On the aberrations of two-level Orthogonal Arrays with removed runs

Sulle aberrazioni di Orthogonal Arrays binari con punti rimossi

Roberto Fontana and Fabio Rapallo

Abstract We consider binary Orthogonal Arrays and we analyze the aberrations of the fractions obtained by the deletion of $p = 1,2$ or 3 design points. Some explicit formulae are given for $p = 1$ and some examples are presented in the other cases.
Abstract *A partire da Orthogonal Arrays a due livelli, studiamo le aberrazioni delle frazioni ottenute tramite rimozione di* $p = 1,2$ *o* 3 *punti sperimentali. Deriviamo alcune formule esplicite nel caso* $p = 1$*, e per gli altri casi presentiamo alcuni esempi.*

Key words: Fractional factorial designs, Word-Length Pattern

1 Introduction

The theory of Orthogonal Arrays (OAs) has a long history and represents a major research topic for both methodology and applied Statistics. The need for efficient experimental designs has led to the definition of several criteria for the choice of the design points. All such criteria aim to produce the best estimates of the relevant parameters for a given sample size. As general references for OAs, the reader can refer to [5].

In particular, we consider only binary designs under the Minimum Aberration (MA) criterion, but we focus on the following problem. In several situations, it is hard to define *a priori* a fixed sample size. For example, budget constraints or time limitations may occur after the definition of the design, or even when the experiments are running, thus leading to an incomplete design. In such a situation, it is relevant not only to choose an OA with good properties, but also to define an order

---

Roberto Fontana
DISMA, Politecnico di Torino, e-mail: roberto.fontana@polito.it

Rapallo Fabio
DISIT, Università del Piemonte Orientale e-mail: fabio.rapallo@uniupo.it

of the design points, so that the experimenter can stop the sequence of runs and loose as little information as possible. While OAs with added runs are well studied, see for instance [2], less has been done in the case of OAs with removed runs. The order of the runs of an OA and the OAs with removed runs are both interesting topics, but the first one is mainly studied with the aim of minimizing the changes of the factor levels, see e.g. [8], while the second one is only considered in the framework of D-optimality, see [1]. In [7] and [8] several papers, describing practical problems where OAs with missing runs play a major role, are listed.

In this work, we consider the MA criterion for binary OAs and we study the behavior of the aberrations when p points are removed from an OA, for small values of p. The MA criterion is based on the Word-Length Pattern (WLP) introduced in [4]. In practice, the WLP is used to discriminate among different designs $\mathscr{F}_1, \ldots, \mathscr{F}_d$ by looking at the lexicographic minimum of the vector

$$A_{\mathscr{F}_i} = (A_0(\mathscr{F}_i) = 1, A_1(\mathscr{F}_i), \ldots, A_m(\mathscr{F}_i)) \ , \quad i = 1, \ldots, d.$$

2 Orthogonal Arrays and aberrations

Let us consider an experiment with m 2-level factors. The full factorial design is $\mathscr{D} = \{-1, 1\}^m$. We briefly recall here the basic definitions concerning Orthogonal Arrays and aberrations. For details refer to [3].

Definition 1. A fraction $\mathscr{F}$ is a multiset $(\mathscr{F}_*, f_*)$ whose underlying set of elements $\mathscr{F}_*$ is contained in $\mathscr{D}$ and f_* is the multiplicity function $f_* : \mathscr{F}_* \to \mathbb{N}$ that for each element in $\mathscr{F}_*$ gives the number of times it belongs to the multiset $\mathscr{F}$.

We recall that the underlying set of elements $\mathscr{F}_*$ is the subset of $\mathscr{D}$ that contains all the elements of $\mathscr{D}$ that appear in $\mathscr{F}$ at least once. We denote the number of elements of the fraction $\mathscr{F}$ by $\#\mathscr{F}$, with $\#\mathscr{F} = \sum_{x \in \mathscr{F}_*} f_*(x)$.

To describe the counting function of a fraction, we follow the theory in [3]. The simple terms of the form X_j, i.e., the j-th component function which maps a point $x = (x_1, \ldots, x_m)$ of $\mathscr{D}$ to its j-th component,

$$X_j \colon \mathscr{D} \ni (x_1, \ldots, x_m) \longmapsto x_j \in \{-1, 1\}$$

and the interactions $X^\alpha = X_1^{\alpha_1} \cdot \ldots \cdot X_m^{\alpha_m}$, $\alpha \in L = \{0, 1\}$ i.e., the monomial functions of the form

$$X^\alpha : \mathscr{D} \ni (x_1, \ldots, x_m) \mapsto x_1^{\alpha_1} \cdot \ldots \cdot x_m^{\alpha_m}$$

are a basis of all the real functions defined over $\mathscr{D}$. We use this basis to represent the counting function of a fraction according to the following definition.

