The Annals of Statistics
1995, Vol. 23, No. 6, 2102–2115

ON THE EXISTENCE OF SATURATED AND NEARLY SATURATED ASYMMETRICAL ORTHOGONAL ARRAYS[1]

BY RAHUL MUKERJEE AND C. F. JEFF WU

Indian Institute of Management and University of Michigan

We develop a combinatorial condition necessary for the existence of a saturated asymmetrical orthogonal array of strength 2. This condition limits the choice of integral solutions to the system of equations in the Bose–Bush approach and can thus strengthen considerably the Bose–Bush approach as applied to a symmetrical part of such an array. As a consequence, several nonexistence results follow for saturated and nearly saturated orthogonal arrays of strength 2. One of these leads to a partial settlement of an issue left open in a paper by Wu, Zhang and Wang. Nonexistence of a class of saturated asymmetrical orthogonal arrays of strength 4 is briefly discussed.

1. Introduction. An asymmetrical (or mixed-level) orthogonal array OA$(N, s_1^{m_1} \cdots s_\gamma^{m_\gamma}, \sigma)$ of strength σ is an $N \times m$ matrix, $m = m_1 + \cdots + m_\gamma$, in which m_i columns have $s_i (\geq 2)$ symbols such that for any σ columns all possible combinations of symbols appear equally often [Rao (1973)]. A symmetrical orthogonal array OA(N, s^m, σ) is defined analogously. Because of the wide applicability of orthogonal arrays [for example, as optimal fractional factorial plans; see Cheng (1980)], their existence problem, for given values of the parameters, is of both theoretical and practical interest. While the literature in this direction appears to be reasonably rich in the symmetric case, not many results, apart from the one given by an extension of Rao's (1947) bound, are as yet available in the asymmetric case; see Wu, Zhang and Wang [(1992), hereafter abbreviated as WZW] and Wang and Wu (1992) for more details.

In connection with the existence problem of symmetrical orthogonal arrays, it is useful to find a good upper bound for m, given N, s and σ. Bose and Bush (1952) provided one such bound, which we call the BB bound, and subsequently there have been several bounds in the coding-theoretic literature [MacWilliams and Sloane (1977)]. For studying the existence of an asymmetrical OA$(N, s_1^{m_1} \cdots s_\gamma^{m_\gamma}, \sigma)$, one approach is to apply bounds for symmetrical arrays to OA$(N, s_1^{m_1}, \sigma)$ (by ignoring the $m_2 + \cdots + m_\gamma$ columns with $s_2, \dots, s_\gamma$ symbols) and similarly to OA$(N, s_2^{m_2}, \sigma)$ and so on. As pointed

Received April 1993; revised February 1995.
[1]Research supported by the Natural Sciences and Engineering Research Council of Canada while the first author was visiting the University of Waterloo, and NSF Grant DMS-94-04300. The work of the first author was also supported by a grant from the Centre for Management and Development Studies, IIM, Calcutta.
AMS 1991 *subject classifications*. Primary 62K15, 05B15.
Key words and phrases. Bose–Bush bound, Delsarte theory.

out by WZW, this approach does not seem to produce sharp results. For $\gamma = 2$, $s_1 = s$ and $s_2 = s^r$, s a prime power, WZW constructed a class of $\mathrm{OA}(N, s^{m_1}(s^r)^{m_2}, 2)$ with $N = s^k$, $k = rq + p$, $0 \le p \le r - 1$, $m_1(s-1) + m_2(s^r - 1) = s^k - 1$ and $m_2 = 1, \dots, B_1$, where

$$B_1 = (s^k - s^{r+p})/(s^r - 1) + 1. \tag{1.1}$$

The case of $p = 0$ is trivial because $B_1 = (s^k - 1)/(s^r - 1)$ and therefore cannot be further increased. For the rest of the paper we only consider $p \ge 1$. The question is: Can there be more than B_1 columns with s^r symbols? By applying the BB bound to $\mathrm{OA}(N, (s^r)^{m_2}, 2)$, WZW obtained B_2 as an upper bound for m_2, where

$$B_2 = \frac{s^k - s^p}{s^r - 1} - \theta_0 - 1,$$

where θ_0 is the integer part of θ,

$$\theta = \left[\tfrac{1}{4} + s^r(s^r - s^p)\right]^{1/2} - \left(s^r - s^p + \tfrac{1}{2}\right).$$

They also showed that $B_2 \ge B_1$ and the equality holds iff $p = 1$ and $s = 2$. They then conjectured that the BB bound is not sharp for general values of $p = 1$ and $s \ge 3$, or $p \ge 2$.

The arrays considered in WZW are saturated in the sense that they leave no degree of freedom for error estimation. It is known that the Delsarte (1973) theory provides a powerful tool for studying the existence of saturated symmetrical orthogonal arrays; see Noda (1979), Hong (1986), Mukerjee and Kageyama (1994) and the references therein. This motivates us to consider a similar approach for investigating the existence of saturated orthogonal arrays in the asymmetrical case. For a saturated asymmetrical orthogonal array of strength 2 we obtain in Lemma 1 a necessary condition for its existence. By using the condition and other combinatorial techniques, we prove in Theorem 1 that the B_1 value in (1.1) cannot be further improved in the case of $p = 1$ and general r and s. We employ Lemma 1 in Section 2 also to study the maximum for m_2 in a saturated $\mathrm{OA}(4s^2, 2^{m_1}s^{m_2}, 2)$ for odd s. Some results on the existence of nearly saturated orthogonal arrays of strength 2 are presented in Section 3. Section 4, which deals with arrays of strength 4, shows the nonexistence of a saturated $\mathrm{OA}(N, s_1^{m_1}s_2^{m_2}, 4)$ with $N \le 1000$, $m_1 + m_2 \ge 5$, $2 \le s_1 < s_2 \le 7$.

