Statistica Sinica **27** (2017), 1347-1369
doi:https://doi.org/10.5705/ss.202015.0161

GENERATORS FOR NONREGULAR 2^{k-p} DESIGNS

Robert W. Mee

University of Tennessee and Nankai University

Abstract: There exist nonregular two-level designs with run sizes a power of 2. Many of these designs have a defining relation. This article characterizes nonregular two-level fractional factorial designs according to three types. First, there are designs that can be constructed using generators that are linear combinations of orthogonal interactions from a subset of the factors. All possible generators for 16- and 32-run designs are identified. A second type of orthogonal two-level designs has partial replication, which requires adding one or more dummy factors to obtain generators. Intermediate to these two types are orthogonal designs that have no partial replication, but require augmentation in order to obtain generators. This classification and subsequent insight benefit the construction and characterization of nonregular designs. Designs of the first type have a defining relation that is easily produced from the generators. For the other cases, generators are useful for obtaining the indicator function or the extended word length pattern. Given familiarity with regular two-level fractional factorial designs, this article can serve as a bridge to understanding nonregular fractions.

Key words and phrases: Defining relation, extended word length pattern, fractional factorial, indicator function, orthogonal array, partial replication.

1. Introduction

Let OA(n, 2^k) denote an n-run orthogonal array with k two-level factors. The number of non-isomorphic OA(16, 2^k) is known for each $k = 3, \ldots, 15$; the counts are 3, 5, 11, 27, 55, 80, 87, 78, 58, 36, 18, 10, 5, respectively (Xu and Deng (2005)). The five OA(16, 2^{15}) come from the five Hadamard matrices of order 16 (Hall (1961)) and all 16-run orthogonal arrays for $k < 15$ can be obtained as a projection from one or more of these five matrices, where projection means that some columns are omitted. Johnson and Jones (2011) studied these OA(16, 2^k) in detail for $k = 6$, 7, and 8, showing how all can be constructed using generators. This construction is only available for OA(n, 2^k) when n is a power of 2; the standard notation 2^{k-p} will be used to denote orthogonal two-level fractions with such run sizes. Since Johnson and Jones (2011) have detailed the 2^{6-2}, 2^{7-3}, and 2^{8-4} designs, this article will emphasize larger designs, where the possible generator constructions are more varied. Kharaghani and Tayfeh-Rezaie (2013) determined that there are 13,710,027 Hadamard matrices of order 32, an astronomical increase over the 16-run case where only five exist. Schoen, Eendebak and Nguyen (2010) estimated that a complete enumeration of all OA(32, 2^k) would require 50,000 years of computing time by their method. Given so many possibilities, understanding their structure will surely benefit the search for attractive 32-run and larger designs, according to various metrics. Recently, Eendebak and Schoen (2017) have succeeded in enumerating all isomorphic classes of OA(32, 2^7), and this article will focus much attention on these 2^{7-2} designs.

Each 2^{k-p} design is classified as either regular or nonregular. For regular 2^{k-p} designs, pairs of main effect and interaction contrasts have correlations of 0, 1, or -1. Unreplicated regular 2^{k-p} fractions necessarily project into a full factorial in $b = k-p$ *basic* factors. The remaining factors are generated using p individual basic-factor-interactions. The defining relation is constructed from these p generators, and represents the $2^p - 1$ interactions that are identically $+1$. For example, consider the 2^{7-2} design where the letters A–E denote the basic factors and the $p = 2$ remaining factors are generated as $F = ABCD$ and $G = CDE$. The defining relation for this 2^{7-2} is $I = ABCDF = CDEG = ABEFG$, since $ABCDF \times CDEG = ABC^2D^2EFG = ABEFG$. This design has resolution IV, with word length pattern $wlp = (A_3, \ldots, A_7) = (0, 1, 2, 0, 0)$, where A_q denotes the number of q-factor interactions in the defining relation. The $2^p - 1 = 3$ interactions appearing in the defining relation are identically $+1$ for all 32 treatment combinations in this quarter fraction. The four-factor interaction $CDEG$ creates three pairs of aliased two-factor interactions, $CD = EG$, $CE = DG$, and $CG = DE$. As a result, one cannot estimate the full two-factor interaction model based only on this design. Let $\mathbf{X}$ denote the model matrix for the full two-factor interaction model, with its $r = 1 + k(k+1)/2$ columns; for $k = 7$, $r = 29$. Due to the three singularities caused by $CDEG = 1$, rank($\mathbf{X}$) = 26 for this design. Further details for regular 2^{k-p} designs are described in experimental design books, such as (Wu and Hamada, 2009, Chap. 4), (Montgomery, 2012, Chap. 8), and Mee and Dean (2015).

Nonregular 2^{k-p} designs differ from regular fractions in that some pairs of main effect and interaction contrasts have correlations with magnitude between 0 and 1. Despite this difference, if at least one subset of b factors forms a full 2^b factorial, the full factorial model matrix for this subset of factors forms an $n \times n$ matrix with orthogonal columns and the remaining p factors can be written as linear combinations of basic-factor-interactions. This implies that all such designs can be expressed by a set of p generators and/or by a defining relation. Such designs are the focus of Section 2. Consider as an example the 2^{7-2} design with

$$\begin{aligned} F &= 0.5(ABC + ABCD + ABCE - ABCDE) = 0.5ABC(1 + D + E - DE), \\ G &= 0.5(ADE + ABDE + CDE - BCDE) \quad\;\; = 0.5DE(A + AB + C - BC). \end{aligned} \tag{1.1}$$

Just as the defining relation for a regular fraction is constructed from its generators, the same is possible here. The defining relation for this nonregular design can be written as:

$$\begin{aligned} I &= 0.5ABCF(1 + D + E - DE) = 0.5DEG(A + AB + C - BC) \\ &= 0.25FG(A - C - AB - AD - AE - BC + CD + CE + CDE \\ &\quad - ADE + BCE + ABE + BCD + ABD + ABDE + BCDE), \end{aligned} \tag{1.2}$$

where the expression following the last equal sign is the generalized interaction, $0.5ABCF(1 + D + E - DE) \times 0.5DEG(A + AB + C - BC)$.

Providing generators and a defining relation for nonregular 2^{k-p} designs will help those familiar only with regular 2^{k-p} fractions understand nonregular designs. For instance, defining relations for 2^{k-p} designs display all the J-characteristics of a design (Tang (2001)). Inspecting (1.2), we see that this design has no perfect correlations among the main effect and interaction columns, since no term in (1.2) has a coefficient of ± 1. The presence of $0.25FGA$ and $-0.25FGC$ in (1.2) reveals that the design does not project into an equally replicated 2^3 in $\{A, F, G\}$ or $\{C, F, G\}$. So this design is only strength 2, while its generalized resolution, as defined by Deng and Tang (1999), is $3+(1-0.25) = 3.75$. Both the generalized word length pattern (Tang and Deng (1999)) and the extended word length pattern (Li, Lin and Ye (2003)) can be obtained from (1.2). This design is inferior to the minimum aberration regular 2^{7-2} in terms of its generalized resolution, but the two-factor interaction model is estimable from this design; that is, its rank$(\mathbf{X}) = 29$.

