Chapter 9

Nonregular factorial and supersaturated designs

Hongquan Xu, Department of Statistics, University of California, Los Angeles, CA 90095, U.S.A. Email: hqxu@stat.ucla.edu (April 23, 2014)

9.1 Introduction

Fractional factorial designs are classified into two broad types: *regular* designs and *nonregular* designs. Regular designs are constructed through defining relations among factors; they are introduced in Chapter 1 (Section 1.7) and fully described in Chapter 7. These designs have a simple alias structure in that any two factorial contrasts are either orthogonal or fully aliased. The run sizes are always a power of two, three or another prime, and thus the "gaps" between possible run sizes increase exponentially as the power increases.

Plackett and Burman (1946) first gave a large collection of two-level nonregular designs whose run sizes are not a power of two. These designs are often referred to as the Plackett-Burman designs in the literature, and belong to a wide class of orthogonal arrays (Rao 1947).

Plackett-Burman designs and other nonregular designs are widely used in various screening experiments for their run size economy and flexibility (Wu and Hamada (2009)). They fill the gaps between regular designs in terms of various run sizes and are flexible in accommodating various combinations of factors with different numbers of levels. Unlike regular designs, nonregular designs may exhibit complex alias structures, that is, a large number of effects may be neither orthogonal nor fully aliased, which makes it difficult to identify and interpret significant effects. For this reason, nonregular designs have traditionally been used to estimate factor main effects only but not their interactions. However, in many practical situations it is often questionable whether the interaction effects are negligible.

Hamada and Wu (1992) went beyond the traditional approach and proposed an analysis strategy for nonregular designs in which some interactions could be entertained and estimated through their complex alias structure. They pointed out that ignoring interactions can lead to (i) important effects being missed, (ii) unimportant effects being erroneously detected, and (iii) estimated effects having reversed signs resulting in incorrectly recommended factor levels. Their pioneering work motivated the recent studies in design properties, optimality criteria, construction, and analysis of nonregular designs.

Supersaturated designs are factorial designs whose run sizes are too small to allow estimation of all factorial effects of interest. They have become increasingly popular in the last two decades because of their potential for reducing the number of runs. Broadly speaking, supersaturated designs are a special class of nonregular designs; some of the optimality criteria and results developed for nonregular designs can easily be extended to the setting of supersaturated designs. Since supersaturated designs are typically used to estimate main effects only, the problems associated with supersaturated designs are relatively simpler than those for other nonregular designs so in this chapter we emphasize nonregular fractions.

The remainder of the introduction gives some basic concepts and definitions. An *orthogonal array* (OA) of n runs, k factors, s levels and strength t, denoted by $OA(n, s^k, t)$, is an $n \times k$ matrix in which each column has s symbols or levels and for any t columns all possible s^t combinations of symbols appear equally often as a row in the $n \times t$ subarray. A regular s^{k-q} design of resolution r is an $OA(n = s^{k-q}, s^k, t = r - 1)$, but not every OA with these

parameters is a regular design. Further let $OA(n, s_1^{k_1} \times \cdots s_m^{k_m}, t)$ denote a mixed-level OA of strength t with k_i columns of s_i levels for $i = 1, \ldots, m$. Hedayat, Sloane and Stufken (1999) gave a comprehensive account of theory and applications of OAs.

OAs of strength 2, such as Plackett-Burman designs, allow all the main effects to be estimated independently and they are universally optimal for the main effects model (Cheng 1980). A necessary condition for the existence of an $OA(n, s_1^{k_1} \times \cdots \times s_m^{k_m}, 2)$ is that $n - 1 \geq \sum_{i=1}^m k_i(s_i - 1)$. A design is called saturated if $n - 1 = \sum_{i=1}^m k_i(s_i - 1)$ and supersaturated if $n - 1 < \sum_{i=1}^m k_i(s_i - 1)$. In the literature, OAs of strength 2 are also called orthogonal designs or OAs without mentioning the strength explicitly. For convenience, an OA of strength 1 is also called a *balanced design*, where every level appears equally often for each factor.

9.2 Examples of Nonregular Designs

Example 1. Consider an experiment reported by Vander Heyden et al. (1999) who used the high-performance liquid chromatography (HPLC) method to study the assay of ridogrel and its related compounds in ridogrel oral film-coated tablet simulations. The researcher used a 12-run Plackett-Burman design to evaluate the importance of eight factors on several responses. One response was the percentage recovery of main compound, ridogrel. The eight factors were pH of the buffer (A), column manufacturer (B), column temperature (D), percent of organic solvent in the mobile phase at the start of the gradient (E), percent of organic solvent in the mobile phase at the end of the gradient (F), flow of the mobile phase (H), detection wavelength (I), and concentration of the buffer (J). Table 9.1 gives the design matrix and the observed data. Fitting a main effects model, we get

$$\hat{y} = 101.04 + 0.34A - 0.22B - 0.36D - 0.56E + 0.44F - 0.01H + 0.26I - 0.31J, \quad (9.1)$$

where each factor has two levels coded as $+1$ and -1 for $+$ and $-$, respectively. This model has $R^2 = 0.78$ with $\hat{\sigma} = 1.045$ on 3 degrees of freedom. The most significant factors are E and F with p-values of 0.16 and 0.24, respectively. The researchers concluded there was no significant relationship between any of the factors and this response because none of the

Table 9.1: Design and Data for the HPLC Experiment

Run	A	B	C	D	E	F	G	H	I	J	K	MC
1	+	+	+	−	+	+	−	+	−	−	−	101.6
2	+	+	−	+	−	−	−	+	+	+	−	101.7
3	+	−	+	+	−	+	−	−	−	+	+	101.6
4	+	−	−	−	+	+	+	−	+	+	−	101.9
5	+	−	+	−	−	−	+	+	+	−	+	101.8
6	−	+	+	+	−	+	+	−	+	−	−	101.1
7	−	+	−	−	−	+	+	+	−	+	+	101.1
8	−	−	−	+	+	+	−	+	+	−	+	101.6
9	−	−	+	+	+	−	+	+	−	+	−	98.4
10	−	+	+	−	+	−	−	−	+	+	+	99.7
11	+	+	−	+	+	−	+	−	−	−	+	99.7
12	−	−	−	−	−	−	−	−	−	−	−	102.3

Note: Columns C, G and K were not used in the experiment.

effects are significant at the 10% level.

For Plackett-Burman designs, main effects are partially aliased with two-factor interactions (2fi's). Non-negligible 2fi's could bias the estimates of the main effects. Phoa, Wong and Xu (2009) reanalyzed the data and found one very significant interaction. The interaction EF is more significant than the main effects E and F. They found the following model

$$\hat{y} = 101.04 - 0.56E + 0.44F - 0.30H + 0.88EF, \tag{9.2}$$

where E and F are the percentages of organic solvent in the mobile phase at the start and the end of the gradient, respectively, and H is the flow of the mobile phase. This model has $R^2 = 0.96$, indicating a good fit. In the model (9.2), H is significant at the 5% level (p-value=0.012) and E, F and EF are significant at the 1% level.

Example 2. Consider an experiment reported by Groten et al. (1996, 1997) who performed a 4-week oral/inhalatory study in which the toxicity of combinations of nine compounds was examined in male Wistar rats. The nine compounds were formaldehyde (A), dichloromethane (B), aspirin (C), cadmium chloride (D), stannous chloride (E), loperamide (F), spermine (G), butylated hydroxyanisole (H), and di-ethylhexyl (J). Their experiment used a regular 2^{9-5} design with design generators $E = ABCD, F = AD, G = AE, H = AC, J = AB$. For each factor, the low level corresponds to no compound. One of the responses measured was aspartate aminotransferase (ASAT) activity. The design and data are given in Table 9.2.

Table 9.2: Design and data for the chemical toxicity experiment

Run	A^*	A	B	C	D	E	F	G	H	J	ASAT
1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	70
2	1	−1	1	1	−1	1	1	−1	−1	−1	71
3	−1	−1	1	1	1	−1	−1	1	−1	−1	86
4	1	1	−1	−1	1	1	1	1	−1	−1	75
5	−1	−1	1	−1	−1	−1	1	1	1	−1	65
6	1	1	−1	1	−1	1	−1	1	1	−1	70
7	−1	1	−1	1	1	−1	1	−1	1	−1	96
8	−1	−1	1	−1	1	1	−1	−1	1	−1	65
9	−1	−1	−1	1	−1	−1	1	1	−1	1	77
10	−1	1	1	−1	−1	1	−1	1	−1	1	71
11	1	1	1	−1	1	−1	1	−1	−1	1	88
12	−1	−1	−1	1	1	1	−1	−1	−1	1	80
13	1	1	1	1	−1	−1	−1	−1	1	1	68
14	−1	−1	−1	−1	−1	1	1	−1	1	1	69
15	1	−1	−1	−1	1	−1	−1	1	1	1	72
16	1	1	1	1	1	1	1	1	1	1	82

Note: Columns A–J form a regular 2^{9-5} design; columns A^* and B–J form a nonregular 2^{9-5} design.

Ignoring three-factor or higher-order interactions, the alias relations among main effects and 2fi's are $A = BJ = CH = DF = EG$, $B = AJ, C = AH, D = AF, E = AG, F = AD, G = AE, H = AC, J = AB$, $BC = DG = EF = HJ, BD = CG = EH = FJ, BE = CF = DH = GJ$, $BF = CE = DJ = GH, BG = CD = EJ = FH, BH = CJ = DE = FG$. The researchers believed that formaldehyde (A) did not interact with other compounds, so the main effects of B–J can be estimated with confidence. The estimate of A is fully aliased with four 2fi's.

Groten et al. (1996, 1997) first analyzed the main effects, and then analyzed the significant main effects together with their 2fi's in a subsequent analysis. They concluded that C, D, E, F, DE and DF were significant effects and obtained the following model

$$\hat{y} = 75.31 + 3.44C + 5.19D - 2.44E + 2.56F - 2.56DE + 2.19DF.$$

Recall that DF is fully aliased with A. When two effects are fully aliased in a regular design, it is impossible to distinguish between them based on the data only in the analysis. When a 2fi is aliased with a main effect, we often assume the 2fi is negligible. However, Groten et

al. (1996, 1997) ignored the main effect of formaldehyde (A) and concluded that the DF interaction was significant in their model, based on their expert opinion. This contradicts the conventional statistical practice. This problem could be avoided if a nonregular design had been used.

Phoa, Xu and Wong (2009) demonstrated that we could estimate A and DF simultaneously if a nonregular design had been used. Consider a new design in which formaldehyde has level settings as column A^* and all other factors have the same level settings as before. Column A^* differs from column A in the runs 2, 7, 10 and 15, where the high and low levels are switched. For the new design, A^* is not fully aliased with any 2fi, but is partially aliased with 16 2fi's with correlation ± 0.5. Here, and throughout this chapter, correlation is defined as $1/n$ times the inner product of the columns corresponding to the factorial effects. For example, the correlation between A^* and CD is -0.5 while the correlation between A^* and DF is 0.5. Since A^* is not fully aliased with any 2fi's, we can separate A^* and DF. One problem is that we do not have data from the new design. If formaldehyde (A) were negligible as Groten et al. (1996, 1997) suggested, the changes we made in the levels for formaldehyde (A) would not affect the responses; so it is reasonable to use the same data for the new design. Then we can estimate A^* and DF simultaneously, yielding the following fitted model:

$$\hat{y} = 75.31 + 3.44C + 5.19D - 2.44E + 2.56F - 2.56DE + 3.46DF - 2.54A^* - 1.94H,$$

where all effects are significant at the 5% level. The model has $R^2 = 0.96$. The estimates of A^* and DF have opposite signs and the estimate of DF becomes larger (3.46 vs 2.19) in the modified design. One possible interpretation is that formaldehyde (A) was important and the effect of DF could be underestimated in the original design when the main effect of A was ignored. For more discussions on this experiment see Phoa, Xu and Wong (2009).

The advantage of the nonregular design is that it is possible to estimate partially aliased effects without adding extra runs. The disadvantage is that the analysis becomes more complicated.

