Strong orthogonal arrays and associated Latin hypercubes for computer experiments

BY YUANZHEN HE

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

heyuanzhen@gmail.com

AND BOXIN TANG

Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

boxint@sfu.ca

SUMMARY

This paper introduces, constructs and studies a new class of arrays, called strong orthogonal arrays, as suitable designs for computer experiments. A strong orthogonal array of strength t enjoys better space-filling properties than a comparable orthogonal array in all dimensions lower than t while retaining the space-filling properties of the latter in t dimensions. Latin hypercubes based on strong orthogonal arrays of strength t are more space-filling than comparable orthogonal array-based Latin hypercubes in all g dimensions for any $2 \leqslant g \leqslant t - 1$.

Some key words: Orthogonal array-based Latin hypercube; Space-filling design; (t, m, s)-net.

1. INTRODUCTION

Computer experiments are concerned with building a statistical surrogate model based on the data from running a computer code, and designing such an experiment is a crucial step in this process of model building. Space-filling designs have been widely accepted as appropriate for this purpose. A space-filling design is any design that strews its points in the design region in some uniform fashion. Uniformity may be evaluated using a distance criterion (Johnson et al., 1990) or a discrepancy criterion (Fang & Mukerjee, 2000). For a high-dimensional input space, it is more fruitful to consider designs that are space-filling in lower dimensional projections. The idea of Latin hypercube designs is to achieve the maximum uniformity in all one-dimensional projections (McKay et al., 1979). Orthogonal array-based Latin hypercubes (Tang, 1993) carry this idea further, giving designs that, in addition to being Latin hypercubes, achieve uniformity in t-dimensional margins when orthogonal arrays of strength t are employed.

In this paper, we introduce, construct and study a new class of arrays, strong orthogonal arrays, for computer experiments. A strong orthogonal array of strength t does as well as a comparable orthogonal array in t-dimensional projections, but the former achieves uniformity on finer grids than the latter in all g-dimensional projections for any g less than t. Consequently, Latin hypercubes constructed from a strong orthogonal array of strength t are more space-filling than comparable orthogonal array-based Latin hypercubes in all g-dimensional projections for any $2 \leqslant g \leqslant t - 1$. The concept of strong orthogonal arrays is motivated by the notion of nets from quasi-Monte Carlo methods (Niederreiter, 1987). The formulation of this new concept has two advantages. Firstly, strong orthogonal arrays are more general than nets in terms of run sizes; and secondly, strong orthogonal arrays are defined in the form and language that are familiar to design practitioners and researchers. This not only makes existing results from nets more accessible to the design community but also allows us to obtain new designs and theoretical results.

2. Introducing strong orthogonal arrays

2·1. *Notation and background*

An $n \times m$ matrix A where the jth column has s_j levels $\{0, 1, \ldots, s_j - 1\}$ is said to be an orthogonal array of size n, m factors, and strength t if for any $n \times t$ submatrix of A, every possible level combination occurs with the same frequency. Such an array is denoted by $\mathrm{OA}(n, m, s_1 \times \cdots \times s_m, t)$. When $s_1 = \cdots = s_m = s$, the array is symmetric and denoted by $\mathrm{OA}(n, m, s, t)$. Orthogonal arrays play a prominent role in both statistical and combinatorial design, and have become the backbone of designs for multifactor experiments. For detailed accounts on orthogonal arrays, we refer to Dey & Mukerjee (1999) and Hedayat et al. (1999).

An elementary interval in base s is an interval in $[0, 1)^m$ of form $\prod_{j=1}^{m}[c_j/s^{d_j}, (c_j + 1)/s^{d_j})$ where nonnegative integers c_j and d_j satisfy $0 \leqslant c_j < s^{d_j}$. For $0 \leqslant w \leqslant k$, a (w, k, m)-net in base s is a set of s^k points in $[0, 1)^m$ such that every elementary interval in base s of volume s^{w-k} contains exactly s^w points. Nets and related sequences were first defined by Sobol' (1967) for $s = 2$ and later by Niederreiter (1987) for general s.