Concatenating smaller orthogonal arrays is another method for constructing nonregular 2^{k-p} designs. For instance, consider the three designs in Table 1, which were each constructed by concatenating a 12-run array and a 20-run array. The 20-run array shared by both Designs I and III is the minimum G-aberration design found by Xu and Deng (2005), an OA(20, 2^7) that is not a projection of any Hadamard design. To this we add different versions of the OA(12, 2^7). If we use the first 12 rows in Table 1, we get a design for which the first five columns form a full 2^5. So Table 1's Design I can be constructed by concatenation, or by using generators for F and G. However, neither Design II nor III projects into

Table 1. Three nonregular OA(32, 2^7) created by concatenation.

Design I: First 32 rows (Type 1) Design III: Last 32 rows (Type 3)							Design II: (Type 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	−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	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	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	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	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	−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	−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	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	−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	−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	−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	−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	−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
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	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	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	−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	−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	−1	−1	1	−1	1							
1	−1	−1	1	−1	−1	1							
−1	−1	−1	1	1	1	−1							

a full 2^5. Although Design II has 32 distinct rows, each 5-factor projection has 28 or fewer distinct rows. That Design III contains no 2^5 factorial projection is immediately obvious, since the 2^{7-2} only has 30 distinct rows. Design III is formed by the minimum G-aberration OA(20, 2^7) plus the last 12 rows, which were obtained by reversing the sign for columns 2-6 from the first 12 rows. The last two rows of the OA(20, 2^7) match the first two rows of the second OA(12, 2^7). So Design III has 2 degrees of freedom for pure error, yet it still supports estimation of the two-factor interaction model, which requires a minimum of $r = 29$ distinct rows.

We may refer to the 2^{k-p} designs as of one of three types, as illustrated in Table 1. Type 1 designs can be constructed using generators that are linear combinations of basic factor interactions. All unreplicated regular fractions and seemingly most nonregular fractions are Type 1. Generators and the corresponding defining relation provide a basis both for construction and for understanding these designs. We label as Type 2 all 2^{k-p} designs that have $2^b = n$ distinct rows, but no b-factor projection has n distinct rows. Type 3 designs are those with partial replication. We use this numbering because Type 2 designs are intermediate to the other two types. Like Type 1 designs, Type 2 designs have no replication, but like Type 3 designs, they have no subset of b factors to serve as a basis. When dropping a factor from a Type 1 design, the projection may be Type 1, 2, 3, or even a fully replicated $2^{(k-1)-p}$ design; that is, the projection may or may not have a b-factor basis and if it has no basis, it may or may not have replication. The reverse is not possible. A Type 3 design cannot project to Type 1 or 2, since it necessarily has some replication. Similarly, projections of Type 2 designs cannot be Type 1, but may be Type 3. Design Types 2 and 3 are discussed separately in Sections 3 and 4, respectively. While for both types it is necessary to add one or more dummy factors to create an orthogonal basis for generators, the possible relevance of partial replication (Liao and Chai (2004)) justifies their separate treatment.

All OA(16, 2^k) with $k \geq 8$ are Type 1. For $n = 16$ and $5 \leq k \leq 7$, 78 are Type 1, 1 is Type 2, 11 are Type 3, while 3 are completely replicated fractions.

Eendebak and Schoen (2017) report enumerating all 530,469,996 isomorphic classes of 2^{7-2} orthogonal designs, 395,932,754 (75%) of which have rank($\mathbf{X}$) = 29. The three Table 1 designs are among this number, but their D-efficiencies, $\det(\mathbf{X}^T\mathbf{X}/32)^{1/29}$, are only 0.7342, 0.5935, and 0.6215, respectively. In contrast, the top 100 2^{7-2} designs in terms of D-efficiency have $0.8156 \leq$ D-eff ≤ 0.8432.

Ninety-seven of these top 100 are Type I designs and the other three are

Type II; thus, none of the top 100 have partial replication. In the next section, we discuss several D-efficient Type I designs, as well as 2^{7-2} designs with rank($\mathbf{X}$) < 29.

2. Type 1: Nonregular 2^{k-p} Designs with Linear Combination Generators

2.1. Generators, defining relations, and word length patterns

Let $\mathbf{H}_n$ denote the order-n Sylvester-type Hadamard matrix (Hedayat, Sloane and Stufken, 1999, p. 149), which may be obtained recursively using $\mathbf{H}_1 = [1]$ and $\mathbf{H}_n = [\mathbf{H}_{n/2},\ \mathbf{H}_{n/2};\ \mathbf{H}_{n/2},\ -\mathbf{H}_{n/2}]$. The first column of $\mathbf{H}_n$ is a column of 1's. Omitting this, we have the regular OA(n, 2^{n-1}) with columns in Yates order (Mee, 2009, pp. 190, 485); denote this saturated main effect design as $\mathbf{S}_n$. Let $\mathbf{O}$ denote any $(n-1) \times k$ matrix such that $\mathbf{O}^T\mathbf{O} = \mathbf{I}_k$, the identity matrix. Then $\mathbf{D} = \mathbf{S}_n\mathbf{O}$ is an $n \times k$ orthogonal design. However, to guarantee that $\mathbf{D}$ is a two-level design we must impose further conditions on $\mathbf{O}$.

For all unreplicated regular 2^{k-p} designs of resolution III or higher and all nonregular 2^{k-p} of strength 2 or higher that project to a full factorial in b factors, where $2^b = n$, we assume, without loss of generality that the basic columns correspond to Yates columns 1, 2, 4, 8, $\ldots$, 2^{b-1} and that each of the first b columns of $\mathbf{O}$ contain a single +1 corresponding to these columns of $\mathbf{S}_n$, with all remaining elements of those b columns being zero. If we reorder the columns of $\mathbf{S}_n$ by moving all b basic columns to the left, then $\mathbf{O}$ may be written as the block diagonal matrix

$$\mathbf{O} = \begin{bmatrix} \mathbf{I}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix},$$

where the matrix $\mathbf{G}$ defines the generators for the remaining $p = k - b$ columns.

Table 2 summarizes five different 2^{7-2} designs with simple generators involving either one or four interactions, one regular and four nonregular. For the regular 2^{7-2} design discussed in Section 1, the columns of $\mathbf{G}$ have a single +1 in the rows corresponding to the interactions $ABCD$ and CDE, respectively.

For nonregular designs having a subset of b columns forming a full unreplicated 2^b, the matrix $\mathbf{G}$ contains at least one column with four or more non-zero elements corresponding to interactions of the basic factors, as we saw in (1.1). The Appendix justifies this assertion. For $\mathbf{G}$ to produce an orthogonal design, we already have that $\mathbf{G}^T\mathbf{G} = \mathbf{I}_p$. We now consider the conditions for $\mathbf{G}$ that ensure levels for $\mathbf{D}$ of ± 1 exclusively. Choose any column g of $\mathbf{G}$, let h denote

Table 2. Generators for five Type 1 2^{7-2} designs with simple generators.