Example 3. Ding et al. (2013) reported an experiment studying a system with Herpes sim-

Table 9.3: Design and data of the antiviral drug experiment

		A	B	C	D	E		Readout	
Run	1	2	3	4	5	6	7	Replicate1	Replicate2
1	0	0	0	0	0	0	0	78.6	81.9
2	0	1	1	1	1	1	1	13.3	16.7
3	0	2	2	2	2	2	2	3.4	3.8
4	1	0	0	1	1	2	2	21.4	25.2
5	1	1	1	2	2	0	0	8.6	4.4
6	1	2	2	0	0	1	1	18.0	27.3
7	2	0	1	0	2	1	2	7.3	2.4
8	2	1	2	1	0	2	0	17.9	23.7
9	2	2	0	2	1	0	1	52.9	54.3
10	0	0	2	2	1	1	0	13.2	8.8
11	0	1	0	0	2	2	1	2.1	4.5
12	0	2	1	1	0	0	2	73.4	73.9
13	1	0	1	2	0	2	1	19.6	14.6
14	1	1	2	0	1	0	2	59.1	41.7
15	1	2	0	1	2	1	0	1.4	2.6
16	2	0	2	1	2	0	1	7.3	4.8
17	2	1	0	2	0	1	2	22.3	24.0
18	2	2	1	0	1	2	0	14.1	18.3

plex virus type 1 (HSV-1) and five antiviral drugs, namely, Interferon-alpha (A), Interferon-beta (B), Interferon-gamma (C), Ribavirin (D), and Acyclovir (E). Their original experiment used a composite design that consists of a 16-run factorial design with 2 levels and an 18-run OA with 3 levels. Two researchers conducted the experiment independently with different random orders, yielding two replicates. Here we look at the 18-run OA only, which corresponds to columns 2–6 of the commonly used $OA(18, 3^7, 2)$; see Table 9.3. For each drug, the 3 levels 0, 1, and 2 correspond to no drug, intermediate drug dosage, and high drug dosage, respectively. The observed data, readout, were the percentage of infected cells after the combination drug treatment.

Following Ding et al. (2013), we use the square root of readout as response. Since this is an $OA(18, 3^5, 2)$, we can fit both linear and quadratic main effects. A main effects model identifies D and E as the most significant drugs, both their linear and quadratic effects are significant. With only 18 runs, we do not have sufficient degrees of freedom to estimate all the interactions among five drugs. Nevertheless, we can perform stepwise variable selection. We find two significant bilinear effects DE and AC. The results are similar to those obtained by Ding et al. (2013) using the entire 34-run composite design.

9.3 Alias Structure

For regular designs, we can find the alias relationships among factorial effects from defining relations. For nonregular designs, the alias structure is complicated and we need to use the general regression method. Suppose we fit a model

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\theta}_1 + \boldsymbol{\epsilon}, \tag{9.3}$$

where $\mathbf{y}$ is an $n \times 1$ vector of the responses, $\mathbf{X}_1$ is an $n \times p_1$ matrix corresponding to the fitted model, $\boldsymbol{\theta}_1$ is a $p_1 \times 1$ vector of the model parameters, and $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of normal errors. The least squares estimator of $\boldsymbol{\theta}_1$ is

$$\hat{\boldsymbol{\theta}}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}$$

which is unbiased under the model (9.3). However, if the **true** model is

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\theta}_1 + \mathbf{X}_2\boldsymbol{\theta}_2 + \boldsymbol{\epsilon}, \tag{9.4}$$

where $\mathbf{X}_2$ is an $n \times p_2$ matrix corresponding to the additional variables that are not in the fitted model and $\boldsymbol{\theta}_2$ is a $p_2 \times 1$ vector of the additional model parameters. It is easy to show that

$$E(\hat{\boldsymbol{\theta}}_1) = \boldsymbol{\theta}_1 + \mathbf{C}\boldsymbol{\theta}_2,$$

where $\mathbf{C} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$ is the **alias matrix**; see Chapter 7 (Section 7.2.2), Box and Draper (1987), and Wu and Hamada (2009).

Example 4. Consider the HPLC experiment in Example 1. Suppose the model (9.3) contains the intercept and 8 main effects, and the model (9.4) contains all $\binom{8}{2} = 28$ 2fi's besides the main effects. The matrix $\mathbf{X}_1$ is a 12×9 matrix and $\mathbf{X}_2$ is a 12×28 matrix. The alias matrix $\mathbf{C} = 12^{-1}\mathbf{X}_1'\mathbf{X}_2$ is a 9×28 matrix whose elements are $-1/3$, 0, or $1/3$. Except for the first row, each row represents a factor and has exactly seven 0's, corresponding to the seven 2fi's involving the factor. Each main effect is partially aliased with $\binom{7}{2} = 21$ 2fi's (that

do not involve the factor itself) with correlation $1/3$ or $-1/3$. For example, we have

$$\begin{aligned} E(\hat{H}) &= H + \frac{1}{3}(AB - AD - AE - AF + AI - AJ - BD - BE + BF - BI \\ &+ BJ + DE - DF + DI + DJ + EF - EI - EJ - FI - FJ - IJ). \end{aligned} \quad (9.5)$$

If EF is the only significant interaction, this simplifies to $E(\hat{H}) = H + \frac{1}{3}EF$. Under the model (9.2), $H + \frac{1}{3}EF$ is estimated as $-0.30 + \frac{1}{3}(0.88) \approx -0.01$, which agrees well with the estimate of H in the main effects model (9.1). The main effect of H is not significant in the main effects model because it is partially canceled by the significant EF interaction.

Example 5. Consider the chemical toxicity experiment in Example 2. We again look at the alias relations between the main effects and 2fi's. In this case, $\mathbf{X}_1$ is a 16×10 matrix and $\mathbf{X}_2$ is a 16×36 matrix. It is easy to see that $\mathbf{X}_1'\mathbf{X}_1 = 16\mathbf{I}$ so the alias matrix $\mathbf{C} = 16^{-1}\mathbf{X}_1'\mathbf{X}_2$ is a 10×36 matrix. For the regular design with column A, we have $E(\hat{A}) = A + BJ + CH + DF + EG$, $E(\hat{B}) = B + AJ$, etc. These are indeed the same as the alias relations among the main effects and 2fi's. For the nonregular design with column A^*, we have $E(\hat{A}^*) = A^* + 0.5(BC + DG + EF + HJ + BF + CE + DJ + GH - BG - CD - EJ - FH + BJ + CH + DF + EG)$, $E(\hat{B}) = B + 0.5(A^*C + A^*F - A^*G + A^*J)$, etc. For the regular design, A is fully aliased with four 2fi's with correlation 1, and the other eight main effects are each fully aliased with one 2fi (involving factor A) with correlation 1. For the nonregular design, A^* is partially aliased with 16 2fi's with correlation ± 0.5, and the other eight main effects are each partially aliased with four 2fi's (involving factor A^*) with correlation ± 0.5.

9.4 Optimality Criteria

The main objective of the construction of optimal nonregular designs is to minimize the aliasing of higher-order interactions on the main effects. The minimum aberration (MA) criterion, defined in Chapter 1 (Section 1.7.2) and Chapter 7 (Section 7.2.4), is the standard criterion for comparing regular designs. This criterion can be extended to nonregular designs. This section introduces some of these extensions.

9.4.1 Generalized minimum aberration

For a factorial design $\mathbf{D}$ with n runs and k factors, the full ANOVA model is

$$\mathbf{y} = \mathbf{X}_0\theta_0 + \mathbf{X}_1\boldsymbol{\theta}_1 + \cdots + \mathbf{X}_k\boldsymbol{\theta}_k + \boldsymbol{\epsilon}, \tag{9.6}$$

where $\mathbf{y}$ is the vector of n observations, θ_0 the general mean, $\boldsymbol{\theta}_j$ the vector of jth-order factorial effects, $\mathbf{X}_0$ the vector of 1's, $\mathbf{X}_j$ the matrix of contrast coefficients for $\boldsymbol{\theta}_j$, and $\boldsymbol{\epsilon}$ the vector of independent random errors. Note that jth-order factorial effects represent main effects when $j = 1$ and interactions when $j \geq 2$. Here we consider only the cases where the contrast coefficient of a j-factor interaction is the product of its corresponding contrast coefficients of j main effects. As in Xu and Wu (2001), the main effect contrasts are normalized so that they have the same length $\sqrt{n}$ for a balanced design. In particular, for a two-level factor, the contrast vector of its main effect is coded as $(-1, 1)$; for a three-level factor, the contrast vectors of the linear and quadratic main effects are coded as $(-\sqrt{3/2}, 0, \sqrt{3/2})$ and $(1/\sqrt{2}, -\sqrt{2}, 1/\sqrt{2})$, respectively.

For $j = 1, \ldots, k$, Xu and Wu (2001) defined A_j, a function of $\mathbf{X}_j$, to measure the overall aliasing between all jth-order factorial effects and the general mean. Specifically, let $\mathbf{X}_j = [x_{il}^{(j)}]$ and define

$$A_j(\mathbf{D}) = n^{-2}\mathbf{1}'\mathbf{X}_j\mathbf{X}_j'\mathbf{1} = n^{-2}\sum_{l=1}^{n_j}\left(\sum_{i=1}^{n} x_{il}^{(j)}\right)^2, \tag{9.7}$$

where $\mathbf{1}$ is the $n \times 1$ vector of ones and n_j is the number of all jth-order factorial effects. The value of A_j is independent of the choice of the orthonormal contrasts used. The vector $(A_1, \ldots, A_k)$ is called the *generalized wordlength pattern*, because for a 2-level regular design, A_j is the number of words of length j. The generalized minimum aberration (GMA) criterion (Xu and Wu 2001) is to sequentially minimize $A_1, A_2, A_3, \ldots$. A design that does this is said to have GMA.

Example 6. Consider two 2-level designs with 4 runs and 3 factors in Table 9.4. The first design $\mathbf{D}_1$ is a regular 2^{3-1} design with defining relation $I = ABC$ and the second design $\mathbf{D}_2$ is called a one-factor-at-a-time design. Table 9.4 also shows elements of the corresponding $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ matrices. For $\mathbf{D}_1$, we have $A_1 = (0^2+0^2+0^2)/4^2 = 0$, $A_2 = (0^2+0^2+0^2)/4^2 = 0$,

Table 9.4: Two 2-level designs with 4 runs and 3 factors

(a) $\mathbf{D}_1$: a regular 2^{3-1} design

	$\mathbf{X}_1$			$\mathbf{X}_2$			$\mathbf{X}_3$
Run	A	B	C	AB	AC	BC	ABC
1	+	+	+	+	+	+	+
2	+	−	−	−	−	+	+
3	−	+	−	−	+	−	+
4	−	−	+	+	−	−	+
Sum	0	0	0	0	0	0	4

(b) $\mathbf{D}_2$: one-factor-at-a-time design

	$\mathbf{X}_1$			$\mathbf{X}_2$			$\mathbf{X}_3$
Run	A	B	C	AB	AC	BC	ABC
1	+	+	+	+	+	+	+
2	−	+	+	−	−	+	−
3	−	−	+	+	−	−	+
4	−	−	−	+	+	+	−
Sum	−2	0	2	2	0	2	0

Table 9.5: A 3^{3-1} design

	$\mathbf{X}_1$			$\mathbf{X}_2$									$\mathbf{X}_3$							
$A\,B\,C$	A		B		C		$A\times B$				$A\times C$				$B\times C$				$A\times B\times C$							
0 0 0	−1	1	−1	1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	−1	1	1	−1	1	−1	−1	1
0 1 1	−1	1	0	−2	0	−2	0	2	0	−2	0	2	0	−2	0	0	0	4	0	0	0	−4	0	0	0	4
0 2 2	−1	1	1	1	1	1	−1	−1	1	1	−1	−1	1	1	1	1	1	1	−1	−1	−1	−1	1	1	1	1
1 0 1	0	−2	−1	1	0	−2	0	0	2	−2	0	0	0	4	0	2	0	−2	0	0	0	0	0	−4	0	4
1 1 2	0	−2	0	−2	1	1	0	0	0	4	0	0	−2	−2	0	0	−2	−2	0	0	0	0	0	0	4	4
1 2 0	0	−2	1	1	−1	1	0	0	−2	−2	0	0	2	−2	−1	1	−1	1	0	0	0	0	2	−2	2	−2
2 0 2	1	1	−1	1	1	1	−1	1	−1	1	1	1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1
2 1 0	1	1	0	−2	−1	1	0	−2	0	−2	−1	1	−1	1	0	0	2	−2	0	0	2	−2	0	0	2	−2
2 2 1	1	1	1	1	0	−2	1	1	1	1	0	−2	0	−2	0	−2	0	−2	0	−2	0	−2	0	−2	0	−2
Sum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−3	−3	3	−9	3	−9	9	9
Scale	a	b	a	b	a	b	a^2	$a\,b$	$a\,b$	b^2	a^2	$a\,b$	$a\,b$	b^2	a^2	$a\,b$	$a\,b$	b^2	a^3	$a^2\,b$	$a^2\,b$	$a\,b^2$	$a^2\,b$	$a\,b^2$	$a\,b^2$	b^3

$a = \sqrt{3/2}$ and $b = 1/\sqrt{2}$

and $A_3 = 4^2/4^2 = 1$. For $\mathbf{D}_2$, we have $A_1 = [(-2)^2+0^2+2^2]/4^2 = 0.5$, $A_2 = [2^2+0^2+2^2]/4^2 = 0.5$, and $A_3 = 0^2/4^2 = 0$. Since $A_1(\mathbf{D}_1) < A_1(\mathbf{D}_2)$, $\mathbf{D}_1$ has less aberration than $\mathbf{D}_2$; so the regular design $\mathbf{D}_1$ is preferred to the one-factor-at-a-time design $\mathbf{D}_2$ with respect to the GMA criterion. This agrees with the well known result that factorial designs are more efficient than one-factor-at-a-time designs.

