

Construction results for strong orthogonal arrays of strength three

CHENLU SHI* and BOXIN TANG**

Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.
E-mail: * chenlus@sfu.ca; ** boxint@sfu.ca

Strong orthogonal arrays were recently introduced as a class of space-filling designs for computer experiments. The most attractive are those of strength three for their economical run sizes. Although the existence of strong orthogonal arrays of strength three has been completely characterized, the construction of these arrays has not been explored. In this paper, we provide a systematic and comprehensive study on the construction of these arrays, with the aim at better space-filling properties. Besides various characterizing results, three families of strength-three strong orthogonal arrays are presented. One of these families deserves special mention, as the arrays in this family enjoy almost all of the space-filling properties of strength-four strong orthogonal arrays, and do so with much more economical run sizes than the latter. The theory of maximal designs and their doubling constructions plays a crucial role in many of theoretical developments.

Keywords: computer experiment; doubling and projection; maximal design; second order saturated design; space-filling design

1. Introduction

Computer models are powerful tools that enable researchers to study complex systems in natural sciences, engineering, social sciences and humanities. When a computer program representing a complex model is expensive to run, it is desirable to build a more economical version of the model. Computer experiments are concerned with building these so-called surrogate models. This type of models is built using data consisting of a set of inputs and the corresponding outputs from the computer program. A vital step in the process of designing such an experiment is the selection of the right kinds of inputs to use. A widely accepted type of designs used for computer experiments is that of space-filling designs (Santner, Williams and Notz [24]; Fang, Li and Sudjianto [7]). Broadly speaking, a space-filling design refers to any design that scatters its points in the design region in some sort of uniform fashion.

There are several ways to obtain space-filling designs. An intuitively appealing approach is to use criteria based on distances or discrepancies. See Johnson, Moore and Ylvisaker [14] and Fang, Lin, Winker and Zhang [8] for early work, and Moon, Dean and Santner [19], Lin and Kang [16], Wang, Xiao and Xu [26] for more recent developments. Orthogonality also plays an important role in the quest for space-filling designs (Ye [28]; Bingham, Sitter and Tang [2]; Georgiou and Efthimiou [9]; Liu and Liu [17]). Use of orthogonal arrays to generate space-filling designs has a long history and goes back to McKay, Beckman and Conover [18], Owen [22] and Tang [25].

Motivated by (t, m, s) -nets from quasi-Monte Carlo (Niederreiter [21]), He and Tang [11] introduced and studied *strong orthogonal arrays* (SOAs), which are more space-filling in low dimensions than comparable ordinary orthogonal arrays. SOAs of strength three are most useful because of their economical run sizes. They are the subject of study in this paper.

Although the existence problem of strength-three SOAs has been completely solved (He and Tang [12]), the same cannot be said about the construction of such arrays. We undertake in this paper a systematic and comprehensive study on the construction of strength-three SOAs. We do not want just any array; we want an array that possesses more space-filling properties whenever possible. More specifically, we aim at constructing strength-three SOAs that enjoy some of the space-filling properties that only strength-four SOAs can offer. Besides various characterizing results, we present three families of strength-three SOAs. The arrays in one of these families enjoy almost all of the space-filling properties of strength-four strong orthogonal arrays, and do so with much more economical run sizes than the latter. Section 2 introduces these better strength-three SOAs and presents illustrative examples of three families of strength-three SOAs. Characterizations and constructions will be studied under various scenarios in Section 3. We conclude the paper with a discussion in Section 4.

2. Background, preliminaries and examples

2.1. Background

We use $\text{OA}(n, m, s_1 \times \cdots \times s_m, t)$ to denote an orthogonal array of n runs, with m factors and having strength t , such that the j th factor has s_j levels taken from $\{0, 1, \dots, s_j - 1\}$. The array is symmetric if $s_1 = \cdots = s_m = s$ and asymmetric otherwise. A simple notation $\text{OA}(n, m, s, t)$ is used for the symmetric case. Orthogonal arrays provide very useful designs in many scientific and technological investigations. Hedayat, Sloane and Stufken [13] is devoted entirely to them; Dey and Mukerjee [6] and Cheng [5] also contain abundant sources of information on orthogonal arrays.

Point sets and sequences from quasi-Monte Carlo (Niederreiter [21]) have long been recognized as useful in design of experiments – see Bates, Buck, Riccomagno and Wynn [1], Owen [23], and Haaland and Qian [10]. Inspired by the combinatorial characterization of (t, m, s) -nets by Lawrence [15] and Mullen and Schmid [20], He and Tang [11] introduced *strong orthogonal arrays* (SOAs) as space-filling designs for computer experiments. SOAs are more general than (t, m, s) -nets, and they are formulated in design language and thus more accessible to design researchers.

An $n \times m$ matrix with entries from $\{0, 1, \dots, s^t - 1\}$ is called an SOA of n runs, m factors, s^t levels and strength t if any subarray of g columns for any g with $1 \leq g \leq t$ can be collapsed into an $\text{OA}(n, g, s^{u_1} \times \cdots \times s^{u_g}, g)$ for any positive integers u_1, \dots, u_g with $u_1 + \cdots + u_g = t$, where collapsing s^t levels into s^{u_j} levels is according to $[a/s^{t-u_j}]$ for $a = 0, 1, \dots, s^t - 1$. We use $\text{SOA}(n, m, s^t, t)$ to denote such an array.

SOAs of strength three are most useful, since SOAs of strength two are no more space-filling than ordinary orthogonal arrays of strength two and SOAs of strength four or higher are prohibitively expensive. The present article focuses on this class of most useful arrays.

2.2. Preliminaries

The existence of strength-three SOAs has been completely characterized in He and Tang [12], but the construction of these arrays has not been explored. The objectives of this paper are to provide a systematic and comprehensive study on the construction of strength-three SOAs. We want to find those designs that are most space-filling within this class of arrays. Our approach is to identify and construct those strength-three SOAs that enjoy some of the space-filling properties that only strength-four SOAs can offer.