2. Saturated orthogonal arrays of strength 2.

We begin by considering an $\mathrm{OA}(N, s_1^{m_1}s_2^{m_2}, 2)$ which is saturated in the sense that $N - 1 = m_1(s_1 - 1) + m_2(s_2 - 1)$. We derive a necessary condition for the existence of such an array.

For $i = 1, 2$, let 1_i be the $s_i \times 1$ vector with all elements unity and $P_i = [\mathbf{p}_i(1), \dots, \mathbf{p}_i(s_i)]$ be an $(s_i - 1) \times s_i$ matrix such that the $s_i \times s_i$ matrix $(s_i^{-1/2}1_i, P_i')'$ is orthogonal. Let A be a saturated $\mathrm{OA}(N, s_1^{m_1}s_2^{m_2}, 2)$. Without loss of generality, suppose the first m_1 columns of A have the symbols $1, 2, \ldots, s_1$ and the last m_2 columns have the symbols $1, 2, \ldots, s_2$. Thus

$$(2.1)\quad A = \begin{bmatrix} \alpha_{111} & \alpha_{112} & \cdots & \alpha_{11m_1} & \alpha_{211} & \alpha_{212} & \cdots & \alpha_{21m_2} \\ \alpha_{121} & \alpha_{122} & \cdots & \alpha_{12m_1} & \alpha_{221} & \alpha_{222} & \cdots & \alpha_{22m_2} \\ & & \vdots & & & & \vdots & \\ \alpha_{1N1} & \alpha_{1N2} & \cdots & \alpha_{1Nm_1} & \alpha_{2N1} & \alpha_{2N2} & \cdots & \alpha_{2Nm_2} \end{bmatrix},$$

where $\alpha_{ijk} \in \{1, 2, \ldots, s_i\}$, $i = 1, 2$, $1 \le j \le N$, $1 \le k \le m_i$.

Let ε be an $N \times 1$ vector with each element $N^{-1/2}$. For $i = 1, 2$, let A_i^* be a matrix of order $N \times (m_i(s_i - 1))$ defined as

$$A_i^* = \{s_i/N\}^{1/2}\left[\mathbf{p}_i'(\alpha_{ijk})\right], \qquad 1 \le j \le N, 1 \le k \le m_i.$$

Finally, define

$$(2.2)\qquad A^* = [\,\varepsilon \quad A_1^* \quad A_2^*\,].$$

Since A is an orthogonal array, it is not hard to see that $A^{*\prime}A^* = I_N$, the $N \times N$ identity matrix. From the saturation condition, A^* is an $N \times N$ square matrix. Hence $A^*A^{*\prime} = I_N$, that is, the scalar product of any two distinct rows of A^* must vanish. This leads to the following key condition (2.4).

For $1 \le j$, $u \le N$, $j \ne u$, consider the jth and uth rows of A^*. Since their scalar product is zero, we have

$$(2.3)\qquad \frac{1}{N} + \sum_{i=1}^{2} \frac{s_i}{N} \sum_{k=1}^{m_i} \mathbf{p}_i'(\alpha_{ijk})\mathbf{p}_i(\alpha_{iuk}) = 0.$$

However, by the definition of P_i, for $i = 1, 2$,

$$\sum_{k=1}^{m_i} \mathbf{p}_i'(\alpha_{ijk})\mathbf{p}_i(\alpha_{iuk}) = \sum_{k=1}^{m_i}\left[\delta(\alpha_{ijk}, \alpha_{iuk}) - \frac{1}{s_i}\right] = \Delta_i^{(ju)} - \frac{m_i}{s_i},$$

where $\Delta_i^{(ju)} = \sum_{k=1}^{m_i}\delta(\alpha_{ijk}, \alpha_{iuk})$, and

$$\begin{aligned} \delta(\alpha_{ijk}, \alpha_{iuk}) &= 1, &&\text{if } \alpha_{ijk} = \alpha_{iuk}, \\ &= 0, &&\text{otherwise.} \end{aligned}$$

Note that $\Delta_i^{(ju)}$ can be interpreted as the number of coincidences between the jth and uth rows of the submatrix of A given by its s_i-symbol columns. Such numbers of coincidences play a crucial role in the Delsarte theory for symmetric orthogonal arrays.

The relation (2.3) now simplifies to

$$s_1\Delta_1^{(ju)} + s_2\Delta_2^{(ju)} = m_1 + m_2 - 1$$

and we have the following lemma.

LEMMA 1. *Consider any two distinct rows of a saturated orthogonal array* $\mathrm{OA}(N, s_1^{m_1}s_2^{m_2}, 2)$. *For* $i = 1, 2$, *let* Δ_i *be the number of coincidences between* *these two rows arising from the s_i-symbol columns. Then Δ_1 and Δ_2 are nonnegative integers satisfying $\Delta_1 \le m_1$, $\Delta_2 \le m_2$,*

$$s_1\Delta_1 + s_2\Delta_2 = m_1 + m_2 - 1. \tag{2.4}$$

Lemma 1 can be easily extended to a saturated OA($N, s_1^{m_1}s_2^{m_2} \cdots s_\gamma^{m_\gamma}, 2$) to yield the necessary condition $s_1\Delta_1 + \cdots + s_\gamma\Delta_\gamma = m_1 + \cdots + m_\gamma - 1$, the notational system being obvious. Our approach, based on Lemma 1, for studying the existence of a saturated OA($N, s_1^{m_1}s_2^{m_2}, 2$) is summarized below. First we employ (2.4) to find all possible integral-valued solutions for (Δ_1, Δ_2) in the range $0 \le \Delta_i \le m_i$, $i = 1, 2$, and thus find the set Ω of all possible values of Δ_1. Consider now the first row of the subarray given by the s_1-symbol columns. Among the other rows of this subarray, let there be τ_ω rows having ω coincidence with the first row, where $\omega \in \Omega$. Then, as in the derivation of the BB bound,

$$\sum_{\omega\in\Omega} \binom{\omega}{i}\tau_\omega = \binom{m_1}{i}(Ns_1^{-i} - 1), \qquad i = 0, 1, 2. \tag{2.5}$$

