design is obtained. This is because levels switches in rows affect the interrelation among the factors in an experimental design.

Table 8: Normalized Hadamard matrix 55

+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−
+	+	+	+	+	+	−	−	−	−	−	−	+	+	+	−	−	−	+	+	+	−	−	−
+	+	+	+	+	+	−	−	−	−	−	−	−	−	−	+	+	+	−	−	−	+	+	+
+	+	+	−	−	−	+	+	+	−	−	−	+	−	−	+	+	−	+	+	−	+	−	−
+	+	+	−	−	−	+	+	+	−	−	−	−	+	+	−	−	+	−	−	+	−	+	+
+	−	−	+	+	−	+	+	−	+	−	−	+	+	+	+	+	+	−	−	−	−	−	−
+	−	−	+	+	−	+	+	−	+	−	−	−	−	−	−	−	−	+	+	+	+	+	+
+	+	−	+	−	−	+	−	−	−	+	+	+	+	−	+	−	−	+	−	−	−	+	+
+	+	−	+	−	−	−	+	−	−	+	+	−	+	−	−	+	+	−	+	+	+	−	−
+	−	+	−	+	−	+	−	−	−	+	+	−	−	+	+	−	+	+	−	+	+	−	−
+	−	+	−	+	−	−	+	−	−	+	+	+	−	+	−	+	−	−	+	−	−	+	+
+	−	−	+	−	+	+	−	+	−	+	−	+	−	+	−	+	−	−	−	+	+	+	−
+	−	−	−	+	+	−	+	+	−	+	−	+	+	−	−	−	+	+	−	−	+	−	+
+	+	−	−	−	+	+	−	−	+	+	−	−	−	+	−	+	+	+	+	−	−	−	+
+	−	+	−	−	+	−	+	−	+	+	−	−	+	−	+	+	−	+	−	+	−	+	−
+	+	−	−	+	−	−	−	+	+	+	−	+	−	−	+	−	+	−	+	+	−	+	−
+	−	+	+	−	−	−	−	+	+	+	−	−	+	+	+	−	−	−	+	−	+	−	+
+	−	−	+	−	+	−	+	+	−	−	+	−	−	+	+	−	+	+	+	−	−	+	−
+	−	−	−	+	+	+	−	+	−	−	+	−	+	−	+	+	−	−	+	+	−	−	+
+	+	−	−	−	+	−	+	−	+	−	+	+	−	+	+	−	−	−	−	+	+	−	+
+	−	+	−	−	+	+	−	−	+	−	+	+	+	−	−	−	+	−	+	−	+	+	−
+	+	−	−	+	−	−	−	+	+	−	+	−	+	+	−	+	−	+	−	−	+	+	−
+	−	+	+	−	−	−	−	+	+	−	+	+	−	−	−	+	+	+	−	+	−	−	+

B Details of 24-run designs

Table 9: Top-3 two-level designs with 24 runs according to the GA4 classifier and the mixed A4 classifier.

ID	A_3	$F_3(24,16,8)$			A_4	$F_4(24,16,8)$			A_5	$F_5(16,8)$		A_6	$F_6(16,8)$		Rank	PEC_5	PIC_5
4.1	0.00	0	0	0	0.11	0	0	1	-	-	-	-	-	-	11	-	-
4.2	0.00	0	0	0	1.00	1	0	0	-	-	-	-	-	-	8	-	-
4.3	0.11	0	0	1	0.00	0	0	0	-	-	-	-	-	-	11	-	-
5.1	0.00	0	0	0	0.56	0	0	5	0.00	0	0	0.00	0	0	16	1	0.86812
5.2	0.11	0	0	1	0.33	0	0	3	0.11	0	1	0.00	0	0	16	1	0.90220
5.3	0.22	0	0	2	0.11	0	0	1	0.00	0	0	0.00	0	0	16	1	0.92601
5.4	0.22	0	0	2	0.11	0	0	1	0.44	1	0	0.00	0	0	16	1	0.93901
6.1	0.00	0	0	0	1.67	0	0	15	0.00	0	0	0.00	0	0	18	1	0.86812
6.2	0.00	0	0	0	1.67	0	0	15	0.00	0	0	0.44	1	0	17	1	0.86812
6.3	0.22	0	0	2	1.00	0	0	9	0.44	0	4	0.00	0	0	22	1	0.90049
6.4	0.22	0	0	2	1.22	0	0	11	0.22	0	2	0.00	0	0	21	1	0.88139
7.1	0.00	0	0	0	3.89	0	0	35	0.00	0	0	0.44	1	0	19	1	0.86812
7.2	0.44	0	0	4	2.33	0	0	21	1.33	0	12	0.22	0	2	24	1	0.89668
7.3	0.44	0	0	4	2.56	0	0	23	0.89	0	8	0.44	1	0	22	1	0.88825
8.1	0.00	0	0	0	7.78	0	0	70	0.00	0	0	1.78	4	0	20	1	0.86812
8.2	0.67	0	0	6	5.78	0	0	52	2.00	0	18	1.11	1	6	23	1	0.87690
8.3	0.78	0	0	7	5.89	0	0	53	1.56	0	14	1.11	1	6	24	1	0.86945
9.1	0.00	0	0	0	14.00	0	0	126	0.00	0	0	5.33	12	0	21	1	0.86812
9.2	1.11	0	0	10	10.89	0	0	98	3.11	0	28	3.56	4	16	24	1	0.86782
9.3	1.22	0	0	11	10.89	0	0	98	3.00	0	27	3.56	4	16	24	1	0.86420
10.1	0.00	0	0	0	23.33	0	0	210	0.00	0	0	13.33	30	0	22	1	0.86812
10.2	1.56	0	0	14	18.67	0	0	168	5.33	0	48	9.33	12	36	24	1	0.86493
10.3	2.00	0	0	18	18.67	0	0	168	4.00	0	36	9.33	12	36	23	1	0.85458
11.1	0.00	0	0	0	36.67	0	0	330	0.00	0	0	29.33	66	0	23	1	0.86812
11.2	2.00	0	0	18	30.00	0	0	270	8.67	0	78	21.33	30	72	24	1	0.86428
11.3	3.11	0	0	28	27.33	0	0	246	10.67	0	96	21.33	30	72	24	0.99134	0.85082

continued on next page

Table 9 (continued)