The next result is taken from He and Tang [11] and is needed in the paper.

Lemma 1. *An SOA($n, m, s^3, 3$), say D , exists if and only if there exist three arrays $A = (a_1, \dots, a_m)$, $B = (b_1, \dots, b_m)$ and $C = (c_1, \dots, c_m)$ such that (a_i, a_j, a_u) , (a_i, a_j, b_j) and (a_i, b_i, c_i) are OA($n, 3, s, 3$) s for all $i \neq j$, $i \neq u$ and $j \neq u$. These arrays are related through $D = s^2A + sB + C$.*

An SOA($n, m, s^3, 3$) is collapsible into an OA($n, 3, s, 3$) in any three-dimension, and thus achieves stratifications on $s \times s \times s$ grids in all three-dimensions. In any two-dimension, an SOA($n, m, s^3, 3$) can be collapsed into an OA($n, 2, s \times s^2, 2$) and an OA($n, 2, s^2 \times s, 2$), and it therefore achieves stratifications on $s \times s^2$ and $s^2 \times s$ grids in all two-dimensions.

An SOA($n, m, s^4, 4$) is more space-filling, achieving

- (α) stratifications on $s^2 \times s^2$ grids in all two-dimensions,
- (β) stratifications on $s^2 \times s \times s$, $s \times s^2 \times s$ and $s \times s \times s^2$ grids in all three-dimensions, and
- (γ) stratifications on $s^3 \times s$ and $s \times s^3$ grids in all two-dimensions.

Our goals are to construct strength-three SOAs that enjoy some or all of properties α , β and γ . The following provides a basis for later construction results. Its proof is similar to that of Proposition 2 of He and Tang [11].

Proposition 1. *An SOA($n, m, s^3, 3$), as characterized in Lemma 1 through A , B and C , has*

- (i) *property α if and only if (a_i, b_i, a_j, b_j) is an OA($n, 4, s, 4$) for all $i \neq j$,*
- (ii) *property β if and only if (a_i, a_j, a_u, b_u) is an OA($n, 4, s, 4$) for all $i \neq j$, $i \neq u$ and $j \neq u$,*
and
- (iii) *property γ if and only if (a_i, a_j, b_j, c_j) is an OA($n, 4, s, 4$) for all $i \neq j$.*

An SOA($n, m, s^4, 4$) has two more space-filling properties: (δ) stratifications on a set of s^4 intervals in all one-dimensions and (ϵ) stratifications on $s \times s \times s \times s$ grids in all four-dimensions. These two properties are not very interesting for strength-three SOAs. Property δ is not interesting at all because Latin hypercubes based on an SOA($n, m, s^3, 3$) can achieve the maximum stratifications in all one-dimensions. Property ϵ is not very interesting because an SOA($n, m, s^3, 3$) with this property implies the existence of an OA($n, m, s, 4$), requiring a large run size n for a given number m of factors.

The remainder of the paper is devoted to the construction of strength-three SOAs with some or all of properties α , β and γ , by making use of regular 2^{m-p} designs. To our surprise, the results

are far richer and more insightful than we have initially anticipated. The theory of maximal designs and doubling constructions as studied by Chen and Cheng [4] plays a crucial role in many of our theoretical developments. In return, our construction results also bring some new insights into the theory of doubling.

2.3. Three examples

To help the reader appreciate the general theoretical results to be presented in Section 3, we provide three examples in this subsection.

Example 1. Consider the following array (transposed):

```

7 3 6 2 7 3 6 2 4 0 5 1 4 0 5 1 5 1 4 0 5 1 4 0 6 2 7 3 6 2 7 3
7 7 2 2 5 5 0 0 6 6 3 3 4 4 1 1 5 5 0 0 7 7 2 2 4 4 1 1 6 6 3 3
7 5 6 4 3 1 2 0 4 6 5 7 0 2 1 3 7 5 6 4 3 1 2 0 4 6 5 7 0 2 1 3
7 7 4 4 5 5 6 6 2 2 1 1 0 0 3 3 7 7 4 4 5 5 6 6 2 2 1 1 0 0 3 3
7 5 6 4 5 7 4 6 6 4 7 5 4 6 5 7 3 1 2 0 1 3 0 2 2 0 3 1 0 2 1 3
7 1 0 6 3 5 4 2 4 2 3 5 0 6 7 1 5 3 2 4 1 7 6 0 6 0 1 7 2 4 5 3
7 1 2 4 7 1 2 4 2 4 7 1 2 4 7 1 5 3 0 6 5 3 0 6 0 6 5 3 0 6 5 3
7 3 2 6 5 1 0 4 4 0 1 5 6 2 3 7 3 7 6 2 1 5 4 0 0 4 5 1 2 6 7 3
7 1 4 2 3 5 0 6 2 4 1 7 6 0 5 3 3 5 0 6 7 1 4 2 6 0 5 3 2 4 1 7
    
```

As one can easily verify, this is an SOA(32, 9, 8, 3). But it is a special SOA(32, 9, 8, 3), as any subarray of two columns becomes an OA(32, 2, 4, 2) when the eight levels are collapsed into four levels according to 0, 1 → 0; 2, 3 → 1; 4, 5 → 2; 6, 7 → 3. Because of this, the array achieves stratifications on 4 × 4 grids in all two-dimensions, that is, it possesses property α.

