NA SATURATED FRACTIONAL REPLICATE PLANS

by

W. T. Federer and B. K. Pal[1]

BU-709-M[2] June 1980

ABSTRACT

A class of saturated fractional replicate designs are presented. They are denoted as nonorthogonal array, NA, designs. The method of construction is described and plans are given for various numbers of runs, factors, and levels of factors. A comparison of the designs is made with the lowest possible determinant of the design matrix and with an upper bound. The upper bound is not usually achievable.

---

[1] Statistical Quality Control Unit, Indian Statistical Institute, Bangalore, India.

[2] In the Mimeo Series of the Biometrics Unit, Cornell University, Ithaca, New York.

NA SATURATED FRACTIONAL REPLICATE PLANS

by

W. T. Federer and B. K. Pal[1]

BU-709-M[2] June 1980

1. Introduction

In many industrial experiments it is necessary to utilize as few runs as possible because of cost, material, or time constraints. This means that saturated or nearly saturated fractional replicates will be required for these situations. Orthogonal array designs for n^2 runs, n+1 factors, each as n levels and denoted as $OA[n^2,n+1,(n)]$, are variance optimal designs. However, OA's are not available for all situations. Different levels of factors and sometimes interactions between pairs of factors may be of interest to the experimenter. Hence, OA designs may not be suitable or may require too many runs.

To fill this need, a class of saturated fractional replicates, which are nonorthogonal array designs, NA designs, has been constructed. The construction procedures make use of latin square design, Youden design, orthogonal latin square designs, and balanced Youden designs theory. Also, use is made of results developed by Anderson and Federer (1973) to establish upper and lower bounds on the value of the determinant of the design matrix. The design plans given are compared with these bounds.

2. Nonorthogonal Arrays for Industrial Experimentation

To illustrate the class of nonorthogonal array treatment designs presented here, we first consider an example. Suppose that it is desired to conduct an

---

[1] Statistical Quality Control Unit, Indian Statistical Institute, Bangalore, India.

[2] In the Mimeo Series of the Biometrics Unit, Cornell University, Ithaca, New York.

experiment using two factors at four levels and one factor at two levels in eight runs. Note that an orthogonal array would require 16 runs. Two such designs which allow solutions for the eight parameters of this design are:

One-at-a-time plan

run	Treatment level for a	b	c
1	0	0	0
2	1	0	0
3	0	1	0
4	0	2	0
5	0	3	0
6	0	0	1
7	0	0	2
8	0	0	3

NA[8,3,(2,4,4)]

run	Treatment level for a	b	c
1	0	0	0
2	0	1	1
3	0	2	2
4	0	3	3
5	1	0	1
6	1	1	2
7	1	2	3
8	1	3	0

where NA[r = number of runs, f = number of factors, (n_i = number of levels of each factor)] is used to denote the nonorthogonal array used with r runs, f factors, and n_i levels of a factor. A third plan would be to replace the levels of factor c in runs 5 to 8 with levels 1, 0, 3, 2. This last design is not a connected design in that solutions for the parameters are not possible, and hence will not be considered further.

Two possible linear model response equations for this example are a cell-means model and a factorial main effect model as follows

$$Y_{hij} = \mu_{hij} + \epsilon_{hij} \tag{2.1}$$

and

$$Y_{hij} = \mu + \alpha_h + \beta_i + \gamma_j + \epsilon_{hij} \quad , \tag{2.2}$$

where μ_{hij} is the mean of combination hij and is estimated by Y_{hij} for these saturated main effect plans, μ is an effect common to all observations, α_h is the effect of the h^{th} level of factor a, β_i is the effect of the i^{th} level of factor b, γ_j is the effect of the j^{th} level of factor c, and the ϵ_{hij} are independently and identically distributed with zero mean and common variance σ_ϵ^2. Equation (2.2) is over-parameterized. Therefore, we reparameterize it as follows for factors a, b, and c:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} = \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}.$$