2·2. *Strong orthogonal arrays*

Before giving a formal definition, let us first look at an example. Consider the transposed array

$$
\begin{matrix}
4 & 7 & 5 & 6 & 7 & 7 & 5 & 4 & 4 & 6 & 5 & 6 & 3 & 0 & 2 & 1 & 0 & 0 & 2 & 3 & 3 & 1 & 2 & 1 \\
0 & 6 & 5 & 3 & 6 & 7 & 5 & 1 & 0 & 2 & 4 & 3 & 7 & 1 & 2 & 4 & 1 & 0 & 2 & 6 & 7 & 5 & 3 & 4 \\
0 & 3 & 4 & 7 & 3 & 6 & 5 & 5 & 1 & 2 & 0 & 6 & 7 & 4 & 3 & 0 & 4 & 1 & 2 & 2 & 6 & 5 & 7 & 1 \\
0 & 6 & 1 & 6 & 7 & 3 & 4 & 5 & 5 & 3 & 0 & 2 & 7 & 1 & 6 & 1 & 0 & 4 & 3 & 2 & 2 & 4 & 7 & 5 \\
0 & 2 & 4 & 3 & 6 & 7 & 1 & 4 & 5 & 7 & 1 & 2 & 7 & 5 & 3 & 4 & 1 & 0 & 6 & 3 & 2 & 0 & 6 & 5 \\
0 & 2 & 0 & 6 & 3 & 6 & 5 & 1 & 4 & 7 & 5 & 3 & 7 & 5 & 7 & 1 & 4 & 1 & 2 & 6 & 3 & 0 & 2 & 4 \\
0 & 3 & 0 & 2 & 6 & 3 & 4 & 5 & 1 & 6 & 5 & 7 & 7 & 4 & 7 & 5 & 1 & 4 & 3 & 2 & 6 & 1 & 2 & 0 \\
0 & 7 & 1 & 2 & 2 & 6 & 1 & 4 & 5 & 3 & 4 & 7 & 7 & 0 & 6 & 5 & 5 & 1 & 6 & 3 & 2 & 4 & 3 & 0 \\
0 & 7 & 5 & 3 & 2 & 2 & 4 & 1 & 4 & 7 & 1 & 6 & 7 & 0 & 2 & 4 & 5 & 5 & 3 & 6 & 3 & 0 & 6 & 1 \\
0 & 6 & 5 & 7 & 3 & 2 & 0 & 4 & 1 & 6 & 5 & 3 & 7 & 1 & 2 & 0 & 4 & 5 & 7 & 3 & 6 & 1 & 2 & 4 \\
1 & 3 & 5 & 7 & 7 & 3 & 1 & 1 & 5 & 3 & 5 & 7 & 6 & 4 & 2 & 0 & 0 & 4 & 6 & 6 & 2 & 4 & 2 & 0.
\end{matrix}
\tag{1}
$$

This array has the following interesting properties.
(i) The array becomes an $\mathrm{OA}(24, 11, 2, 3)$ after the eight levels are collapsed into two levels according to

$$
[a/4] =
\begin{cases}
0, & a = 0, 1, 2, 3, \\
1, & a = 4, 5, 6, 7,
\end{cases}
$$

where $[x]$ denotes the largest integer not exceeding x.

(ii) Any subarray of two columns can be collapsed into an $\mathrm{OA}(24, 2, 2 \times 4, 2)$ as well as an $\mathrm{OA}(24, 2, 4 \times 2, 2)$, where collapsing into two levels is done by $[a/4]$ and collapsing into four levels is done using $[a/2]$.

(iii) Any subarray of one column is an $\mathrm{OA}(24, 1, 8, 1)$.

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

With this definition, the array in (1) is an $\mathrm{SOA}(24, 11, 8, 3)$. Since an $\mathrm{SOA}(n, m, s^t, t)$ can be collapsed into an $\mathrm{OA}(n, m, s, t)$, as seen by taking $g = t$ and $u_1 = \cdots = u_t = 1$ in Definition 1, we must have $n = \lambda s^t$ for some integer λ. We call this λ the index of the strong orthogonal array in the same way as that of an orthogonal array.