Example 2. The array (transposed) below

```

7 6 6 7 6 7 7 6 3 2 2 3 2 3 3 2 1 0 0 1 0 1 1 0 5 4 4 5 4 5 5 4
7 0 4 3 4 3 7 0 3 4 0 7 0 7 3 4 1 6 2 5 2 5 1 6 5 2 6 1 6 1 5 2
7 4 2 1 4 7 1 2 3 0 6 5 0 3 5 6 1 2 4 7 2 1 7 4 5 6 0 3 6 5 3 0
7 2 0 5 6 3 1 4 3 6 4 1 2 7 5 0 1 4 6 3 0 5 7 2 5 0 2 7 4 1 3 6
7 4 6 5 2 1 3 0 3 0 2 1 6 5 7 4 1 2 0 3 4 7 5 6 5 6 4 7 0 3 1 2
7 2 4 1 0 5 3 6 3 6 0 5 4 1 7 2 1 4 2 7 6 3 5 0 5 0 6 3 2 7 1 4
7 6 2 3 0 1 5 4 3 2 6 7 4 5 1 0 1 0 4 5 6 7 3 2 5 4 0 1 2 3 7 6
7 0 0 7 2 5 5 2 3 4 4 3 6 1 1 6 1 6 6 1 4 3 3 4 5 2 2 5 0 7 7 0
    
```

is an SOA(32, 8, 8, 3), which has properties both α and β. The array has property β since any subarray of three columns becomes an OA(32, 3, 4 × 2 × 2, 3), an OA(32, 3, 2 × 4 × 2, 3), or an OA(32, 3, 2 × 2 × 4, 3) when the levels of one factor are collapsed into four levels, and the levels of the other two factors are collapsed into two levels. Collapsing eight levels into two levels is according to 0, 1, 2, 3 → 0; 4, 5, 6, 7 → 1.

Example 3. The following array (transposed)

```

7 6 6 7 6 7 7 6 2 3 3 2 3 2 2 3 0 1 1 0 1 0 0 1 5 4 4 5 4 5 5 4
7 0 4 3 4 3 7 0 2 5 1 6 1 6 2 5 0 7 3 4 3 4 0 7 5 2 6 1 6 1 5 2
7 4 2 1 4 7 1 2 2 1 7 4 1 2 4 7 0 3 5 6 3 0 6 5 5 6 0 3 6 5 3 0
7 2 0 5 6 3 1 4 2 7 5 0 3 6 4 1 0 5 7 2 1 4 6 3 5 0 2 7 4 1 3 6
7 4 6 5 2 1 3 0 2 1 3 0 7 4 6 5 0 3 1 2 5 6 4 7 5 6 4 7 0 3 1 2
7 2 4 1 0 5 3 6 2 7 1 4 5 0 6 3 0 5 3 6 7 2 4 1 5 0 6 3 2 7 1 4
7 6 2 3 0 1 5 4 2 3 7 6 5 4 0 1 0 1 5 4 7 6 2 3 5 4 0 1 2 3 7 6
    
```

is an SOA(32, 7, 8, 3) with all of the properties α , β and γ . It has property γ because any subarray of two columns becomes an OA(32, 2, 2×8 , 2) or an OA(32, 2, 8×2 , 2) when the levels of one factor are collapsed into two levels.

The arrays in Examples 1, 2 and 3 enjoy some or all of properties α , β and γ . Strength-four SOAs automatically have these properties, but for 32 runs, an SOA(32, m , 16, 4) can be constructed only for $m \leq 3$ (He and Tang [11], Theorem 1).

3. Construction results

We now concentrate on SOA($n, m, s^3, 3$)s with $s = 2$, and consider their constructions using regular 2^{m-p} designs. This means that in applying Lemma 1 to obtain an SOA, the columns of A , B and C are all selected from a saturated regular two-level design. With k independent factors, one can obtain a saturated design of $n = 2^k$ runs for $m = n - 1$ factors. Let S denote this saturated design. If S is viewed as a collection of columns, then any subset of S is a design of resolution III or higher. Designs with repeated columns have resolution II and are also needed in presenting our construction results.

Two-level designs are commonly and conveniently studied with the two levels denoted by ± 1 . When A , B and C all have two levels ± 1 , we can make them to have levels 0, 1 by $(A + 1)/2$, $(B + 1)/2$ and $(C + 1)/2$. This implies that $D = 4A + 2B + C$ in Lemma 1 should be replaced by

$$D = 2A + B + C/2 + 7/2. \tag{1}$$

3.1. Designs with property α

When strength-three SOAs are constructed using regular 2^{m-p} designs, the following result can be obtained.

Theorem 1. *If an SOA($n, m, 8, 3$) is to be constructed using regular A , B and C with their columns selected from a saturated design S , then it has property α if and only if A is of resolution IV or higher and (A, B, B') has resolution III or higher, where $B' = (b'_1, \dots, b'_m)$ with $b'_j = a_j b_j$.*

Proof. According to Lemma 1 and Proposition 1(i), we only need to show that (a_i, b_i, a_j, b_j) where $i \neq j$ has strength four if and only if (A, B, B') has resolution III or higher. That (a_i, b_i, a_j, b_j) has strength four means that the four columns are independent, thus without any defining words among them. That (A, B, B') has resolution III or higher simply says that (A, B, B') has no repeated columns. Therefore, to prove Theorem 1, it remains to assert that (a_i, b_i, a_j, b_j) where $i \neq j$ has no defining words if and only if (A, B, B') has no repeated columns. This assertion can be easily verified. \square

Designs A , B and B' in Theorem 1 produce a collection of mutually exclusive triplets $(a_j, b_j, b'_j = a_j b_j)$ with $j = 1, \dots, m$. For a different purpose, such mutually exclusive triplets

are also wanted in Wu [27]. There is, however, a major difference, being that we require A to have resolution IV or higher whereas there is no such a requirement in Wu [27]. This implies that the method of Wu [27] would not help us to construct A , B and B' in Theorem 1. New methods have to be developed for this purpose.

Theorem 1 leads to an important result on the number of factors in an $\text{SOA}(n, m, 8, 3)$ with property α .