Example 7. Consider the 3-level design given in Table 9.5. This design is a regular 3^{3-1} design with defining relation $I = ABC^2$ and has one word of length 3, i.e., ABC^2; for details of 3-level regular designs see Chapter 1 (Section 1.7.1) and Chapter 7 (Section 7.10). Table 9.5 also shows the orthogonal polynomial contrasts. From the definition (9.7), we have

$A_1(\mathbf{D}) = A_2(\mathbf{D}) = 0$ and

$$\begin{aligned} A_3(\mathbf{D}) &= [(-3a^3)^2 + (-3a^2b)^2 + (3a^2b)^2 + (-9ab^2)^2 \\ &+ (3a^2b)^2 + (-9ab^2)^2 + (9ab^2)^2 + (9b^3)^2]/9^2, \end{aligned}$$

with $a = \sqrt{3/2}$ and $b = 1/\sqrt{2}$. This simplifies to $A_3(\mathbf{D}) = 2$.

For a regular 2-level design, as in Example 6, the generalized wordlength pattern is the same as the wordlength pattern defined in Chapter 7, Section 7.2.4. For a regular s-level design, $A_j(\mathbf{D})$ defined in (9.7) is the total degrees of freedom associated with words of length j in the generating relation, that is, $A_j(\mathbf{D})$ is $s-1$ times the number of words of length j. Example 7 illustrates this for $s = 3$. Hence GMA reduces to minimum aberration for regular designs. The *minimum G_2-aberration* criterion, proposed by Tang and Deng (1999), for 2-level designs is a special case of the GMA criterion.

Suppose $\mathbf{D}$ is a balanced design, i.e., an OA of strength 1. For each factor, each symbol appears the same number of times. It is easy to see that $\mathbf{X}_1'\mathbf{1}$ is a vector of 0s so $A_1(\mathbf{D}) = 0$. Let $\mathbf{X}_1'\mathbf{X}_1 = (a_{ij})$. Then a_{ij}/n is the correlation between the ith and jth columns of $\mathbf{X}_1$. Since each contrast coefficient of a 2-factor interaction is the product of its corresponding two main effect contrast coefficients, $\mathbf{X}_2'\mathbf{1}$ is a column vector whose elements are a_{ij}, so $A_2(\mathbf{D}) = n^{-2}\mathbf{1}'\mathbf{X}_2\mathbf{X}_2'\mathbf{1} = n^{-2}\sum_{i<j} a_{ij}^2$. That is, $A_2(\mathbf{D})$ measures the overall aliasing among all possible main effects. For an OA of strength 2, $A_2(\mathbf{D}) = 0$; the reverse is also true. Xu and Wu (2001) showed the following important property regarding the generalized wordlength pattern.

Theorem 1. *A design $\mathbf{D}$ is an OA of strength t if and only if $A_j(\mathbf{D}) = 0$ for $1 \le j \le t$.*

Therefore, following the GMA criterion, among all possible designs, we shall choose balanced designs and among them choose orthogonal designs with maximum strength.

Example 8. Consider choosing five columns from the commonly used $OA(18, 3^7, 2)$ given in Table 9.3. There are 21 possible choices. For illustration, consider three choices. Let $\mathbf{D}_1$ be the design formed by columns 2 to 6, $\mathbf{D}_2$ be the design formed by columns 1, 4–7, and $\mathbf{D}_3$ be the design formed by columns 1–5, respectively. The generalized wordlength patterns

for the three designs are $(0, 0, 5, 7.5, 0)$, $(0, 0, 6.5, 4.5, 1.5)$, and $(0, 0, 7, 3.5, 2)$, respectively. Hence, $\mathbf{D}_1$ is the best according to the GMA criterion. It can be easily verified that $\mathbf{D}_1$ has GMA among all 21 5-factor designs.

The GMA criterion has a sound statistical justification. Suppose we fit the model

$$\mathbf{y} = \mathbf{X}_0\theta_0 + \mathbf{X}_1\boldsymbol{\theta}_1 + \boldsymbol{\epsilon}, \tag{9.8}$$

which contains the general mean and main effects. For a balanced design, the least squares estimator of main effects $\boldsymbol{\theta}_1$, $\hat{\boldsymbol{\theta}}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}$, is unbiased if (9.8) is correct. However, under the full model (9.6),

$$E(\hat{\boldsymbol{\theta}}_1) = \boldsymbol{\theta}_1 + \mathbf{C}_2\boldsymbol{\theta}_2 + \cdots + \mathbf{C}_k\boldsymbol{\theta}_k,$$

where $\mathbf{C}_j = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_j$ is the alias matrix between the main effects and j-factor interactions. The estimation of the main effects is biased or contaminated by (non-negligible) interactions. A good design should have a small contamination. For a matrix $\mathbf{C} = (c_{ij})$, let $\|\mathbf{C}\|^2 = \sum_{i,j}|c_{ij}|^2$ be its squared norm. The value $\|\mathbf{C}_j\|^2$ is an overall measure of the *contamination* of j-factor interactions on the estimation of main effects. It can be shown that $\|\mathbf{C}_j\|^2$ is independent of the choice of orthonormal contrasts. In the spirit of the *hierarchical ordering principle* discussed by Wu and Hamada (2009, Section 4.6), a good design should sequentially minimize $\|\mathbf{C}_j\|^2$ for $j = 2, \ldots, k$. Xu and Wu (2001) showed that if all k factors have s levels,

$$\|\mathbf{C}_j\|^2 = (j+1)A_{j+1} + j(s-2)A_j + (k-j+1)(s-1)A_{j-1} \text{ for } j = 2, \ldots, k.$$

It is easy to see that sequentially minimizing $A_3, A_4, \ldots$ is equivalent to sequentially minimizing $\|\mathbf{C}_2\|^2, \|\mathbf{C}_3\|^2, \ldots$.

Theorem 2. *The GMA criterion sequentially minimizes the contamination of j-factor interactions on the estimation of main effects for $j = 2, 3 \ldots$.*

This result was first proved by Tang and Deng (1999) for 2-level designs and extended by

Xu and Wu (2001) for general designs.

An issue of the GMA criterion is computation. It is cumbersome to compute the generalized wordlength pattern according to the definition of $A_j(\mathbf{D})$ in (9.7). Fortunately, Xu and Wu (2001) provided an efficient method for computing the generalized wordlength patterns via coding theory. An alternative approach is to use the minimum moment aberration to be discussed in the next subsection.

The *Hamming distance* between two vectors $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ is the number of places where they differ, i.e., the number of l's such that $x_l \neq y_l$. For an $n \times k$ design $\mathbf{D}$, let $d_{ij}(\mathbf{D})$ be the Hamming distance between rows i and j, and $B_l(\mathbf{D}) = n^{-1}|\{(i,j) : d_{ij}(\mathbf{D}) = l, i, j = 1, \ldots, n\}|$ for $l = 0, 1, \ldots, k$. In coding theory, the vector $(B_0(\mathbf{D}), B_1(\mathbf{D}), \ldots, B_k(\mathbf{D}))$ is the *distance distribution*. It is obvious that $\sum_{l=0}^{k} B_l(\mathbf{D}) = n$. Xu and Wu (2001) showed that the generalized wordlength patterns are linear combinations of the distance distributions, and vice verse.

Theorem 3. *For an $n \times k$ design $\mathbf{D}$ with s levels and $j = 0, 1, \ldots, k$,*

$$A_j(\mathbf{D}) = n^{-1} \sum_{i=0}^{k} P_j(i; k, s) B_i(\mathbf{D}), \tag{9.9}$$

$$B_j(\mathbf{D}) = n s^{-k} \sum_{i=0}^{k} P_j(i; k, s) A_i(\mathbf{D}), \tag{9.10}$$

where $P_j(x; k, s) = \sum_{i=0}^{j} (-1)^i (s-1)^{j-i} \binom{x}{i} \binom{k-x}{j-i}$ are the Krawtchouk polynomials.

The equations (9.9) and (9.10) are known as the generalized *MacWilliams identities*, which play a pivotal role in the theoretical development of nonregular designs.

Two designs are called (combinatorially) *isomorphic* if one design can be obtained from the other by permutations of rows, columns and levels in the columns. Isomorphic designs have the same generalized wordlength pattern.

Cheng and Ye (2004) proposed another extension of the minimum aberration criterion. The Cheng and Ye extension is intended for designs with quantitative factors and depends

on the contrasts used in the model. On the other hand, the Xu and Wu extension described in this section is intended for designs with qualitative factors and does not depend on the contrasts used in the model.

9.4.2 Minimum moment aberration

Based on coding theory, Xu (2003) proposed the minimum moment aberration criterion for assessing nonregular designs. For an $n \times k$ design $\mathbf{D}$ with s levels and a positive integer m, define the mth power moment to be

$$K_m(\mathbf{D}) = [n(n-1)/2]^{-1} \sum_{1 \le i < j \le n} [\delta_{ij}(\mathbf{D})]^m , \tag{9.11}$$

where $\delta_{ij}(\mathbf{D}) = k - d_{ij}(\mathbf{D})$ is the number of coincidences between rows i and j. The *minimum moment aberration* criterion proposed by Xu (2003) is to sequentially minimize the power moments $K_1, K_2, \ldots$.

The power moments measure the similarity among runs (i.e., rows). The first and second power moments measure the average and variance of the similarity among runs. Minimizing the power moments makes runs to be as dissimilar as possible. As we will see, the power moments also measure the non-orthogonality among columns, and orthogonal designs have small power moments.

Note that the computation of the power moments involves the numbers of coincidences between rows. By applying generalized MacWilliams identities and Pless power moment identities, two fundamental results in coding theory (MacWilliams and Sloane 1977, Chapter 5), Xu (2003) showed that the power moments K_m defined in (9.11) are linear combinations of the generalized wordlength patterns $A_1, \ldots, A_m$ defined in (9.7).

Theorem 4. *For an $n \times k$ design $\mathbf{D}$ with s levels and $m = 1, 2, \ldots, k$,*

$$K_m(\mathbf{D}) = c_m A_m(\mathbf{D}) + c_{m-1} A_{m-1}(\mathbf{D}) + \ldots + c_1 A_1(\mathbf{D}) + c_0, \tag{9.12}$$

where $c_m = m! s^{-m} n(n-1)^{-1}$ and c_i are constants depending only on i, n, k, s for $i < m$.

It is not difficult to see now that sequentially minimizing $K_1, K_2, \ldots$ is equivalent to sequentially minimizing $A_1, A_2 \ldots$. Therefore, the minimum moment aberration is equivalent to the GMA.

The minimum moment aberration provides a useful tool for efficient computation and theoretical development. For an $n \times k$ design with s levels, the complexity of computing the generalized wordlength pattern according to the definition (9.7) is $O(ns^k)$ whereas the complexity of computing k power moments is $O(n^2k^2)$. The saving in computation is tremendous when the number of factors k is large. This observation led to successful algorithmic constructions of mixed-level OAs (Xu (2002)), a catalog of 3-level regular designs (Xu (2005b)), and blocked regular designs with minimum aberration (Xu and Lau (2006)). As a theoretical tool, Xu (2003) developed a unified theory for nonregular and supersaturated designs. Xu and Lau (2006) and Xu (2006) further used the concept of minimum moment aberration to develop a theory for blocked regular designs and constructed minimum aberration blocked regular designs with 32, 64 and 81 runs.

For mixed-level designs, Xu (2003) suggested to use weights to reflect the importance of the factors. For a design $\mathbf{D} = (x_{il})$, assign weight $w_l > 0$ to the lth column and let

$$\delta_{ij}(\mathbf{D}) = \sum_{l=1}^{k} w_l \delta(x_{il}, x_{jl}) \tag{9.13}$$

be the weighted coincidence number between the ith and jth rows, where $\delta(x, y) = 1$ if $x = y$ and 0 otherwise. Then define $K_m(\mathbf{D})$ as in (9.11). A special choice for a mixed-level design is to choose weight proportional to its number of levels, say, $w_l = s_l$. For this choice, Xu (2003) showed that if $\mathbf{D}$ is an OA of strength t, the identity in (9.12) holds for $m = 1, \ldots, t + 1$. Therefore, the minimum moment aberration is weakly equivalent to the GMA for mixed-level designs.