The notion of strong orthogonal arrays is motivated by nets and in fact the former include the latter as special cases. To see this, consider the case where $\lambda = s^w$ for some integer w. Then it is not hard to see that the set of $n = s^{w+t}$ points, $(x_{i1}, \ldots, x_{im})$ for $i = 1, \ldots, n$, obtained from $A = (a_{ij})$, an SOA(n, m, s^t, t), via $x_{ij} = (a_{ij} + 0.5)/s^t$ is a $(w, w+t, m)$-net in base s. Conversely, suppose that a set of $n = s^k$ points $(x_{i1}, \ldots, x_{im})$ for $i = 1, \ldots, n = s^k$ in $[0, 1)^m$ is a (w, k, m)-net in base s. Then the array $A = (a_{ij})$ given by $a_{ij} = [x_{ij}s^{k-w}]$ is an SOA$(s^k, m, s^{k-w}, k - w)$. We document this in a proposition.

PROPOSITION 1. *If $\lambda = s^w$ for integer w, then the existence of an SOA$(\lambda s^t, m, s^t, t)$ is equivalent to that of a (w, k, m)-net in base s where $k = w + t$.*

As strong orthogonal arrays are defined without restricting the index to be a power of s, they provide a much broader concept than (w, k, m)-nets. In § 3, we will discuss several families of such arrays that would not be possible by just considering (w, k, m)-nets. Strong orthogonal arrays are formulated in design language, and as such they will help promote the application of nets to computer experiments and should also facilitate cross-fertilization between experimental design, combinatorics and quasi-Monte Carlo methods.

2·3. *Latin hypercubes based on strong orthogonal arrays*

A Latin hypercube of n runs for m factors is an $n \times m$ matrix in which each column is a permutation of $0, 1, \ldots, n - 1$. Consider the index λ in an SOA(n, m, s^t, t) where $n = \lambda s^t$. If $\lambda = 1$, then the SOA(n, m, s^t, t) itself is a Latin hypercube. When $\lambda \geqslant 2$, the strong orthogonal array is not a Latin hypercube but can be easily converted into one by expanding the s^t levels into $n = \lambda s^t$ levels. For each column, this is done by replacing the λ entries for level j by any permutation of $j\lambda, j\lambda + 1, \ldots, (j + 1)\lambda - 1$ for $j = 0, 1, \ldots, s^t - 1$. The Latin hypercubes generated this way are called strong orthogonal array-based Latin hypercubes.

An SOA(n, m, s^t, t) promises a stratification of design points on an $s^{u_1} \times \cdots \times s^{u_g}$ grid for any positive u_j with $u_1 + \cdots + u_g = t$ when projected onto any $g \leqslant t$ dimensions. A Latin hypercube based on an SOA(n, m, s^t, t) further maximizes such a stratification for one-dimensional projections while retaining all the projection properties the strong orthogonal array can offer.

Starting with an OA(n, m, s, t), one can construct an orthogonal array-based Latin hypercube by, for each column, replacing the n/s entries for level j by any permutation of $jn/s, jn/s + 1, \ldots, (j + 1)n/s - 1$ for $j = 0, 1, \ldots, s - 1$ (Tang, 1993). Because of the simplicity of the method and/or the availability of online orthogonal arrays, orthogonal array-based Latin hypercubes have received considerable attention in the literature of computer experiments. The rest of the section discusses how Latin hypercubes based on strong orthogonal arrays compare with those based on orthogonal arrays.