Proposition 2. *If an $\text{SOA}(n, m, 8, 3)$ with property α , as characterized in Theorem 1, exists, then it must hold that $m \leq 5n/16$.*

Proof. By Theorem 1, design A has resolution IV or higher. A resolution IV or higher design is either even or even/odd. A design is even if the lengths of the words in its defining relation are all even, and it is even/odd if its defining relation contains words of both even and odd lengths. If A is even/odd, then we have that $m \leq 5n/16$ (Butler [3]; Chen and Cheng [4]). Now suppose that A is even. Then A is a subset of a saturated resolution IV design, say Q , which has form $Q = (e_k, e_k Q_0)$, where e_1, \dots, e_k are independent factors and Q_0 consists of e_1, \dots, e_{k-1} , and all their interaction columns. Clearly, $S = (Q_0, Q)$. For each $j = 1, \dots, m$, as $a_j = b_j b'_j$, one of b_j and b'_j must be in Q_0 and the other in Q . Since A does not share any column with B or B' , we therefore have that $2m \leq n/2$, with $n/2$ being the number of columns in Q . This shows that $m \leq n/4$ if A is even. The proof is completed. \square

The proof of Proposition 2 actually reveals that, in order to construct an $\text{SOA}(n, m, 8, 3)$ with property α , if an even A is used, it is impossible to obtain more than $n/4$ factors, and one has to consider an even/odd A in order to break this barrier.

We next present a recursive construction of designs A , B and B' needed in Theorem 1. Recall that $B' = (b'_1, \dots, b'_m)$ is determined by $A = (a_1, \dots, a_m)$ and $B = (b_1, \dots, b_m)$ via $b'_j = a_j b_j$. Let A_k, B_k and B'_k , based on k independent factors e_1, \dots, e_k , satisfy the condition in Theorem 1 that A_k is of resolution IV or higher and (A_k, B_k, B'_k) is of resolution III or higher. Then A_{k+2}, B_{k+2} and B'_{k+2} , based on $k + 2$ independent factors e_1, \dots, e_{k+2} , can be constructed to satisfy the requirement in Theorem 1. This is done by defining

$$\begin{aligned} A_{k+2} &= (A_k, e_{k+1}A_k, e_{k+2}A_k, e_{k+1}e_{k+2}A_k), \\ B_{k+2} &= (B_k, e_{k+2}B_k, e_{k+1}e_{k+2}B_k, e_{k+1}B_k). \end{aligned} \tag{2}$$

Then $B'_{k+2} = (B'_k, e_{k+1}e_{k+2}B'_k, e_{k+1}B'_k, e_{k+2}B'_k)$. It is straightforward to verify that A_{k+2} has resolution IV or higher and $(A_{k+2}, B_{k+2}, B'_{k+2})$ has resolution III or higher.

Essentially, A_{k+2}, B_{k+2} and B'_{k+2} are obtained by doubling A_k, B_k and B'_k twice, respectively. However, their columns are re-arranged in order for $(A_{k+2}, B_{k+2}, B'_{k+2})$ to have resolution III or higher. The above construction of A_{k+2} and B_{k+2} from A_k and B_k gives a recursive construction of an $\text{SOA}(n, m, 8, 3)$ with property α .

Proposition 3. *Suppose that an $\text{SOA}(n, m, 8, 3)$ with property α is available. Then an $\text{SOA}(4n, 4m, 8, 3)$ with property α can be constructed.*

Table 1. Designs A and B for constructing an SOA(128, 40, 8, 3) with property α

A				B			
e_1	e_1e_6	e_1e_7	$e_1e_6e_7$	$e_2e_4e_6$	$e_3e_4e_5e_7$	$e_2e_3e_5e_6$	$e_3e_4e_5e_6$
e_2	e_2e_6	e_2e_7	$e_2e_6e_7$	$e_1e_3e_6$	$e_1e_3e_7$	$e_3e_4e_5$	e_3e_4
e_3	e_3e_6	e_3e_7	$e_3e_6e_7$	$e_1e_4e_5$	$e_1e_2e_3e_5$	$e_1e_2e_3$	e_1e_4
e_4	e_4e_6	e_4e_7	$e_4e_6e_7$	e_1e_2	e_2e_3	e_1e_3	$e_1e_3e_4$
$e_1e_2e_3e_4$	$e_1e_2e_3e_4e_6$	$e_1e_2e_3e_4e_7$	$e_1e_2e_3e_4e_6e_7$	$e_1e_3e_5e_6e_7$	$e_1e_2e_5e_7$	$e_1e_4e_5e_6$	$e_2e_4e_5e_6$
e_1e_5	$e_1e_5e_6$	$e_1e_5e_7$	$e_1e_5e_6e_7$	$e_2e_4e_7$	$e_2e_4e_5e_7$	$e_1e_3e_4e_5e_6$	$e_1e_2e_3e_5e_6$
e_2e_5	$e_2e_5e_6$	$e_2e_5e_7$	$e_2e_5e_6e_7$	$e_1e_4e_6$	$e_1e_4e_7$	$e_3e_4e_6$	$e_2e_3e_4e_5e_6$
e_3e_5	$e_3e_5e_6$	$e_3e_5e_7$	$e_3e_5e_6e_7$	e_2e_4	$e_1e_2e_3e_7$	$e_2e_3e_5$	$e_1e_2e_6$
e_4e_5	$e_4e_5e_6$	$e_4e_5e_7$	$e_4e_5e_6e_7$	$e_2e_3e_4$	$e_1e_3e_5$	$e_1e_2e_5$	$e_1e_2e_4e_5$
$e_1e_2e_3e_4e_5$	$e_1e_2e_3e_4e_5e_6$	$e_1e_2e_3e_4e_5e_7$	$e_1e_2e_3e_4e_5e_6e_7$	$e_2e_3e_6e_7$	$e_2e_3e_5e_7$	$e_1e_3e_5e_6$	$e_2e_3e_6$

For $k = 4$ and $n = 16$, one can easily see that $A_4 = (e_1, e_2, e_3, e_4, e_1e_2e_3e_4)$ and $B_4 = (e_3e_4, e_1e_4, e_1e_2, e_2e_3, e_1e_3)$ satisfy the requirement. For $k = 7$ and $n = 128$, the required A_7 and B_7 have been found to have $m = 40$ factors, and are presented in Table 1.