9.4.3 Generalized resolution and projection properties for 2-level designs

The resolution, defined in Chapter 1 (Section 1.7.2) and Chapter 7 (Section 7.2.4), is an

important concept for regular designs. To define the generalized resolution, we need to consider projections of a design. For an $n \times k$ design $\mathbf{D}$, a j-factor projection design is the $n \times j$ submatrix representing the j factors. For clarity, we use $\mathbf{d}$ to denote a projection.

Suppose that $\mathbf{d}$ is an $n \times j$ matrix (d_{il}) with two levels denoted by -1 and $+1$, and let

$$\rho(\mathbf{d}) = \frac{1}{n} \sum_{i=1}^{n} d_{i1} \times \cdots \times d_{ij}. \tag{9.14}$$

The quantity $\rho(\mathbf{d})$ is called a *design moment* in the response surface design literature. If $\mathbf{d}$ is a balanced design, $\rho(\mathbf{d})$ is the correlation between the main effect of a column and the interaction involving the other $j-1$ columns. For illustration, consider the 12-run Plackett-Burman design in Table 9.1. For $\mathbf{d} = \{A, B\}$, $\rho(\mathbf{d}) = 0$ since A and B are orthogonal. For $\mathbf{d} = \{A, B, C\}$, $\rho(\mathbf{d}) = -1/3$ is the correlation between the main effect A (or B or C) and the 2fi BC (or AC or AB). For $\mathbf{d} = \{A, B, C, D\}$, $\rho(\mathbf{d}) = -1/3$ is the correlation between the main effect A (or B or C or D) and 3fi BCD (or ACD or ABD or ABC), as well as the correlation between 2fi's AB and CD (or AC and BD, or AD and BC).

The quantity $|\rho(\mathbf{d})|$ is called the normalized J-characteristics by Tang and Deng (1999) or aliasing index by Cheng, Li and Ye (2004) and Phoa and Xu (2009) because $0 \leq |\rho(\mathbf{d})| \leq 1$. When $|\rho(\mathbf{d})| > 0$, the k columns in $\mathbf{d}$ form a *word* of length k with aliasing index $|\rho(\mathbf{d})|$. A word is called *complete* if $|\rho(\mathbf{d})| = 1$ or *partial* if $|\rho(\mathbf{d})| < 1$. When $\rho(\mathbf{d}) = 0$, the k columns in $\mathbf{d}$ do not form a word.

It is not difficult to see that if $\mathbf{D}$ is a two-level regular design then $|\rho(\mathbf{d})| = 0$ or 1 where $\mathbf{d}$ is a projection of $\mathbf{D}$. Ye (2004) showed that the reverse is also true. Therefore, for a nonregular design, there always exists some projection $\mathbf{d}$ such that $0 < |\rho(\mathbf{d})| < 1$. It can be shown that $A_j(\mathbf{D})$, as defined in (9.7), can be expressed as

$$A_j(\mathbf{D}) = \sum_{|\mathbf{d}|=j} \rho^2(\mathbf{d}), \tag{9.15}$$

where the summation is over all j-factor projections $\mathbf{d}$.

Suppose that r is the smallest integer such that $A_r(\mathbf{D}) > 0$. Then the *generalized resolu-*

tion (Deng and Tang (1999)) is defined to be

$$\tilde{R} = \tilde{R}(\mathbf{D}) = r + \delta, \text{ where } \delta = 1 - \max_{|\mathbf{d}|=r} |\rho(\mathbf{d})|, \tag{9.16}$$

where the maximization is over all projections $\mathbf{d}$ with r factors. Grömping and Xu (2014) recently extended the generalized resolution to orthogonal arrays with mixed levels.

Example 9. Consider the 12-run Plackett-Burman design in Table 9.1. It is an OA of strength 2, so $A_1 = A_2 = 0$ but $A_3 > 0$. It is straightforward to verify that $|\rho(\mathbf{d})| = 1/3$ for any 12×3 subdesign $\mathbf{d}$. So the generalized resolution is $\tilde{R} = 3 + (1 - 1/3) = 11/3$.

Like the GMA, the generalized resolution has a good statistical justification. For a 2-level OA with strength 2, maximizing the generalized resolution is equivalent to minimizing the maximum bias of any individual 2fi on the estimation of the main effects (Deng and Tang 2002). In contrast, the GMA criterion minimizes the overall bias of all 2fi's on the estimation of the main effects.

The generalized resolution has a nice geometric property. It is easy to see that for an $OA(n, 2^k, t)$, $\rho(\mathbf{d}) = 0$ for any projection $\mathbf{d}$ with t factors or fewer and therefore $r \leq \tilde{R} < r+1$ where $r = t + 1$. If $\delta > 0$, any projection with $r = t + 1$ factors contains at least $n\delta/2^r$ copies of a full 2^r factorial (Deng and Tang (1999)). This result was first proved by Cheng (1995). Cheng (1995, 1998) and Bulutoglu and Cheng (2003) further studied some hidden projection properties of nonregular designs; see Xu, Phoa and Wong (2009) for a review.

Box and Tyssedal (1996) defined a design to be of *projectivity p* if the projection onto every subset of p factors contains a full factorial design, possibly with some points replicated. It follows from these definitions that an OA of strength t is of projectivity at least t. A regular design of resolution r has projectivity $r-1$ while a nonregular design of generalized resolution $r + \delta$ has projectivity at least r if $\delta > 0$. A result of Cheng (1995) implies that, as long as the run size n is not a multiple of 2^{t+1}, an $OA(n, 2^k, t)$ with $k \geq t + 2$ has projectivity $t + 1$, even though the strength is only t.

Example 10. Consider the two 2^{9-5} designs in the chemical toxicity experiment given in Table 9.2. Both designs have the same generalized wordlength pattern $(0, 0, 4, 14, 8, 0, 4, 1, 0)$.

For the regular design, there are four words of length 3 with aliasing index 1; for the nonregular design, there are 16 words of length 3 with aliasing index 0.5. The regular design has generalized resolution 3.0 and projectivity 2 while the nonregular design has generalized resolution 3.5 and projectivity 3.

Deng and Tang (1999) went beyond the generalized resolution and defined another version of the generalized minimum aberration, which we refer to as the *minimum G-aberration* criterion for clarity. Roughly speaking, the minimum G-aberration criterion always chooses a design with the smallest confounding frequency among designs with maximum generalized resolution. Formally, the minimum G-aberration criterion is to sequentially minimize the components in the confounding frequency vector

$$\text{CFV} = [(f_{11}, \ldots, f_{1n}); (f_{21}, \ldots, f_{2n}); \ldots; (f_{k1}, \ldots, f_{kn})],$$

where f_{ji} denotes the frequency of j-factor projections $\mathbf{d}$ with $|\rho(\mathbf{d})| = 1 + (1 - i)/n$.

The minimum G_2-aberration criterion, proposed by Tang and Deng (1999) and mentioned in Section 9.4.1, is a relaxed version of the minimum G-aberration criterion. For 2-level regular designs both criteria reduce to the traditional minimum aberration criterion. However, these two criteria can result in selecting different nonregular designs. Minimum G-aberration nonregular designs always have maximum generalized resolution whereas minimum G_2-aberration nonregular designs may not. This is in contrast to the regular case where minimum aberration regular designs always have maximum resolution among all regular designs.

9.4.4 Projection aberration, estimation capacity and design efficiency

The GMA criterion cannot distinguish designs when they have the same generalized wordlength pattern. It is useful to examine projections by using a criterion like the minimum G-aberration criterion. There are $\binom{k}{p}$ projected designs with p factors. Each of these designs has an A_p value, which is referred to as the *projected* A_p value, to distinguish the *overall* A_p value calculated from the whole k-factor design. The frequency of the projected A_p values

Table 9.6: Overall and projected A_3 values

	Overall	Frequency of projected A_3			
Design	A_3	1/2	2/3	1	2
$\mathbf{D}_1$	10	20	0	0	0
$\mathbf{D}_2$	13	16	0	3	1
$\mathbf{D}_3$	13	14	0	6	0

is called the *p-dimensional projection frequency*. For an $OA(n, s^k, t)$, when projecting onto any t factors, we always get a full factorial design. So it is sufficient to consider projection frequencies with $p = t + 1$, $t + 2$ and so on. The larger the projected A_p values are, the more severe the aliasing is. One approach is to sequentially minimize the $(t+1)$-dimensional projection frequency starting from the largest projected A_{t+1} value. If there is a tie, one can further compare the $(t + 2)$-dimensional projection frequency and so on. This criterion is referred to as the *projection aberration criterion* by Xu, Cheng and Wu (2004) and can serve as an extension of the minimum G-aberration for general designs.

Example 11. Consider choosing six columns from the commonly used $OA(18, 3^7, 2)$ given in Table 9.3. There are seven possible choices. For illustration, consider three choices. Let $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$ be the resulting designs from omitting the first, second and third columns, respectively. The generalized wordlength patterns for the three designs are $(0, 0, 10, 22.5, 0, 7)$, $(0, 0, 13, 13.5, 9, 4)$, and $(0, 0, 13, 13.5, 9, 4)$, respectively. Hence, $\mathbf{D}_1$ is the best according to the GMA criterion. Note that $\mathbf{D}_2$ and $\mathbf{D}_3$ have the same generalized wordlength pattern but they have different projection patterns. The frequencies of projected A_3 values are listed in Table 9.6. Among the three designs, $\mathbf{D}_2$ is the worst under the projection aberration criterion since one of its 3-factor projections has projected $A_3 = 2$; $\mathbf{D}_1$ is again the best because all its 3-factor projections have projected $A_3 = 0.5$.

Xu and Deng (2005) proposed another projection aberration criterion. When considering the projection frequency, they replaced $A_j(\mathbf{d})$, as defined in (9.7), with $K_j(\mathbf{d})$, as defined in (9.11), for a j-factor projection $\mathbf{d}$. They referred to the resulting criterion as the *moment aberration projection* criterion. The moment aberration projection criterion works the same as the the projection aberration criterion based on the A_j values. Both criteria will often select the same best design. However, the moment aberration projection criterion can distin-

guish more designs than the projection aberration criterion based on the A_j values, even for the 2-level case. The concept of moment aberration projection turns out to be very useful in the algorithmic construction of regular designs; see Xu (2005b, 2009).

For regular designs, Cheng, Steinberg and Sun (1999) justified the minimum aberration criterion by showing that it is a good surrogate for some model-robustness criteria. Following their approach, Cheng, Deng and Tang (2002) considered the situation where (i) the main effects are of primary interest and their estimates are required and (ii) the experimenter would like to have as much information about 2fi's as possible, under the assumption that higher-order interactions are negligible. Without knowing which 2fi's are significant, they considered the set of models containing all of the main effects and f 2fi's for $f = 1, 2, 3, \ldots$. Let E_f be the number of estimable models and D_f be the average of D-efficiencies of all models that contain main effects plus f 2fi's. Cheng, Deng and Tang (2002) showed that two-level GMA designs tend to have large E_f and D_f values, especially for small f; therefore, the GMA criterion provides a good surrogate for the traditional model-dependent efficiency criteria. Ai, Li and Zhang (2005) and Mandal and Mukerjee (2005) extended their approach to mixed-level designs.

9.4.5 Uniformity and space-filling property

Uniformity or space-filling is a desirable design property for physical and computer experiments (Fang, Li and Sudjianto (2006)). Various discrepancies have been used to assess the space-filling property or uniformity (Fang and Wang (1994), Fang, Lin, Winker and Zhang (2000)). These discrepancies all have geometrical meanings and can be interpreted as the difference between the empirical distribution of points in the design and the uniform distribution. Among them, the centered L_2-discrepancy, proposed by Hickernell (1998), is the most frequently used. For an $n \times k$ design $\mathbf{D} = (x_{il})$ over the unit cube $[0, 1]^k$, the squared

centered L_2-discrepancy (CD) has an analytic expression as follows:

$$\begin{aligned}[CD(\mathbf{D})]^2 &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{l=1}^{k}\left(1+\frac{1}{2}\left|x_{il}-\frac{1}{2}\right|+\frac{1}{2}\left|x_{jl}-\frac{1}{2}\right|-\frac{1}{2}\left|x_{il}-x_{jl}\right|\right)\\ &\quad -\frac{2}{n}\sum_{i=1}^{n}\prod_{l=1}^{k}\left(1+\frac{1}{2}\left|x_{il}-\frac{1}{2}\right|-\frac{1}{2}\left|x_{il}-\frac{1}{2}\right|^2\right)+\left(\frac{13}{12}\right)^k. \qquad (9.17)\end{aligned}$$

The centered L_2-discrepancy is defined over the unit cube $[0,1]^k$, but the s levels in a factorial design are normally denoted as $0, 1, \ldots, s-1$. Thus as it is often done in the literature, whenever the centered L_2-discrepancy is calculated, level i ($i = 0, 1, \ldots, s-1$) should first be transformed to $(2i+1)/(2s)$. Note that this is a useful relationship only when the s levels of each factor actually represent equally spaced values of an underlying continuous variable.