ID	A_3	$F_3(24,16,8)$			A_4	$F_4(24,16,8)$			A_5	$F_5(16,8)$		A_6	$F_6(16,8)$		Rank	PEC_5	PIC_5
12.1	0.00	0	0	0	55.00	0	0	495	0.00	0	0	58.67	132	0	24	1	0.86812
12.2	4.00	0	0	36	41.67	0	0	375	17.33	0	156	42.67	60	144	24	0.99242	0.85153
12.3	5.00	0	0	45	36.67	0	0	330	26.67	12	192	29.33	30	144	24	0.98485	0.84696
12.4	5.00	0	0	45	36.67	0	0	330	26.67	30	120	29.33	66	0	24	0.96212	0.82790
13.1	10.00	0	0	90	41.67	1	0	366	53.33	24	384	56.00	12	456	24	0.96503	0.82092
13.2	10.00	0	0	90	41.67	1	0	366	53.33	24	384	56.00	48	312	24	0.96503	0.82062
13.3	10.00	0	0	90	41.67	1	0	366	53.33	42	312	56.00	30	384	24	0.95105	0.80828
13.4	10.00	0	0	90	41.67	1	0	366	53.33	60	240	56.00	84	168	24	0.88112	0.75622
13.5	10.67	0	0	96	40.33	1	0	354	54.67	24	396	54.67	48	300	24	0.93473	0.79409
13.6	10.89	0	0	98	41.00	0	4	353	50.00	6	426	62.44	16	498	24	0.97824	0.82254
13.7	8.00	0	2	64	48.33	0	0	435	46.67	12	372	57.33	60	276	24	0.99456	0.84857
13.8	6.00	0	6	30	55.00	0	0	495	40.00	30	240	58.67	132	0	24	0.98834	0.84833
13.9	6.00	1	0	45	55.00	0	0	495	40.00	30	240	58.67	132	0	24	0.94172	0.81619
14.1	15.11	0	0	136	53.22	1	4	454	81.11	8	698	109.33	28	872	24	0.96054	0.80565
14.2	15.11	0	0	136	52.78	1	7	438	81.33	10	692	111.11	22	912	24	0.95804	0.80491
14.3	15.11	0	0	136	52.56	1	14	408	81.78	16	672	111.11	26	896	24	0.95504	0.80212
14.4	12.00	0	4	92	61.00	1	0	540	80.00	24	624	98.67	90	528	24	0.98701	0.83853
14.5	12.00	0	12	60	61.00	0	6	525	80.00	60	480	98.67	162	240	24	0.96903	0.82370
14.6	12.00	0	12	60	61.00	0	6	525	80.00	60	480	98.67	162	240	24	0.97702	0.82904
14.7	12.00	0	12	60	61.00	1	0	540	80.00	60	480	98.67	162	240	24	0.96503	0.82042
15.1	20.00	0	0	180	67.67	2	21	507	124.44	12	1072	193.33	72	1452	24	0.95538	0.80025
15.2	20.00	0	0	180	67.67	2	21	507	124.44	12	1072	193.33	72	1452	24	0.95738	0.80178
15.3	20.00	0	0	180	67.67	2	23	499	124.44	12	1072	193.33	72	1452	24	0.95438	0.79985
15.4	20.00	0	0	180	67.67	3	13	530	124.44	13	1068	193.33	42	1572	24	0.94838	0.79525
15.5	18.11	0	3	151	73.00	3	0	630	125.33	13	1076	178.67	60	1368	24	0.97269	0.81988
15.6	18.11	0	8	131	73.00	1	4	632	125.33	40	968	178.67	114	1152	24	0.96703	0.81516
15.7	18.11	0	8	131	73.00	1	4	632	125.33	40	968	178.67	114	1152	24	0.97003	0.81738
15.8	18.11	0	8	131	73.00	1	4	632	125.33	40	968	178.67	114	1152	24	0.97003	0.81738
15.9	18.11	0	8	131	73.00	1	4	632	125.33	40	968	178.67	114	1152	24	0.96903	0.81661
15.10	18.11	0	8	131	73.00	1	4	632	125.33	40	968	178.67	114	1152	24	0.97303	0.81937
15.11	18.11	0	8	131	73.00	1	4	632	125.33	40	968	178.67	114	1152	24	0.97303	0.81937
15.12	18.11	0	8	131	73.00	3	0	630	125.33	40	968	178.67	114	1152	24	0.96870	0.81657

continued on next page

Table 9 (continued)