Combining the construction of A_k and B_k for $k = 4$ and 7 with Propositions 2 and 3, we obtain another major result.

Theorem 2. *With the exception of $k = 5$, we have that*

- (i) *an SOA($n = 2^k, m, 8, 3$) with property α can be constructed for every $k \geq 4$ and for $m = 5n/16$ factors, and*
- (ii) *the SOA($n, 5n/16, 8, 3$) with property α given in (i) has the maximum number of factors.*

For $k = 5$, the maximum number m of factors for desired A and B is 9. This follows from Wu [27] or by a direct check on a computer. We have that

$$A_5 = (e_1, e_2, e_3, e_4, e_5, e_1e_2e_3, e_1e_2e_4, e_1e_2e_5, e_1e_3e_4e_5),$$

$$B_5 = (e_4e_5, e_3e_5, e_1e_4, e_2e_3, e_1e_3, e_1e_2e_4e_5, e_1e_5, e_3e_4, e_1e_2),$$

based on which an SOA(32, 9, 8, 3) with property α can be constructed. This array was displayed earlier in Example 1.

In addition to A and B , to construct an SOA($n, m, 8, 3$) as in (1), design C is also required. For the SOAs discussed in this subsection, design C can be trivially obtained. For given $A = (a_1, \dots, a_m)$ and $B = (b_1, \dots, b_m)$, one can use $C = (c_1, \dots, c_m)$ by taking c_j to be any column other than a_j, b_j and a_jb_j for $j = 1, \dots, m$.

3.2. Designs with property β

For this class of arrays, the following characterization can be obtained.

Theorem 3. *If an SOA($n, m, 8, 3$) is to be constructed using regular A, B and C , then it has property β if and only if A is of resolution IV or higher, $(B, B') \subseteq \bar{A}$ and (B, B') does not contain any interaction column involving two factors from A , where $\bar{A} = S \setminus A$.*

Proof. We only need to give a proof for the sufficiency part, as will be seen, all of our arguments are reversible. According to Lemma 1 and Proposition 1(ii), we need to show that (a_i, a_j, a_u, b_u) does not have a defining word for any distinct i, j, u . The three columns a_i, a_j, a_u do not form a word of length three because A has resolution IV or higher, Columns a_i, a_j, b_u do not form a word of length three because B does not contain an interaction column involving two factors from A . Columns a_i, a_u, b_u do not form a word of length three because $(B, B') \subseteq \bar{A} = S \setminus A$. Finally, a_i, a_j, a_u, b_u together cannot give a word of length four, since B' has no column that is an interaction involving two factors from A . We have thus completed the proof. \square

Note that (B, B') in Theorem 3 is allowed to have repeated columns. Based on Theorem 3 and the theory of doubling (Chen and Cheng [4]), the next result can be established.

Proposition 4. *If an SOA($n, m, 8, 3$) with property β , as characterized in Theorem 3, exists, we must have that $m \leq n/4$.*

Proof. The basic idea of the proof is by contradiction. We will show that if $m \geq n/4 + 1$, then it is impossible to find designs A and B that satisfy the conditions as required in Theorem 3.

Now suppose $m \geq n/4 + 1$. As A has resolution IV or higher, it is either maximal or a projection design of a maximal design (Chen and Cheng [4], Proposition 3.1). If A is maximal, then the two-factor interactions of A use up all the columns in \bar{A} . Therefore, there does not exist $B \subseteq \bar{A}$ that has no interaction column involving two factors of A .

Suppose that A is a projection design of a maximal design of resolution IV or higher, that is, A is obtained by deleting columns from a maximal design. Let W be this maximal design with w columns. Obviously, $w \geq m + 1$. Then by Theorems 3.4 and 3.5 of Chen and Cheng [4], we must have that $w = n/4 + 2^j$, where j is an integer satisfying $w \geq m + 1$, and that W can be obtained by doubling, repeatedly j times, a maximal design of $n/2^j$ runs for $n/2^{j+2} + 1$ factors. For W , consider its alias sets that do not contain main effects. Let l_1, \dots, l_f where $f = 2^k - 1 - w$ be the sizes of these alias sets. Then Theorems 2.2 and 3.2 of Chen and Cheng [4] imply that $l_i \geq 2^j$ for all $i = 1, \dots, f$. When one column is deleted from W , the size of each of these alias set decreases at most by one, since the two-factor interactions associated with this deleted column are mutually orthogonal and thus no more than one of them can belong to the same alias set. Therefore, if less than 2^j columns are deleted from W , then none of these alias sets will become empty. Recall that A is obtained from W by deleting $w - m$ columns and $w - m = n/4 + 2^j - m < 2^j$ as $m > n/4$. This means that each of these alias sets contains at least one two-factor interaction of A . Because of this, the columns of B in Theorem 3 can only be selected from $W \setminus A$ for otherwise B will contain a two-factor interaction of A . Because W has resolution IV or higher, the columns of B' must be all outside of W , and thus are two-factor interactions of A . This contradicts to the requirement for B' . The proof is finally completed. \square

A construction of designs A and B required in Theorem 3 is now presented. Again, let e_1, \dots, e_k be independent factors. Let P_0 consist of e_3, \dots, e_k and all of their interactions, and

let $P = (I, P_0)$ where I is the all-ones column. Then $S = (P_0, e_1P, e_2P, e_1e_2P)$. Now take $A = e_1P$ and $B = e_2P$. It can be easily seen that such A and B meet the requirements in Theorem 3. Note that A and B have $m = n/4$ factors.