Fang and Mukerjee (2000) found a connection between aberration and uniformity for 2-level regular designs. This connection was extended by Ma and Fang (2001) for general two-level designs. The basic result states that for a two-level $n \times k$ design $\mathbf{D}$, regular or nonregular, the centered L_2-discrepancy can be expressed in terms of its generalized wordlength pattern $A_i(\mathbf{D})$ as follows:

$$[CD(\mathbf{D})]^2 = \left(\frac{13}{12}\right)^k - 2\left(\frac{35}{32}\right)^k + \left(\frac{9}{8}\right)^k\left(1+\sum_{i=1}^{k}\frac{A_i(\mathbf{D})}{9^i}\right).$$

Since the coefficient of $A_i(\mathbf{D})$ decreases exponentially with i, one can anticipate that designs with small $A_i(\mathbf{D})$ for small values of i should have small $CD(\mathbf{D})$; in other words, GMA designs tend to be uniform over the design region.

However, the result cannot be generalized to multi-level designs directly, as level permutations of one or more factors can alter the centered L_2-discrepancy, but keep the generalized wordlength pattern unchanged. By considering level permutations of three-level designs, Tang, Xu and Lin (2012) established a relationship between average centered L_2-discrepancy and generalized wordlength pattern. Tang and Xu (2013) generalized the relationship to designs with arbitrary levels. Zhou and Xu (2014) further generalized their

results for any discrepancy defined by a reproducing kernel and showed that GMA designs have good space-filling properties on average in terms of distance as well.

Hickernell and Liu (2002) showed that the GMA criterion could be defined and generalized using discrepancy. Tang (2001) and Ai and Zhang (2004a) showed that GMA designs have good low-dimensional projection properties.

9.5 Construction Methods

The construction of optimal nonregular designs is challenging for two simple reasons: (i) nonregular designs do not have a unified mathematical description and (ii) the class of nonregular designs is much larger than the class of regular designs.

9.5.1 Algorithmic methods

A simple strategy for constructing GMA or other optimal designs is to search over all possible projection designs from existing OAs, such as Plackett-Burman designs, or Hadamard matrices. A Hadamard matrix of order n is an $n \times n$ matrix of entries $+1$ and -1 whose columns (and rows) are orthogonal to each other; see Chapter 7 of Hedayat et al. (1999). A Hadamard matrix is said to be normalized if all of the entries of the first row and column are $+1$. From a normalized Hadamard matrix of order n, one obtains a saturated $OA(n, 2^{n-1}, 2)$ by deleting the first column. Deng and Tang (2002) presented a catalog of GMA designs by searching over Hadamard matrices of order 16, 20, and 24; Xu and Deng (2005) searched optimal designs over all Hadamard matrices of order 16 and 20 and 3-level designs from 68 saturated $OA(27, 3^{13}, 2)$'s. A limitation of this strategy is that we could miss the optimal design in some cases because the optimal design cannot be expressed as such a projection.

Much effort has been devoted to the complete enumeration of all non-isomorphic designs with a small number of runs. Sun, Li and Ye (2002) proposed an algorithm for sequentially constructing non-isomorphic orthogonal designs. They successfully enumerated all 2-level

OAs with 12, 16, and 20 runs. An important result is that all 16-run OAs are projections of one of the five 16-run Hadamard matrices. However, such a result does not hold for 20-run designs. Ye et al. (2008) presented a complete set of combinatorially non-isomorphic OAs of types $OA(12, 2^k3^1, 2), OA(18, 3^k, 2), OA(18, 2^13^k, 2)$, and $OA(20, 2^k5^1, 2)$. Schoen (2009) also presented all OAs with 18 runs.

Schoen, Eendebak and Nguyen (2010) proposed a general algorithm which can also handle mixed-level designs. They successfully enumerated most non-trivial mixed-level OAs up to 28 runs with strength 2, 64 runs with strength 3, and 168 runs with strength 4. They completely enumerated all 24-run OAs with strength 2, and 28-run OAs up to 7 columns. The number of non-isomorphic designs $OA(28, 2^k, 2)$ is 4, 7, 127, 17,826, and 5,882,186, respectively, for $k = 3, 4, 5, 6, 7$.

Algorithmic constructions are typically limited to small run sizes (≤ 32) or small number of factors due to the existence of a large number of designs and the difficulty of determining whether two designs are isomorphic or not. Algebraic or combinatorial methods are necessary to construct larger designs. A good construction method is the quaternary code method introduced by Xu and Wong (2007).

9.5.2 Quaternary code designs

A quaternary code (QC) is a linear subspace over $Z_4 = \{0, 1, 2, 3\}$ (mod 4), the ring of integers modulus 4. A key device is the so-called Gray map:

$$\phi : 0 \to (0, 0), 1 \to (0, 1), 2 \to (1, 1), 3 \to (1, 0), \tag{9.18}$$

which maps each symbol in Z_4 to a pair of symbols in Z_2. Let $\mathbf{G}$ be an $a \times b$ matrix with elements from Z_4 and let $\mathbf{C}$ consist of all possible linear combinations of the row vectors of $\mathbf{G}$ over Z_4. Applying the Gray map to $\mathbf{C}$, one obtains a $4^a \times 2b$ matrix or a two-level design with 4^a runs and $2b$ factors, denoted by $\mathbf{D}$. Although $\mathbf{C}$ is linear over Z_4, $\mathbf{D}$ may or may not be linear over Z_2. From $\mathbf{D}$, we can construct a two-level design with 2^{2a+1} runs and $4b$

Table 9.7: An Example of Quaternary Code and Nonregular Design

(a) Quaternary code **C**

Run	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	1	1	2	1	3
3	0	2	2	0	2	2
4	0	3	3	2	3	1
5	1	0	2	1	1	1
6	1	1	3	3	2	0
7	1	2	0	1	3	3
8	1	3	1	3	0	2
9	2	0	0	2	2	2
10	2	1	1	0	3	1
11	2	2	2	2	0	0
12	2	3	3	0	1	3
13	3	0	2	3	3	3
14	3	1	3	1	0	2
15	3	2	0	3	1	1
16	3	3	1	1	2	0

(b) Nonregular design **D**

Run	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	1	1	1	0	1	1	0
3	0	0	1	1	1	1	0	0	1	1	1	1
4	0	0	1	0	1	0	1	1	1	0	0	1
5	0	1	0	0	1	1	0	1	0	1	0	1
6	0	1	0	1	1	0	1	0	1	1	0	0
7	0	1	1	1	0	0	0	1	1	0	1	0
8	0	1	1	0	0	1	1	0	0	0	1	1
9	1	1	0	0	0	0	1	1	1	1	1	1
10	1	1	0	1	0	1	0	0	1	0	0	1
11	1	1	1	1	1	1	1	1	0	0	0	0
12	1	1	1	0	1	0	0	0	0	1	1	0
13	1	0	0	0	1	1	1	0	1	0	1	0
14	1	0	0	1	1	0	0	1	0	0	1	1
15	1	0	1	1	0	0	1	0	0	1	0	1
16	1	0	1	0	0	1	0	1	1	1	0	0

factors via the *doubling method* as follows:

$$\mathbf{D}^* = \begin{pmatrix} \mathbf{D} & \mathbf{D} \\ \mathbf{D} & \mathbf{D}+1 \end{pmatrix} \pmod 2. \tag{9.19}$$

Example 12. Consider a 2×6 matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 & 3 \end{bmatrix}.$$

All linear combinations of the two rows of $\mathbf{G}$ form a 16×6 linear code $\mathbf{C}$ over Z_4. Applying the Gray map, we obtain a 16×12 matrix $\mathbf{D} = \phi(\mathbf{C})$, which is a 2^{12-8} design. See Table 9.7 for the matrices $\mathbf{C}$ and $\mathbf{D}$. It is straightforward to verify that $\mathbf{D}$ has generalized resolution 3.5; therefore, it is a nonregular design. Moreover, the 32×24 matrix $\mathbf{D}^*$ obtained via doubling as in (9.19) is a 2^{24-19} design and has generalized resolution 3.5 too. For comparison, in both cases the best regular design of the same size has resolution 3.

Example 13. Consider a 4×8 matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 3 \end{bmatrix}.$$

All linear combinations of the rows of $\mathbf{G}$ over Z_4 form a 256×8 quaternary linear code $\mathbf{C}$. Applying the Gray map, we obtain a 256×16 matrix $\mathbf{D} = \phi(\mathbf{C})$, which is isomorphic to the (extended) Nordstrom-Robinson code (Xu 2005a). The resulting design $\mathbf{D}$ is an $OA(256, 2^{16}, 5)$ with many remarkable properties: It has generalized resolution 6.5 and projectivity 7. For comparison, for a regular design to achieve the same resolution and projectivity, it would require at least 512 runs. For more statistical properties and results from the Nordstrom-Robinson code, see Xu (2005a).

Lemma 1. *Let* $\mathbf{G}$ *be an* $a \times b$ *matrix over* Z_4, $\mathbf{C}$ *be the quaternary linear code generated by* $\mathbf{G}$ *and* $\mathbf{D} = \phi(\mathbf{C})$ *be the binary image. Then* $\mathbf{D}$ *is an OA of strength* 2 *if and only if* $\mathbf{G}$ *satisfies the following conditions:*

(i) it does not have any column containing entries 0 *and* 2 *only,*

(ii) none of the columns is a multiple of another column over Z_4*.*

Xu and Wong (2007) further showed that if the two conditions in Lemma 1 are satisfied, then $\mathbf{D}$ has generalized resolution at least 3.5 and $\mathbf{G}$ has a maximum of $(4^a - 2^a)/2$ columns. Such a matrix can be constructed as follows:

1. Write down all possible columns of a elements over Z_4.
2. Delete columns that do not contain any 1's.
3. Delete columns whose first non-zero and non-two entries are 3's.

Theorem 5. *For an integer* $a > 1$*, let* $\mathbf{G}$ *be the generator matrix obtained from the above procedure. Then the binary image* $\mathbf{D}$ *generated by* $\mathbf{G}$ *has* 4^a *rows,* $4^a - 2^a$ *columns, and*

generalized resolution 3.5. *The double* $\mathbf{D}^*$ *of* $\mathbf{D}$ *has* 2^{2a+1} *rows,* $2^{2a+1} - 2^{a+1}$ *columns, and generalized resolution* 3.5.

Note that the nonregular designs constructed in Theorem 5 have generalized resolution 3.5. It is known that a regular 2-level design with n runs and k factors has resolution at most 3 when $k > n/2$; see Chapter 7, Section 7.4.1. Therefore, nonregular designs constructed from quaternary codes have higher resolution than corresponding regular designs when resolution 4 designs do not exist.

Since QC designs are linear over Z_4, we can enumerate QC designs sequentially in a similar manner as enumerating regular designs. Xu and Wong (2007) developed a sequential algorithm, similar to those by Chen, Sun and Wu (1993) and Xu (2005b). They also presented a collection of nonregular designs with 32, 64, 128 and 256 runs and up to 64 factors, many of which are better than regular designs of the same size in terms of resolution, aberration and projectivity.

The linear structure of a quaternary code also facilitates the derivation and analytical study of properties of QC designs. Phoa and Xu (2009) studied quarter-fraction QC designs which are defined by a generator matrix that consists of an identity matrix plus an extra column. They showed that the generalized resolution, generalized wordlength pattern and projectivity can be calculated in terms of the frequencies of the numbers 1, 2 and 3 that appear in the extra column.

Specifically, consider an $a \times (a+1)$ generator matrix $\mathbf{G} = (\mathbf{v}, \mathbf{I}_a)$, where $\mathbf{v}$ is an $a \times 1$ column vector over Z_4 and $\mathbf{I}_a$ is the $a \times a$ identity matrix. The binary image $\mathbf{D}$ generated by $\mathbf{G}$ is a $2^{(2a+2)-2}$ design. It is easy to verify that the identity matrix $\mathbf{I}_a$ generates a full 2^{2a} design; therefore, the properties of $\mathbf{D}$ depend on the column $\mathbf{v}$ only. For $i = 0, 1, 2, 3$, let f_i be the number of times that number i appears in column $\mathbf{v}$. Phoa and Xu (2009) showed that the number of words of $\mathbf{D}$, their lengths and aliasing indexes can be expressed in terms of the frequency f_i.

Theorem 6. *The* $2^{(2a+2)-2}$ *design* $\mathbf{D}$ *generated by* $\mathbf{G} = (\mathbf{v}, \mathbf{I}_a)$ *has* 1 *complete word of length* $2f_1 + 2f_3 + 2$ *and* $2/\rho^2$ *words of length* $f_1 + 2f_2 + f_3 + 1$ *with aliasing index* $\rho = 2^{-\lfloor (f_1+f_3)/2 \rfloor}$,

where $\lfloor x \rfloor$ *is the integer value of* x.

