ON THE ABERRATIONS OF MIXED LEVEL ORTHOGONAL ARRAYS WITH REMOVED RUNS

ROBERTO FONTANA AND FABIO RAPALLO

Abstract. Given an Orthogonal Array we analyze the aberrations of the sub-fractions which are obtained by the deletion of some of its points. We provide formulae to compute the Generalized Word-Length Pattern of any sub-fraction. In the case of the deletion of one single point, we provide a simple methodology to find which the best sub-fractions are according to the Generalized Minimum Aberration criterion. We also study the effect of the deletion of 1, 2 or 3 points on some examples. The methodology does not put any restriction on the number of levels of each factor. It follows that any mixed level Orthogonal Array can be considered.

1. Introduction

The theory of Orthogonal Arrays (OAs) has a long history which began with Combinatorics and is today a major research topic in both methodological and applied Statistics. For instance, OAs are used in industrial experimentation to determine the optimum mix of factors to predict a response variable. 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 [10], [4] and [11].

An important object associated to a design is its Generalized Word-Length Pattern (GWLP). The GWLP is used to discriminate among different designs through the Generalized Minimum Aberration (GMA) criterion: given two designs \( \mathcal{F}_1 \) and \( \mathcal{F}_2 \) with \( m \) factors, the corresponding GWLPs are two vectors

\[
A_{\mathcal{F}_i} = (A_0(\mathcal{F}_i) = 1, A_1(\mathcal{F}_i), \ldots, A_m(\mathcal{F}_i)) \quad i = 1, 2.
\]

The GMA criterion is defined as the sequential minimization of these GWLPs: \( \mathcal{F}_1 \) is better than \( \mathcal{F}_2 \) if there exists \( j \) such that \( A_0(\mathcal{F}_1) = A_0(\mathcal{F}_2), \ldots, A_j(\mathcal{F}_1) = A_j(\mathcal{F}_2) \) and \( A_{j+1}(\mathcal{F}_1) < A_{j+1}(\mathcal{F}_2) \). The GMA criterion is usually applied to Orthogonal Arrays (OA), see [10]. The GMA criterion was introduced in [8] for two-level designs and then extended to non-regular multilevel designs in [10].

While for binary designs the meaning of the GWLP is simple, since it counts the number of defining equations with a given length, in the multilevel case the interpretation of the GWLP is in general less intuitive. Nevertheless, the GWLP is a measure of the aliasing among the simple terms and interaction terms of the design, see e.g. [9].

2010 Mathematics Subject Classification. 62K15.

Key words and phrases. Generalized Minimum Aberration criterion; Generalized Word-Length Pattern; Incomplete designs; Orthogonal Arrays.
In this paper, we consider fractional factorial designs under the GMA criterion, and we focus on the order of the runs, motivated by the following practical problem. In several situations, it is hard to fix the sample size a priori. 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 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. Some results in this direction are introduced in [1] and in [13], but in those papers the goal of the analysis was to maximize the $D$-optimality, and therefore those works consider a model-based environment. Some examples of two-level OAs with removed runs is discussed in [15]. A different approach to the problem of the order of the runs is presented in [14] for the two-level case. Although the goal of the analysis is different Wang and Jan discuss several practical problems involving OAs with missing runs and include some useful references.

In this work, we consider the GMA criterion for binary OAs and we study the behavior of the aberrations when $p$ points are removed from an OA, for small values of $p$. Since the GWLP does not depend on the coding of the factor levels, in [12] the use of the complex coding is suggested to express the basis of the polynomial complex functions over a design, and therefore the counting function. The choice of the complex coding is particularly useful for simplifying the expressions of the aberrations and of the GWLP. Moreover, the coefficients of the counting function can be expressed in terms of the counts of the levels appearing in each simple or interaction term.

In particular, we write the GWLP of a fraction through a row-wise decomposition, showing that the GWLP of a fraction depends only on the mutual position of the design points inside the grid of the corresponding full factorial design. This approach is alternative to the classical decomposition of the GWLP in terms of the aberrations and allows us to easily compute the GWLP for sub-fractions, i.e. fractions with removed runs.

The paper is structured as follows. After a brief summary of the basic definitions concerning counting functions, aberrations, GWLP, and OAs (Sect. 2), we introduce some formulae to decompose the GWLP of a fraction (Sect. 3) and we use such formulae to analyze several examples of OAs with removed runs (Sect. 4). Moreover, a recursive formula for the two-level case is introduced to further simplify the computations (Sect. 5). Finally, some pointers to future works are briefly discussed (Sect. 6).

2. Fractions, counting functions and aberration

In this section we collect some relevant definitions and results on fractions of factorial designs and their representation through polynomial counting functions to fix the notation and to make explicit some major formulae to be used later. In our presentation we use the complex coding of the factor levels, as done in, e.g., [12], although most of the results remains valid for any choice of an orthonormal choice of the coding. Moreover, we briefly summarize some properties of OAs in order to highlight the connections between orthogonality and the coefficients of the counting function. In particular, we recall the results expressing the aberrations
and the GWLP of a fractional design in terms of the coefficients of the polynomial counting function.

Let us consider an experiment with \( m \) factors \( D_j, j = 1, \ldots, m \). Let us code the \( s_j \) levels of the factor \( D_j \) by the \( s_j \)-th roots of the unity

\[
D_j = \{\omega_{s_j}^{(s_j)}, \ldots, \omega_{s_j}^{(s_j)}\},
\]

where \( \omega_{s_j}^{(s_j)} = \exp\left(\frac{2\pi i}{s_j} k\right) \), \( k = 0, \ldots, s_j - 1, j = 1, \ldots, m \).