ID	A_3	$F_3(24,16,8)$			A_4	$F_4(24,16,8)$			A_5	$F_5(16,8)$		A_6	$F_6(16,8)$		Rank	PEC_5	PIC_5
16.1	24.89	0	0	224	90.22	4	32	648	181.33	0	1632	305.78	64	2496	24	0.94872	0.79405
16.2	24.89	0	0	224	90.22	4	40	616	181.33	0	1632	305.78	96	2368	24	0.94780	0.79331
16.3	24.89	0	0	224	90.22	4	40	616	181.33	0	1632	305.78	96	2368	24	0.94780	0.79331
16.4	24.89	0	0	224	90.22	4	64	520	181.33	0	1632	305.78	192	1984	24	0.94689	0.79297
16.5	24.44	0	4	204	91.11	6	0	766	181.33	8	1600	304.00	44	2560	24	0.96062	0.80489
16.6	24.44	0	6	196	91.11	2	8	770	181.33	34	1496	304.00	92	2368	24	0.96062	0.80435
16.7	24.44	0	6	196	91.11	2	8	770	181.33	34	1496	304.00	92	2368	24	0.95559	0.80051
16.8	24.44	0	6	196	91.11	2	8	770	181.33	34	1496	304.00	92	2368	24	0.95559	0.80085
16.9	24.44	0	6	196	91.11	3	3	781	181.33	19	1556	304.00	73	2444	24	0.95994	0.80418
16.10	24.44	0	6	196	91.11	3	3	781	181.33	19	1556	304.00	73	2444	24	0.96268	0.80614
16.11	24.44	0	6	196	91.11	3	3	781	181.33	19	1556	304.00	73	2444	24	0.96268	0.80614
16.12	24.44	0	6	196	91.11	3	3	781	181.33	19	1556	304.00	73	2444	24	0.95994	0.80418
17.1	31.11	0	0	280	115.11	2	83	686	254.22	0	2288	487.11	342	3016	24	0.95637	0.79787
17.2	31.11	0	0	280	115.11	2	83	686	254.22	0	2288	487.11	342	3016	24	0.95669	0.79808
17.3	31.11	0	0	280	115.11	3	77	701	254.22	0	2288	487.11	342	3016	24	0.95507	0.79703
17.4	31.11	0	0	280	115.11	3	77	701	254.22	0	2288	487.11	342	3016	24	0.95427	0.79635
17.5	31.11	0	0	280	115.11	3	83	677	254.22	0	2288	487.11	340	3024	24	0.95006	0.79333
17.6	31.11	0	0	280	115.11	3	83	677	254.22	0	2288	487.11	340	3024	24	0.95184	0.79469
18.1	37.33	0	0	336	148.00	9	81	927	352.00	0	3168	730.67	438	4824	24	0.94643	0.79089
18.2	37.33	0	0	336	148.00	9	81	927	352.00	0	3168	730.67	438	4824	24	0.94993	0.79346
18.3	37.33	0	0	336	148.00	12	16	1160	352.00	0	3168	730.67	80	6256	24	0.93207	0.77966
18.4	37.33	0	0	336	148.00	12	20	1144	352.00	0	3168	730.67	72	6288	24	0.92694	0.77601
19.1	45.33	0	0	408	185.33	4	38	1480	472.00	0	4248	1082.67	136	9200	24	0.93369	0.77816
19.2	45.33	0	0	408	185.33	4	38	1480	472.00	0	4248	1082.67	136	9200	24	0.93387	0.77828
19.3	45.33	0	0	408	185.33	4	38	1480	472.00	0	4248	1082.67	136	9200	24	0.93455	0.77885
19.4	45.33	0	0	408	185.33	4	38	1480	472.00	0	4248	1082.67	136	9200	24	0.93404	0.77848
19.5	45.33	0	0	408	185.33	4	40	1472	472.00	0	4248	1082.67	128	9232	24	0.93627	0.78006
19.6	45.33	0	0	408	185.33	4	40	1472	472.00	0	4248	1082.67	128	9232	24	0.93352	0.77809
19.7	45.33	0	0	408	185.33	4	80	1312	472.00	56	4024	1082.67	336	8400	24	0.92948	0.77531
20.1	53.33	0	0	480	231.67	5	48	1848	629.33	0	5664	1546.67	192	13152	24	0.93421	0.77857
20.2	53.33	0	0	480	231.67	5	160	1400	629.33	0	5664	1546.67	1120	9440	24	0.94763	0.78942

continued on next page

Table 9 (continued)

ID	A_3	$F_3(24,16,8)$			A_4	$F_4(24,16,8)$			A_5	$F_5(16,8)$		A_6	$F_6(16,8)$		Rank	PEC_5	PIC_5
20.3	53.33	0	0	480	231.67	5	160	1400	629.33	0	5664	1546.67	1120	9440	24	0.95227	0.79290
20.4	53.33	0	0	480	231.67	9	90	1644	629.33	0	5664	1546.67	492	11952	24	0.94543	0.78718
20.5	53.33	0	0	480	231.67	9	90	1644	629.33	0	5664	1546.67	492	11952	24	0.94253	0.78493
21.1	63.33	0	0	570	285.00	0	0	2565	816.00	0	7344	2176.00	0	19584	24	0.90226	0.75259
21.2	63.33	0	0	570	285.00	5	48	2328	816.00	48	7152	2176.00	192	18816	24	0.92643	0.77066
21.3	63.33	0	0	570	285.00	5	160	1880	816.00	160	6704	2176.00	1120	15104	24	0.93489	0.77852
21.4	63.33	0	0	570	285.00	5	160	1880	816.00	160	6704	2176.00	1120	15104	24	0.93135	0.77587
22.1	73.33	0	0	660	348.33	0	0	3135	1056.00	0	9504	2992.00	0	26928	24	0.90226	0.75259
22.2	73.33	0	0	660	348.33	10	99	2649	1056.00	84	9168	2992.00	546	24744	24	0.93860	0.77953
22.3	73.33	0	0	660	348.33	10	99	2649	1056.00	84	9168	2992.00	546	24744	24	0.94087	0.78134
22.4	73.33	0	0	660	348.33	15	56	2776	1056.00	64	9248	2992.00	240	25968	24	0.92869	0.77223
22.5	73.33	0	0	660	348.33	15	56	2776	1056.00	64	9248	2992.00	240	25968	24	0.92641	0.77064
23.1	84.33	0	0	759	421.67	0	0	3795	1349.33	0	12144	4048.00	0	36432	24	0.90226	0.75259
23.2	84.33	0	66	495	421.67	0	330	2475	1349.33	1056	7920	4048.00	3168	23760	24	0.91925	0.76914
23.3	84.33	1	12	702	421.67	5	60	3510	1349.33	96	11760	4048.00	288	35280	24	0.92781	0.77243

C Details of 28-run designs

Table 10: Top-3 two-level designs from with 28 runs according to the GA4 classifier (4–27 factors) and the mixed A4 classifier (4–14 factors).