Theorem 4. *With the above choice of A and B , we have that*

- (i) *an SOA($n = 2^k, m, 8, 3$) with property β can be constructed for $m = n/4$ factors, and*
- (ii) *the SOA($n, n/4, 8, 3$) given in (i) has the maximum number of factors among all SOA($n, m, 8, 3$) with property β .*

Theorem 4 is a direct consequence of Proposition 4 and Theorem 3. The remark about C at the end of Section 3.1 equally applies here.

3.3. Designs with more than one of properties α, β and γ

The problem of constructing SOA($n, m, 8, 3$)s with property α or β has been completely solved in the previous two subsections. We now consider the construction of SOA($n, m, 8, 3$)s with property γ . SOA($n, m, 8, 3$)s with property γ achieve stratifications on 2×8 and 8×2 grids in all two-dimensions. Unlike those with property α or β , a simple characterization as in Theorems 1 or 3 is not available for arrays with property γ , but a sufficient condition can be given.

Proposition 5. *Suppose that an SOA($n, m, 8, 3$) is to be constructed using regular A, B and C . If A and B satisfy that (i) A is resolution IV or higher, (ii) $(B, B') \subseteq \bar{A}$ shares no common column with $A^{(2)}$ where $A^{(2)}$ collects all the two-factor interactions of A , and (iii) the set of distinct columns in $(B, B', A^{(2)})$ is a subset of \bar{A} but is unequal to \bar{A} , then an SOA($n, m, 8, 3$) with property γ can be constructed. This array also has property β .*

Proof. Take $C = (c_1, \dots, c_m)$ where $c_j = c$ is a column in \bar{A} but not any column of $(B, B', A^{(2)})$. Then we can easily check that (a_i, a_j, b_j, c_j) is an orthogonal array of strength four for any $i \neq j$. This shows that array D constructed in (1) is an SOA($n, m, 8, 3$) with property γ . This array also has property β because A and B meet the requirements in Theorem 3. □

Let us go back to $S = (P_0, e_1P, e_2P, e_1e_2P)$ where $P = (I, P_0)$ and P_0 consists of e_3, \dots, e_k and all their interactions. If we take $A = e_1P_0$ and $B = e_2P_0$, then the conditions in Proposition 5 are all met.

Proposition 6. *An SOA($n = 2^k, m, 8, 3$) with properties β and γ can be constructed for $m = n/4 - 1$ factors.*

Proposition 6 inspires us to consider SOA($n, m, 8, 3$)s with all properties α, β and γ . It is intriguing and somewhat surprising that they can be constructed for $m = n/4 - 1$ factors.

Let X, Y, Z be three copies of P_0 , which have the same set of columns as P_0 but with their columns ordered differently. Suppose that they can be found such that $x_jy_j = z_j$ where

x_j, y_j, z_j are the j th column of X, Y, Z , respectively, for $j = 1, \dots, m = n/4 - 1$. Then $S = (P_0, e_1, e_2, e_1e_2, e_1X, e_2Y, e_1e_2Z)$. Now if we take

$$A = e_1X \quad \text{and} \quad B = e_2Y.$$

Then all the conditions in Theorems 1 and 3, and Proposition 5 are satisfied.

Theorem 5. *The above construction gives an SOA($n, n/4 - 1, 8, 3$) with all three properties α, β and γ .*

If we take

$$A = (e_1, e_1X) \quad \text{and} \quad B = (e_2, e_2Y)$$

instead, then Theorems 1 and 3 both apply.

Corollary 1. *An SOA($n, m, 8, 3$) with properties α and β can be constructed for $m = n/4$ factors.*

It remains to establish the existence of X, Y and Z satisfying that $x_j y_j = z_j$. Suppose that X_k, Y_k and Z_k have this property and are based on k independent factors e_1, \dots, e_k . Then X_{k+2}, Y_{k+2} and Z_{k+2} can be constructed recursively to have the same property as follows:

$$\begin{aligned} X_{k+2} &= (X_k, e_{k+1}, e_{k+1}X_k, e_{k+2}, e_{k+2}X_k, e_{k+1}e_{k+2}, e_{k+1}e_{k+2}X_k), \\ Y_{k+2} &= (Y_k, e_{k+2}, e_{k+2}Y_k, e_{k+1}e_{k+2}, e_{k+1}e_{k+2}Y_k, e_{k+1}, e_{k+1}Y_k), \\ Z_{k+2} &= (Z_k, e_{k+1}e_{k+2}, e_{k+1}e_{k+2}Z_k, e_{k+1}, e_{k+1}Z_k, e_{k+2}, e_{k+2}Z_k). \end{aligned}$$

For $k = 2$, X_k, Y_k, Z_k are given by $X_2 = (e_1, e_2, e_1e_2)$, $Y_2 = (e_2, e_1e_2, e_1)$ and $Z_2 = (e_1e_2, e_1, e_2)$. For $k = 3$, they are given by $X_3 = (e_1, e_2, e_1e_2, e_3, e_1e_3, e_2e_3, e_1e_2e_3)$, $Y_3 = (e_1e_2e_3, e_1e_3, e_2, e_1, e_2e_3, e_3, e_1e_2)$ and $Z_3 = (e_2e_3, e_1e_2e_3, e_1, e_1e_3, e_1e_2, e_2, e_3)$. We thus obtain the next result.