Design Description	Generators	Defining Relation	Gen. res.; rank($\mathbf{X}$) *gwlp*
1 Min aber.	$F = ABCD$ $G = CDE$	$I = ABCDF$ $= CDEG$ $= ABEFG$	4; 26 (0,1,2,0)
2 Min G-aber.	$F = ABCD$ $G = 0.5E(AB + AC$ $+BD - CD)$	$I = ABCDF$ $= 0.5EG(AB + AC$ $+BD - CD)$ $= 0.5EFG(CD$ $+BD + AC - AB)$	4.5; 28 (0,1,2,0)
3 D-eff=0.7967	$F = 0.5ABC(1 + D$ $+E - DE)$ $G = 0.5DE(A + AB$ $+C - BC)$	See Equation (1.2)	3.75; 29 (1/8, 9/8, 11/8, 3/8)
4	$F = 0.5DE(B + C$ $+AC - AB)$ $G = 0.5BE(C + D$ $+AD - AC)$	$I = 0.5DEF(B + C$ $+AC - AB)$ $= 0.5BEG(C + D$ $+AD - AC)$ $= 0.5CFG(B + D$ $+AB - AD)$	4.5; 26 (0, 1.5, 1.5, 0)
5 $B_4 = 0$	$F = ABCDE$ $G = 0.5A(B + C$ $+D - BCD)$	$I = ABCDEF$ $= 0.5AG(B + C$ $+D - BCD)$ $= 0.5EFG(BC$ $+BD + CD - 1)$	3.5; 28 (1, 0, 1, 1)

Gen. res., Generalized resolution; *gwlp*, generalized word length pattern = $(B_3, \ldots, B_k)$

the non-zero elements of g, and let $\mathbf{S}_h$ denote the columns of $\mathbf{S}_n$ corresponding to these elements of h. Define $e = [e_1, \ldots, e_n]^T$, where $e_i = \pm 1$ $(i = 1, \ldots, n)$, $e_1 + \ldots + e_n = 0$, and let $\mathcal{E}$ be the set of all such vectors. For the generated column produced by g (or h) to have levels exclusively ± 1 requires that $\mathbf{S}_h h = e$ for some $e \in \mathcal{E}$. A solution h exists if and only if e lies in the column space of $\mathbf{S}_h$.

For all OA(16, 2^k) that project to a full 2^4 factorial, every generator involves either 1 or 4 interactions. This was confirmed by a complete enumeration of all possible $e \in \mathcal{E}$. In every case of four interactions, $\mathbf{S}_h$ is a replicated 2^{4-1} design of resolution IV. As in (1.1), each basic factor appears two or four times and the product of the four interactions is -1 for all n treatment combinations. The eight distinct rows of this 2^{4-1} can be written as $[\mathbf{E}; -\mathbf{E}]$, where $\mathbf{E} = [1, 1, 1, -1; 1, 1, -1, 1; 1, -1, 1, 1; -1, 1, 1, 1]$. Clearly, if $\mathbf{E}h$ is a vector of ± 1's, then so is $-\mathbf{E}h$. Since $\mathbf{E}^{-1}$ exists, so does h. The solution h is a length-4 vector of ± 0.5's with an odd number of $+$'s and $-$'s, or it is a single 1 or -1. As is true for every case, such solutions h sum to ± 1 and $h^T h = 1$. The Appendix lists the classes of four-interaction generators for orthogonal arrays with $n = 16$ and 32. For $n = 32$, we will see that other generators h involving more than four interactions are possible.

Table 2's Designs 2–5 are nonregular 2^{7-2}. Design 2, with $F = ABCD$ and $G = 0.5E(AB + AC + BD - CD)$, is the minimum G-aberration design (Deng and Tang (1999)), isomorphic to Schoen and Mee (2012) Design 7.16; see also Xu (2005). G-aberration is based on the extended word length pattern, which for nonregular designs in this section can be obtained directly from the defining relation. The defining relation for this nonregular design is

$$I = ABCDF = 0.5EG(AB+AC+BD-CD) = 0.5EFG(CD+BD+AC-AB).$$

This defining relation contains four 4-factor interactions and four 5-factor interactions with coefficients ± 0.5, plus $ABCDF$ with coefficient 1. Thus, the extended word length pattern is

$$ewlp = \{f_{4.5} = 4; f_{5.0} = 1, f_{5.5} = 4\}, \tag{2.1}$$

where $f_{t.s}$ is the number of t-factor interactions summing to $\pm(1 - 0.s)n$. Since the lowest-order count is $f_{4.5}$, the generalized resolution is 4.5. Minimum G-aberration is based on sequentially minimizing the lower-order counts in $ewlp$. Design 2 has minimum G-aberration because every other 2^{7-2} has either lower generalized resolution or $f_{4.5} > 4$. Even though this design avoids having a 4-factor interaction with a coefficient of 1, it does not permit estimation of the two-factor interaction model, since $\text{rank}(\mathbf{X}) = 28$. The singularity is revealed in the defining relation; just multiply $I = 0.5EG(AB + AC + BD - CD)$ by EG. (Every other strength 3 2^{7-2} has $\text{rank}(\mathbf{X}) \leq 26$.)

Designs 3 and 4 both involve two generators that are linear combinations of 3-factor and higher-order interactions. For Design 3, with generators given by (1.1), the generalized interaction appearing in the last two lines of the defining relation (1.2) has 16 terms, each with coefficient ± 0.25. Thus there are 24 interactions that do not sum to zero, 8 with a coefficient of ± 0.5 and 16 with a coefficient of ± 0.25. Design 3's extended word length pattern is

$$ewlp = (f_{3.75} = 2; f_{4.5} = 3, f_{4.75} = 6; f_{5.5} = 4, f_{5.75} = 6; f_{6.5} = 1, f_{6.75} = 2).$$

Just as minimum aberration sequentially minimizes the wlp, minimum G_2 aberration (Tang and Deng (1999)) sequentially minimizes a design's generalized word length pattern ($gwlp$), where $gwlp$ is a vector $(B_3, \ldots, B_k)$, with $B_t =$

$\sum_s f_{t.s}(1-0.s)^2$ replacing A_t. For example,

$$gwlp = \{2(0.25)^2, 3(0.5)^2+6(0.25)^2, 4(0.5)^2+6(0.25)^2, 1(0.5)^2+2(0.25)^2\}$$
$$= (0.125, 1.125, 1.375, 0.375)$$

for Design 3. Butler (2003) provides a simpler means to compute *gwlp* using moments of the row coincidence matrix. While Table 2's Design 3, with D-eff = 0.7967, is better than all Table 1 designs, this is not among the more D-efficient designs found by Eendebak and Schoen (2017), which we will study in the next subsection.

In contrast to Design 3, the product of Design 4's two 4-interaction generators produces eight pairs of interactions, four pairs with opposite signs that cancel and four pairs with matching sign that result in another 4-interaction combination with coefficients ± 0.5. Defining relations for nonregular designs are not generally unique. Though the individual terms are unique, corresponding to all the interactions that do not sum to zero, they will appear in different groups if other factors are chosen as the basic factors. Table 2 shows one representation for Design 4's defining relation. However, if we choose $\{A, B, E, F, G\}$ as the basic factors, the four terms involving C appear in one group, those involving D appear in the second group, and those involving CD appear as the third group:

$$\begin{aligned} I &= 0.5BCG(E - AE + F + AF) \\ &= 0.5BDE(F + G + AG - AF) = 0.5CDF(E + AE + G - AG). \end{aligned} \tag{2.2}$$

Only nine of Design 4's 21 five-factor projections produce a full 2^5. With $p = 2$, the defining relation is short enough to see which pairs of factors can be generators. For example, A cannot be a generated factor, since it appears in only six of the 12 terms in (2.2), not eight.