In the case of strength two, a Latin hypercube based on an SOA$(n, m, s^2, 2)$ achieves a stratification on an $s \times s$ grid in any two dimensions and so does a Latin hypercube based on an OA$(n, m, s, 2)$. Thus, Latin hypercubes based on arrays OA$(n, m, s, 2)$ do as well as those based on arrays SOA$(n, m, s^2, 2)$ in terms of filling two-dimensional projections. For higher strength, strong orthogonal array-based Latin hypercubes are better. Let us look at the case of strength three. Consider two Latin hypercubes, one based on an OA$(n, m, s, 3)$ and the other based on an SOA$(n, m, s^3, 3)$. Both achieve a stratification on an $s \times s \times s$ grid in any three-dimensional projection. But the Latin hypercube from the SOA$(n, m, s^3, 3)$ achieves stratifications on $s^2 \times s$ and $s \times s^2$ grids in any two-dimensional projection, while the one from the OA$(n, m, s, 3)$ promises only a stratification on an $s \times s$ grid for two dimensions.

This gain comes with a price. The existence of an SOA(n, m, s^t, t) implies the existence of an OA(n, m, s, t) but the converse is not true in general. The next section examines the construction of strong orthogonal arrays from ordinary orthogonal arrays. We will see that an SOA$(n, m - 1, s^3, 3)$ can be constructed from an OA$(n, m, s, 3)$. Losing only one column in the case of strength three is a small price to pay for being able to have strong orthogonal arrays. For strength $t \geqslant 4$, the price gets higher as more columns have to be sacrificed to obtain strong orthogonal arrays from ordinary orthogonal arrays.

3. CONSTRUCTING STRONG ORTHOGONAL ARRAYS

The main result of this section is Theorem 1, stating that strong orthogonal arrays can be constructed from ordinary orthogonal arrays.

THEOREM 1. *If an* $OA(n, m, s, t)$ *exists, then an* $SOA(n, m', s^t, t)$ *can be constructed where*

$$m' = \begin{cases} [m/e], & \text{if } t = 2e \text{ is even,} \\ [(m-1)/e], & \text{if } t = 2e+1 \text{ is odd.} \end{cases}$$

Theorem 1 follows from combining Proposition 2 and Lemma 1, to be presented in the remainder of this section. Lawrence (1996) introduced the concept of a generalized orthogonal array. Our definition below is given in a slightly different format.

An $n \times (tm)$ matrix B with entries from $\{0, 1, \ldots, s-1\}$, where the tm columns b_{ij} are put into m groups of t columns each, $B = (B_1, \ldots, B_m)$ and $B_i = (b_{i1}, \ldots, b_{it})$, is called a generalized orthogonal array of size n, m constraints, s levels and strength t if the matrix B^* consisting of t columns b_{ij} where $i = i_1, \ldots, i_g$ and $j = 1, \ldots, u_i$ is an orthogonal array of strength t for any $1 \leqslant g \leqslant t$, any $1 \leqslant i_1 < \cdots < i_g \leqslant m$ and any positive integers u_k with $u_1 + \cdots + u_g = t$. The matrix B^* selects the first u_1 columns from B_{i_1}, and the first u_2 columns from B_{i_2}, and so on. We use $GOA(n, m, s, t)$ to denote such an array.

For an ordinary orthogonal array $OA(n, m, s, t)$, any submatrix of t columns is an orthogonal array of strength t. In a generalized orthogonal array $GOA(n, m, s, t)$, which has a total of mt columns, only a special collection of submatrices of t columns is required to have this property. Within each of the m groups of columns, selection of columns must be done in an orderly fashion and no column may be skipped. This is the reason why generalized orthogonal arrays are also called ordered orthogonal arrays (Mullen & Schmid, 1996). Our presentation is in line with ordered orthogonal arrays.

We give a simple example illustrating the concept. Let a_1, a_2, a_3, a_4 be the four columns of an $OA(8, 4, 2, 3)$. Then $B = \{(a_1, a_4, a_2); (a_2, a_4, a_3); (a_3, a_4, a_1)\}$ is a $GOA(8, 3, 2, 3)$ where we have $B_1 = (a_1, a_4, a_2)$, $B_2 = (a_2, a_4, a_3)$ and $B_3 = (a_3, a_4, a_1)$.

Consider a submatrix of B consisting of columns $b_{11}, \ldots, b_{m1}$. Clearly, this submatrix is an $OA(n, m, s, t)$, which can be seen by choosing $u_1 = \cdots = u_t = 1$ for $g = t$ and any $1 \leqslant i_1 < \cdots < i_t \leqslant m$. This shows that the existence of a $GOA(n, m, s, t)$ implies the existence of an $OA(n, m, s, t)$. So we must have $n = \lambda s^t$ for some integer λ, which is called the index of the generalized orthogonal array.