Theorem 6. *For any $k \geq 2$, X_k, Y_k, Z_k can be constructed to satisfy that $x_j y_j = z_j$.*

3.4. Three families of strength-three SOAs

Various characterizing and construction results have been presented in Sections 3.1, 3.2 and 3.3. We now highlight three families of strength-three SOAs and summarize their constructions. This should be helpful to those readers who are mainly interested in the final products, namely better strength-three SOAs, and are less concerned with the theoretical characterizations of these arrays.

The first family is given by Theorem 2, which asserts that an SOA($n = 2^k, m, 8, 3$) can be constructed to have property α for $m = 5n/16$ with the exception for $k = 5$ in which case $m = 9$ instead of $m = 10$. The construction of this array D is via (1) that requires arrays A, B and C . Arrays A and B can be constructed recursively by (2), with A_4 and B_4 given right after

Proposition 3, and A_7 and B_7 given in Table 1. Once $A = (a_1, \dots, a_m)$ and $B = (b_1, \dots, b_m)$ are available, array $C = (c_1, \dots, c_m)$ can easily be obtained by taking c_j to be any column other than a_j, b_j and $a_j b_j$ for $j = 1, \dots, m$. Obviously, any array given by deleting columns from the $SOA(n = 2^k, 5n/16, 8, 3)$ still has property α . But if $m \leq n/4$, better arrays exist, which we discuss next.

The second family of strength-three SOAs has properties both α and β . This is given by Corollary 1, which says that an $SOA(n = 2^k, m, 8, 3)$ with properties both α and β can be constructed for $m = n/4$ factors. If one needs an array with less than $n/4$ factors, one can simply delete some columns from the $SOA(n = 2^k, n/4, 8, 3)$ and the resulting array still has properties both α and β . This is unnecessary, however, due to the availability of the third family of strength-three SOAs.

The third family of strength-three SOAs enjoys all properties α, β and γ . According to Theorem 5, we can construct an $SOA(n = 2^k, m, 8, 3)$ with all properties α, β and γ for $m = n/4 - 1$ factors. Immediately available is an $SOA(n = 2^k, m, 8, 3)$ with all properties α, β and γ for $m < n/4 - 1$, which can be obtained by deleting columns from the one for $m = n/4 - 1$.

We have seen that the first, second and third families of strength-three SOAs enjoy increasingly better space-filling properties.

The constructions for the second and third families of SOAs are very much related, of which we now give a summary. The saturated design S based on k independent factors e_1, \dots, e_k can be written as

$$S = (P_0, e_1, e_2, e_1 e_2, e_1 X, e_2 Y, e_1 e_2 Z),$$

where P_0 consists of e_3, \dots, e_k and all their interactions, and X, Y and Z are three copies of P_0 obtained by permuting the columns of P_0 such that $x_j y_j = z_j$ with x_j, y_j and z_j being the j th column of X, Y and Z , respectively. The existence of X, Y and Z having this property is guaranteed by Theorem 6, with their constructions provided right before that theorem.

Recall the construction of strength-three SOAs using (1), which needs three arrays A, B and C . The second family of SOAs chooses

$$A = (e_1, e_1 X) \quad \text{and} \quad B = (e_2, e_2 Y),$$

with array $C = (c_1, \dots, c_m)$ given by taking c_j to be any column other than a_j, b_j and $a_j b_j$ where a_j and b_j are the j th columns of A and B , respectively. The third family of SOAs uses

$$A = e_1 X \quad \text{and} \quad B = e_2 Y,$$

along with array $C = (c_1, \dots, c_m)$ given by taking $c_j = e_1$ for all $j = 1, \dots, m$.

Table 2 provides a numerical comparison of the maximum numbers of factors for three families of SOAs of strength three and SOAs of strength four.

Theorem 5 is a remarkable result, as $SOA(n, m, 8, 3)$ s with all three properties α, β and γ achieve stratifications on (i) $4 \times 4, 8 \times 2$ and 2×8 grids in all two-dimensions and (ii) $4 \times 2 \times 2, 2 \times 4 \times 2$ and $2 \times 2 \times 4$ grids in all three-dimensions. The only property they do not have, when compared with SOAs of strength four, is to achieve stratifications on $2 \times 2 \times 2 \times 2$ grids in four-dimensions. SOAs of strength four are very expensive to construct. According to He and Tang ([11], Theorem 1), $SOA(n, m, 16, 4)$ can be constructed for up to $m = \lfloor M(k)/2 \rfloor$ factors where

Table 2. Maximum numbers of factors SOAs of strength three and four

k	$n = 2^k$	Strength three			Strength four
		Family 1	Family 2	Family 3	
4	16	5	4	3	2
5	32	9	8	7	3
6	64	20	16	15	4
7	128	40	32	31	5
8	256	80	64	63	8

$M(k)$ is the maximum number of factors for resolution V designs with $n = 2^k$ runs. Note that $M(k)/2$ is much smaller than $n/4 - 1$ as $M(k)$ is in the order of $O(\sqrt{n})$.

4. Discussion

This paper introduces and constructs several families of strength-three SOAs that enjoy some of the space-filling properties of strength-four SOAs. Various characterizing and construction results are presented. The theory of maximal designs and their doubling constructions plays a crucial role in many of the theoretical arguments. Strength-three SOAs constructed in this paper should provide very useful space-filling designs for computer experiments.

The present paper focuses on the construction using regular 2^{m-p} designs. One interesting direction worth pursuing is to consider the construction using regular s^{m-p} designs. Some results are possible although it seems unlikely that we can obtain many rich results as what has been done using 2^{m-p} designs. Another research direction is to examine the use of non-regular two-level designs. We feel that some of the constructions in this paper could be generalized to include non-regular situations. This deserves further investigation.

Acknowledgements

The research is supported by Natural Sciences and Engineering Research Council of Canada.