Definition 2. The counting function R of a fraction $\mathscr{F}$ is a polynomial defined over $\mathscr{D}$ so that for each $x \in \mathscr{D}$, $R(x)$ equals the number of appearances of x in the fraction. A $0-1$ valued counting function is called an indicator function of a single-replicate fraction $\mathscr{F}$. We denote by c_α the coefficients of the representation of R on $\mathscr{D}$ using the monomial basis $\{X^\alpha,\ \alpha \in L\}$:

$$R(x) = \sum_{\alpha \in L} c_\alpha X^\alpha(x),\ x \in \{-1,1\}^m,\ c_\alpha \in \mathbb{R}.$$

Among the fractions of a full factorial design 2^m, we consider Orthogonal Arrays.

Definition 3. A fraction $\mathscr{F}$ factorially projects onto the I-factors, $I = \{i_1,\ldots,i_k\} \subset \{1,\ldots,m\}$, $i_1 < \ldots < i_k$, if the projection $\pi_I(\mathscr{F})$ is a full factorial design or a multiple of a full factorial design, i.e., the multiset $(\{-1,1\}^m, f_*)$ where the multiplicity function f_* is constant over $\{-1,1\}^m$.

Definition 4. A fraction $\mathscr{F}$ is an Orthogonal Array (OA) of strength t if it factorially projects onto any I-factors with $\#I = t$.

The connections between the OAs and the counting function are given in the proposition below.

Proposition 1. *A fraction is an OA of strength t if and only if all the coefficients* $c_\alpha,\ \alpha \neq 0 \equiv (0,\ldots,0)$ *of the counting function up to the order t are* 0*.*

Definition 5. The Word-Length Pattern (WLP) of a fraction $\mathscr{F}$ of the full factorial design $\mathscr{D}$ is the vector $A_{\mathscr{F}} = (A_0(\mathscr{F}), A_1(\mathscr{F}),\ldots,A_m(\mathscr{F}))$, where

$$A_j(\mathscr{F}) = \sum_{|\alpha|_0 = j} a_\alpha \quad j = 0,\ldots,m,$$

$$a_\alpha = \left(\frac{c_\alpha}{c_0}\right)^2,$$

$|\alpha|_0$ is the number of non-null elements of α, and $c_0 := c_{(0,\ldots,0)} = \#\mathscr{F}/\#\mathscr{D}$.

In the definition above, the number a_α is the aberration of the term X^α.

3 The effect on the WLP of the removal of one, two or three points

In this section we study the effect on the WLP of the removal of one, two or three points from an OA of strength t. With respect to the removal of one point outlined in Sect. 3.1 we obtain an analytical expression for the first $t+1$ terms of the WLP.

3.1 One removed point

Let us consider an orthogonal array $\mathscr{F}$ with n runs, m 2-level factors and strength t. The WLP of $\mathscr{F}$ is

$$A_{\mathscr{F}} = (A_0(\mathscr{F}) = 1, A_1(\mathscr{F}) = 0, \ldots, A_t(\mathscr{F}) = 0, A_{t+1}(\mathscr{F}), \ldots, A_m(\mathscr{F}))\,.$$

Let us consider a point $e = (e_1, \ldots, e_m) \in \mathscr{F}$ and the fraction $\mathscr{F}_e$ that contains all the points of $\mathscr{F}$ apart from e, $\mathscr{F}_e = \mathscr{F} \setminus \{e\}$. Let us denote by $R = \sum_\alpha c_\alpha X^\alpha$ the counting function of $\mathscr{F}$ and by $R_{\{e\}} = \sum_\alpha c_\alpha^{(e)} X^\alpha$ the counting function of the fraction made by the single point e. We can write

$$R_{\{e\}}(x_1, \ldots, x_m) = \frac{1}{2^m}(1 + e_1 x_1) \cdot \ldots \cdot (1 + e_m x_m)\,.$$