If the system of equations (2.5) fails to admit a nonnegative integral-valued solution for τ_ω, $\omega \in \Omega$, then the nonexistence of the saturated OA($N, s_1^{m_1}s_2^{m_2}, 2$) follows. In fact, often one does not even have to utilize all the equations in (2.5) to prove nonexistence. It may be remarked that if one works with Δ_2, instead of Δ_1, then the resulting system of equations becomes equivalent to (2.5) and hence yields identical results. Extension of this approach to $\gamma \ge 3$ is straightforward: (i) replace (2.4) by $\sum s_j\Delta_j = \sum m_j - 1$ and (ii) replaces (2.5) by $\gamma - 1$ analogous systems of equations representing the first $\gamma - 1$ subarrays given by the s_j-symbol columns, $j = 1, \dots, \gamma - 1$.

The following result, in continuation of Theorem 2 in WZW, can be proved using the approach outlined above. This partially settles a question left open in Section 5 of their paper.

THEOREM 1. *For any prime power s and arbitrary positive integers r and q ($r \ge 2$, $q \ge 1$), a saturated asymmetrical orthogonal array* OA($s^{rq+1}, s^{m_1}(s^r)^{m_2}, 2$) *exists if and only if m_1 and m_2 are nonnegative integers satisfying*

$$m_1(s-1) + m_2(s^r - 1) = s^{rq+1} - 1 \tag{a}$$

and

$$0 \le m_2 \le \frac{s^{rq+1} - s^{r+1}}{s^r - 1} + 1. \tag{b}$$

PROOF. The "if" part is proved in Theorem 2 of WZW. It remains to prove the "only if" part. For $q = 1$, this is an immediate consequence of the fact that s^{r+1} is not an integral multiple of s^{2r}. We therefore prove the "only if" part for $q \geq 2$. The necessity of (a) is obvious. To prove the necessity of (b), assume an OA$(s^{rq+1}, s^{m_1}(s^r)^{m_2}, 2)$ exists for

$$m_2 = \frac{s^{rq+1} - s^{r+1}}{s^r - 1} + 1 + \xi \tag{2.6a}$$

and [cf. (a)]

$$\begin{aligned} m_1 &= \frac{1}{s-1}\left[s^{rq+1} - 1 - (s^r - 1)\left\{\frac{s^{rq+1} - s^{r+1}}{s^r - 1} + 1 + \xi\right\}\right] \\ &= 1 + (s - \xi - 1)\frac{s^r - 1}{s - 1} = s^r - \xi\frac{s^r - 1}{s - 1}, \end{aligned} \tag{2.6b}$$

where ξ is a positive integer satisfying

$$1 \leq \xi \leq s - 1, \tag{2.7}$$

as $m_1 \geq 1$.

Then with $s_1 = s$ and $s_2 = s^r$ in Lemma 1, Δ_1 and Δ_2, as defined there, must satisfy

$$\begin{aligned} s\Delta_1 + s^r\Delta_2 = m_1 + m_2 - 1 &= (s - \xi - 1)\frac{s^r - 1}{s - 1} + \frac{s^{rq+1} - s^{r+1}}{s^r - 1} + 1 + \xi \\ &= (s - \xi - 1)(s + s^2 + \cdots + s^{r-1}) + \frac{s(s^{rq} - 1)}{s^r - 1}, \end{aligned}$$

that is,

$$\begin{aligned} \Delta_1 + s^{r-1}\Delta_2 &= (s - \xi - 1)\frac{s^{r-1} - 1}{s - 1} + \frac{s^{rq} - 1}{s^r - 1} \\ &= \left(s^{r-1} - \xi\frac{s^{r-1} - 1}{s - 1}\right) + \left\{(s^r) + \cdots + (s^r)^{q-1}\right\}. \end{aligned}$$

Hence,

$$\begin{aligned} \Delta_1 &= s^{r-1} - \xi\left(\frac{s^{r-1} - 1}{s - 1}\right) + s^r v - s^{r-1}\Delta_2. \\ &= s^{r-1}(sv + 1 - \Delta_2) - \xi\left(\frac{s^{r-1} - 1}{s - 1}\right), \end{aligned}$$

where $v = 1 + t + t^2 + \cdots + t^{q-2}$, $t = s^r$. From (2.6b) we also have

$$\Delta_1 \leq m_1 = (s - \xi)s^{r-1} - \xi\left(\frac{s^{r-1} - 1}{s - 1}\right),$$

which implies that all possible values of Δ_1 can be expressed as $js^{r-1} - \xi(s^{r-1} - 1)/(s - 1)$ ($= \Delta_{1j}$, say) with $1 \leq j \leq s - \xi$. Note that $j = 0$ would imply $\Delta_1 < 0$, which is impossible.

Consider now the first row of the orthogonal array. Among the other rows, let there be f_j rows having Δ_{1j} coincidences with the first row arising from the s-symbol columns, $1 \le j \le s - \xi$. Then as in (2.5),

$$\sum_{j=1}^{s-\xi} f_j = s^{rq+1} - 1, \tag{2.8a}$$

$$\sum_{j=1}^{s-\xi} \left\{ js^{r-1} - \xi\left(\frac{s^{r-1}-1}{s-1}\right)\right\} f_j = m_1(s^{rq} - 1). \tag{2.8b}$$

By multiplying (2.8a) by $\{s^{r-1} - \xi(s^{r-1} - 1)/(s - 1)\}$, subtracting that from (2.8b) and then simplifying using (2.6b), one obtains

$$\begin{aligned}\sum_{j=1}^{s-\xi} (j-1)s^{r-1}f_j &= \left[s^r - \xi\left(\frac{s^r-1}{s-1}\right)\right](s^{rq}-1)\\ &\quad -(s^{rq+1}-1)\left\{s^{r-1} - \xi\left(\frac{s^{r-1}-1}{s-1}\right)\right\}\\ &= s^{r-1} - s^r - \xi(s^{rq} - s^{r-1})\\ &= -\xi s^{rq} - (s-1-\xi)s^{r-1}.\end{aligned} \tag{2.9}$$

Clearly, the left-hand side of (2.9) is nonnegative, while, by (2.7), the right-hand side of (2.9) is negative. Thus we reach a contradiction and this completes the proof of the theorem. □

It follows from Theorem 1 and the discussion in Section 1 that under the setup of Theorem 1, the construction procedure in WZW cannot be improved upon in the sense that one cannot accommodate more s^r-level columns than they have done. In particular, for $r = 2$ (which implies $p \le 1$) the work of WZW completely settles the problem considered in their paper.