ID	A_3	$F_3(12;4)$		A_4	$F_4(20;12;4)$			A_5	$F_5(24;16;8)$			A_6	$F_6(24;16;8)$			Rank	PEC_5	PIC_5
4.1	0.08	0	4	0.02	0	0	1	0.00	0	0	0	0	0	0	0	11	-	-
4.2	0.08	0	4	0.18	0	1	0	0.00	0	0	0	0	0	0	0	11	-	-
4.3	0.08	0	4	0.51	1	0	0	0.00	0	0	0	0	0	0	0	11	-	-
5.1	0.20	0	10	0.10	0	0	5	0.00	0	0	0	0	0	0	0	16	1	0.93544
5.2	0.20	0	10	0.10	0	0	5	0.08	0	0	1	0	0	0	0	16	1	0.94075
5.3	0.20	0	10	0.10	0	0	5	0.33	0	1	0	0.00	0	0	0	16	1	0.94091
5.4	0.20	0	10	0.27	0	1	4	0.00	0	0	0	0.00	0	0	0	16	1	0.89597
5.5	0.20	0	10	0.27	0	1	4	0.08	0	0	1	0.00	0	0	0	16	1	0.89344

continued on next page

Table 10 (continued)

ID	A_3	$F_3(12;4)$		A_4	$F_4(20;12;4)$			A_5	$F_5(24;16;8)$			A_6	$F_6(24;16;8)$			Rank	PEC_5	PIC_5
5.6	0.20	0	10	0.27	0	1	4	0.33	0	1	0	0.00	0	0	0	16	1	0.89620
5.7	0.20	0	10	0.43	0	2	3	0.00	0	0	0	0.00	0	0	0	16	1	0.85196
5.8	0.20	0	10	0.43	0	2	3	0.08	0	0	1	0.00	0	0	0	16	1	0.84978
6.1	0.41	0	20	0.31	0	0	15	0.00	0	0	0	0.73	1	0	0	22	1	0.93544
6.2	0.41	0	20	0.31	0	0	15	0.24	0	0	3	0.33	0	1	0	22	1	0.93635
6.3	0.41	0	20	0.31	0	0	15	0.49	0	0	6	0.08	0	0	1	22	1	0.93726
6.4	0.41	0	20	0.47	0	1	14	0.08	0	0	1	0.33	0	1	0	22	1	0.92259
6.5	0.41	0	20	0.47	0	1	14	0.24	0	0	3	0.33	0	1	0	22	1	0.92233
6.6	0.41	0	20	0.47	0	1	14	0.33	0	0	4	0.08	0	0	1	22	1	0.92268
6.7	0.41	0	20	0.63	0	2	13	0.16	0	0	2	0.08	0	0	1	22	1	0.90891
6.8	0.41	0	20	0.63	0	2	13	0.24	0	0	3	0.00	0	0	0	22	1	0.90860
6.9	0.41	0	20	0.63	0	2	13	0.33	0	0	4	0.08	0	0	1	22	1	0.90948
7.1	0.71	0	35	0.88	0	1	34	1.55	0	2	11	0.41	0	0	5	28	1	0.93102
7.2	0.71	0	35	1.04	0	2	33	1.22	0	0	15	0.57	0	1	3	28	1	0.92515
7.3	0.71	0	35	1.04	0	2	33	1.39	0	0	17	0.41	0	0	5	28	1	0.92553
7.4	0.71	0	35	1.04	0	2	33	1.22	0	1	11	0.57	0	1	3	28	1	0.92494
7.5	0.71	0	35	1.20	0	3	32	1.06	0	0	13	0.57	0	1	3	28	1	0.91913
7.6	0.71	0	35	1.20	0	3	32	1.22	0	0	15	0.41	0	0	5	28	1	0.91946
7.7	0.71	0	35	1.20	0	3	32	1.22	0	0	15	0.41	0	1	1	28	1	0.91910
7.8	0.71	0	35	1.20	0	3	32	1.06	0	1	9	0.57	0	1	3	28	1	0.91885
7.9	0.71	0	35	1.20	0	3	32	1.06	0	2	5	0.57	0	1	3	28	1	0.91877
8.1	1.14	0	56	2.90	0	9	61	3.27	0	0	40	0.65	0	0	8	28	1	0.91075
8.2	1.14	0	56	2.90	0	9	61	3.27	0	1	36	0.65	0	0	8	28	1	0.91032
8.3	1.14	0	56	2.90	0	9	61	3.27	0	2	32	0.65	0	0	8	28	1	0.91054
8.4	1.14	0	56	3.06	0	10	60	2.94	0	0	36	0.82	0	0	10	28	1	0.90730
8.5	1.14	0	56	3.06	0	10	60	2.94	0	0	36	0.82	0	1	6	28	1	0.90736
8.6	1.14	0	56	3.06	0	10	60	3.27	0	0	40	0.33	0	0	4	28	1	0.90777
8.7	1.14	0	56	3.22	0	11	59	2.61	0	0	32	0.98	0	0	12	28	1	0.90436
8.8	1.14	0	56	3.22	0	11	59	2.61	0	0	32	0.98	0	1	8	28	1	0.90430
8.9	1.14	0	56	3.22	0	11	59	2.61	0	0	32	0.98	0	2	4	28	1	0.90431
8.10	1.14	0	56	2.90	1	6	63	3.27	0	1	36	0.65	0	0	8	28	1	0.90907
8.11	1.14	0	56	2.90	1	6	63	3.27	0	3	28	0.65	0	0	8	28	1	0.90856
8.12	1.14	0	56	2.90	1	6	63	3.27	0	4	24	0.65	0	0	8	28	1	0.90870

continued on next page

Table 10 (continued)