Since $\mathbf{D}$ has a complete word of length $2(f_1+f_3)+2$, its projectivity is at most $2(f_1+f_3)+1$. The following theorem from Phoa and Xu (2009) shows that the projectivity of $\mathbf{D}$ is not affected by the partial words.

Theorem 7. *Suppose that* $\mathbf{D}$ *is the* $2^{(2a+2)-2}$ *design generated by* $\mathbf{G} = (\mathbf{v}, \mathbf{I}_a)$.

(a) *If* $f_2 > 0$, *the projectivity of* $\mathbf{D}$ *is* $2(f_1 + f_3) + 1$.

(b) *If* $f_2 = 0$ *and* $f_1 + f_3 > 0$, *the projectivity of* $\mathbf{D}$ *is* $2(f_1 + f_3) - 1$.

Based on these theoretical results, Phoa and Xu (2009) constructed optimal quarter-fraction QC designs under the maximum resolution, GMA and maximum projectivity criteria. These optimal QC designs are often better than regular designs of the same size in terms of the design criterion. The GMA QC designs have the same aberration as the minimum aberration regular designs, and frequently with larger resolution and projectivity. A maximum projectivity QC design is often different from a minimum aberration or maximum resolution design but can have much larger projectivity than a minimum aberration regular design. They further showed that some of these QC designs have GMA and maximum projectivity among all possible designs. Zhang et al. (2011) and Phoa et al. (2012) investigated 1/8th and 1/16th fraction QC designs which are defined by a generator matrix that consists of an identity matrix plus two additional columns. Phoa (2012) further studied 1/64th fraction QC designs.

9.6 Optimality Results

It is infeasible to search over all possible designs in many situations. Theoretical results are extremely useful to determine whether a design is optimal under the GMA or other criteria. Xu (2003) gave several sufficient conditions for a design to have GMA among all possible designs using the concept of minimum moment aberration.

One sufficient condition is that the numbers of coincidences between distinct rows are constant or differ by at most one.

Theorem 8. *Design* $\mathbf{D}$ *has GMA among all possible designs if the differences among all* $\delta_{ij}(\mathbf{D})$*,* $i < j$*, do not exceed one.*

In other words, a design has GMA if its design points are equally or nearly equally spaced over the design region. As an example, the 12-run Plackett-Burman design in Table 9.1 has GMA because the numbers of coincidences between any two distinct rows are 5. It is easy to see that deleting any column yields an $OA(12, 2^{10}, 2)$, which has GMA among all possible designs too.

An important class of designs that satisfy the conditions in Theorem 8 are saturated OAs of strength 2. Mukerjee and Wu (1995) showed the following result.

Lemma 2. *The numbers of coincidences between any distinct pair of rows of a saturated* $OA(n, s^k, 2)$ *are constant; specifically,* $\delta_{ij}(\mathbf{D}) = (n - s)/(s^2 - s)$ *for any* $i \neq j$*.*

Another sufficient condition relates to projections of a design.

Theorem 9. *Design* $\mathbf{D}$ *has GMA among all possible designs if* $\mathbf{D}$ *is an* $OA(n, s^k, t)$ *and there are no repeated runs in any* $(t + 1)$*-factor projection.*

For example, consider the $OA(18, 3^6, 2)$ given by columns 2–7 in Table 9.3. It is easy to verify that its projection onto any three columns does not have repeated runs. Thus, this design (and any of its projections) has GMA among all possible designs.

Another general technique for constructing optimal designs is linear programming, which employs the generalized MacWilliams identities (9.9) and (9.10). The linear programming technique has been used to establish bounds on the maximum size of a code for given length and distance in coding theory (MacWilliams and Sloane 1977, Section 17.4) and bounds on the minimum size of an OA for given number of factors and strength (Hedayat et al. 1999, Section 4.5). Xu (2005a) used linear programming to show that several nonregular designs derived from the Nordstrom-Robinson code have GMA among all possible designs. The following result is from Xu (2005a).

Theorem 10. *Any regular 2^{k-2} minimum aberration design has GMA among all possible designs.*

Butler (2003, 2004) presented a number of construction results that allow 2-level GMA designs to be found for many of the cases with n =16, 24, 32, 48, 64 and 96 runs. Butler (2005) further developed theoretical results and presented methods that allow GMA designs to be constructed for more than two levels. A key tool used by Butler (2003, 2004, 2005) is some identities that link the generalized wordlength patterns with moments of the inner products or Hamming distances between the rows; see also Chapter 7, Section 7.6. These identities can be derived easily from the generalized Pless power moment identities developed by Xu (2003).

9.7 Supersaturated Designs

The study of supersaturated designs (SSDs) dates back to Satterthwaite (1959) and Booth and Cox (1962). The former suggested the use of random balanced designs and the latter proposed an algorithm to construct systematic SSDs. Many methods have been proposed for constructing two-level SSDs in the last two decades after Lin (1993) and Wu (1993). The early construction methods use Hadamard matrices or balanced incomplete block designs; see, among others, Lin (1993), Wu (1993), Nguyen (1996), Cheng (1997), and Tang and Wu (1997). Early algorithmic construction includes Lin (1995), Nguyen (1996), and Li and Wu (1997). Chapter 1 (Section 1.7.4) gives some other references. Georgiou (2014) gave a review of construction methods and provided many additional references.

Lin (1993) used half fractions of Hadamard matrices to construct two-level SSDs. First obtain a saturated $OA(n, 2^{n-1}, 2)$ from a normalized Hadamard matrix of order n (by deleting the first column which is a column of ones). For example, we obtain the 12×11 Plackett-Burman design in Table 9.1 from a normalized Hadamard matrix of order 12. Now for each column, half of the entries are 1 and the other half are -1. Use any column as the branching column and take those rows whose entries are 1 in the branching column. Deleting the branching column yields an SSD with $n/2$ runs and $n-2$ columns (provided that the

resulting design has no repeated columns).

Wu (1993) proposed another construction method by utilizing partial aliasing of 2fi's among Plackett-Burman designs or Hadamard matrices. Consider the 12×11 design matrix in Table 9.1 again for illustration. There are $\binom{11}{2} = 45$ 2fi's. None of the 45 2fi's are fully aliased with the original 11 columns or other 2fi's. By combining the 45 columns with the original 11 columns, we obtain an SSD with 12 runs and 66 columns.

A popular criterion in the SSD literature is the $E(s^2)$ criterion [Booth and Cox (1962) and Lin (1993)]. For a balanced $n \times k$ design $\mathbf{D}$ with two levels denoted by 1 and -1,

$$E(s^2) = \sum_{1 \le i \le j \le k} s_{ij}^2 / [k(k-1)/2],$$

with $s_{ij} = \mathbf{c}_i' \mathbf{c}_j$, where $\mathbf{c}_i$ and $\mathbf{c}_j$ are the ith and jth columns of $\mathbf{D}$. Nguyen (1996) and Tang and Wu (1997) independently derived the following lower bound for two-level SSDs with n runs and k factors:

$$E(s^2) \ge n^2(k-n+1)/[(k-1)(n-1)]. \tag{9.20}$$

The GMA criterion can be used to assess general SSDs, including mixed-level designs. Following the discussion of $A_2(\mathbf{D})$ in Section 9.4.1, it is easy to see that $E(s^2) = 2n^2 A_2(\mathbf{D})/[k(k-1)]^2$ for two-level SSDs. Therefore, the GMA criterion can be viewed as a refinement of the $E(s^2)$ criterion and the general optimality results on the GMA criterion can be applied to the SSDs directly. For example, using the connection between the minimum moment aberration and GMA, Xu and Wu (2005) obtained the following result regarding multi-level SSDs, which include many previous results as special cases.

Theorem 11. *For an SSD* $\mathbf{D}$ *with* n *runs and* k *factors at* s *levels,*

$$A_2(\mathbf{D}) \ge \frac{k(s-1)(ks-k-n+1)}{2(n-1)} + \frac{(n-1)s^2\eta(1-\eta)}{2n}, \tag{9.21}$$

where $\eta = k(n-s)/(ns-s) - \lfloor k(n-s)/(ns-s) \rfloor$ *is the fractional part of* $k(n-s)/(ns-s)$. *The lower bound is achieved if and only if the numbers of coincidences,* $\delta_{ij}(\mathbf{D})$, *differ by at most one for all* $i < j$. *Furthermore, an SSD achieving the lower bound is optimal under the*

GMA criterion.

When applied to two-level SSDs, Theorem 11 improves the lower bound (9.20) whenever $\eta > 0$. Butler et al. (2001) and Bulutoglu and Cheng (2004) gave further improvements on the lower bound (9.20) for two-level SSDs.

Many optimal SSDs that achieve the lower bound in Theorem 11 can be derived from saturated OAs. A key property of saturated OAs, stated in Lemma 2, is that the numbers of coincidences, $\delta_{ij}(\mathbf{D})$, are constant for any pair of rows $i \neq j$.

Tang and Wu (1997) proposed construction of optimal two-level SSDs by juxtaposing saturated OAs derived from Hadamard matrices. This method can be easily extended to construct optimal multi-level SSDs. Suppose $\mathbf{D}_1, \ldots, \mathbf{D}_f$ are f saturated $OA(n, s^k, 2)$ with $k = (n-1)/(s-1)$. Let $\mathbf{D} = \mathbf{D}_1 \cup \cdots \cup \mathbf{D}_f$ be the $n \times fk$ array obtained by *column* juxtaposition, which may have duplicated or fully aliased columns. It is evident that $\delta_{ij}(\mathbf{D})$ are constant for any $i < j$. Then by Theorem 11, $\mathbf{D}$ is an optimal SSD under the GMA criterion. The conclusion may no longer be valid if repeated columns are removed.

When $n = s^2$, Lemma 2 implies that the numbers of coincidences between any two rows are equal to 1 for a saturated $OA(n, s^k, 2)$. Then removing any number of orthogonal columns from $\mathbf{D}$ also results in an optimal SSD under GMA, because the resulting design has the property that the numbers of coincidences between any two rows differ by at most one. In particular, for any k, the lower bound in Theorem 11 is tight when $n = s^2$.

The half fraction method of Lin (1993) can be easily extended to construct multi-level SSDs as follows. Taking any column of a saturated $OA(n, s^k, 2)$ as the branching column, we obtain s fractions according to the levels of the branching column. After removing the branching column, the fractions have the properties that all columns are balanced and the numbers of coincidences between any two rows are constant. The *row* juxtaposition of any f fractions form an SSD with fn/s rows and $k-1$ columns, of which the numbers of coincidences between any two rows differ by at most one. By Theorem 11, such a design is optimal under GMA. For $n = s^2$, any subdesign is also optimal, because the numbers of coincidences between any two rows are either 0 or 1.

For any prime power s and integer $p > 0$, a saturated $OA(n, s^k, 2)$ exists where $n = s^p$ and $k = (n-1)/(s-1)$. The following result is from Xu and Wu (2005).

Theorem 12. *Let s be a prime power. There exists an optimal s-level $n \times k$ SSD under the GMA criterion that achieves the lower bound in Theorem 11 when any of the following conditions hold:*

(i) $n = s^p$ and $k = f(s^p - 1)/(s-1)$ where $p > 0$ and $f > 0$ are integers.

(ii) $n = fs^{p-1}$ and $k = (s^p - 1)/(s-1) - 1$ where $p > 0$ and $0 < f < s$ are integers.

(iii) $n = s^2$ and any integer $k > 0$.

(iv) $n = fs$ and any integer $0 < k \leq s$ where $0 < f < s$ is an integer.

The above optimal SSDs may contain fully aliased columns, which are not useful in practice. To further distinguish designs with the same (overall) A_2 values, we consider their 2-factor projections and apply the generalized resolution for 2-level designs or the projection aberration idea in general; see Sections 9.4.3 and 9.4.4. To maximize the generalized resolution is equivalent to minimize the maximum absolute correlation between any two columns, or the $\max(s^2) = \max_{i<j} s_{ij}^2$ criterion in the literature. While there are abundant results on the $E(s^2)$ criterion or its extensions, there are relatively few results on the $\max(s^2)$ criterion. Cheng and Tang (2001) studied the maximum number of factors that an SSD can have under the constraint on $\max(s^2)$.

Xu and Wu (2005) presented explicit construction methods that produce optimal SSDs without fully aliased columns using linear and quadratic functions over finite fields. The construction method was closely related to the Addelman and Kempthorne (1961) construction method of $OA(2s^p, s^k, 2)$ with $k = 2(s^p - 1)/(s-1)$. Numerical comparisons for small 3-, 4- and 5-level SSDs indicate that their algebraic method produces good SSDs.