PROPOSITION 2. *Let* $B = \{(b_{11}, \ldots, b_{1t}); \ldots; (b_{m1}, \ldots, b_{mt})\}$ *be a* $GOA(n, m, s, t)$. *Then* $A = (a_1, \ldots, a_m)$ *is an* $SOA(n, m, s^t, t)$, *where*

$$a_i = \sum_{j=1}^{t} b_{ij} s^{t-j}. \tag{2}$$

Conversely, if $A = (a_1, \ldots, a_m)$ *is an* $SOA(n, m, s^t, t)$, *then* $B = \{(b_{11}, \ldots, b_{1t}); \ldots; (b_{m1}, \ldots, b_{mt})\}$ *is a* $GOA(n, m, s, t)$, *where the* b_{ij} *are uniquely determined by* a_i *as given in* (2).

Proposition 2 implies in particular that the existence of a strong orthogonal array is equivalent to that of a generalized orthogonal array. Lawrence (1996) showed that the existence of a net is equivalent to the existence of a special generalized orthogonal array. Without much modification, his method can also be used to establish Proposition 2. For the benefits of design researchers, a version of the proof using design language is given in the Appendix.

LEMMA 1. *If an* $OA(n, m, s, t)$ *exists, then a* $GOA(n, m', s, t)$ *can be constructed where*

$$m' = \begin{cases} [m/e], & \text{if } t = 2e \text{ is even,} \\ [(m-1)/e], & \text{if } t = 2e+1 \text{ is odd.} \end{cases}$$

Lemma 1 is due to Lawrence (1996). Let $C = (c_1, \ldots, c_m)$ be an $OA(n, m, s, t)$. We present in detail the construction of a $GOA(n, m', s, t)$, $B = \{(b_{11}, \ldots, b_{1t}); \ldots; (b_{m'1}, \ldots, b_{m't})\}$ where m' is

as in Lemma 1, for the most useful cases $t = 2, 3, 4, 5$, which should be sufficient for the reader to see the generality of the method.

For $t = 2$, we have $m' = m$. The b_{ij} in $B = \{(b_{11}, b_{12}); \ldots; (b_{m1}, b_{m2})\}$ are obtained by $(b_{11}, \ldots, b_{m1}) = (c_1, \ldots, c_m)$ and $(b_{12}, \ldots, b_{m2}) = (c_2, \ldots, c_m, c_1)$.

For $t = 3$, we have $m' = m - 1$. The b_{ij} in $B = \{(b_{11}, b_{12}, b_{13}); \ldots; (b_{m'1}, b_{m'2}, b_{m'3})\}$ are given by $(b_{11}, \ldots, b_{m'1}) = (c_1, \ldots, c_{m'})$, $(b_{12}, \ldots, b_{m'2}) = (c_m, \ldots, c_m)$, and $(b_{13}, \ldots, b_{m'3}) = (c_2, \ldots, c_{m'}, c_1)$.

For $t = 4$, we have $m' = [m/2]$. We obtain $B = \{(b_{11}, b_{12}, b_{13}, b_{14}); \ldots; (b_{m'1}, b_{m'2}, b_{m'3}, b_{m'4})\}$ by setting $(b_{11}, \ldots, b_{m'1}) = (c_1, \ldots, c_{m'})$, $(b_{12}, \ldots, b_{m'2}) = (c_{m'+1}, \ldots, c_{2m'})$, $(b_{13}, \ldots, b_{m'3}) = (c_{m'+2}, \ldots, c_{2m'}, c_{m'+1})$, and $(b_{14}, \ldots, b_{m'4}) = (c_2, \ldots, c_{m'}, c_1)$.