Then we estimate $\mu = \mu' + A_0 + B_0 + C_0$, A_1, B_1, B_2, B_3, C_1, C_2, and C_3. The particular contrasts selected were the Helmert polynomial contrasts. These were selected here because the theory developed by Anderson and Federer (1973) made use of this set of contrasts in obtaining upper and lower bounds on the determinants of the incidence matrix. As these authors state, the theory could have been developed in an analogous manner for any set of orthogonal contrasts. Thus, our response equation (2.2) now becomes:

$$Y_{hij} = \mu + A_h + B_i + C_j + \epsilon_{hij} \tag{2.3}$$

$h = 1,2$, $i = 1,2,3$, and $j = 1,2,3$. Solving the following set of equations results in solutions for the parameters as follows for the second design:

$$\begin{bmatrix} Y_{000} \\ Y_{011} \\ Y_{022} \\ Y_{033} \\ Y_{101} \\ Y_{112} \\ Y_{123} \\ Y_{130} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 0 & -2 & 1 & 0 & -2 & 1 \\ 1 & 1 & 0 & 0 & -3 & 0 & 0 & -3 \\ 1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 0 & -2 & 1 \\ 1 & -1 & 0 & -2 & 1 & 0 & 0 & -3 \\ 1 & -1 & 0 & 0 & -3 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ A_1 \\ B_1 \\ B_2 \\ B_3 \\ C_1 \\ C_2 \\ C_3 \end{bmatrix} \tag{2.4}$$

and the solutions are:

$$\begin{bmatrix} \hat{\mu} \\ \hat{A}_1 \\ \hat{B}_1 \\ \hat{B}_2 \\ \hat{B}_3 \\ \hat{C}_1 \\ \hat{C}_2 \\ \hat{C}_3 \end{bmatrix} = \frac{1}{24} \begin{bmatrix} 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 & -3 & -3 & -3 & -3 \\ 3 & -9 & 3 & 3 & 9 & -3 & -3 & -3 \\ 3 & -1 & -5 & 3 & 1 & 5 & -3 & -3 \\ 3 & 1 & -1 & -3 & -1 & 1 & 3 & -3 \\ 9 & -3 & -3 & -3 & -9 & 3 & 3 & 3 \\ 1 & 5 & -3 & -3 & -1 & -5 & 3 & 3 \\ -1 & 1 & 3 & -3 & 1 & -1 & -3 & 3 \end{bmatrix} \begin{bmatrix} Y_{000} \\ Y_{011} \\ Y_{022} \\ Y_{033} \\ Y_{101} \\ Y_{112} \\ Y_{123} \\ Y_{130} \end{bmatrix} . \tag{2.5}$$

Equation (2.5) in matrix form is $\underline{\hat{\beta}} = X^{-1}\underline{Y}$ and (2.4) is $X\underline{\hat{\beta}} = \underline{Y}$.

In a similar fashion, matrix equations like (2.4) and (2.5) could be constructed for design 1, the one-at-a-time plan. To compare the designs for variance efficiency, we note that the $X\underline{\hat{\beta}} = \underline{Y}$ in regression, that $(X'X)^{-1}\sigma_\epsilon^2$ is the variance covariance matrix of the vector $\underline{\hat{\beta}}$. Hence, to minimize $(X'X)^{-1}$

we maximize X'X, or for saturated fractions, if we maximize the absolute value of the determinant $|X|$, we minimize its inverse. Now, the value of the determinant of the 8×8 matrix X in (2.4) is 4608 and of the analogous X for design 1 is 1152. Thus, the second design has $\frac{1}{4}$ the variance of the first design.

From Anderson and Federer (1973), we note that design one has the minimum value possible among all connected saturated fractional replicates. That is, it is least-optimal with respect to variance optimality. Also, from their paper we note that the upper bound on the value of this determinant is 9216. We should note that no plan may exist which achieves the upper bound. In our case, the upper bound appears unachievable and design two appears to be the best that can be constructed.

The method of constructing $NA[r,f,(n_1,n_2,\cdots)]$ designs involves use of latin square design, Youden design, orthogonal latin square designs, and balanced Youden designs theory (see, e.g., Hedayat, Seiden, and Federer (1973)). For one factor at two levels and two other factors at n levels one may use any two rows of an $n \times n$ latin square design. For one factor at three levels, and the other three factors at n levels, one may use the same three rows of two orthogonal latin squares, or even two nonorthogonal squares where the rows are selected in such a manner as to approach or achieve variance balance between the columns and the symbols in each square and between the symbols of the two squares. To illustrate, consider $NA[21,4,(3,7,7,7)]$ as follows:

row	columns of square 1							columns of square 2						
	0	1	2	3	4	5	6	0	1	2	3	4	5	6
0	1	2	3	4	5	6	0	1	2	3	4	5	6	0
1	2	3	4	5	6	0	1	3	4	5	6	0	1	2
2	4	5	6	0	1	2	3	0	1	2	3	4	5	6
3	0	1	2	3	4	5	6	6	0	1	2	3	4	5
4	3	4	5	6	0	1	2	5	6	0	1	2	3	4
5	5	6	0	1	2	3	4	2	3	4	5	6	0	1
6	6	0	1	2	3	4	5	4	5	6	0	1	2	3

The rows of a pair of orthogonal latin squares of order n=7 have been arranged in such a fashion that the first three rows of both squares form Youden designs; likewise, the first four rows form Youden designs. It is not possible to form Youden designs for five rows, but it is possible for six (see, e.g., Federer (1970)). Construction of orthogonal arrays arranged such that as much variance balance as possible is achieved between the columns and symbols of any square and between symbols of any two squares, results in efficient, if not the most efficient, fractional replicate plans for saturated designs.