9.8 Analysis Strategies

We begin with a discussion of analysis of nonregular factorial experiments and towards the end give references on analysis for SSDs.

The analysis strategy proposed by Hamada and Wu (1992) consists of three steps:

Step 1. Entertain all the main effects and interactions that are orthogonal to the main effects. Use standard analysis methods such as ANOVA and half-normal plots to select significant effects.

Step 2. Entertain the significant effects identified in the previous step and 2fi's that include at least one factor that has a significant main effect. Identify significant effects using a forward selection regression procedure.

Step 3. Entertain the significant effects identified in the previous step and all the main effects. Identify significant effects using a forward selection regression procedure.

Iterate between Steps 2 and 3 until the selected model stops changing. Note that the traditional analysis of Plackett-Burman or other nonregular designs ends at Step 1.

Hamada and Wu (1992) based their analysis strategy on two empirical principles, *effect sparsity* and *effect heredity* (Wu and Hamada (2009, Section 4.6)). Effect sparsity implies that only few main effects and even fewer 2fi's are relatively important in a factorial experiment. Effect heredity means that in order for an interaction to be included, at least one of the main effects associated with its parent factors should be included. In other words, effect heredity excludes models that contain an interaction but none of its parent main effects. Hamada and Wu (1992) wrote that the strategy works well when both principles hold and the correlations between partially aliased effects are small to moderate. The effect sparsity assumption suggests that only a few iterations will be required.

Using this procedure, Hamada and Wu (1992) reanalyzed data from three real experiments: a cast fatigue experiment using a 12-run Plackett-Burman design with seven 2-level factors, a blood glucose experiment using an 18-run mixed-level OA with one 2-level and

seven 3-level factors, and a heat exchange experiment using a 12-run Plackett-Burman design with ten 2-level factors. They demonstrated that the traditional main effects analysis was limited and the results were misleading. Phoa, Xu and Wong (2009) gave three more real examples to show the importance of considering interactions for screening experiments.

Hamada and Wu (1992) discussed limitations of their analysis strategy and provided solutions. Wu and Hamada (2009, Chapter 8) suggested some further extensions such as the use of all subset variable selection if possible.

Example 14. Consider the HPLC experiment in Example 1. The traditional main effects analysis shows that the two most important factors are E and F. The model that consists of only the main effects of E and F has $R^2 = 0.41$. Using the Hamada-Wu analysis strategy, we find a significant EF interaction in step 2. Adding EF to E and F increases R^2 from 0.41 to 0.89. In step 3, we further identify factor H, which is missed in the traditional approach, as significant at the 5% level. We repeat steps 2 and 3 iteratively until no more new significant effects are identified and the model does not change anymore. When this happens, we stop the procedure and report the final model, which is given in (9.2).

Box and Meyer (1993) proposed a Bayesian approach by considering all the possible explanations (models including interactions) of the data from a screening experiment and identifying those that fit the data well. The assumptions for prior distributions in their approach are as follows:

1. Each factor has independent prior probability π being active.
2. All effects from a model are assigned independent prior normal distributions with mean 0 and variance $\gamma^2\sigma^2$.
3. A noninformative prior distribution is employed for experimental noise σ.

The prior probability of a model with f active factors is $\pi^f(1-\pi)^{k-f}$ for a k-factor design. The model with f active factors includes main effects for each active factor and all of their interactions (up to any desired order). The parameter γ captures the magnitude of the effects relative to experimental noise σ. Box, Hunter and Hunter (2005) suggested to choose

$\pi = 0.25$ and γ between 2 and 3, based on a survey of a number of published analyses of factorial experiments. The results are not very sensitive to moderate changes in π and γ when active factors are present.

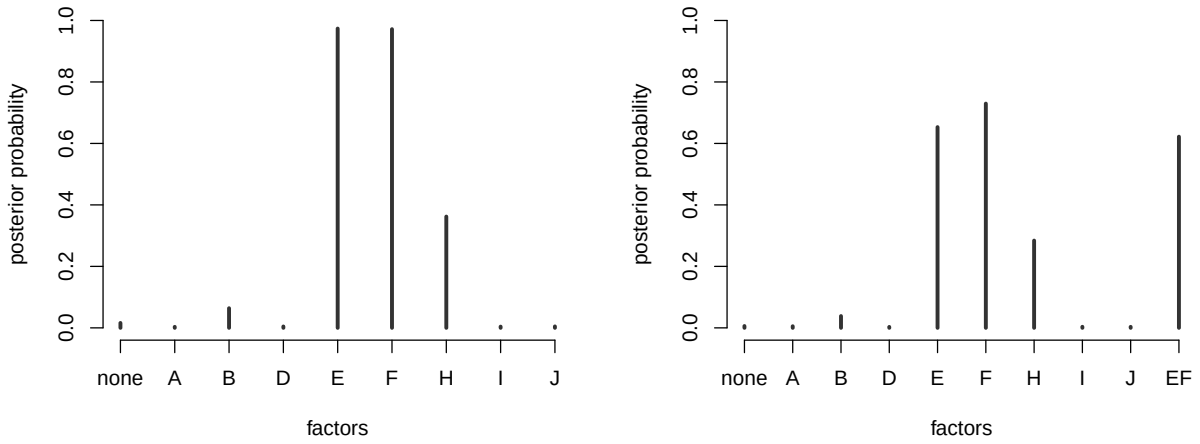
A Bayesian framework is used to assign posterior probabilities to all the models considered; see Box and Meyer (1993) or Box, Hunter and Hunter (2005). These posterior probabilities are then accumulated to marginal posterior probabilities for each factor. A factor which has a relatively high posterior probability implies that either its main effect or an interaction involving it or both are important.

Example 15. We analyze the HPLC experiment in Example 1 via the Bayesian approach. The posterior probability plot in Figure 9.1 (left) shows the marginal posterior probabilities for each factor with $\pi = 0.25$ and $\gamma = 2$. The posterior probability is high for factors E and F, moderate for factor H and small for other factors. This suggests that factors E, F and H are active. However, the marginal posterior probabilities do not tell which 2fi's are significant. Since the frequentist approach identifies the EF interaction as significant, we perform a second Bayesian analysis by treating the EF interaction as a new factor. The resulting posterior probability plot in Figure 9.1 (right) shows that EF is as significant as factors E and F. Factor H is also significant, but not as significant as E, F and EF. The finding is consistent with the frequentist approach.

Chipman, Hamada and Wu (1997) proposed a more sophisticated Bayesian approach for analyzing data with complex aliasing. They employed a Gibbs sampler to perform an efficient stochastic search of the model space, whereas Box and Meyer (1993) evaluated all possible models, which could require intensive computation for large data sets. In addition, Chipman et al. (1997) carefully implemented the effect sparsity and effect heredity principles with hierarchical models. They further introduced two types of effect heredity: weak and strong heredity. Under weak heredity, a 2-factor interaction is important only if at least one of its component factors is significant, while under strong heredity, both of its component factors have to be significant.

Yuan, Joseph and Lin (2007) proposed an efficient variable selection approach based on the least angle regression (LARS) algorithm of Efron et al. (2004). They modified the LARS

Figure 9.1: The posterior probability plot for the HPLC Experiment (left) with the original factors and (right) with the EF interaction.



algorithm so that heredity principles can be taken into account in the variable selection process.

When all factors are quantitative, it is natural to consider a polynomial model to explore the relationship between the response and factors. For k quantitative factors, denoted by $x_1, \ldots, x_k$, the second-order model is

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{1=i<j}^{k} \beta_{ij} x_i x_j + \boldsymbol{\epsilon}, \tag{9.22}$$

where $\beta_0, \beta_i, \beta_{ii}, \beta_{ij}$ are unknown parameters and $\boldsymbol{\epsilon}$ is the error term. For the pure quadratic terms β_{ii} to be estimated, all the factors must have more than two levels. The second-order model (9.22) has $p = (k+1)(k+2)/2$ parameters. When the run size n is less than p, we cannot estimate all the parameters in (9.22). A traditional approach is to assume that the bilinear (or interaction) terms β_{ij} are negligible and fit a model with the linear and pure quadratic terms only, which is the main effects model for three-level designs. However, non-negligible interaction terms will bias the estimate of linear and pure quadratic terms. A better approach is to use the Hamada-Wu strategy and perform variable selection guided

by the effect heredity principle.

Cheng and Wu (2001) proposed an alternative analysis strategy in order to achieve the dual purpose of factor screening and response surface exploration using a single design. Their analysis strategy has two stages:

Stage 1. Perform factor screening and identify important factors.

Stage 2. Fit a second-order model for the factors identified in stage 1.

Various screening analyses can be utilized in stage 1, such as the conventional ANOVA or half-normal plots of the main effects, which include both linear and pure quadratic terms for 3-level factors. Their analysis strategy again assumes that effect sparsity and effect heredity principles hold. They reanalyzed a PVC insulation experiment reported by Taguchi (1987) that used a regular 27-run design with nine 3-level factors. They identified a significant linear-by-linear interaction effect which was missed by Taguchi. Xu, Cheng and Wu (2004) gave another example which uses an 18-run OA with one 2-level factor and seven 3-level factors.

Finally, for SSDs we typically consider the main effects only with the assumption that all 2fi's are negligible. In principle, any variable selection procedures can be used for analyzing SSDs. Many analysis strategies have been used to analyze SSDs. The list includes forward stepwise regressions (Lin 1993, Westfall et al. 1998), all subsets regressions (Abraham et al. 1999), Bayesian variable selections (Chipman et al. 1997), penalized least squares (Li and Lin 2003), partial least squares (Zhang et al. 2007), the Dantzig selector (Phoa, Pan and Xu 2009), and many others.

9.9 Concluding Remarks

We give an overview of recent developments in nonregular fractional and supersaturated designs in this chapter. In summary, nonregular designs are more flexible, require smaller numbers of runs, and have better statistical properties than regular designs. Yet the analysis

of nonregular designs is more complicated due to the partial aliasing among the effects. Xu, Phoa and Wong (2009) highlighted some future directions of research, from applications and analysis of nonregular designs, to construction of good nonregular designs with large run sizes and optimality results with respect to the generalized resolution.

One underlying assumption for the GMA criterion in Section 9.4.1 is that factor levels are regarded as nominal symbols. This is appropriate for experiments with qualitative factors where there is no ordering among the levels. However, for experiments with quantitative factors, polynomial models such as (9.22) or other models are often used to describe the relationship between the response and the factors. In these circumstances, permuting levels for one or more factors can lead to designs with different geometrical structures and statistical properties. An important question is how the levels should be permuted for quantitative factors. There are some recent developments on this topic; see Tang, Xu and Lin (2012), Tang and Xu (2013, 2014), and Zhou and Xu (2014). However, more work needs to be done.

Acknowledgments

The research was supported by National Science Foundation grant DMS-1106854.

References

1. Abraham, B., Chipman, H., Vijayan, K., 1999. Some risks in the construction and analysis of supersaturated designs. *Technometrics*, 41, 135-141.

2. Addelman, S. and Kempthorne, O. (1961). Some main-effect plans and orthogonal arrays of strength two. *Ann. Math. Statist.* **32** 1167–1176.

3. Ai, M. Y. and Zhang, R. C. (2004). Projection justification of generalized minimum aberration for asymmetrical fractional factorial designs. *Metrika*, **60**, 279-285.

4. Ai, M. Y., Li, P. F. and Zhang, R. C. (2005). Optimal criteria and equivalence for nonregular fractional factorial designs. *Metrika*, **62**, 73-83.

5. Booth, K. H. V. and Cox, D. R. (1962). Some systematic supersaturated designs. *Technometrics* **4** 489–495.

6. Box, G. E. P. and Draper, N. R. (1987). *Empirical Model-building and Response Surfaces*. New York: Wiley.

7. Box, G. E. P. and Meyer, R. D. (1993). Finding the active factors in fractionated screening experiments. *J. Quality Technology*, **25**, 94–105.

8. Box, G. E. P. and Tyssedal, J. (1996). Projective properties of certain orthogonal arrays. *Biometrika*, **83**, 950–955.

9. Box, G. E. P., Hunter, W. G. and Hunter, J. S. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery*, 2nd ed. New York: Wiley.

10. Bulutoglu, D. A. and Cheng, C. S. (2003). Hidden projection properties of some nonregular fractional factorial designs and their applications. *Ann. Statist.*, **31**, 1012-1026.

11. Bulutoglu, D. A. and Cheng, C. S. (2004). Construction of $E(s^2)$-optimal supersaturated designs. *Ann. Statist.* **32** 1662–1678.

12. Butler, N. A. (2003). Minimum aberration construction results for nonregular two-level fractional factorial designs. *Biometrika*, **90**, 891–898.