ID	A_3	$F_3(12;4)$		A_4	$F_4(20;12;4)$			A_5	$F_5(24;16;8)$			A_6	$F_6(24;16;8)$			Rank	PEC_5	PIC_5
8.13	1.14	0	56	2.90	2	3	65	3.27	0	4	24	0.65	0	0	8	28	1	0.90707
9.1	1.71	0	84	5.51	0	18	108	6.61	0	0	81	2.45	0	0	30	28	1	0.90738
9.2	1.71	0	84	5.51	0	18	108	6.69	0	4	66	2.29	0	0	28	28	1	0.90735
9.3	1.71	0	84	5.51	0	18	108	6.61	0	9	45	2.45	0	0	30	28	1	0.90722
9.4	1.71	0	84	5.67	0	19	107	6.61	0	4	65	2.12	0	0	26	28	1	0.90564
9.5	1.71	0	84	5.84	0	20	106	5.88	0	4	56	2.78	0	0	34	28	1	0.90384
9.6	1.71	0	84	5.84	0	20	106	6.04	0	4	58	2.61	0	0	32	28	1	0.90382
9.7	1.71	0	84	5.84	0	20	106	6.37	0	5	58	1.96	0	0	24	28	1	0.90385
9.8	1.71	0	84	5.51	1	15	110	6.69	0	6	58	2.29	0	0	28	28	1	0.90648
9.9	1.71	0	84	5.51	2	12	112	6.69	0	8	50	2.29	0	0	28	28	1	0.90573
10.1	2.45	0	120	11.47	0	44	166	8.90	0	2	101	6.78	0	1	79	28	1	0.89314
10.2	2.45	0	120	11.47	0	44	166	9.31	0	2	106	6.37	0	1	74	28	1	0.89300
10.3	2.45	0	120	11.63	0	45	165	8.82	0	2	100	6.69	0	1	78	28	1	0.89198
10.4	2.45	0	120	11.80	0	46	164	8.33	0	0	102	7.35	0	3	78	28	1	0.89108
10.5	2.45	0	120	11.80	0	46	164	8.41	0	1	99	7.10	0	1	83	28	1	0.89097
10.6	2.45	0	120	11.80	0	46	164	8.65	0	1	102	6.37	0	0	78	28	1	0.89108
10.7	2.45	0	120	10.82	2	34	174	9.80	0	0	120	7.35	0	0	90	28	1	0.89642
10.8	2.45	0	120	10.82	2	34	174	9.80	0	0	120	7.35	0	4	74	28	1	0.89629
10.9	2.45	0	120	11.14	2	36	172	9.39	0	1	111	7.10	0	2	79	28	1	0.89415
10.10	2.45	0	120	10.49	6	20	184	11.43	0	16	76	5.06	0	0	62	28	1	0.89621
11.1	3.37	0	165	21.43	0	90	240	8.49	0	0	104	20.73	0	14	198	28	1	0.87942
11.2	3.37	0	165	21.43	0	90	240	11.10	0	0	136	16.00	0	0	196	28	1	0.87935
11.3	3.37	0	165	21.43	0	90	240	11.27	0	0	138	16.00	0	0	196	28	1	0.87937
11.4	3.37	0	165	18.82	3	65	262	14.86	0	0	182	14.69	0	6	156	28	1	0.88890
11.5	3.37	0	165	21.43	3	81	246	10.61	0	0	130	16.65	0	18	132	28	1	0.87818
11.6	3.37	0	165	18.82	6	56	268	14.86	0	0	182	14.69	0	6	156	28	1	0.88794
11.7	3.37	0	165	18.82	12	38	280	14.86	0	0	182	14.69	0	18	108	28	1	0.88606
12.1	4.49	0	220	32.14	0	135	360	19.76	0	0	242	32.00	0	0	392	28	1	0.87934
12.2	4.49	0	220	32.14	0	135	360	25.47	0	0	312	32.00	0	0	392	27	1	0.87913
12.3	4.49	0	220	28.22	6	93	396	25.47	0	0	312	29.39	0	12	312	28	1	0.88859
12.4	4.49	0	220	31.49	6	113	376	18.94	0	0	232	33.31	0	36	264	28	1	0.87961
12.5	4.49	0	220	29.20	15	72	408	24.98	0	0	306	26.45	0	0	324	28	1	0.88414

continued on next page

Table 10 (continued)