Plans for various NA[$r,f,(n_1,n_2,\cdots)$] designs are presented in Table 1. The NA[21,4,(3,7,7,7)] plan given above for the first three rows is from Table 1 where the first factor is rows, the second factor is columns, the third factor is symbols of square one, and the fourth factor is symbols of square two above. The third plan presented for NA[8,3,(2,4,4)] is also obtained directly from Table 1. No special ordering of rows of latin squares of orders 2, 3, and 5 is required. A plan is given for NA[12,3,(2,6,6)]. For a plan of NA[18,4,(3,6,6,6)], one would proceed as follows:

row	column of square 1						column of square 2					
	0	1	2	3	4	5	0	1	2	3	4	5
0	1	2	3	4	5	0	4	5	0	1	2	3
1	2	3	4	5	0	1	0	1	2	3	4	5
2	5	0	1	2	3	4	3	4	5	0	1	2

to obtain the 18 combinations: 0014, 0125, 0230, 0341, 0452, 0503, 1020, 1131, 1242, 1353, 1404, 1515, 2053, 2104, 2215, 2320, 2431, 2542. An almost balanced incomplete block design arrangement in columns is achieved with each square, and there is a partially balanced arrangement of symbols between unordered pairs of the two squares (Hedayat, Seiden, and Federer (1973)). It may be possible to achieve a more nearly balanced arrangement by selecting a different latin square than the above cyclic one. However, the above appears to be best for the square selected. A similar procedure to that outlined above was utilized to obtain plans for factors with eight and nine levels given in Table 1.

It should be noted that the above NA[18,4,(3,6,6,6)] could also be a plan for four factors at three levels, and for three factors at two levels with three of the interaction effects being obtainable. Thus, we could write the above as NA[18,7,(3,2×3,2×3,2×3)] to indicate that factor a has three levels, that factors b and c are in all possible combinations, factors d and e are in all possible combinations, and that factors f and g are in all possible combinations allowing estimation of the A effect, the B, C, and B × C effects, the E, F and E × F effects, and the G, H, and G × H effects. Also, the NA[18,3,(2,9,9)] plan could be used for one factor at two levels, two factors at three levels each and their interaction, and one factor at nine levels.

	0	0	0	0
	0	1	1	1
	0	2	2	2
	1	0	1	2
	1	1	2	0
NA[6,3,(2,3,3)]	1	2	0	1
	2	0	2	1
	2	1	0	2
OA[9,4,(3)]	2	2	1	0

	0	0	0	0	0
	0	1	1	1	1
	0	2	2	2	2
	0	3	3	3	3
	1	0	1	3	2
	1	1	0	2	3
	1	2	3	1	0
NA[8,3,(2,4,4)]†	1	3	2	0	1
	2	0	2	1	3
	2	1	3	0	2
	2	2	0	3	1
NA[12,4,(3,4,4,4)]	2	3	1	2	0
	3	0	3	2	1
	3	1	2	3	0
	3	2	1	0	3
OA[16,5,(4)]	3	3	0	1	2

	0	0	1	1	1	1	1	1
	0	1	2	2	2	2	2	2
	0	2	3	3	3	3	3	3
	0	3	4	4	4	4	4	4
	0	4	5	5	5	5	5	5
	0	5	6	6	6	6	6	6
	0	6	0	0	0	0	0	0
	1	0	2	3	4	5	6	0
	1	1	3	4	5	6	0	1
	1	2	4	5	6	0	1	2
	1	3	5	6	0	1	2	3
	1	4	6	0	1	2	3	4
	1	5	0	1	2	3	4	5
NA[14,3,(2,7,7)]	1	6	1	2	3	4	5	6
	2	0	4	0	3	6	2	5
	2	1	5	1	4	0	3	6
	2	2	6	2	5	1	4	0
	2	3	0	3	6	2	5	1
	2	4	1	4	0	3	6	2
	2	5	2	5	1	4	0	3
NA[21,4,(3,7,7,7)]	2	6	3	6	2	5	1	4
	3	0	0	6	5	4	3	2
	3	1	1	0	6	5	4	3
	3	2	2	1	0	6	5	4
	3	3	3	2	1	0	6	5
	3	4	4	3	2	1	0	6
	3	5	5	4	3	2	1	0
NA[28,5,(4,7,7,7,7)]	3	6	6	5	4	3	2	1