Table 2's Design 5 is akin to a resolution III* design (Draper and Lin (1990)) in that its defining relation involves 3-factor interactions but no 4-factor interactions. Since $ABCD = EF$, $G = 0.5(AB + AC + AD - EF)$; this is the only singularity for the two-factor interaction model, so rank($\mathbf{X}$) = 28. However, this design, with all 4-factor interaction contrasts summing to 0, is an excellent choice for building a small central composite design (Angelopoulos, Evangelaras and Koukouvinos (2009)).

Which 2^{7-2} design is preferred depends on the model one anticipates needing to describe the factors' effects on the response. If the full two-factor interaction model is to be fit, then one of the strength 2, D-efficient arrays in Table 3 (d1, d2, or d3) is recommended, or even a non-orthogonal, D-optimal design. If instead, one anticipates only a few active two-factor interactions, then the strength 3, mimimum G-aberration design, with generalized resolution 4.5 and a single linear dependency in $\mathbf{X}$, is recommended. For a detailed discussion of design choice, see Mee, Schoen and Edwards (2017).

For 2^{k-p} designs of size 32 or larger, the columns of $\mathbf{G}$ may involve more than four interactions and do not necessarily have all non-zero elements of the same magnitude. Design I from Table 1 is just such an example, as are all the more D-efficient Type I 2^{7-2} (Eendebak and Schoen (2017)). We now turn to examining what generators are possible for $n = 32$ before examining these D-efficient designs.

2.2. All possible generators for OA(32, 2^k)

To determine all possible generators for OA(32, 2^k), I generated every possible vector e of length 32, fixing only the last element at $+1$, to find every possible g satisfying $\mathbf{S}_{32}^* g = e$ for some $e \in \mathcal{E}$. Here, $\mathbf{S}_{32}^*$ is the 32×26 matrix obtained by deleting the five basic columns from $\mathbf{S}_{32}$. Of the $31!/(15!16!) = 300{,}540{,}195$ e vectors, 403,990 permitted a solution g. This computation verified that only six varieties of generators g exist for $n = 32$: (i) 25 zeros and a single ± 1; (ii) 22 zeros and four ± 0.5; (iii) 18 zeros, seven ± 0.25, and one ± 0.75; (iv) 16 zeros, eight ± 0.25, and two ± 0.5; (v) 13 zeros, 12 ± 0.25, and one ± 0.5; and (vi) 10 zeros and 16 ± 0.25. Variety (v) was the most common by far, accounting for 67% of the generators; varieties (iv)–(vi) together accounted for 98%.

Only 17 of the 530,469,996 isomorphic classes of 2^{7-2} have strength 3 (generalized resolution of 4 or higher). That this count is so modest, in contrast to the astronomical number of strength 2 OA(32, 2^k) is a reflection that the classes of generators that involve more than one interaction but no two-factor interactions is very limited (see the Appendix). For 32-run arrays, varieties (iii)–(vi) cannot be used for strength 3, since they always involve at least one two-factor interaction.

Generator varieties (i) and (ii) appear in Table 2's designs, while the 97 Type I designs among the top 100 D-efficient designs have only the generator varieties (iii)–(vi). The top 100 designs take on 18 different D-efficiency values. Table 3 lists examples of the first three of these, plus two other cases that involve the rarer generator varieties. We denote these D-efficient 2^{7-2} designs as d1, d2, d3, d7, and d13, according to the rank of their D-efficiency value. For each design, we list D-efficiency, the maximum VIF (= the largest diagonal element of $(\mathbf{X}^T\mathbf{X}/n)^{-1}$), and the generators.

Table 3's Design d1 has variety (iv) and (v) generators, with the 0.5 coefficients associated with 4- and 5-factor interactions. This design is one of five D-optimal designs, with D-eff = 0.8432. All five have identical distributions of VIFs, with a maximum = 2.76; they simply differ on which effects receive the larger VIFs. This is true even though one of the five designs is a Type 2 design.

The second-best D-efficiency among 2^{7-2} designs is 0.8415; there are four such designs, and all have maximum VIF = 2.36. There are three designs with the third-best D-efficiency value of 0.836; these designs have the best maximum VIF (= 1.96) of all 100 D-efficient designs. Both generators are variety (v) and nine interactions are shared. No length 3 word contains the factor E, so E's main effect has a VIF = 1. The design's ewlp = ($f_{3.75} = 8$, $f_{4.75} = 12$, $f_{5.5} = 4$, $f_{5.75} = 8$, $f_{6.75} = 4$).

Table 3 lists Designs d7 and d13 to show cases with the rarer generator varieties (iii) and (vi), respectively. Factor E has the largest VIF of 4.89 for Design d7, because 7 of the 10 three-factor interactions that appear in the defining relation involve E.

Generator Varieties (iv)–(vi) also appear in a 2^{12-7} design provided by Bulutoglu and Kaziska (2009) as their counterexample refuting two conjectures regarding maximal orthogonal arrays. Here, it is impossible to add a 13^{th} two-level factor that is orthogonal to these 12 columns. Their 2^{12-7} maximal design projects to a full 2^5 in just one of its 792 5-factor projections: Columns 1, 2, 7, 11, 12. The generator matrix, which utilizes all 26 interactions for the seven generators, is provided as supplementary material. Studying the sets of generators for other maximal designs Bulutoglu and Kaziska (2009) found would likely provide insights regarding why it is not possible to add additional factors. That every row of $\mathbf{G}$ has a nonzero entry is a necessary but not sufficient condition.

Variety (ii) generators, as in (1.1), involve four interaction columns that together form a replicated resolution IV 2^{4-1}, while the eight interaction columns in Variety (iii) generators (e.g., Table 3's design d7) form a replicated resolution IV 2^{8-4}. However, the columns of $\mathbf{S}_h$ for Variety (iv)–(vi) generators often form resolution III fractions, which we illustrate via the 7 generators (Columns 3–6 and 8–10) of Bulutoglu and Kaziska (2009) 2^{12-7} maximal array. Variety (iv): $\mathbf{S}_h$ for Column 9 is a replicated 2^{10-6} fraction with $A_3 = 8$ and $A_4 = 18$; $\mathbf{S}_h$ for Column 3 is a foldover of the 2^{10-6} fraction, so $A_3 = 0$ and $A_4 = 18$. Variety (v): $\mathbf{S}_h$ for Column 4 is a 2^{13-8} fraction with $A_3 = 14$ and $A_4 = 23$; $\mathbf{S}_h$ for Columns 5, 6, and 10 is a 2^{13-8} fraction with $A_3 = 10$ and $A_4 = 23$. Variety (vi):

Table 3. Generators for five D-efficient Type 1 2^{7-2} designs.