ID	A_3	$F_3(12;4)$		A_4	$F_4(20;12;4)$			A_5	$F_5(24;16;8)$			A_6	$F_6(24;16;8)$			Rank	PEC_5	PIC_5
12.6	4.49	0	220	28.22	18	57	420	25.47	0	0	312	29.39	0	36	216	28	1	0.88608
13.1	5.84	0	286	46.43	0	195	520	32.82	0	0	402	59.43	0	0	728	28	1	0.87933
13.2	5.84	0	286	46.43	0	195	520	41.39	0	0	507	59.43	0	0	728	27	1	0.87913
14.1	7.43	0	364	65.00	0	273	728	52.00	0	0	637	104.00	0	0	1274	28	1	0.87931
15.1	14.02	29	426	65.57	21	168	1176	100.33	0	30	1109	164.49	0	0	2015	28	0.99767	0.87134
15.2	14.18	30	425	65.57	21	168	1176	99.51	0	31	1095	164.49	0	0	2015	28	0.99600	0.86913
15.3	15.49	38	417	59.37	15	148	1202	108.00	0	36	1179	165.47	0	11	1983	28	0.99734	0.86892
15.4	15.49	38	417	59.37	15	148	1202	108.16	0	44	1149	164.98	0	9	1985	28	0.99667	0.86838
15.x	15.98	41	414	57.41	0	181	1184	108.16	0	12	1277	173.63	0	62	1879	28	1	0.87081
16.1	17.31	36	524	87.43	28	224	1568	145.63	0	44	1608	263.18	0	0	3224	28	0.99725	0.87075
16.2	19.10	47	513	79.92	21	199	1600	155.27	0	54	1686	263.51	0	15	3168	28	0.99679	0.86767
16.3	19.10	47	513	79.92	21	199	1600	155.27	0	56	1678	263.51	0	15	3168	28	0.99679	0.86769
16.4	19.10	47	513	79.92	21	199	1600	155.27	0	60	1662	263.51	0	15	3168	28	0.99679	0.86768
16.5	19.27	48	512	79.76	21	198	1601	154.45	0	50	1692	264.33	0	19	3162	28	0.99634	0.86675
16.6	19.27	48	512	79.76	21	198	1601	154.45	0	52	1684	264.33	0	19	3162	28	0.99611	0.86654
16.x	20.73	57	503	74.69	0	230	1590	156.41	0	28	1804	275.92	0	81	3056	28	1	0.86631
17.1	23.18	57	623	105.06	28	262	2090	218.20	0	76	2369	407.84	0	24	4900	28	0.99628	0.86704
17.2	23.51	59	621	105.06	28	262	2090	216.24	0	64	2393	407.84	0	24	4900	28	0.99612	0.86540
17.3	23.51	59	621	105.06	28	262	2090	216.24	0	72	2361	407.84	0	24	4900	28	0.99612	0.86549
17.4	23.51	59	621	105.06	28	262	2090	216.24	0	80	2329	407.84	0	24	4900	28	0.99548	0.86491
17.5	23.67	60	620	104.73	28	260	2092	215.84	0	84	2308	408.82	0	44	4832	28	0.99644	0.86503
17.x	25.96	74	606	96.57	0	294	2086	219.43	0	46	2504	427.10	0	117	4764	28	1	0.86342
18.1	29.71	80	736	128.73	28	322	2710	304.98	0	96	3352	623.35	0	60	7396	28	0.99568	0.86317
18.2	29.71	80	736	128.73	28	322	2710	304.98	0	96	3352	623.35	0	76	7332	28	0.99568	0.86318
18.3	29.71	80	736	128.73	28	322	2710	304.98	0	104	3320	623.35	0	76	7332	28	0.99568	0.86324
18.4	29.71	80	736	128.73	30	316	2714	304.98	0	96	3352	623.35	0	68	7364	28	0.99498	0.86278
18.5	29.88	81	735	128.57	28	321	2711	304.08	0	116	3261	624.08	0	88	7293	28	0.99615	0.86296
18.x	31.67	92	724	123.35	0	373	2687	302.37	0	68	3432	641.96	0	171	7180	28	1	0.86184
19.1	36.59	103	866	159.10	28	406	3442	413.39	0	128	4552	923.92	0	164	10662	28	0.99561	0.86097

continued on next page

Table 10 (continued)