	0	0	0	0	0	0
	0	1	1	1	1	1
	0	2	2	2	2	2
	0	3	3	3	3	3
	0	4	4	4	4	4
	1	0	1	2	3	4
	1	1	2	3	4	0
	1	2	3	4	0	1
	1	3	4	0	1	2
NA[10,3,(2,5,5)]	1	4	0	1	2	3
	2	0	2	4	1	3
	2	1	3	0	2	4
	2	2	4	1	3	0
	2	3	0	2	4	1
NA[15,4(3,5,5,5)]	2	4	1	3	0	2
	3	0	3	1	4	2
	3	1	4	2	0	3
	3	2	0	3	1	4
	3	3	1	4	2	0
NA[20,5(4,5,5,5,5)]	3	4	2	0	3	1
	4	0	4	3	2	1
	4	1	0	4	3	2
	4	2	1	0	4	3
	4	3	2	1	0	4
OA[25,6,(5)]	4	4	3	2	1	0

	0	0	0
	0	1	1
	0	2	2
	0	3	3
	0	4	4
	1	0	1
	1	1	2
	1	2	3
	1	3	4
	1	4	5
NA[12,3,(2,6,6)]	1	5	0

	0	0	0	5
	0	1	1	6
	0	2	2	7
	0	3	3	0
	0	4	4	1
	0	5	5	2
	0	6	6	3
	0	7	7	4
	1	0	1	7
	1	1	2	0
	1	2	3	1
	1	3	4	2
	1	4	5	3
	1	5	6	4
	1	6	7	5
NA[16,3,(2,8,8)]	1	7	0	6

	4	0	3	5	0	2	4	6
	4	1	4	6	1	3	5	0
	4	2	5	0	2	4	6	1
	4	3	6	1	3	5	0	2
	4	4	0	2	4	6	1	3
	4	5	1	3	5	0	2	4
NA[35,6,(5,7,7,7,7,7)]	4	6	2	4	6	1	3	5
	5	0	5	2	6	3	0	4
	5	1	6	3	0	4	1	5
	5	2	0	4	1	5	2	6
	5	3	1	5	2	6	3	0
	5	4	2	6	3	0	4	1
	5	5	3	0	4	1	5	2
NA[42,7,(6,7,7,7,7,7,7)]	5	6	4	1	5	2	6	3
	6	0	6	4	2	0	5	3
	6	1	0	5	3	1	6	4
	6	2	1	6	4	2	0	5
	6	3	2	0	5	3	1	6
	6	4	3	1	6	4	2	0
	6	5	4	2	0	5	3	1
OA[49,8,(7)]	6	6	5	3	1	6	4	2

	2	0	3	4
	2	1	4	5
	2	2	5	6
	2	3	6	7
	2	4	7	0
	2	5	0	1
	2	6	1	2
NA[24,4,(3,8,8,8)]	2	7	2	3

	0	0	0	5
	0	1	1	6
	0	2	2	7
	0	3	3	8
	0	4	4	0
	0	5	5	1
	0	6	6	2
	0	7	7	3
	0	8	8	4
	1	0	1	7
	1	1	2	8
	1	2	3	0
	1	3	4	1
	1	4	5	2
	1	5	6	3
	1	6	7	4
	1	7	8	5
NA[18,3,(2,9,9)]	1	8	0	6
	2	0	3	4
	2	1	4	5
	2	2	5	6
	2	3	6	7
	2	4	7	8
	2	5	8	0
	2	6	0	1
	2	7	1	2
NA[27,4,(3,9,9,9)]	2	8	2	3

† This is not a connected design.

Table 1. Nonorthogonal arrays, NA, and orthogonal arrays, OA, saturated fractional replicates for various numbers of factors and levels.