For $t = 5$, we have $m' = [(m - 1)/2]$. We obtain $B = \{(b_{11}, \ldots, b_{15}); \ldots; (b_{m'1}, \ldots, b_{m'5})\}$ by letting $(b_{11}, \ldots, b_{m'1}) = (c_1, \ldots, c_{m'})$, $(b_{12}, \ldots, b_{m'2}) = (c_{m'+1}, \ldots, c_{2m'})$, $(b_{13}, \ldots, b_{m'3}) = (c_m, \ldots, c_m)$, $(b_{14}, \ldots, b_{m'4}) = (c_{m'+2}, \ldots, c_{2m'}, c_{m'+1})$, and $(b_{15}, \ldots, b_{m'5}) = (c_2, \ldots, c_{m'}, c_1)$.

Various strong orthogonal arrays can be constructed using Theorem 1. Here we indicate three families of strong orthogonal arrays that cannot be obtained from nets. The Addelman–Kempthorne construction gives an $\mathrm{OA}\{2s^k, 2(s^k - 1)/(s - 1) - 1, s, 2\}$ for odd prime power s, from which we obtain an $\mathrm{SOA}\{2s^k, 2(s^k - 1)/(s - 1) - 1, s^2, 2\}$. If a Hadamard matrix of order 4λ exists, implying the existence of an $\mathrm{OA}(4\lambda, 4\lambda - 1, 2, 2)$, then Theorem 1 allows an $\mathrm{SOA}(4\lambda, 4\lambda - 1, 4, 2)$ to be constructed. Again, if a Hadamard matrix of order 4λ exists, which implies the existence of an $\mathrm{OA}(8\lambda, 4\lambda, 2, 3)$, then we can construct an $\mathrm{SOA}(8\lambda, 4\lambda - 1, 8, 3)$. The $\mathrm{SOA}(24, 11, 8, 3)$ in § 2·2 is a member of this family.

4. Discussion and further results

For a given set of parameters, some orthogonal arrays are better than others; the generalized minimum aberration can be used to select an optimal orthogonal array (Tang & Deng, 1999; Xu & Wu, 2001). The same issue exists for strong orthogonal arrays. The generalized minimum aberration should still be useful for selecting strong orthogonal arrays. Exactly how this can be done is not straightforward, and it would be of interest to obtain some results in this direction. Orthogonal Latin hypercubes are useful when regression type models are considered (Ye, 1998), and they can also be viewed as stepping stones to space-filling designs (Lin et al., 2009). Strong orthogonal array-based Latin hypercubes are tailored to achieve low-dimensional space-filling. When $s^{[t/2]}$ is not too small, an $\mathrm{SOA}(n, m, s^t, t)$ should perform well in terms of orthogonality because of its stratification on $s^{[t/2]} \times s^{t-[t/2]}$ grids in two dimensions. It would be nice if this could be quantified in some fashion, and we leave the problem for future research.

Let $h(n, s, t)$ denote the largest m for an $\mathrm{SOA}(n, m, s^t, t)$ to exist. The design literature has rich results on $f(n, s, t)$, the largest m for an $\mathrm{OA}(n, m, s, t)$ to exist; see Hedayat et al. (1999). Knowing how $h(n, s, t)$ depends on $f(n, s, t)$ is practically useful in addition to providing theoretical insights. We conclude the paper with an investigation into this problem. We first have $f(n, s, t) \geqslant h(n, s, t)$, since an $\mathrm{OA}(n, m, s, t)$ can be constructed from an $\mathrm{SOA}(n, m, s^t, t)$ by level collapsing. Theorem 1 implies that $h(n, s, t) \geqslant [f(n, s, t)/e]$ for even $t = 2e$ and $h(n, s, t) \geqslant [\{f(n, s, t) - 1\}/e]$ for odd $t = 2e + 1$. For $t = 3$, we obtain $f(n, s, 3) \geqslant h(n, s, 3) \geqslant f(n, s, 3) - 1$. We now show that a stronger result holds for $s = 2$.