ID	A_3	$F_3(12;4)$		A_4	$F_4(20;12;4)$			A_5	$F_5(24;16;8)$			A_6	$F_6(24;16;8)$			Rank	PEC_5	PIC_5
19.2	36.59	103	866	159.10	28	406	3442	413.39	0	130	4544	923.92	0	156	10694	28	0.99579	0.86107
19.3	36.59	103	866	159.10	28	406	3442	413.39	0	146	4480	923.92	0	148	10726	28	0.99561	0.86091
19.4	36.59	103	866	159.10	30	400	3446	413.39	0	120	4584	923.92	0	164	10662	28	0.99510	0.86057
19.5	36.59	103	866	159.10	30	400	3446	413.39	0	128	4552	923.92	0	148	10726	28	0.99510	0.86070
19.6	36.59	103	866	159.10	30	400	3446	413.39	0	136	4520	923.92	0	132	10790	28	0.99493	0.86037
19.7	36.76	104	865	158.78	28	404	3444	412.24	0	148	4458	926.20	0	162	10698	28	0.99596	0.86078
19.x	37.90	111	858	156.00	0	471	3405	409.47	0	97	4628	938.12	0	247	10504	28	1	0.86102
20.1	44.00	127	1013	196.84	30	510	4305	547.92	0	157	6084	1329.31	0	280	15164	28	0.99542	0.85923
20.2	44.00	127	1013	196.84	30	510	4305	547.92	0	177	6004	1329.31	0	248	15292	28	0.99542	0.85914
20.3	44.16	128	1012	196.02	12	559	4274	547.27	0	204	5888	1334.69	0	366	14886	28	0.99845	0.86129
20.4	44.16	128	1012	196.18	28	512	4305	547.10	0	180	5982	1333.71	0	283	15206	28	0.99607	0.85941
20.x	44.82	132	1008	195.04	0	589	4256	544.00	128	6152	356	1340.08	0	0	14992	28	1	0.86037
21.1	52.12	153	1177	240.67	34	624	5327	715.10	0	261	7716	1878.86	0	400	21416	28	0.99568	0.85815
21.2	52.12	153	1177	240.67	36	618	5331	715.10	0	241	7796	1878.86	0	384	21480	28	0.99548	0.85799
21.3	52.12	153	1177	240.67	36	618	5331	715.10	0	257	7732	1878.86	0	388	21464	28	0.99577	0.85829
21.4	52.29	154	1176	240.67	8	702	5275	713.80	0	227	7836	1878.53	0	499	21016	28	0.99833	0.86000
21.x	52.61	156	1174	240.18	0	723	5262	711.92	176	8017	480	1881.80	0	0	21132	28	1.00000	0.85980
22.1	60.98	181	1359	293.12	10	851	6454	920.98	0	292	10114	2591.10	0	702	28933	28	0.99848	0.85949
22.2	60.98	181	1359	293.12	11	848	6456	920.98	0	288	10130	2591.10	0	664	29085	28	0.99822	0.85940
22.3	60.98	181	1359	293.12	11	848	6456	920.98	0	290	10122	2591.10	0	672	29053	28	0.99829	0.85944
22.4	60.98	181	1359	293.12	11	848	6456	920.98	0	305	10062	2591.10	0	693	28969	28	0.99859	0.85957
22.5	60.98	181	1359	293.12	12	845	6458	920.98	0	285	10142	2591.10	0	693	28969	28	0.99806	0.85913
22.6	60.98	181	1359	293.12	12	845	6458	920.98	0	292	10114	2591.10	0	730	28821	28	0.99837	0.85943
22.x	61.31	183	1357	292.80	0	879	6436	918.45	226	10347	650	2593.63	0	0	29172	28	1	0.85919
23.1	70.59	211	1560	354.43	0	1064	7791	1172.65	0	291	13201	3509.14	0	879	39471	28	1.00000	0.85904
23.2	70.59	211	1560	354.43	7	1043	7805	1172.65	0	361	12921	3509.14	0	938	39235	28	0.99875	0.85906
23.3	70.59	211	1560	354.43	7	1043	7805	1172.65	0	370	12885	3509.14	0	940	39227	28	0.99863	0.85901
23.4	70.59	211	1560	354.43	7	1043	7805	1172.65	0	374	12869	3509.14	0	936	39243	28	0.99866	0.85903
23.5	70.59	211	1560	354.43	8	1040	7807	1172.65	0	354	12949	3509.14	0	937	39239	28	0.99845	0.85885
23.6	70.59	211	1560	354.43	8	1040	7807	1172.65	0	357	12937	3509.14	0	938	39235	28	0.99878	0.85916
23.7	70.59	211	1560	354.43	8	1040	7807	1172.65	0	358	12933	3509.14	0	922	39299	28	0.99837	0.85882

continued on next page

Table 10 (continued)

ID	A_3	$F_3(12;4)$		A_4	$F_4(20;12;4)$			A_5	$F_5(24;16;8)$			A_6	$F_6(24;16;8)$			Rank	PEC_5	PIC_5
24.1	80.82	242	1782	425.18	0	1276	9350	1479.84	0	368	16656	4680.16	0	1172	52644	28	1	0.85885
24.2	80.82	242	1782	425.18	10	1246	9370	1479.84	0	463	16276	4680.16	0	1251	52328	28	0.99847	0.85871
24.3	80.82	242	1782	425.18	11	1243	9372	1479.84	0	457	16300	4680.16	0	1267	52264	28	0.99885	0.85899
24.4	80.82	242	1782	425.18	11	1243	9372	1479.84	0	459	16292	4680.16	0	1264	52276	28	0.99882	0.85897
24.5	80.82	242	1782	425.18	11	1243	9372	1479.84	0	460	16288	4680.16	0	1259	52296	28	0.99864	0.85880
25.1	92.00	276	2024	506.00	0	1518	11132	1848.00	0	462	20790	6160.00	0	1540	69300	28	1.00000	0.85866
25.2	92.00	276	2024	506.00	14	1476	11160	1848.00	0	568	20366	6160.00	0	1662	68812	28	0.99864	0.85862
25.3	92.00	276	2024	506.00	14	1476	11160	1848.00	0	571	20354	6160.00	0	1659	68824	28	0.99872	0.85870
25.4	92.00	276	2024	506.00	14	1476	11160	1848.00	0	572	20350	6160.00	0	1658	68828	28	0.99870	0.85869
25.5	92.00	276	2024	506.00	15	1473	11162	1848.00	0	562	20390	6160.00	0	1650	68860	28	0.99881	0.85872
25.6	92.00	276	2024	506.00	15	1473	11162	1848.00	0	563	20386	6160.00	0	1649	68864	28	0.99866	0.85866
25.7	92.00	276	2024	506.00	15	1473	11162	1848.00	0	569	20362	6160.00	0	1673	68768	28	0.99868	0.85866
26.1	104.00	312	2288	598.00	0	1794	13156	2288.00	0	572	25740	8008.00	0	2002	90090	28	1	0.85866
26.2	104.00	312	2288	598.00	18	1740	13192	2288.00	0	704	25212	8008.00	0	2158	89466	28	0.99863	0.85863
26.3	104.00	312	2288	598.00	20	1734	13196	2288.00	0	692	25260	8008.00	0	2122	89610	28	0.99859	0.85858
27.1	117.00	351	2574	702.00	0	2106	15444	2808.00	0	702	31590	10296.00	0	2574	115830	28	1.00000	0.85866
27.2	117.00	351	2574	702.00	24	2034	15492	2808.00	0	846	31014	10296.00	0	2718	115254	28	0.99855	0.85855
27.3	117.00	351	2574	702.00	27	2025	15498	2808.00	0	864	30942	10296.00	0	2736	115182	28	0.99900	0.85863

References

Angelopoulos, P., Evangalaras, H., Koukouvinos, C., and Lappas, E. (2007). An effective step-down algorithm for the construction and the identification of nonisomorphic orthogonal arrays. *Metrika*, 66:139–149.