Theorem 1 illustrates a situation where use of (2.5) produces a sharper result than that of the BB bound as applied to a symmetrical part of an asymmetrical orthogonal array. In general, in the present context, application of (2.5) will always be at least as powerful as that of the BB bound. This is because equation (2.4) may provide more specific information about the possible values of Δ_i than the trivial fact $0 \le \Delta_i \le m_i$. Consequently, (2.5) will always be at least as strong as the corresponding system of equations that yields the BB bound.

REMARK 1. Notwithstanding the previous remark, there are situations where use of (2.5) fails to produce a result better than the BB bound. The following example serves as an illustration.

EXAMPLE 1. We consider a saturated OA$(256, 2^{m_1}8^{m_2}, 2)$. By Theorem 2 in WZW, such an array exists for m_1, m_2 satisfying $m_1 \ge 1$, $m_1 + 7m_2 = 255$, $m_2 \le 33$, while, as noted in their Table 3, application of the BB bound to the eight-symbol subarray $OA(256, 8^{m_2}, 2)$ yields $m_2 \le 34$. Thus the question of existence of $OA(256, 2^{17}8^{34})$ remains open. As will be seen below, even the application of (2.5) fails to settle this issue. If an $OA(256, 2^{17}8^{34}, 2)$ exists, then by (2.4) one must have $2\Delta_1 + 8\Delta_2 = 50$, so that the possible values of Δ_2 are 2, 3, 4, 5 and 6, since $\Delta_1 \le 17$, $\Delta_2 \le 34$. Consider now the first row of the array. Among the other rows, let there be f_j rows having j coincidences with the first row arising from the eight-symbol columns, $2 \le j \le 6$. Then following (2.5),

$$\begin{aligned} f_2 + f_3 + f_4 + f_5 + f_6 &= 255, \\ 2f_2 + 3f_3 + 4f_4 + 5f_5 + 6f_6 &= 1054, \\ f_2 + 3f_3 + 6f_4 + 10f_5 + 15f_6 &= 1683. \end{aligned} \tag{2.10}$$

The system of equations (2.10), however, has many nonnegative integral-valued solutions, for example, one particular solution is given by $(f_2, f_3, f_4, f_5, f_6) = (0, 17, 187, 51, 0)$. It can also be seen that a similar approach base on Δ_1 leads to equations equivalent to (2.10). Thus, in this example, use of (2.5) fails to produce a result better than that given by the BB bound.

REMARK 2. In general, in situations studied in Theorem 2 of WZW that are not covered by our Theorem 1 (e.g., the case $p \ge 2$), use of our approach does not seem to be helpful in the sense that it fails to produce inconsistent equations. However, as the following example illustrates, there can be other situations where our approach is useful.

EXAMPLE 2. Consider a saturated $OA(4s^2, 2^{m_1}s^{m_2}, 2)$ where s (≥ 3) is odd and m_1, m_2 are positive integers satisfying

$$m_1 + (s-1)m_2 = 4s^2 - 1. \tag{2.11}$$

Application of the BB bound to its s-symbol subarray yields

$$\begin{aligned} &m_2 \le 16, && \text{if} \quad s = 3, \\ &m_2 \le 23, && \text{if} \quad s = 5, \\ &m_2 \le 4s + 2, && \text{if} \quad s > 7. \end{aligned} \tag{2.12}$$

We first consider $s = 3$ and $s = 5$. By employing (2.4) and (2.5) we will prove the nonexistence of saturated $OA(36, 2^3 3^{16}, 2)$, $OA(36, 2^5 3^{15}, 2)$ and $OA(100, 2^7 5^{23}, 2)$.

If an $OA(36, 2^5 3^{15}, 2)$ exists, then by (2.4), Δ_1 and Δ_2 must satisfy $2\Delta_1 + 3\Delta_2 = 19$ so that the possible value of Δ_1 ($\le m_1$) are 2 and 5. As before, among the rows of the orthogonal array other than the first row, let there be f_j rows having j coincidences with the first row ($j = 2, 5$) arising from the two-symbol columns. Then $f_2 + f_5 = 35$, $2f_2 + 5f_5 = 85$ and $f_2 + 10f_5 = 80$ [cf. (2.10)], with the unique solution $f_2 = 30$, $f_5 = 5$. Without loss of generality, let the first row of the two-symbol subarray be 11111. Since $f_5 = 5$, there are five other rows of this subarray which equal 11111. Since $f_2 = 30$, there are 30 rows of this subarray having exactly two 1's and three 2's. Without loss of generality (by rearranging the columns if necessary), let one of these 30 rows be 11222. Since any two distinct rows of the two-symbol subarray must have exactly two or five coincidences, it is not hard to check that each of these 30 rows must be 11222. However, then the subarray is not a two-symbol orthogonal array at all. This contradiction shows the nonexistence of $\mathrm{OA}(36, 2^5 3^{15}, 2)$. We omit the proofs for the nonexistence of $\mathrm{OA}(36, 2^3 3^{16}, 2)$ and $\mathrm{OA}(100, 2^7 5^{23}, 2)$ because they are similar but simpler. In consideration of (2.12) and the nonexistence of these three arrays, it now follows that $m_2 \le 4s + 2$ for each odd s (≥ 3).