Theorem 2. *We have $h(n, 2, 3) = f(n, 2, 3) - 1$.*

The proof of Theorem 2 is given in the Appendix. Theorem 2 does not require the existence of relevant Hadamard matrices. However, if a Hadamard matrix of order 4λ exists, then $f(8\lambda, 2, 3) = 4\lambda$ and thus $h(8\lambda, 2, 3) = 4\lambda - 1$.

Acknowledgement

APPENDIX

Proof of Proposition 2

Let $B = \{(b_{11}, \ldots, b_{1t}); \ldots; (b_{m1}, \ldots, b_{mt})\}$ be a GOA(n, m, s, t) with index λ. Let $a_i = \sum_{j=1}^{t} b_{ij} s^{t-j}$. Then it can be verified that $A = (a_1, \ldots, a_m)$ is an SOA(n, m, s^t, t). We do this for $t = 3$ for ease of presentation. The same idea applies to the general t. When $t = 3$, we obtain

$$a_i = b_{i1}s^2 + b_{i2}s + b_{i3}. \tag{A1}$$

Since B is a GOA$(n, m, s, 3)$, we have that $B_i = (b_{i1}, b_{i2}, b_{i3})$ is an OA$(n, 3, s, 3)$. Thus, each 3-tuple (x_1, x_2, x_3) where $x_j = 0, 1, \ldots, s - 1$ occurs λ times as a row of $B_i = (b_{i1}, b_{i2}, b_{i3})$. The equation $x = x_1 s^2 + x_2 s + x_3$ provides a one-one correspondence between the set of integers $x = 0, 1, \ldots, s^3 - 1$ and the set of 3-tuples (x_1, x_2, x_3) where $x_j = 0, 1, \ldots, s - 1$. Recalling equation (A1), we see that each integer $0 \leqslant x \leqslant s^3 - 1$ occurs λ times in vector a_i. This takes care of the requirement for $g = 1$ in Definition 1. We next show that the array (a_1, a_2) can be collapsed into an OA$(n, 2, s^2 \times s, 2)$. When collapsed into s^2 levels, a_1 becomes $b_{11}s + b_{12}$ and when collapsed into s levels, a_2 becomes b_{21}. As (b_{11}, b_{12}, b_{21}) is an OA$(n, 3, s, 3)$, it easily follows that $(b_{11}s + b_{12}, b_{21})$ is an OA$(n, 2, s^2 \times s, 2)$. Finally, we need to show that (a_1, a_2, a_3) can be collapsed into an OA$(n, 3, s, 3)$, which is immediate if one observes that when collapsed into s levels, a_i becomes b_{i1}.

Conversely, let $A = (a_1, \ldots, a_m)$ be an SOA(n, m, s^t, t) with index λ. Every integer $0 \leqslant x \leqslant s^t - 1$ can be uniquely written as $x = \sum_{j=1}^{t} x_j s^{t-j}$ for some x_j with $0 \leqslant x_j \leqslant s - 1$. Applying this fact to every component of vector a_i, we obtain $a_i = \sum_{j=1}^{t} b_{ij} s^{t-j}$ for unique vectors $b_{i1}, \ldots, b_{it}$, all with entries from $0, 1, \ldots, s - 1$. Then it can be verified that $B = \{(b_{11}, \ldots, b_{1t}); \ldots; (b_{m1}, \ldots, b_{mt})\}$ is a GOA(n, m, s, t) with index λ, which we will do for $t = 3$. Since $A = (a_1, \ldots, a_m)$ is an SOA(n, m, s^t, t) with index λ, each integer $0 \leqslant x \leqslant s^t - 1$ occurs λ times in a_i, which implies that each 3-tuple (x_1, x_2, x_3) where $x_j = 0, 1, \ldots, s - 1$ occurs λ times as a row of $B_i = (b_{i1}, b_{i2}, b_{i3})$ due to the unique representation $x = \sum_{j=1}^{t} x_j s^{t-j}$. Thus, $B_i = (b_{i1}, b_{i2}, b_{i3})$ is an OA$(n, 3, s, 3)$. That (a_1, a_2) can be collapsed into an OA$(n, 2, s^2 \times s, 2)$ means that $(b_{11}s + b_{12}, b_{21})$ is an OA$(n, 2, s^2 \times s, 2)$, implying that (b_{11}, b_{12}, b_{21}) is an OA$(n, 3, s, 3)$. Finally, that (b_{11}, b_{21}, b_{31}) is an OA$(n, 3, s, 3)$ follows from the fact that, when collapsed into s levels, (a_1, a_2, a_3) becomes (b_{11}, b_{21}, b_{31}).