13. Butler, N. A. (2004). Minimum G_2-aberration properties of two-level foldover designs. *Statist. Probab. Lett.*, **67**, 121–132.

14. Butler, N. A. (2005). Generalised minimum aberration construction results for symmetrical orthogonal arrays. *Biometrika*, **92**, 485-491.

15. Butler, N. A., Mead, R., Eskridge, K. M. and Gilmour, S. G. (2001). A general method of constructing $E(s^2)$-optimal supersaturated designs. *J. R. Stat. Soc. Ser. B* **63** 621–632.

16. Chen, J., Sun, D. X. and Wu, C. F. J. (1993). A catalogue of two-level and three-level fractional factorial designs with small runs. *Internat. Statist. Rev.*, **61**, 131–145.

17. Cheng, C. S. (1980). Orthogonal arrays with variable numbers of symbols. *Ann. Statist.*, **8**, 447–453.

18. Cheng, C. S. (1995). Some projection properties of orthogonal arrays. *Ann. Statist.*, **23**, 1223–1233.

19. Cheng, C. S. (1997). $E(s^2)$-optimal supersaturated designs. *Statist. Sinica* **7** 929–939.

20. Cheng, C. S. (1998). Some hidden projection properties of orthogonal arrays with strength three. *Biometrika*, **85**, 491–495.

21. Cheng, C. S., Deng, L. Y. and Tang, B. (2002). Generalized minimum aberration and design efficiency for nonregular fractional factorial designs. *Statist. Sinica*, **12**, 991–1000.

22. Cheng, C. S., Steinberg, D. M. and Sun, D. X. (1999). Minimum aberration and model robustness for two-level fractional factorial designs. *J. Roy. Statist. Soc. Ser. B*, **61**, 85–93.

23. Cheng, S. W. and Wu, C. F. J. (2001). Factor screening and response surface exploration (with discussion). *Statist. Sinica*, **11**, 553–604.

24. Cheng, S. W. and Ye, K. Q. (2004). Geometric isomorphism and minimum aberration for factorial designs with quantitative factors. *Ann. Statist.*, **32**, 2168–2185.

25. Cheng, S. W., Li, W. and Ye, K. Q. (2004). Blocked nonregular two-level factorial designs. *Technometrics*, **46**, 269–279.

26. Chipman, H., Hamada, M. and Wu, C.F.J. (1997). A Bayesian variable-selection approach for analyzing designed experiments with complex aliasing. *Technometrics*, **39**, 372–381.

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

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

29. Ding, X., Xu, H., Hopper, C., Yang, J., and Ho, C.-M. (2013). Use of fractional factorial designs in antiviral drug studies. *Quality and Reliability Engineering International*, **29**, 299–304. DOI:10.1002/qre.1308

30. Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression. *Ann. Statist.*, **32**, 407-499.

31. Fang, K. T. and Mukerjee, R. (2000). A connection between uniformity and aberration in regular fractions of two-level factorials. *Biometrika*, **87**, 193–198.

32. Fang, K. T. and Wang, Y. (1994). *Number-Theoretic Methods in Statistics*. London: Chapman and Hall.

33. Fang, K. T., Li, R. and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. London: Chapman and Hall/CRC.

34. Fang, K. T., Lin, D. K. J., Winker, P. and Zhang, Y. (2000). Uniform design: Theory and application. *Technometrics*, **42**, 237-248.

35. Georgiou, S. D. (2014). Supersaturated designs: A review of their construction and analysis. *Journal of Statistical Planning and Inference*, **144**, 92–109.

36. Grömping, U. and Xu, H. (2014). Generalized resolution for orthogonal arrays. *Ann. Statist.*, in press.

37. Groten, J.P., Schoen, E.D. and Feron, V.J. (1996). Use of factorial designs in combination toxicity studies. *Food and Chemical Toxicology*, 34, 1083-1089.

38. Groten, J.P., Schoen, E.D., Kuper, C.F., van Bladeren, P.J., Van Zorge, J.A. and Feron, V.J. (1997). Subacute toxicity of a mixture of nine chemicals in rats: detecting interactive effects with a fractionated two-level factorial design. *Fundamental and Applied Toxicology*, 36, 13-29.

39. Hamada, M. and Wu, C. F. J. (1992). Analysis of designed experiments with complex aliasing. *J. Quality Technology*, **24**, 130–137.

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

41. Hickernell, F. J. (1998). A generalized discrepancy and quadrature error bound. *Math. Comp.*, **67**, 299–322.

42. Hickernell, F. J. and Liu, M. Q. (2002). Uniform designs limit aliasing. *Biometrika*, **89**, 893–904.

43. Li, R. and Lin, D.K.J. (2003). Analysis methods for supersaturated design: some comparisons. *Journal of Data Science*, 1, 249-260.

44. Li, W. W. and Wu, C. F. J. (1997). Columnwise-pairwise algorithms with applications to the construction of supersaturated designs. *Technometrics* **39** 171–179.

45. Lin, D. K. J. (1993). A new class of supersaturated designs. *Technometrics*, **35**, 28–31.

46. Lin, D. K. J. (1995). Generating systematic supersaturated designs. *Technometrics* **37** 213–225.

47. Ma, C. X. and Fang, K. T. (2001). A note on generalized aberration in factorial designs. *Metrika*, **53**, 85–93.

48. MacWilliams, F. J. and Sloane, N. J. A. (1977). *The Theory of Error-correcting Codes*. Amsterdam: North-Holland.

49. Mandal, A. and Mukerjee, R. (2005). Design efficiency under model uncertainty for nonregular fractions of general factorials. *Statist. Sinica*, **15**, 697-707.

50. Mukerjee, R. and Wu, C. F. J. (1995). On the existence of saturated and nearly saturated asymmetrical orthogonal arrays. *Ann. Statist.* **23** 2102–2115.

51. Nguyen, N.-K. (1996). An algorithmic approach to constructing supersaturated designs. *Technometrics* **38** 69–73.

52. Phoa, F. K. H. (2012). A code arithmetic approach for quaternary code designs and its application to (1/64)th-fractions. *Ann. Statist.*, **40**, 3161–3175.

53. Phoa, F. K. H. and Xu, H. (2009). Quarter-fraction factorial designs constructed via quaternary codes. *Ann. Statist.*, **37**, No. 5A, 2561–2581.

54. Phoa, F. K. H., Mukerjee, R., and Xu, H. (2012). One-Eighth- and One-Sixteenth-Fraction Quaternary Code Designs With High Resolution. *J. Statist. Plann. Inference*, 142, 1073-1080.

55. Phoa, F. K. H., Pan, Y.-H. and Xu, H. (2009). Analysis of supersaturated designs via the Dantzig selector. *J. Statist. Plann. Inference*, **139**, 2362–2372.

56. Phoa, F. K. H., Wong, W. K., and Xu, H. (2009). The Need of Considering the Interactions in the Analysis of Screening Designs. *Journal of Chemometrics*, 23, 545-553.

57. Phoa, F. K. H., Xu, H., and Wong, W. K. (2009). The Use of Nonregular Fractional Factorial Designs in Combination Toxicity Studies. *Food and Chemical Toxicology*, 47, 2183-2188.

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

59. Rao, C. R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *J. Roy. Statist. Soc. Ser. B*, **9**, 128–139.

60. Satterthwaite, F. E. (1959). Random balance experimentation. *Technometrics* **1** 111–137.

61. Schoen, E. D. (2009). All orthogonal arrays with 18 runs. *Quality and Reliability Engineering International*, **25**, 467–480.

62. Schoen, E. D., Eendebak, P. T. and Nguyen, M. V. M. (2010). Complete enumeration of pure-level and mixed-level orthogonal arrays. *J. Combin. Designs*, 18: 123140.

63. Sun, D. X., Li, W. and Ye, K. Q. (2002). An algorithm for sequentially constructing nonisomorphic orthogonal designs and its applications. Technical report SUNYSB-AMS-02-13, Department of Applied Mathematics and Statistics, SUNY at Stony Brook.

64. Tang, B. (2001). Theory of J-characteristics for fractional factorial designs and projection justification of minimum G_2-aberration. *Biometrika*, **88**, 401–407.

65. Tang, B. and Deng, L. Y. (1999). Minimum G_2-aberration for non-regular fractional factorial designs. *Ann. Statist.*, **27**, 1914–1926.

66. Tang, B. and Wu, C. F. J. (1997). A method for constructing supersaturated designs and its $E(s^2)$ optimality. *Canad. J. Statist.* **25** 191–201.

67. Tang, Y. and Xu, H. (2013). An effective construction method for multi-level uniform designs. *J. Statist. Plann. Inference*, **143**, 1583-1589. doi: 10.1016/j.jspi.2013.04.009

68. Tang, Y. and Xu, H. (2014). Permuting Regular Fractional Factorial Designs for Screening Quantitative Factors. *Biometrika*, in press. doi: 10.1093/biomet/ast073

69. Tang, Y., Xu, H., and Lin, D. K. J. (2012). Uniform fractional factorial designs. *Ann. Statist.*, 40, 891-907.

70. Vander Heyden, Y., Jimidar, M., Hund, E., Niemeijer, N., Peeters, R., Smeyers-Verbeke, J., Massart, D.L., and Hoogmartens, J. (1999). Determination of system suitability limits with a robustness test. *Journal of Chromatography A*, **845**: 145-154.

71. Westfall, P. H., Young, S. S. and Lin, D. K. J. (1998). Forward selection error control in the analysis of supersaturated designs. *Statist. Sinica*, **8**, 101-117.

72. Wu, C. F. J. (1993). Construction of supersaturated designs through partially aliased interactions. *Biometrika*, **80**, 661–669.

73. Wu, C. F. J. and Hamada, M. (2009). *Experiments: Planning, Analysis and Optimization*, 2nd ed., New York: Wiley.

74. Xu, H. (2002). An algorithm for constructing orthogonal and nearly-orthogonal arrays with mixed levels and small runs. *Technometrics*, **44**, 356–368.

75. Xu, H. (2003). Minimum moment aberration for nonregular designs and supersaturated designs. *Statist. Sinica*, **13**, 691–708.

76. Xu, H. (2005a). Some nonregular designs from the Nordstrom and Robinson code and their statistical properties. *Biometrika*, **92**, 385–397.

77. Xu, H. (2005b). A catalogue of three-level regular fractional factorial designs. *Metrika*, **62**, 259–281.

78. Xu, H. (2006). Blocked regular fractional factorial designs with minimum aberration. *Ann. Statist.*, **34**, 2534–2553.

79. Xu, H. (2009). Algorithmic construction of efficient fractional factorial designs with large run sizes. *Technometrics*, **51**, 262–277.

80. Xu, H. and Deng, L. Y. (2005). Moment aberration projection for nonregular fractional factorial designs. *Technometrics*, **47**, 121–131.

81. Xu, H. and Lau, S. (2006). Minimum aberration blocking schemes for two- and three-level fractional factorial designs. *J. Statist. Plann. Inference*, **136**, 4088–4118.

82. Xu, H. and Wong, A. (2007). Two-level nonregular designs from quaternary linear codes. *Statist. Sinica*, **17**, 1191–1213.

83. Xu, H. and Wu, C. F. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Statist.*, **29**, 1066–1077.

84. Xu, H. and Wu, C. F. J. (2005). Construction of optimal multi-level supersaturated designs. *Ann. Statist.*, **33**, 2811–2836.

85. Xu, H., Cheng, S. W. and Wu, C. F. J. (2004). Optimal projective three-level designs for factor screening and interaction detection. *Technometrics*, **46**, 280–292.

86. Xu, H., Phoa, F. K. H., and Wong, W. K. (2009). Recent Developments in Nonregular Fractional Factorial Designs. *Statistics Surveys*, 3, 18-46.

87. Ye, K. Q. (2004). A note on regular fractional factorial designs. *Statist. Sinica*, **14**, 1069–1074.

88. Ye, K. Q., Park, D., Li, W., and Dean, A. M. (2008). Construction and classification of orthogonal arrays with small numbers of runs. *Statistics and Applications*, **6**, (New Series), 5–15.

89. Yuan, M., Joseph, V. R. and Lin, Y. (2007). An efficient variable selection approach for analyzing designed experiments. *Technometrics*, **49**, 430-439.

90. Zhang, Q.Z., Zhang, R.C., and Liu, M.Q. (2007). A method for screening active effects in supersaturated designs. *J. Statist. Plann. Inference*, 137(9), 235-248.

91. Zhang, R.-C., Phoa, F. K. H., Mukerjee, R., and Xu, H. (2011). A Trigonometric Approach to Quaternary Code Designs with Application to One-Eighth and One-Sixteenth Fractions. *Ann. Statist.*, 39, No. 2, 931-955.

92. Zhou, Y.-D. and Xu, H. (2014). Space-filling fractional factorial designs. *J. Amer. Statist. Assoc.*, in press.