The aberration $a_\alpha^{(e)}$ of $R_{\{e\}}$ corresponding to the term α is

$$a_\alpha^{(e)} = \frac{(c_\alpha^{(e)})^2}{(c_0^{(e)})^2} = \frac{(\frac{1}{2^m} e_1^{\alpha_1} \cdot \ldots \cdot e_m^{\alpha_m})^2}{(\frac{1}{2^m})^2} = 1$$

because $(e_1, \ldots, e_m) \in \{-1, 1\}^m$. It follows that the aberration $a_\alpha^{(\mathscr{F}_e)}$ of the fraction $\mathscr{F}_e$ corresponding to the term α, for $1 \le |\alpha|_0 \le t$ is

$$a_\alpha^{(\mathscr{F}_e)} = \frac{(c_\alpha - c_\alpha^{(e)})^2}{(c_0 - c_0^{(e)})^2} = \frac{(c_\alpha^{(e)})^2}{(n/2^m - 1/2^m)^2} = \frac{1}{(n-1)^2}$$

because $c_\alpha = 0$ for $1 \le |\alpha|_0 \le t$. Consequently the terms $A_k(\mathscr{F}_e), k = 1, \ldots, t$ of the WLP of $\mathscr{F}_e$ are

$$A_k(\mathscr{F}_e) = \frac{\binom{m}{k}}{(n-1)^2}, k = 1, \ldots, t\,.$$

This means that $A_k(\mathscr{F}_e), k = 1, \ldots, t$ does not depend on the point e which has been removed. In Sect. 3.2 we study some real examples.

3.2 Examples with one, two or three points

Now we consider the effect on the WLP of the removal of one, two or three points from some OAs. The OAs are taken from the repository publicly available at http://pietereendebak.nl/oapage/, which has been created by Pieter Eendebak and Eric Schoen, see [6].

Let us consider one of the best OAs, with respect to the WLP criterion, in the class of OAs with $m = 5$ 2-level factors, $n = 12$ runs and strength $t = 2$. Let us denote this OA by $\mathscr{F}$ and its points by $e_1, e_2, \ldots, e_{12}$. Writing the runs as columns and the factors as rows, the fraction $\mathscr{F}$ is

$$\mathscr{F} = \begin{array}{c} \begin{array}{cccccccccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 & e_8 & e_9 & e_{10} & e_{11} & e_{12} \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 \end{pmatrix} \end{array}$$

The WLP of $\mathscr{F}$ is $A_{\mathscr{F}} = (1,0,0,A_3(\mathscr{F}) = 1.111, A_4(\mathscr{F}) = 0.5556, A_5(\mathscr{F}) = 0)$.

Now we remove each of the twelve points from $\mathscr{F}$ and we compute the corresponding WLPs. The results are reported in Table 1. We observe that, according to the results of Sect. 3.1, $A_1(\mathscr{F}_e) = 5/(12-1)^2 = 0.041$ and $A_2(\mathscr{F}_e) = 10/(12-1)^2 = 0.083$. It is worth noting that there are two different WLPs. More specifically there are 10 fractions $\mathscr{F}_e$ with $A_3(\mathscr{F}_e) = 1.140$ and 2 fractions $\mathscr{F}_e$ with $A_3(\mathscr{F}_e) = 1.405$.

Table 1 WLPs of the twelve 11-run subsets of $\mathscr{F}$.

point[a]	$A_1(\mathscr{F}_e)$	$A_2(\mathscr{F}_e)$	$A_3(\mathscr{F}_e)$	$A_4(\mathscr{F}_e)$	$A_5(\mathscr{F}_e)$
e_1	0.041	0.083	1.140	0.636	0.008
e_2	0.041	0.083	1.140	0.636	0.008
e_3	0.041	0.083	1.405	0.372	0.008
e_4	0.041	0.083	1.140	0.636	0.008
e_5	0.041	0.083	1.140	0.636	0.008
e_6	0.041	0.083	1.140	0.636	0.008
e_7	0.041	0.083	1.140	0.636	0.008
e_8	0.041	0.083	1.140	0.636	0.008
e_9	0.041	0.083	1.140	0.636	0.008
e_{10}	0.041	0.083	1.405	0.372	0.008
e_{11}	0.041	0.083	1.140	0.636	0.008
e_{12}	0.041	0.083	1.140	0.636	0.008

[a] this column specifies the removed run.

Now, we remove all the possible 66 pairs of points $\{e_i, e_j\}, i, j = 1, \ldots, 12, i < j$ from $\mathscr{F}$. We denote the fraction obtained by removing the points e_i, e_j from $\mathscr{F}$ by $\mathscr{F}_{e_i,e_j}$. We obtain 7 different WLPs which are reported in Table 2.