Next we prove the impossibility of $m_2 = 4s + 1$ and $m_2 = 4s + 2$. First we consider the case $m_2 = 4s + 2$. Then by (2.11), $m_1 = 2s + 1$ and by (2.4), $2\Delta_1 + s\Delta_2 = 6s + 2$, with the only possibilities for Δ_1 given by $\Delta_1 = 1, s + 1, 2s + 1$. Using notation as before, analogously to (2.10),

$$\begin{aligned} f_1 + f_{s+1} + f_{2s+1} &= 4s^2 - 1, \\ f_1 + (s+1) f_{s+1} + (2s+1) f_{2s+1} &= (2s+1)(2s^2 - 1), \\ (s+1) s f_{s+1} + (2s+1) 2s f_{2s+1} &= (2s+1) 2s (s^2 - 1), \end{aligned}$$

with the unique solution $f_1 = 2s + 1$, $f_{s+1} = 4s^2 - 2s - 2$ and $f_{2s+1} = 0$. Without loss of generality, suppose the first row of the two-symbol subarray is $11 \cdots 1$. As $f_1 = 2s + 1$, there are $2s + 1$ rows of this subarray having exactly one 1 and $2s$ 2's. By rearranging columns, if necessary, let the first of these $2s + 1$ rows be, say, $e_1 = 12 \cdots 2$. The second, say e_2, of these $2s + 1$ rows must also have one 1 and $2s$ 2's. However, this is impossible since, as noted above, e_1 and e_2 must have either 1 or $s + 1$ coincidences. Thus the impossibility of $m_2 = 4s + 2$ follows.

We now consider $m_2 = 4s + 1$. Then $m_1 = 3s$ and by (2.4), $2\Delta_1 + s\Delta_2 = 7s$. The possible values of Δ_1 are 0, s, $2s$ and $3s$. As before,

$$\begin{aligned} f_0 + f_s + f_{2s} + f_{3s} &= 4s^2 - 1, \\ s f_s + 2s f_{2s} + 3s f_{3s} &= 3s(2s^2 - 1), \\ s(s-1) f_s + 2s(2s-1) f_{2s} + 3s(3s-1) f_{3s} &= 3s(3s-1)(s^2 - 1). \end{aligned}$$

Combining these equations with coefficients $2s^2$, $-(3s - 1)$ and 1, respectively, we get $2s^2(f_0 + f_{3s}) = -s^2(s-1)(s-2) < 0$, which is impossible.

From the results in the preceding paragraphs, we conclude that $m_2 \le 4s$. Regarding the existence of a saturated $\mathrm{OA}(4s^2, 2^{4s-1} s^{4s}, 2)$, we note that such an array necessarily exists if s (≥ 3) is an odd prime or prime power and a Hadamand matrix of order $4s$ is available, for then one can start with a difference matrix $D_{4s, 4s; s}$ [Dawson (1985); de Launey (1986)] and then employ the construction procedure due to Wang and Wu (1991). For $s = 3$, $\mathrm{OA}(36, 2^{11} 3^{12}, 2)$ has been constructed by Taguchi (1987) based on an $\mathrm{OA}(36, 3^{12}, 2)$ constructed by Seiden (1954). Because it can accommodate a large number of factors with two or three levels, its run size economy makes it a popular candidate for experiments in quality improvement. Our result shows that $m_2 = 12$ is maximum among the arrays $\mathrm{OA}(36, 2^{35-2m_2}3^{m_2}, 2)$.

One can proceed as in Example 2 also to prove the nonexistence of (i) $\mathrm{OA}(108, 2^{107-2x}3^x, 2)$, $x = 49, 50, 51, 52$, and (ii) $\mathrm{OA}(36, 6^1 3^{14} 2^2, 2)$, $\mathrm{OA}(36, 4^1 3^{16}, 2)$ and $\mathrm{OA}(36, 6^2 3^{12} 2^1, 2)$. The result under (i) is particularly significant. Wang (1989) constructed an $\mathrm{OA}(108, 2^{11}3^{48}, 2)$. According to (i), the 48 three-level columns cannot be further improved within the class of saturated arrays.

3. Nearly saturated orthogonal arrays of strength 2. In this section, we show that in certain situations the existence of a nearly saturated orthogonal array implies that of a saturated orthogonal array, which implies that the findings of Section 2 can also be used with reference to nearly saturated orthogonal arrays.

LEMMA 2. *The existence of an* $\mathrm{OA}(N, s_1^{m_1} \cdots s_\gamma^{m_\gamma}, 2)$, *which is nearly saturated in the sense that*

$$\sum_{i=1}^{\gamma} m_i(s_i - 1) = N - 2, \tag{3.1}$$

implies the existence of a saturated $\mathrm{OA}(N, s_1^{m_1} \cdots s_\gamma^{m_\gamma} 2^1, 2)$.

PROOF. For the sake of brevity in presentation, we consider the case $\gamma = 2$, although the proof can be easily extended for general γ. Let A be an $\mathrm{OA}(N, s_1^{m_1} s_2^{m_2}, 2)$ which is nearly saturated in the sense of (3.1). We express A as in (2.1) and construct a matrix A^* as in (2.2). By (3.1), A^* is $N \times (N-1)$ and since A is an orthogonal array, $A^{*\prime}A^* = I_{N-1}$.

Hence there exists an $N \times 1$ vector $h = (h_1, \dots, h_N)'$ such that the $N \times N$ matrix $[A^* \;\; h]$ is orthogonal. Then

$$A^*A^{*\prime} + hh' = I_N.$$

Equating the diagonal elements from both sides of the equation above, by (2.2) and (3.1),

$$N^{-1}(N-1) + h_i^2 = 1, \quad \text{that is,} \quad h_i = \pm N^{-1/2}, \qquad 1 \le i \le N. \tag{3.2}$$

In view of (3.2) and the orthogonality of h to the first column of A^*, in h the number of elements which equal $N^{-1/2}$ must be the same as the number of elements which equal $N^{-1/2}$ (i.e., N must be even). We now add a two-symbol column to A such that for $1 \le i \le N$, if $h_i = N^{-1/2}$, then 1 appears in the ith position of this column, while if $h_i = -N^{-1/2}$, then 2 appears in the ith position of this column. It will be seen that the resulting array, say $\bar{A}$, is an $\mathrm{OA}(N, s_1^{m_1} s_2^{m_2} 2^1, 2)$.