Design	D-eff.	Max VIF	Generators
d1	0.8432	2.76	$F = 0.25(AB + AC + AD - AE + CD$ $-ACD + BCD + CDE$ $+2ACDE - ABCE - ABDE$ $+BCDE - ABCDE)$ $G = 0.25B(A - D + AC + AE - CD + DE$ $+2ADE + ACE + CDE - 2ACDE)$
d2	0.8415	2.36	$F = 0.25D(2E + C + AC - BC - CE$ $+2ABE - ABC + ACE - BCE + ABCE)$ $G = 0.25(AC - AE + BE + CE + DE + ABC$ $-ACD - ACE + 2ABCE - ABCD$ $-ABDE + BCDE + ABCDE)$
d3	0.8360	1.96	$F = 0.25(BD - AB - AC + ABE + ACD$ $+ACE - BCD + BDE + 2ABDE + ABCD$ $-ACDE - BCDE + ABCDE)$ $G = 0.25(AB + AD - BC - ABE + ACD$ $+ADE - BCD + BCE + 2ABDE - ABCD$ $+ACDE + BCDE - ABCDE)$
d7	0.8265	4.89	$F = 0.25E(A - B - C - D - 3ABC$ $+ABD + ACD - BCD)$ $G = 0.25(AC + AE - BC - BE + CD$ $+DE + 3ABCD - ABDE)$
d13	0.8181	3.00	$F = 0.25(AD - AC - BC - BD - CD - CE - DE$ $+ACD - BCD + CDE - ABCD - ABCE$ $+ABDE + ACDE + BCDE - ABCDE)$ $G = 0.25(AB - AE - BD - CE - DE$ $-ABC + BCD + BCE - 2ABDE + ABCE$ $-ACDE + BCDE - ABCDE)$

$\mathbf{S}_h$ for Column 8 is a 2^{16-11} fraction with $A_3 = 20$ and $A_4 = 60$. For the cases where generators having similar coefficients correspond to two different designs, the aliasing among two-factor interactions are equivalent.

The defining relation for Bulutoglu and Kaziska (2009) 2^{12-7} is long, with $2^7 - 1 = 127$ terms. Its *ewlp* begins as ($f_{3.5} = 9$, $f_{3.75} = 98$; $f_{4.5} = 16$, $f_{4.75} = 216$; …). Because there are no complete words, the generalized interactions get longer as more generators are multiplied. For example, the product of all seven generators cannot have any interactions of length less than 7.

2.3. One 16-run example

For $n = 16$, all generators involve one or four interactions. This is true for all OA(16, 2^k), not only those considered by Johnson and Jones (2011). For example, consider the 2^{8-4} design with basic factors A–D and generators $E = 0.5D(C + B + AC - AB)$, $F = 0.5B(D + C + AD - AC)$, $G = 0.5BC(1 + D + A - AD)$, and $H = 0.5CD(1 + B + AB - A)$. From these $p = 4$ generators, the defining relation is

$$
\begin{aligned}
I &= 0.5DE(C+B+AC-AB) && = 0.5BF(D+C+AD-AC)\\
&= 0.5BCG(1+D+A-AD) && = 0.5CDH(1+B+AB-A)\\
&= 0.5CEF(D+B-AD+AB) && = 0.5EG(C+BD-AC+ABD)\\
&= 0.5DFG(1+C-A+AC) && = 0.5BEH(1+C+A-AC)\\
&= 0.5FH(C+BD+AC-ABD) && = 0.5GH(D+B+AD-AB)\\
&= 0.5EFG(1+B+A-AB) && = 0.5EFH(1+D+AD-A)\\
&= 0.5EGH(C+D+AC-AD) && = 0.5FGH(B+C+AB-AC)\\
&= BCDEFGH.
\end{aligned}
\tag{2.3}
$$

This defining relation reveals only three of 21 linear dependencies for the full two-factor interaction model: $2DE = C+B+AC-AB$, $2BF = D+C+AD-AC$ and $2GH = D+B+AD-AB$. Eighteen more singularities involving main effects and two-factor interactions are implicit in (2.3). This design projects to a full 2^4 in 21 different subsets of four factors. If we used a different set of basic factors, the defining relation would reveal other linear dependencies, but the same 14 3-factor interactions, 28 4-factor interactions, and 14 5-factor interactions would appear, together with the 7-factor interaction.

2.4. Generators and defining relations for two large strength 4 arrays

Here we examine two strength-4 designs that have higher generalized resolution than the comparable regular 2^{k-p} designs. The examples are the strength-4 2^{15-8} (Xu (2005); (Mee, 2009, p. 286)) and the strength-4 2^{19-11} ((Hedayat, Sloane and Stufken, 1999, Sec. 10.4); (Mee, 2009, p. 286f)). Examining the generators for these designs helps one to appreciate their structure.

Xu (2005) describes the family of designs generated by the Nordstrom-Robinson codes. The minimum G-aberration OA(32, 2^7) (see Table 2's Design 2) is from this family. Here we focus on the remarkable strength-4 2^{15-8}; this design has generalized resolution 5.5, whereas regular resolution V designs with 128 runs can accommodate no more than 11 factors. Mee (2009) describes the 2^{15-8} design as the concatenation of eight regular 2^{15-11} with sign changes to the generators. Alternatively, the design can be constructed by taking a full factorial in factors A–G and then using the generator matrix given in the supplementary materials. Adding factor $H = ABCEFG$ makes a regular, resolution VII 2^{8-1}. The remaining seven generators use 16 (of 20) four-factor interactions involving D and 12 (of 15) five-factor interactions involving D, the basic factor not appearing in the generator for H. No interaction appears in more than one generator. The 4- and 5-factor interactions not appearing in the $\mathbf{S}_{128}$ matrix that involve D are

$$ABDE, ACDF, BCDG, DEFG, ABDFG, ACDEG, BCDEF. \tag{2.4}$$

The coefficients in the defining relation are all ± 0.5 for interactions involving 5, 6, 9 or 10 factors and +1 for interactions of length 7, 8, and 15. Note the defining relation's symmetry.

A strength-4 2^{19-11} design may be constructed by replicating the strength-4 2^{15-8}, adding an 8^{th} basic column Q, and including three more generators as follows:

$$\begin{aligned}
R &= 0.5(ABFGQ + ACFQ + BCEFQ - EFGQ) \\
&= 0.5FQ(ABG + AC + BCE - EG), \\
S &= 0.5(ACEGQ + BCEFQ - ABEQ + EFGQ) \\
&= 0.5EQ(ACG + BCF - AB + FG), \\
T &= 0.5(ABEQ + ACFQ - BCGQ + EFGQ) \\
&= 0.5Q(ABE + ACF - BCG + EFG). \qquad (2.5)
\end{aligned}$$

The generators for R, S, and T involve all the basic factors for the 2^{15-8} except D. If one replaces D with Q in each of the interactions in (2.4), one obtains the seven interactions that appear in (2.5). This 2^{19-11} design has generalized resolution of only 5, with the first counts in the ewlp being ($f_5 = 6$, $f_{5.5} = 264$). The six words of length 5 in the defining relation are $BNPQR$, $CLPQS$, $DHPQT$, $EMPST$, $FKPRT$, and $GJPRS$, each of which arises as the generalized interaction of the generators for P and three other factors.