Finally, we remove all the possible 220 subsets of three points $\{e_i, e_j, e_k\}, i, j, k = 1, \ldots, 12, i < j < k$ from $\mathscr{F}$. We denote the fraction obtained by removing the points e_i, e_j, e_k from $\mathscr{F}$ by $\mathscr{F}_{e_i,e_j,e_k}$. We obtain 12 different WLPs which are reported in Table 3.

The results show that WLPs depend on the *points* which are removed from a given OA and not simply on *their number*. Table 1 demonstrates that even when just one-single point is deleted there are two different WLPs. From Table 1 we can also conclude that the worst fractions in terms of WLPs are $\mathscr{F}_{e_3}$ and $\mathscr{F}_{e_{10}}$. However if we remove two points simultaneously, further investigations revealed that the best fraction in terms of WLP was $\mathscr{F}_{e_3,e_{10}}$ because $A_1(\mathscr{F}_{e_3,e_{10}}) = 0$.

Table 2 The different WLPs obtained by removing 2 points e_i and e_j from $\mathscr{F}$.

N^a	$A_1(\mathscr{F}_{e_i,e_j})$	$A_2(\mathscr{F}_{e_i,e_j})$	$A_3(\mathscr{F}_{e_i,e_j})$	$A_4(\mathscr{F}_{e_i,e_j})$	$A_5(\mathscr{F}_{e_i,e_j})$
1	0	0.4	1.6	0.2	0
10	0.04	0.24	1.2	0.68	0.04
20	0.08	0.16	1.2	0.76	0
10	0.08	0.16	1.52	0.44	0
10	0.12	0.16	1.12	0.76	0.04
10	0.12	0.16	1.44	0.44	0.04
5	0.16	0.24	1.12	0.68	0

[a] N is the number of fractions with the given WLP

Table 3 The different WLPs obtained by removing 3 points e_i, e_j and e_k from $\mathscr{F}$.

N^a	$A_1(\mathscr{F}_{e_i,e_j,e_k})$	$A_2(\mathscr{F}_{e_i,e_j,e_k})$	$A_3(\mathscr{F}_{e_i,e_j,e_k})$	$A_4(\mathscr{F}_{e_i,e_j,e_k})$	$A_5(\mathscr{F}_{e_i,e_j,e_k})$
30	0.062	0.321	1.309	0.852	0.012
10	0.062	0.321	1.704	0.457	0.012
10	0.062	0.42	1.21	0.753	0.111
10	0.062	0.519	1.704	0.259	0.012
30	0.16	0.222	1.21	0.951	0.012
50	0.16	0.222	1.605	0.556	0.012
10	0.16	0.321	1.111	0.852	0.111
10	0.16	0.321	1.506	0.457	0.111
20	0.16	0.42	1.21	0.753	0.012
10	0.259	0.222	1.407	0.556	0.111
20	0.259	0.321	1.111	0.852	0.012
10	0.259	0.321	1.506	0.457	0.012

[a] N is the number of fractions with the given WLP

References

1. Butler, N.A., Ramos, V.M.: Optimal additions to and deletions from two-level orthogonal arrays. J. R. Stat. Soc. Ser. B. Stat. Methodol. **69**(1), 51–61 (2007)
2. Chatzopoulos, S.A., Kolyva-Machera, F., Chatterjee, K.: Optimality results on orthogonal arrays plus p runs for s^m factorial experiments. Metrika **73**(3), 385–394 (2011)
3. Fontana, R., Pistone, G., Rogantin, M.P.: Classification of two-level factorial fractions. J. Statist. Plann. Inference **87**(1), 149–172 (2000)
4. Fries, A., Hunter, W.G.: Minimum aberration 2^{k-p} designs. Technometrics **22**(4), 601–608 (1980)
5. Hedayat, A.S., Sloane, N.J.A., Stufken, J.: Orthogonal Arrays: Theory and Applications. Springer New York, New York (2012)
6. Schoen, E.D., Eendebak, P.T., Nguyen, M.V.: Complete enumeration of pure-level and mixed-level orthogonal arrays. Journal of Combinatorial Designs **18**(2), 123–140 (2010)
7. Street, D.J., Bird, E.M.: D-optimal orthogonal array minus t run designs. J. Stat. Theory Practice pp. 1–20 (2018). DOI 10.1080/15598608.2018.1441081
8. Wang, P., Jan, H.: Designing two-level factorial experiments using Orthogonal Arrays when the run order is important. The Statistician **44**(2), 379–388 (1995)