Belcher-Novosad, S. and Ingram, D. (2003). Identifying minimum G-aberration designs from Hadamard matrices of order 28. *Journal of the Arkansas Academy of Sciences*, 57:202–205.

Bulutoglu, D. A. and Margot, F. (2008). Classification of orthogonal arrays by integer programming. *Journal of Statistical Planning and Inference*, 138:654–666.

Bulutoglu, D. A. and Ryan, K. J. (2015). Algorithms for finding generalized minimum aberration designs. *Journal of Complexity*, 31:577–589.

Deng, L. Y. and Tang, B. (1999). Generalized resolution and minimum aberration criteria for Plackett-Burman and other nonregular factorial designs. *Statistica Sinica*, 9:1071–1082.

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

Dürig, T. and Fassihi, A. R. (1993). Identification of stabilizing and destabilizing effects of excipient-drug interactions in solid dosage form design. *International Journal of Pharmaceutics*, 97:161–170.

El Ati-Hellal, M., Hellal, F., Dachraoui, M., and Hedhili, A. (2007). Plackett–Burman designs in the pretreatment of macroalgae for Pb, Cr and Al determination by GF-AAS. *C. R. Chimie*, 10:839–849.

Evangelaras, H., Georgiou, S., and Koukouvinos, C. (2004). Evaluation of inequivalent projections of Hadamard matrices of order 24. *Metrika*, 59:51–73.

Evangelaras, H., Koukouvinos, C., and Lappas, E. (2007). Further contributions to nonisomorphic two level orthogonal arrays. *Journal of Statistical Planning and Inference*, 137:2080–2086.

Hedayat, A. S., Sloane, N. J. A., and Stufken, J. (1999). *Orthogonal arrays: Theory and Applications.* Springer, New York NY, USA.

Ingram, D. and Tang, B. (2005). Minimum G aberration design construction and design tables for 24 runs. *Journal of Quality Technology*, 37:101–114.

Kimura, H. (1994a). Characterization of Hadamard matrices of order 28 with Hall sets. *Discrete Mathematics*, 128:257–268.

Kimura, H. (1994b). Classification of Hadamard matrices of order 28. *Discrete Mathematics*, 133:171–180.

Kimura, H. and Ohmori, O. (1986). Construction of Hadamard matrices of order 28. *Graphs and Combinatorics*, 2:247–257.

Li, W. and Nachtsheim, C. J. (2000). Model-robust factorial designs. *Technometrics*, 42:345–352.

Li, Y., Deng, L. Y., and Tang, B. (2004). Design catalog based on minimum G-aberration. *Journal of Statistical Planning and Inference*, 124:219–230.

Lin, D. K. J. and Draper, N. R. (1993). Generating alias relationships for two-level Plackett and Burman designs. *Computational Statistics and Data Analysis*, 15:147–157.

Loeppky, J. L., Sitter, R. R., and Tang, B. (2007). Nonregular designs with desirable projection properties. *Technometrics*, 49:454–467.

Mee, R. W. (2009). *A comprehensive guide to factorial two-level experimentation.* Springer-Verlag, New York, NY, USA.

Montgomery, D. C. (2009). *Design and analysis of experiments.* 7th edition, Wiley, New York.

Plackett, R. L. and Burman, J. P. (1946). The design of optimum multifactorial experiments. *Biometrika*, 33:305–325.

Rao, C. R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *Journal of the Royal Statistical Society Supplement*, 9:128–139.

Schoen, E. D., Eendebak, P. T., and Nguyen, M. V. M. (2010). Complete enumeration of pure-level and mixed-level orthogonal arrays. *Journal of Combinatorial Designs*, 18:123–140.

Schoen, E. D. and Mee, R. W. (2012). Two-level designs of strength 3 and up to 48 runs. *Journal of the Royal Statistical Society Series C*, 61:163–174.

Sloane, N. J. A. (2015a). A library of Hadamard matrices. http://www. neilsloane.com/Hadamard/.

Sloane, N. J. A. (2015b). The on-line encyclopedia of integer sequences. http://oeis.org.

Sun, D. X., Li, W., and Ye, K. Q. (2008). An algorithm for sequentially constructing nonisomorphic orthogonal designs and its applications. *Statistics and Application*, 6:144–158.

Tang, B. and Deng, L. Y. (1999). Minimum G_2-aberration for nonregular fractional factorial designs. *Annals of Statistics*, 27:1914–1926.

Wu, C. F. J. and Hamada, M. S. (2009). *Experiments: Planning, analysis, and parameter design optimization.* 2nd edition, Wiley, New York, NY, USA.

Yamamoto, S., Fujii, Y., Hyodo, Y., and Yumiba, H. (1995). Classification of orthogonal 24-run 2^5 factorial designs derivable from saturated two-symbol orthogonal arrays of strenth 2, size 24 and index 6. *SUT Journal of Mathematics*, 31:39–53.

Yumiba, H., Hyodo, Y., and Yamamoto, S. (1997). Classification of two-symbol orthogonal arrays of size 24, strength 2, 6 constraints and index 6 derivable from saturated orthogonal arrays having 23 constraints. *SUT Journal of Mathematics*, 33:47–63.