To that effect, it is enough to show that the newly added two-symbol column is orthogonal to each column of A. Without loss of generality, consider the first column of A given by $(\alpha_{111}, \alpha_{121}, \dots, \alpha_{1N1})'$, where $\alpha_{1i1} \in \{1, 2, \dots, s_1\}$, $1 \le i \le N$ [see (2.1)]. For $1 \le j \le s_1$, $u = 1, 2$, let ϕ_{ju} be the frequency of occurrence of the pair (j, u) as a row in the $N \times 2$ subarray of $\bar{A}$ given by the first column of A and the newly added two-symbol column. Since h is orthogonal to each column of A^*,

$$\sum_{i=1}^{N} h_i \mathbf{p}_1(\alpha_{1i1}) = \mathbf{0}, \tag{3.3}$$

where $\mathbf{0}$ is the null vector of order $s_1 - 1$. By our construction, for $1 \le j \le s_1$, the pair (h_i, α_{1i1}) equals $(N^{-1/2}, j)$ for ϕ_{j1} choices of i and $(-N^{-1/2}, j)$ for ϕ_{j2} choices of i. Hence by (3.3),

$$\sum_{j=1}^{s_1} (\phi_{j1} - \phi_{j2}) \mathbf{p}_1(j) = \mathbf{0}.$$

Thus by the definition of the matrix P_1, there exists a constant ϕ_0 such that $\phi_{j1} - \phi_{j2} = \phi_0$ for each j. Recalling the structure of h,

$$\sum_{j=1}^{s_1} \phi_{j1} = \sum_{j=1}^{s_1} \phi_{j2} = \tfrac{1}{2} N.$$

Hence $\phi_0 = 0$ and $\phi_{j1} = \phi_{j2}$ for each j. At the same time, since A is an orthogonal array of strength 2, $\phi_{j1} + \phi_{j2} = N s_1^{-1}$ for each j. Therefore, $\phi_{j1} = \phi_{j2} = N/(2 s_1)$ for each j. Thus it follows that $\bar{A}$ is an OA$(N, s_1^{m_1} s_2^{m_2} 2^1, 2)$. That it is saturated is obvious from (3.1). □

REMARK 3. Lemma 2 shows that an orthogonal array of strength 2, which is nearly saturated in the sense of (3.1), is embedded in a saturated orthogonal array of strength 2. One may wonder whether a similar conclusion holds also for orthogonal arrays which are less nearly saturated. In general, the answer to this question is negative. Consider, for example, an OA$(18, 2^1 3^7, 2)$ which is known to exist [Wang and Wu (1991)]. Here $N = 18$, $\gamma = 2$, $s_1 = 2$, $s_2 = 3$, $m_1 = 1$, $m_2 = 7$ and $\sum m_i(s_i - 1) = N - 3$. It is easy to see that if this orthogonal array is embedded in a 18-run saturated orthogonal array of strength 2, then the saturated array must be an OA$(18, 2^1 3^8, 2)$, which is, however, nonexistent, as an application of the BB bound to its three-symbol subarray shows. This shows that, in general, Lemma 2 cannot be strengthened further. However, the following lemma, whose proof is sketched in the Appendix, presents a partial extension.

LEMMA 3. *Suppose N is not an integral multiple of* 3. *Then the existence of an* OA$(N, s_1^{m_1} \cdots s_\gamma^{m_\gamma}, 2)$ *satisfying*

$$\sum_{i=1}^{\gamma} m_i(s_i - 1) = N - 3 \tag{3.4}$$

implies the existence of a saturated OA$(N, s_1^{m_1} \cdots s_\gamma^{m_\gamma} 2^2, 2)$.

Combining the findings in Section 2 with Lemmas 2 and 3, it is possible to prove the nonexistence of certain nearly saturated asymmetrical orthogonal arrays of strength 2. Thus, by Lemma 2, the nonexistence of saturated orthogonal arrays OA$(36, 2^{35-2x}3^x, 2)$ $(x = 13, 14, 15, 16)$, OA$(108, 2^{107-2x}3^x, 2)$ $(x = 49, 50, 51, 52)$, OA$(36, 6^1 3^{14} 2^2, 2)$ and OA$(36, 6^2 3^{12} 2^1, 2)$, as noted in Section 2, implies the nonexistence of the nearly saturated arrays OA$(36, 2^{34-2x}3^x, 2)$ $(x = 13, 14, 15, 16)$, OA$(108, 2^{106-2x}3^x, 2)$ $(x = 49, 50, 51, 52)$, OA$(36, 6^1 3^{14} 2^1, 2)$ and OA$(36, 6^2 3^{12}, 2)$, respectively. Similarly for odd s (≥ 5), if s is not an integral multiple of 3, then by Lemma 3, the nonexistence of saturated OA$(4s^2, 2^{3s}s^{4s+1}, 2)$ and OA$(4s^2, 2^{2s+1}s^{4s+2}, 2)$ (see Example 2) implies that of OA$(4s^2, 2^{3s-2}s^{4s+1}, 2)$ and OA$(4s^2, 2^{2s-1}s^{4s+2}, 2)$.

4. Saturated orthogonal arrays of strength 4. In this section we briefly consider saturated asymmetrical orthogonal arrays OA$(N, s_1^{m_1}s_2^{m_2}, 4)$, that is, those with

$$\begin{aligned} N - 1 = {} & m_1(s_1 - 1) + m_2(s_2 - 1) + \binom{m_1}{2}(s_1 - 1)^2 \\ & + \binom{m_2}{2}(s_2 - 1)^2 + m_1 m_2 (s_1 - 1)(s_2 - 1). \end{aligned} \tag{4.1}$$

We first indicate an analogue of Lemma 1 with reference to such arrays.