3. Type 2: 2^{k-p} Designs without Partial Replication and no $2^b = n$ Projection

Let b_{max} be the maximum size projection to have equal replication. Thus, for Type 1 designs, $b_{max} = b$, while for Types 2 and 3, $b_{max} < b$. Sort the design by any such set of b_{max} columns and add $d = b - b_{max}$ dummy columns to complete a full 2^b. These columns may serve as an orthogonal basis to create generators for the remaining $k - b_{max}$ columns. There are $2^{b_{max}} \times (2^d)!$ possible n by d

augmentations and, while any one of these may complete a set of orthogonal basis columns, I recommend searching for one that is uncorrelated with all k columns. When this is not possible, as when the original array is maximal, I suggest choosing the augmenting column(s) to have correlations of magnitude 0, $8/n$ or multiples of $8/n$ with the other factors. We now illustrate these ideas with three examples.

Design II in Table 1 has 32 distinct rows but $b_{max} = 4$; columns $D - G$ form an equally replicated 2^4. We need to add a single ± 1 column to create an orthogonal basis, since $d = 1$. By choosing the dummy column, N, orthogonal to all $k = 7$ original columns, we confirm that this Type 2 design is a projection of a Type 1 orthogonal array. In this way, we obtained the following generators for the first three columns:

$$\begin{aligned}
A &= 0.25(FG - FN - DG - EN + 2DEF + DEG + DEN - DFN \\
&\qquad + DGN + EFG - EGN + DFGN + EFGN), \\
B &= 0.25(2DN - EF + EG - FN + GN + 2DFGN - DEF \\
&\qquad + DEG + DFN - DGN), \\
C &= 0.25(2GN + DE - DF + FN - DEG + DEN + DFG - EFN \\
&\qquad + FGN + EFGN + DEGN - DEFN - DEFGN). \qquad (3.1)
\end{aligned}$$

These generators are similar to those for Bulutoglu and Kaziska's maximal 2^{12-7}. The defining relation for this 8-factor design, multiplied by 4, is:

$$\begin{aligned}
4I &= A(FG - DG + 2DEF + DEG + EFG + N(-E - F + DE - DF + DG \\
&\qquad - EG + DFG + EFG)) \\
&= B(-EF + EG - DEF + DEG + N(2D - F + G + DF - DG + 2DFG)) \\
&= C(DE - DF - DEG + DFG + N(F + 2G + DE - EF + FG - DEF \\
&\qquad + DEG + EFG - DEFG)) \\
&= AB(-DG + EF - EG + FG - DEF + EFG + N(E - G + DE + 2EF \\
&\qquad + EG + DFG - EFG)) \\
&= AC(DE + DF - DG - FG + DFG - EFG + N(1 - D + F - G + DE \\
&\qquad + EFG + 2DEFG)) \\
&= BC(DE + DF + 2DG + EF - EG - DEF + DFG + N(1 + D + E - F \\
&\qquad - EG - DFG)) \\
&= ABC(-1 + D + E - F + G + 2EFG + DEFG + N(1 - D + EFG - DE \\
&\qquad + DF - DG)). \qquad (3.2)
\end{aligned}$$

Setting $N = 0$, one obtains the 39 interaction terms of the indicator function (Li, Lin and Ye (2003)) of the original 2^{7-2}. The ewlp = ($f_{3.75} = 7$, $f_{4.5} = 2$, $f_{4.75} = 22$, $f_{5.75} = 6$, $f_{6.5} = 1$, $f_{7.75} = 1$).

There is only one Type 2 OA(32, 2^k) of strength 3; it is a 2^{8-3} design which, when augmented with the factor N, becomes a 2^{9-4} fraction of strength 2, with generators

$$\begin{aligned}
E &= 0.5AB(C+D) + 0.5AN(1-CD),\\
F &= 0.5AC(B+D) + 0.5AN(BD-1),\\
G &= 0.5CD(A+B) + 0.25N(A+D+ACD+ABC-B-C-ABD-BCD),\\
H &= 0.5BD(C-A) + 0.25N(A+B+C+ABC+ABD+BCD-D-ACD).
\end{aligned} \tag{3.3}$$

Dropping N, we have the Type 2, 2^{8-3} fraction with factors A – H. The *ewlp* for this design is simply ($f_{4.5} = 28$). Incidentally, this 2^{8-3} design is a foldover of the only 16-run, Type 2 design.

Not all Type 2, 2^{k-p} designs can be imbedded in Type 1 orthogonal arrays. The Paley OA(32, 2^{31}) is a Type 2 nonregular design in that, though it has 32 distinct rows, every five-factor projection has at least four replicated pairs. Since the design is saturated, any additional factor used to create a 5^{th} basic factor would necessarily be correlated with some of the original 31 columns. Still, one may select four columns that form an equally replicated 2^4 and add a fifth column that creates a 2^5, and has a correlation of either 0, ± 0.25, or ± 0.5 with the other 27 columns in the Paley design. The supplementary materials show the 27 generators that result: 17 of these have 16 ± 0.25, while the other 10 have 12 ± 0.25 and a single ± 0.5. Note that all 10 of the ± 0.5 occur with terms involving the added column, and so do not directly produce aliasing among interaction terms for the 31-factor design. If, instead, we augment Paley's array with a dummy column that has a correlation of ± 0.125 with some of the original 31 columns, the generators will differ from the varieties discussed in Section 2.2; however, the indicator function obtained by dropping all columns involving the dummy column is not impacted by this choice, as all terms with a coefficient of ± 0.125 must disappear.

4. Type 3: Nonregular 2^{k-p} Designs with Partial Replication

Having pure error degrees of freedom can improve the power of tests more than having additional treatment combinations (Liao and Chai (2009)). However, replication increases G_2-aberration. For designs with partial replication, the generalized word length pattern sums to

$$2^p(\sum_{i=1}^{U}\frac{r_i^2}{n})-1, \tag{4.1}$$

where U is the number of distinct rows and $r_1,\ldots,r_U$ are the frequencies for the distinct rows; this result is based on Xu and Wu (2001) equation (8) for $B_0(D)$. If 2^{k-p} design has no replication, then $U=n$ and the factor in parenthesis in (4.1) equals 1, and so the generalized word length pattern sums to 2^{p-1}, just as it does for unreplicated regular fractions. Thus, 2^{k-p} designs having partial replication will generally have worse G_2-aberration than the designs discussed in Sections 2 and 3 that have no replication.

For Table 1's Design III, there are 30 distinct rows and two $r_i=2$. Thus, equation (4.1) equals $4(36/32)-1=3.5$, an increase of 0.5. In general, if all the distinct row frequencies are 1 or 2, then the partial replication increases the gwlp sum by $\nu 2^{p+1}/n$, where $\nu=$ pure error degrees of freedom. For a given value of ν, having row frequencies that differ by at most 1 is preferred; this makes intuitive sense.

OA(n, 2^k) with partial replication are created quite easily by concatenating smaller designs. When combining two arrays, the order of columns for one design can be fixed, while the second array's columns are permuted. For each permutation, every possible subset of columns in the second array may be multiplied by -1, leading to $(k!)2^k$ concatenated designs.