References

- [1] Bates, R.A., Buck, R.J., Riccomagno, E. and Wynn, H.P. (1996). Experimental design and observation for large systems. *J. Roy. Statist. Soc. Ser. B* **58** 77–94, 95–111. [MR1379235](#)
- [2] Bingham, D., Sitter, R.R. and Tang, B. (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika* **96** 51–65. [MR2482134](#) <https://doi.org/10.1093/biomet/asn057>
- [3] Butler, N.A. (2003). Some theory for constructing minimum aberration fractional factorial designs. *Biometrika* **90** 233–238. [MR1966563](#) <https://doi.org/10.1093/biomet/90.1.233>

- [4] Chen, H.H. and Cheng, C.-S. (2006). Doubling and projection: A method of constructing two-level designs of resolution IV. *Ann. Statist.* **34** 546–558. MR2275253 <https://doi.org/10.1214/009053605000000813>
- [5] Cheng, C.S. (2014). *Theory of Factorial Design: Single- and Multi-Stratum Experiments*. Boca Raton: CRC Press.
- [6] Dey, A. and Mukerjee, R. (1999). *Fractional Factorial Plans. Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. MR1679441 <https://doi.org/10.1002/9780470316986>
- [7] Fang, K.-T., Li, R. and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments. Chapman & Hall/CRC Computer Science and Data Analysis Series*. Boca Raton, FL: CRC Press/CRC. MR2223960
- [8] Fang, K.-T., Lin, D.K.J., Winker, P. and Zhang, Y. (2000). Uniform design: Theory and application. *Technometrics* **42** 237–248. MR1801031 <https://doi.org/10.2307/1271079>
- [9] Georgiou, S.D. and Efthimiou, I. (2014). Some classes of orthogonal Latin hypercube designs. *Statist. Sinica* **24** 101–120. MR3183676
- [10] Haaland, B. and Qian, P.Z.G. (2010). An approach to constructing nested space-filling designs for multi-fidelity computer experiments. *Statist. Sinica* **20** 1063–1075. MR2729853
- [11] He, Y. and Tang, B. (2013). Strong orthogonal arrays and associated Latin hypercubes for computer experiments. *Biometrika* **100** 254–260. MR3034340 <https://doi.org/10.1093/biomet/ass065>
- [12] He, Y. and Tang, B. (2014). A characterization of strong orthogonal arrays of strength three. *Ann. Statist.* **42** 1347–1360. MR3226159 <https://doi.org/10.1214/14-AOS1225>
- [13] Hedayat, A.S., Sloane, N.J.A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications. Springer Series in Statistics*. New York: Springer. MR1693498 <https://doi.org/10.1007/978-1-4612-1478-6>
- [14] Johnson, M.E., Moore, L.M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. *J. Statist. Plann. Inference* **26** 131–148. MR1079258 [https://doi.org/10.1016/0378-3758\(90\)90122-B](https://doi.org/10.1016/0378-3758(90)90122-B)
- [15] Lawrence, K.M. (1996). A combinatorial characterization of (t, m, s) -nets in base b . *J. Combin. Des.* **4** 275–293. MR1391810 [https://doi.org/10.1002/\(SICI\)1520-6610\(1996\)4:4<275::AID-JCD5>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1520-6610(1996)4:4<275::AID-JCD5>3.0.CO;2-C)
- [16] Lin, C.D. and Kang, L. (2016). A general construction for space-filling Latin hypercubes. *Statist. Sinica* **26** 675–690. MR3497766
- [17] Liu, H. and Liu, M.-Q. (2015). Column-orthogonal strong orthogonal arrays and sliced strong orthogonal arrays. *Statist. Sinica* **25** 1713–1734. MR3409089
- [18] McKay, M.D., Beckman, R.J. and Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21** 239–245. MR0533252 <https://doi.org/10.2307/1268522>
- [19] Moon, H., Dean, A. and Santner, T. (2011). Algorithms for generating maximin Latin hypercube and orthogonal designs. *J. Stat. Theory Pract.* **5** 81–98. MR2829824 <https://doi.org/10.1080/15598608.2011.10412052>
- [20] Mullen, G.L. and Schmid, W.C. (1996). An equivalence between (t, m, s) -nets and strongly orthogonal hypercubes. *J. Combin. Theory Ser. A* **76** 164–174. MR1406000 <https://doi.org/10.1006/jcta.1996.0098>
- [21] Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods. CBMS-NSF Regional Conference Series in Applied Mathematics* **63**. Philadelphia, PA: SIAM. MR1172997 <https://doi.org/10.1137/1.9781611970081>
- [22] Owen, A.B. (1992). Orthogonal arrays for computer experiments, integration and visualization. *Statist. Sinica* **2** 439–452. MR1187952

- [23] Owen, A.B. (1997). Scrambled net variance for integrals of smooth functions. *Ann. Statist.* **25** 1541–1562. [MR1463564](#) <https://doi.org/10.1214/aos/1031594731>
- [24] Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments*. *Springer Series in Statistics*. New York: Springer. [MR2160708](#) <https://doi.org/10.1007/978-1-4757-3799-8>
- [25] Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J. Amer. Statist. Assoc.* **88** 1392–1397. [MR1245375](#)
- [26] Wang, L., Xiao, Q. and Xu, H. (2018). Optimal maximin L_1 -distance Latin hypercube designs based on good lattice point designs. *Ann. Statist.* **46** 3741–3766. [MR3852667](#) <https://doi.org/10.1214/17-AOS1674>
- [27] Wu, C.-F.J. (1989). Construction of $2^m 4^n$ designs via a grouping scheme. *Ann. Statist.* **17** 1880–1885. [MR1026317](#) <https://doi.org/10.1214/aos/11176347399>
- [28] Ye, K.Q. (1998). Orthogonal column Latin hypercubes and their application in computer experiments. *J. Amer. Statist. Assoc.* **93** 1430–1439. [MR1666638](#) <https://doi.org/10.2307/2670057>

Received October 2018 and revised March 2019