Denote a saturated OA$(N, s_1^{m_1}s_2^{m_2}, 4)$ by A as in (2.1). Define ε, A_1^* and A_2^* as before. Also, for $i = 1, 2$, define the $N \times \left(\binom{m_i}{2}(s_i - 1)^2\right)$ matrices

$$A_{ii}^* = \frac{s_i}{\sqrt{N}}\left[\mathbf{p}_i'(\alpha_{ijk}) \otimes \mathbf{p}_i'(\alpha_{ijl})\right], \qquad 1 \leq j \leq N, 1 \leq k < l \leq m_i,$$

where $\otimes$ denotes Kronecker product. Let A_{12}^* denote the $N \times (m_1 m_2 (s_1 - 1)(s_2 - 1))$ matrix

$$A_{12}^* = \frac{\sqrt{s_1 s_2}}{N}\left[\mathbf{p}_1'(\alpha_{1jk}) \otimes \mathbf{p}_2'(\alpha_{2jl})\right], \qquad 1 \leq j \leq N, 1 \leq k \leq m_1, 1 \leq l \leq m_2.$$

Finally, define

$$A^{**} = [\,\varepsilon \quad A_1^* \quad A_2^* \quad A_{11}^* \quad A_{22}^* \quad A_{12}^*\,],$$

which is a square $N \times N$ matrix because of the saturation condition (4.1). By the definition of an orthogonal array of strength 4, $A^{**\prime}A^{**} = I_N$. Since A^{**} is a square matrix, we have $A^{**}A^{**\prime} = I_N$. Hence proceeding as in the derivation of Lemma 1, we get the following result. A detailed proof can be found in Mukerjee and Wu (1993).

LEMMA 4. *Consider any two distinct rows of a saturated orthogonal array* OA$(N, s_1^{m_1}s_2^{m_2}, 4)$. *Let* Δ_1, Δ_2 *be as in the statement of Lemma* 1. *Then* Δ_1 *and* Δ_2 *are nonnegative integers satisfying* $\Delta_1 \leq m_1$, $\Delta_2 \leq m_2$ *and*

$$\begin{aligned} 0 = {} & 1 - \tfrac{3}{2}m + \tfrac{1}{2}m^2 + \tfrac{1}{2}s_1^2\Delta_1(\Delta_1 - 1) + \tfrac{1}{2}s_2^2\Delta_2(\Delta_2 - 1) \\ & + s_1 s_2 \Delta_1 \Delta_2 - (m - 2)(s_1\Delta_1 + s_2\Delta_2), \end{aligned}$$

where $m = m_1 + m_2$.

To avoid trivalities, let $m(= m_1 + m_2) \geq 5$. We shall now apply Lemma 4 to study the existence of saturated OA$(N, s_1^{m_1} s_2^{m_2}, 4)$ over the range

$$N \leq 1000, \qquad m_1 \geq 1, \qquad m_2 \geq 1, \qquad m(= m_1 + m_2) \geq 5, \qquad 2 \leq s_1 < s_2 \leq 7. \tag{4.2}$$

Observe that the range $N \leq 1000$ should be enough for most practical purposes. First suppose $s_1 = 2$ and $s_2 = 3$. Then the simple fact that N must be an integral multiple of each of the numbers $2^{n_1} 3^{n_2}$, where $0 \leq n_1 \leq m_1$, $0 \leq n_2 \leq m_2$ and $n_1 + n_2 = 4$, eliminates all possibilites other than $m_1 = 30$ and $m_2 = 1$ [$N = 528$, by (4.1)]. Under this situation, by Lemma 4, $\Delta_1 = \frac{1}{2}[30 - 3\Delta_2 \pm \sqrt{30 + 3\Delta_2}\,]$ and no nonnegative integral-valued solution for (Δ_1, Δ_2) is available. In a similar manner, one can show the nonexistence of a saturated OA$(N, s_1^{m_1} s_2^{m_2}, 4)$ over the entire range given by (4.2). The details can be found in Mukerjee and Wu (1993).

APPENDIX

PROOF OF LEMMA 3. For ease in presentation, we consider the case $\gamma = 2$ although it is easy to extend the proof for general γ. Let A be an OA$(N, s_1^{m_1} s_2^{m_2}, 2)$ which satisfies (3.4). From A, construct A^* as in (2.2) and by (3.4) note that A^* is $N \times (N-2)$ satisfying $A^{*\prime} A^* = I_{N-2}$. Hence it is not hard to see that there exists an $N \times 2$ matrix Z with rows, say, $z_1', \dots, z_N'$, such that

$$z_1' = (\beta, 0), \tag{A.1}$$

for some $\beta \geq 0$ and the $N \times N$ matrix $[A^* \quad Z]$ is orthogonal. Then

$$A^* A^{*\prime} + ZZ' = I_N.$$

Equating corresponding elements from both sides of the equation above and using (2.2) and (3.4), we have

$$N^{-1}(N-2) + z_j' z_j = 1, \quad \text{that is,} \quad z_j' z_j = 2/N, 1 \leq j \leq N, \tag{A.2}$$

$$N^{-1}\left\{1 + \left(s_1 \Delta_1^{(ju)} - m_1\right) + \left(s_2 \Delta_2^{(ju)} - m_2\right)\right\} + z_j' z_u = 0, \qquad 1 \leq j \neq u \leq N, \tag{A.3}$$

where $\Delta_1^{(ju)}$ and $\Delta_2^{(ju)}$ are defined in Section 2. By (A.2), (A.3) and the Cauchy–Schwarz inequality,

$$z_j' z_u = g_{ju}/N, \qquad 1 \leq j \neq u \leq N, \tag{A.4}$$

where the g_{ju}'s are integers satisfying

$$|g_{ju}| \leq 2, \qquad 1 \leq j \neq u \leq N. \tag{A.5}$$

Since z_1' is as in (A.1) with $\beta \geq 0$, by (A.2), $z_1' = d(1,0)$, where $d = (2/N)^{1/2}$. Hence by (A.4) and (A.5), the only possibilities for z_j', $2 \leq j \leq N$, are

$$\text{(i)} \quad d(0, \pm 1), \qquad \text{(ii)} \quad d\left(\pm \frac{1}{2}, \pm \frac{\sqrt{3}}{2}\right), \qquad \text{(iii)} \quad d(\pm 1, 0).$$