We now use generators to describe Table 1's Design III, a 2^{7-2} design with partial replication. Design III can be augmented with an 8th column to form a 2^{8-3} fraction of strength 2 for which there exists a five-factor projection forming a full 2^5. To obtain this 8^{th} factor, first we find four columns that form an equally replicated 2^4; $\{B,E,F,G\}$ is one such set. The added column N must have $+1$ and -1 for each pair of identical rows for these basic columns. Of the 2^{16} such columns, choose one that is also orthogonal to A, C and D. Labeling this as N, we obtain the following three generators:

$$\begin{aligned}
A &= 0.25(BEG+EFG-BG-FG)\\
&\quad +0.25N(BEG+BFG+2BE+EG+FG-BF-EF+B-E),\\
C &= 0.25(2BEFG+BEF+EFG+BE+BF-BG-EF)\\
&\quad +0.25N(BEFG-BEG-BFG-EFG+EF+B),\\
D &= 0.25(BEFG+BEF+BFG+EFG-BEG+BE+EF-BF)
\end{aligned}$$

$$+\,0.25N(BEG+EFG-BEF-BF-EG-B+F+G). \tag{4.2}$$

Thus, Design III can be constructed by forming a full factorial in B, E, F, G, and N, adding the generated columns A, C, and D, and then dropping N. So Design III is also a projection of a Type I 2^{8-3}; its defining relation is obtained in the supplementary materials. Setting $N = 0$, add 1, and dividing by 2^p, one obtains the indicator function

$$\begin{aligned}
&[1+0.25AG(BE+EF-B-F)\\
&+0.25C(2BEFG+BE+BF+BEF+EFG-BG-EF)\\
&+0.25AC(-BE+BF-EF+FG-BEF+BEG)\\
&+0.25D(BE+EF+BEF+BFG+EFG+BEFG-BF-BEG)\\
&+0.25AD(G+BE+BF-EF-EG+BEF-BEG-EFG)\\
&+0.25CD(B+E+G-BF+BG-EF+FG+BFG)\\
&+0.25ACD(E-B-BF-BG+EF+EG+FG-2EFG+BEFG)]/4.
\end{aligned} \tag{4.3}$$

From this we obtain $ewlp = (f_{3.75} = 13;\ f_{4.75} = 22;\ f_{5.5} = 1,\ f_{5.75} = 12;\ f_{6.5} = 1,\ f_{7.75} = 1)$ and $gwlp = (13/16,\ 22/16,\ 1,\ 1/4,\ 1/16)$, which sums to 3.5 rather than 3 due to the partial replication.

Pigeon and McAllister (1989) presented a 2^{7-3} design with partial replication. Take the first, second, and fourth columns as basic with labels A, B, and D, since they form an equally replicated 2^3. Then adding a fourth column N that is orthogonal to all the seven main effects, we generate the 2^{7-3} design for A–G using the generators $C = AB$, $E = 0.5D(B - N + AB + AN)$, $F = -AD$, and $G = -0.5D(B + N + AB - AN)$. The defining relation for the augmented design with eight factors has seven terms with coefficient ± 1 that do not involve N,

$$I=ABC=-ADF=-AEG=-BCDF=DEFG=-BCEG=ABCDEFG, \tag{4.4}$$

plus eight linear combinations, such as $-0.5EF(B + AB + N - AN)$, that all include a 3-factor and 4-factor interaction not involving N. Now (4.4) is the defining relation of a regular 2^{7-3}, produced by the three generators $C = AB$, $F = -AD$ and $G = -AE$. However, for the 2^{7-3} design with partial replication considered here, A, B, D, and E do not form a full 2^4 factorial; E is correlated with both BD and ABD. The indicator function for the partially replicated design would include all the terms in (4.4), plus eight 3-factor interactions ($-BDG$,

BFG, BDE, $-BEF$, CDE, $-CDG$, $-CEF$, CFG) and eight 4-factor interactions that come from the augmented design's defining relation, each with a coefficient of 0.5. Thus, $gwlp = (5, 5, 0, 0, 1)$, which sums to 11, as guaranteed by (4.1).

If the maximum row frequency exceeds 2, then $d = b - b_{\max} > 1$, so one would need to augment the design with more than one column to create a set of basic columns. However, for designs constructed by concatenating two unreplicated designs, this would never arise. If the partially replicated design is constructed using parallel flats of size 4, it is common to have four or eight pairs of replicated treatment combinations (Liao and Chai (2004)).

5. Discussion

Generators provide a compact means of representing 2^{k-p} designs. They provide a means for understanding maximal designs, as discussed at the end of Section 2.2. It is hoped that, with further study, the defining relations for (nonregular) Type I 2^{k-p} fractions will lead to a convenient characterization of all relevant linear dependencies involving lower-order terms. In addition, once the classes of eligible generators are fully characterized, they may provide an efficient means of enumerating designs using symbolic computations.

Previous literature (Fontana, Pistone and Rogantin (2000); Ye (2003)) for nonregular designs has made use of indicator functions to characterize nonregular OA$(n, 2^k)$. The indicator function is a polynomial in the factors defined by all the interaction columns with nonzero sums; (4.3) gives the indicator function for Table 1's Design III. The indicator function has been useful for describing regular designs augmented by semifolding (Balakrishnan and Yang (2009); Edwards (2011)) and for foldover of nonregular designs (Li, Lin and Ye (2003)). Butler (2008) used it to describe designs with regular fractions either removed or added. The indicator function exists for any design, but its shortcomings are that it does not easily show the linear dependencies except those due to complete aliasing, its computation may require one to sum each of the $2^k - k - 1$ interaction columns, and it may have so many terms that it is not very useful for understanding the design. The results of this paper show that, when n is a power of 2, all the column sums can be found using symbolic computation with the generators. Thus, the computational shortcoming is mitigated. However, the defining relation is more informative than the indicator function, since it reveals linear dependencies. Thus, for Type I designs, one should utilize the defining relation.

Designs with the best generalized aberration appear to be Type I or Type 2 designs. One approach to proving a general result would be to show that when a Type 3 design is a projection of a Type 1 design, it is not the minimum aberration projection.

This article provides much detail regarding 32-run designs. That strength 3, 32-run orthogonal arrays are so rare and strength 2 arrays are so numerous can be explained in terms of the prevalence of the available generators. All strength 3 OA(32, 2^k) arrays with no replication are Type 1 designs, except for the 2^{8-3} design obtained using (3.3). For $k \leq 8$, there are strength 3 OA(32, 2^k) with partial replication, three each for $k = 7$ and 8. For the 100 most D-efficient 2^{7-2} orthogonal arrays, 97 were Type 1 and three were Type 2. Thus, for 2^{7-2} designs explored here, Type 2 and 3 have been relatively rare. However, there are more than 530 million other strength 2 2^{7-2} fractions, and the prevalence of the three types for this vast set of designs is unknown.

A referee suggested that the structure of generators for nonregular 3^{k-p} might be similarly explored. This is an excellent suggestion that will be investigated elsewhere.

Supplementary Materials

Supplementary materials available online include generator matrices for Bulutoglu and Kaziska (2009) 2^{12-7}, the 2^{15-8} design with generalized resolution 5.5, and the Paley 2^{31-26}. Also included is Pigeon and McAllister (1989) 2^{7-3} with partial replication and MATLAB code for all the examples of this article.

Acknowledgements

The author thanks Hongquan Xu for communicating the result (4.1), and Eric Schoen for the list of OA(32, 2^7).