Proof of Theorem 2

We prove that if an SOA$(n, m, 8, 3)$ exists, then an OA$(n, m + 1, 2, 3)$ can be constructed. Such a result implies that $f(n, 2, 3) \geqslant h(n, 2, 3) + 1$. Theorem 2 then follows as we already know that $h(n, 2, 3) \geqslant f(n, 2, 3) - 1$.

Suppose that an SOA$(n, m, 8, 3)$ exists. Then by Proposition 2, a GOA$(n, m, 2, 3)$ exists and let it be $B = \{(b_{11}, b_{12}, b_{13}); \ldots; (b_{m1}, b_{m2}, b_{m3})\}$. Consider the following $m + 1$ columns $b_{11}, \ldots, b_{m1}, b_{12}$. For convenience, let $c_j = b_{j1}$ for $j = 1, \ldots, m$ and $c_{m+1} = b_{12}$ and consider the array $P = (c_1, \ldots, c_{m+1})$. Because B is a generalized orthogonal array, three columns c_i, c_j, c_k form an orthogonal array of strength 3 for all $1 \leqslant i < j < k \leqslant m$ and for $i = 1$, all $2 \leqslant j \leqslant m$ and $k = m + 1$. Now consider a half fraction of P with m columns, obtained by selecting the rows of $(c_2, \ldots, c_{m+1})$ such that the entries of c_1 are at level 0, and denote this half fraction by Q. Then this Q is an OA$(n/2, m, 2, 2)$ as three columns c_1, c_i, c_j for all $2 \leqslant i < j \leqslant m + 1$ are an orthogonal array of strength 3. It is well known that the existence of an OA$(n/2, m, 2, 2)$ implies the existence of an OA$(n, m + 1, 2, 3)$. This completes the proof.

REFERENCES

DEY, A. & MUKERJEE, R. (1999). *Fractional Factorial Plans.* New York: Wiley.

FANG, K. T. & MUKERJEE, R. (2000). A connection between uniformity and aberration in regular fractions of two-level factorials. *Biometrika* **87**, 193–8.

HEDAYAT, A. S., SLOANE, N. J. A. & STUFKEN, J. (1999). *Orthogonal Arrays: Theory and Applications.* New York: Springer.

Johnson, M., Moore, L. & Ylvisaker, D. (1990). Minimax and maximin distance designs. *J. Statist. Plan. Infer.* **26**, 131–48.

Lawrence, K. M. (1996). A combinatorial characterization of (t, m, s)-nets in base b. *J. Combin. Des.* **4**, 275–93.

Lin, C. D., Mukerjee, R. & Tang, B. (2009). Construction of orthogonal and nearly orthogonal Latin hypercubes. *Biometrika* **96**, 243–7.

McKay, M. D., Beckman, R. J. & 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**, 239–45.

Mullen, G. L. & Schmid, W. Ch. (1996). An equivalence between (T, M, S)-nets and strongly orthogonal hypercubes. *J. Combin. Theory* A **76**, 164–74.

Niederreiter, H. (1987). Point sets and sequences with small discrepancy. *Monatsh. Math.* **104**, 273–337.

Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comp. Math. Math. Phys.* **7**, 784–802.

Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J. Am. Statist. Assoc.* **88**, 1392–7.

Tang, B. & Deng, L. Y. (1999). Minimum G_2-aberration for nonregular fractional factorial designs. *Ann. Statist.* **27**, 1914–26.

Xu, H. & Wu, C. F. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Statist.* **29**, 1066–77.

Ye, K. Q. (1998). Orthogonal column Latin hypercubes and their application in computer experiments. *J. Am. Statist. Assoc.* **93**, 1430–9.