3. Efficiency of the Class of Nonorthogonal Arrays

For the class of saturated fractional replicates given in Table 1, it is possible to obtain an idea of how good these nonorthogonal arrays are. Of course, the orthogonal arrays, OA[r,f,(n)], where each of the factors is at n levels, are fully efficient from a variance viewpoint. One cannot obtain a better fraction. Anderson and Federer (1973) have provided guidelines which allow a saturated fractional replicate to be compared with the worst possible design, i.e., a lower bound for all connected designs, and an upper bound for all fractional replicates. The upper bound may only be achievable for OA[r,f,(n)] designs. They give the lower bound, ℓ, as $\Pi_{i=1}^{f}(s_i!)$. Thus, for the first example in Table 1, the lower bound on the determinant of the matrix X using Helmert polynomials for the parameters, is computed as $2!(3!)^2 = 72$, and for NA[12,4,(3,4,4,4)] as $3!(4!)^3 = 82,944$. The lower bounds are computed for all the NA designs given.

The upper bound for a design may be obtained as the following multiple of the lower bound ℓ:

$$r^{r/2} \Big/ \prod_{i=1}^{f} s_i^{s_i/2} \quad .$$

Thus for NA[6,3,(2,3,3)] and NA[8,3,(2,4,4)], the upper bounds are computed respectively as:

$$\ell 6^{6/2}/2^{2/2}3^{3/2}3^{3/2} = 2^2\ell = 4(72) \quad = 288$$

and

$$\ell 8^{8/2}/2^{2/4}4^{4/2}4^{4/2} = 2^3\ell = 8(1152) = 9216 \ .$$

For the NA[6,3,(2,3,3)] plan given, the determinant of X was 3ℓ with the upper bound being 8ℓ. There is no assurance that the upper bound can be reached except for OA designs, which are impossible for 6 and 8 runs, respectively.

The Anderson-Federer bounds then serve as a criterion for evaluating fractional replicates for all factorials, but particularly for the general asymmetrical factorials where relatively little work has been done.

Plan	Determinant value of bound lower (ℓ)	upper	plan
NA[6,3,(2,3,3)]	$2!3!3! = 72$	4ℓ	$3\ell = 216$
NA[8,3,(2,4,4)]†	$2!4!4! = 1152$	$2^3\ell$	4ℓ
NA[12,4,(3,4,4,4)]	$3!(4!)^3$	$3^{9/2}\ell$	$4^3\ell$
NA[10,3,(2,5,5)]	$2!5!5!$	$2^4\ell$	5ℓ
NA[15,4,(3,5,5,5)]	$3!(5!)^3$	$3^6\ell$	$5^3\ell$
NA[20,5,(4,5,5,5,5)]	$4!(5!)^4$	$4^8\ell$	$5^6\ell$
NA[12,3,(2,6,6)]	$2!6!6!$	$2^5\ell$	6ℓ
NA[18,4,(3,6,6,6)]†	$3!(6!)^3$	$3^{15/2}\ell$	--‡
NA[14,3,(2,7,7)]	$2!7!7!$	$2^6\ell$	7ℓ
NA[21,4,(3,7,7,7)]	$3!(7!)^3$	$3^9\ell$	$7^3\ell$
NA[28,5,(4,7,7,7,7)]	$4!(7!)^4$	$4^{12}\ell$	$7^6\ell$
NA[35,6,(5,7,7,7,7,7)]	$5!(7!)^5$	$5^{15}\ell$	$282,475,249\ell$
NA[42,7,(6,7,7,7,7,7,7)]	$6!(7!)^6$	$6^{18}\ell$	$4.7476(10^{12})\ell$
NA[16,3,(2,8,8)]	$2!8!8!$	$2^7\ell$	8ℓ
NA[24,4,(3,8,8,8)]	$3!(8!)^3$	$3^{21/2}\ell$	$3^2 8^2\ell$
NA[18,3,(2 9 9)]	$2!9!9!$	$2^8\ell$	9ℓ
NA[27,4,(3,9,9,9)]	$3!(9!)^3$	$3^{12}\ell$	1539ℓ

† Plans given in text

‡ Not computed

Table 2. Upper and lower Anderson-Federer bounds for determinant $|X|$ for the plans of Table 1 and the actual value of the determinant $|X|$.

References

Anderson, D. A. and W. T. Federer. [1973]. Representation and construction of main effect plans in terms of (0,1) matrices. BU-499-M in the Mimeo Series of the Biometrics Unit, Cornell University.

Federer, W. T. [1972]. Construction of classes of experimental designs using transversals in Latin squares and Hedayat's sum composition method. Ch. 7 in Statistical Papers in Honor of George W. Snecedor (ed. T. A. Bancroft), Iowa State University Press, Ames, pp. 91-114.

Hedayat, A., E. Seiden and W. T. Federer. [1972]. Some families of designs for multistage experiments: Mutually balanced Youden designs when the number of treatments is prime power or twin primes. I. Annals Math. Stat. 43, 1517-1527.