Appendix

A.1. Possible generators for nonregular designs

Each generator for regular 2^{k-p} designs involves a single interaction. For a Type 1 nonregular 2^{k-p} design, at least one of the generators must involve more than one interaction. It is simple to show that no generators involve only two or three interactions. Every pair of basic-factor-interaction columns of a strength-2 orthogonal array contains the four rows $\mathbf{E}_2 = [1, 1; 1, -1; -1, 1; -1, -1]$. None of the solutions g for $\mathbf{E}_2 g = e$ for $e \in \mathcal{E}$ involve two non-zero entries. For three interaction columns, the unique rows are either a full 2^3 or they form a resolution III 2^{3-1}. Again, complete enumeration shows that the only solutions g have a single non-zero value. Thus, all generators for Type 1 nonregular 2^{k-p} designs involve either one interaction or more than 3 interactions.

A.2. Four-interaction generators for nonregular designs

The conditions for a generator with four interactions are that the generator contains no main effects, lest it produce a factor that is correlated with one of the basic factors, and that the product of the four interactions equals -1. For $n = 16$, there are only four classes of generators:

$$\begin{aligned}
2.0112 &: \ 0.5CD(1 + A + B - AB) \\
1.1122 &: \ 0.5D(A + B + AC - BC) \\
1.1113 &: \ 0.5D(A + B + C - ABC) \\
0.2222 &: \ 0.5(AB + AC + BD - CD)
\end{aligned} \tag{A.1}$$

Each class is identified by a string of counts. For example, 2.0112 indicates that two factors are common to all four terms and the parenthesis has terms with 0, 1, 1, and 2 factors, respectively. Within each class, one can change the signs and permute the letters to obtain other generators within its class. Each of these include at least one two-factor interaction, so every OA(16, 2^k) using one or more of these generators has generalized resolution at most 3.5.

For $n = 32$, there are the four classes in (A.1), plus 14 additional classes of generators, for a total of 18 classes of generators involving just four interactions.

$$\begin{aligned}
3.0112 &: \quad 0.5CDE(1 + A + B - AB) \\
2.1122 &: \quad 0.5DE(A + B + AC - BC) \\
2.1113 &: \quad 0.5DE(A + B + C - ABC) \\
1.2222 &: \quad 0.5E(AB + AC + BD - CD) \\
2.0123 &: \quad 0.5DE(1 + A + BC - ABC) \\
2.0222 &: \quad 0.5DE(1 + AB + AC - BC) \\
1.1223a &: \quad 0.5E(A + BC + CD - ABD) \\
1.1223b &: \quad 0.5E(A + AB + CD - BCD) \\
1.1124 &: \quad 0.5E(A + B + CD - ABCD) \\
1.1133 &: \quad 0.5E(A + B + ACD - BCD) \\
0.2233a &: \quad 0.5(AB + BC + ADE - CDE)
\end{aligned}$$

$$\begin{aligned} 0.2233b: &\quad 0.5(AB + CD + ABE - CDE) \\ 0.2233c: &\quad 0.5(AB + CD + ACE - BDE) \\ 0.2224: &\quad 0.5(AB + CD + DE - ABCE) \end{aligned} \tag{A.2}$$

The first four generator classes for $n = 32$ involve no two-factor interactions. For Type 1 OA(32, 2^k) with generalized resolution of 4.5, these are the only eligible generators, besides the individual 4- and 5-factor interactions.

By complete enumeration for $n = 32$, it was found that there were four other linear combination generators that involve between 8 and 16 interactions, as discussed in Section 2.2. Every one of these involves at least one two-factor interaction and so leads to a strength-2 array. By directly checking all strength 3 OA(32, 2^k), it was confirmed that all Type 1 designs use generators from the first four in (A.2). There are several Type 2 OA(32, 2^k) of strength 3, and one Type 3. These may involve other classes of generators, such as for G and H in (3.3) because they involve two-factor interactions containing the augmented factor, N.

References

Angelopoulos, P., Evangelaras, H. and Koukouvinos, C. (2009). Small, balanced, efficient and near rotatable central composite designs. *J. Stat. Plan. Inf.* **139**, 2010–2013.

Balakrishnan, N. and Yang, P. (2009). De-aliasing effects using semifoldover techniques. *J. Stat. Plan. Inf.* **139**, 3102–3111.

Bulutoglu, D. A. and Kaziska, D. M. (2009). A counterexample to Beder's conjectures about Hadamard matrices. *J. Stat. Plan. Inf.* **139**, 3381–3383.

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

Butler, N. A. (2008). Defining equations for two-level factorial designs. *J. Stat. Plan. Inf.* **138**, 3157–3163.

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.

Draper, N. R. and Lin, D. K. J. (1990). Connections between two-level designs of resolutions III* and V. *Technometrics* **32**, 283–288.

Edwards, D. J. (2011). Optimal semifoldover plans for two-level orthogonal designs. *Technometrics* **53**, 274–284.

Eendebak, P. T. and Schoen, E. D. (2017). Two-level designs to estimate all main effects and two-factor interactions. *Technometrics* **59**, 69–79.

Fontana, R., Pistone, G. and Rogantin, M. (2000). Classification of two-level factorial fractions. *J. Stat. Plan. Inf.* **87**, 149–172.

Hall, M. (1961). Hadamard matrices of order 16. Research Summary No. 36-10, Jet Propulsion Lab., Pasadena, CA.

Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and*

Applications. Springer, New York.

Johnson, M. E. and Jones, B. (2011). Classical design structure of orthogonal designs with six to eight factors and sixteen runs. *Qual. & Rel. Eng. Int.* **27**, 61–70.

Kharaghani, H. and Tayfeh-Rezaie, B. (2013). Hadamard matrices of order 32. *J. Combin. Designs* **21**, 212–221.

Li, W., Lin, D. K. J. and Ye, K. Q. (2003). Optimal foldover plans for two-level nonregular orthogonal designs. *Technometrics* **45**, 347–351.

Liao, C. T. and Chai, F. S. (2004). Partially replicated two-level fractional factorial designs. *Canad. J. of Stat.* **32**, 421–438.

Liao, C. T. and Chai, F. S. (2009). Design and analysis of two-level factorial experiments with partial replication. *Technometrics* **51**, 66–74.

Mee, R. W. (2009). *A Comprehensive Guide to Factorial Two-Level Experimentation.* Springer, New York.

Mee, R. and Dean, A. (2015). Regular fractional factorial designs. In *Handbook of Design and Analysis of Experiments* (Eds. Dean, A., Morris, M., Stufken, J., & Bingham, D). Chapman & Hall/CRC, New York.

Mee, R. W., Schoen, E. D. and Edwards, D. J. (2017). Selecting an orthogonal or non-orthogonal two-level design for screening. *Technometrics*, DOI: 10.1080/00401706.2016.1186562.

Montgomery, D. C. (2012). *Design and Analysis of Experiments.* 8th Ed. Wiley, New York.

Pigeon, J. G. and McAllister, P. R. (1989). A note on partially replicated orthogonal main-effect plans. *Technometrics* **31**, 249–251.

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**, 123–140.

Schoen, E. D. and Mee, R. W. (2012). Two-level designs of strength 3 and up to 48 runs. *Appl. Statist.* **61**, 163–174.

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

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

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

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

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

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

Ye, K. Q. (2003). Indicators functions and its applications in two-level factorial design. *Ann. Statist.* **31**, 984–994.

Department of Business Analytics and Statistics, University of Tennessee, Knoxville, TN 37996-0525, U.S.A.

E-mail: rmee@utk.edu

(Received May 2015; accepted May 2016)