However, by (A.4) and (A.5), Z cannot simultaneously have two rows, one of which is of the form (i) and the other of the form (ii). Hence two cases arise:

Case 1. Each row of Z is of the form

$$d(0,1) \quad \text{or} \quad d(0,-1) \quad \text{or} \quad d(1,0) \quad \text{or} \quad d(-1,0).$$

Case 2. Each row of Z is of the form

$$d\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) \quad \text{or} \quad d\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) \quad \text{or} \quad d\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$$
$$\text{or} \quad d\left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) \quad \text{or} \quad d(1,0) \quad \text{or} \quad d(-1,0).$$

Considering Case 2 first, suppose the vectors listed under this case appear as rows of Z with respective frequencies $b_1, \dots, b_6$. By the definition of Z, its second column has length unity. Therefore, $b_1 + b_2 + b_3 + b_4 = \frac{2}{3}N$, which is impossible because N is not an integral multiple of 3. Thus Case 2 cannot arise and the rows of Z must be as in Case 1. Let then the vectors listed under Case 1 appear as rows of Z with respective frequencies $\lambda_1, \dots, \lambda_4$. Since each column of Z has length unity and is orthogonal to ε, the first column of A^*,

$$\lambda_3 + \lambda_4 = \lambda_1 + \lambda_2 = \tfrac{1}{2}N, \qquad \lambda_3 - \lambda_4 = \lambda_1 - \lambda_2 = 0,$$

so that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \frac{1}{4}N$, that is, N is an integral multiple of 4.

We now add a two-symbol column to A such that for $1 \leq j \leq N$, if $z_j' = d(0,1)$ or $d(1,0)$, then 1 appears in the jth position of this column while if $z_j' = d(0,-1)$ or $d(-1,0)$, then 2 appears in the jth position of this column. As in the proof of Lemma 2, this gives an OA(N, $s_1^{m_1} s_2^{m_2} 2, 2$), which, by (3.4), is nearly saturated in the sense of (3.1). Applying Lemma 2 guarantees the existence of a saturated OA(N, $s_1^{m_1} s_2^{m_2} 2^2, 2$). □

Acknowledgment. We thank the referees for their very constructive suggestions.

REFERENCES

BOSE, R. C. and BUSH, K. A. (1952). Orthogonal arrays of strength two and three. *Ann. Math. Statist.* **23** 508–524.

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

DAWSON, J. E. (1985). A construction for generalized Hadamard matrices $GH(4q, EA(q))$. *J. Statist. Plann. Inference* **11** 103–110.

DE LAUNEY, W. (1986). A survey of generalized Hadamard matrices and difference matrices $D(k, \lambda; G)$ with large k. *Utilitas Math.* **30** 5–29.

DELSARTE, P. (1973). An algebraic approach to the association schemes of coding theory. *Phillips Research Reports Supplement* **10**.

HONG, Y. (1986). On the nonexistence of nontrivial perfect e-codes and tight 2e-designs in Hamming schemes $H(n, q)$ with $e \geq 3$ and $q \geq 3$. *Graphs Combin.* **2** 145–164.

MACWILLIAMS, F. J. and SLOANE, N. J. A. (1977). *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam.

MUKERJEE, R. and KAGEYAMA, S. (1994). On existence of two symbol complete orthogonal arrays. *J. Combin. Theory Ser. A* **66** 176–181.

MUKERJEE, R. and WU, C. F. J. (1993). On the existence of saturated asymmetrical orthogonal arrays. Statistics Technical Report 93-08, Univ. Waterloo.

NODA, R. (1979). On orthogonal arrays of strength 4 achieving Rao's bound. *J. London Math. Soc.* **19** 385–390.

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

RAO, C. R. (1973). Some combinatorial problems of arrays and applications to design of experiments. In *A Survey of Combinatorial Theory* (J. N. Srivastava, ed.) 349–359. North-Holland, Amsterdam.

SEIDEN, E. (1954). On the problem of construction of orthogonal arrays. *Ann. Math. Statist.* **25** 151–156.

TAGUCHI, G. (1987). *System of Experimental Design*. UNIPUB, White Plains, NY.

WANG, J. C. (1989). Orthogonal arrays and nearly orthogonal arrays with mixed levels: construction and applications. Ph.D. Thesis, Univ. Wisconsin, Madison.

WANG, J. C. and WU, C. F. J. (1991). An approach to the construction of asymmetrical orthogonal arrays. *J. Amer. Statist. Assoc.* **86** 450–456.

WANG, J. C. and WU, C. F. J. (1992). Nearly orthogonal arrays with mixed levels and small runs. *Technometrics* **34** 409–422.

WU, C. F. J., ZHANG, R. and WANG, R. (1992). Construction of asymmetrical orthogonal arrays of the type $OA(s^k, s^m(s^{r_1})^{n_1} \cdots (s^{r_t})^{n_t})$. *Statist. Sinica* **2** 203–219.

INDIAN INSTITUTE OF MANAGEMENT
TOKA DIAMOND HARBOR RD.
P.O. BOX 16757
CALCUTTA 70027
INDIA

DEPARTMENT OF STATISTICS
UNIVERSITY OF MICHIGAN
1427 MASON HALL
ANN ARBOR, MICHIGAN 48109-1027