The full factorial design with complex coding is \( D = D_1 \times \cdots \times D_m \). We denote its cardinality by \( \#D = \prod_{j=1}^{m} s_j \). A fraction \( F \) is a multiset \((F_\ast, f_\ast)\) whose underlying set of elements \( F_\ast \) is a subset of \( D \) and \( f_\ast \) is the multiplicity function \( f_\ast : F_\ast \rightarrow \mathbb{N} \) that gives the number of replicates of each design point of \( F_\ast \) in the multiset \( F \).

When a fraction is a single-replicate fraction, i.e., when \( f_\ast(\zeta) = 1 \) for all \( \zeta \in F_\ast \) we will identify \( F \) and \( F_\ast \) with a slight abuse of notation. The number of design points in \( F \) is denoted with \( n = \#F = \sum_{\zeta \in F_\ast} f_\ast(\zeta) \).

A basis of the vector space of the complex functions over \( D \) can be defined as follows. Define the exponent set:

\[
L = \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}
\]

and the projections onto single factors:

\[
X_j : D \ni (\zeta_1, \ldots, \zeta_m) \mapsto \zeta_j \in D_j.
\]

Then, a basis of the complex functions over \( D \) is formed by

\[
\{X^\alpha = X_1^{\alpha_1} \cdots X_m^{\alpha_m} : \alpha \in L\}.
\]

Let \( |\alpha|_0 \) be the number of non-null elements of \( \alpha \). Following the standard terminology in factorial design theory, the monomials \( X^\alpha \) are called factors when \( |\alpha|_0 = 1 \) and interaction terms when \( |\alpha|_0 \geq 2 \). We use this basis to represent the counting function of a fraction according to the following definition.

**Definition 2.1.** The counting function \( R \) of a fraction \( F \) is a complex polynomial defined over \( D \) so that for each \( \zeta \in D \), \( R(\zeta) \) equals the number of appearances of \( \zeta \) in the fraction. A \( 0 - 1 \) valued counting function is called an indicator function of a single-replicate fraction \( F \). We denote by \( c_\alpha \) the coefficients of the representation of \( R \) on \( D \) using the monomial basis \( \{X^\alpha, \alpha \in L\} \):

\[
R(\zeta) = \sum_{\alpha \in L} c_\alpha X^\alpha(\zeta), \ \zeta \in D, \ c_\alpha \in \mathbb{C}.
\]

In the proposition below we summarize some properties of the coefficients of the counting function. The proof of all items can be found in \([12]\). Here \( \overline{z} \) is the complex conjugate of the complex number \( z \).

**Proposition 2.2.** If \( F \) is a fraction of a full factorial design \( D \), \( R = \sum_{\alpha \in L} c_\alpha X^\alpha \)

is its counting function and \( [\alpha - \beta] \) is the \( m \)-tuple made by the componentwise difference in the rings \( \mathbb{Z}_{s_j}, ([\alpha_1 - \beta_1]_{s_1}, \ldots, [\alpha_m - \beta_m]_{s_m}) \), then

1. the coefficients \( c_\alpha \) are given by \( c_\alpha = \frac{1}{\#F} \sum_{\zeta \in F} X^\alpha(\zeta) \);
2. the term \( X^\alpha \) is centered on \( F \), i.e., \( \frac{1}{\#F} \sum_{\zeta \in F} X^\alpha(\zeta) = 0 \) if, and only if, \( c_\alpha = c_{[-\alpha]} = 0 \);
3. the terms \( X^\alpha \) and \( X^\beta \) are orthogonal on \( F \) if and only if, \( c_{[\alpha - \beta]} = 0 \).
Proposition 2.2 has a major application in the representation of the Orthogonal Arrays through the counting function. Recall that, given a subset of indices $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}$, $i_1 < \ldots < i_k$, the projection of a design point $\zeta$ into $D_I := D_{i_1} \times \cdots \times D_{i_k}$ is

$$\pi_I(\zeta) = \zeta_I = (\zeta_{i_1}, \ldots, \zeta_{i_k}) \in D_{i_1} \times \cdots \times D_{i_k}.$$ 

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

Definition 2.3. A fraction $F$ is a (mixed level) Orthogonal Array (OA) of strength $t$ if it factorially projects onto any $I$-factors with $|I| = t$.

From Proposition 2.2 it follows that a fraction factorially projects onto the $I$-factors, $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}$, $i_1 < \ldots < i_k$, if and only if all the coefficients of the counting function involving the $I$-factors only are 0. Thus, 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.

Using the polynomial counting function, [3] provides the following definition of the GWLP $A_F = (A_0(F), \ldots, A_m(F))$ of a fraction $F$.

Definition 2.4. The Generalized Word-Length Pattern (GWLP) of a fraction $F$ of the full factorial design $D$ is the vector $A_F = (A_0(F), A_1(F), \ldots, A_m(F))$, where

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

(1)

$$a_\alpha = \left(\frac{\left\|c_\alpha\right\|_2^2}{c_0}\right)^2,$$

$\left\|z\right\|_2$ is the norm of the complex number $z$, and $c_0 := c_{(0,\ldots,0)} = n/\#D$.

We refer to $a_\alpha$ as the aberration of the interaction $X_\alpha$. Note that $A_0(F) = 1$ for all $F$.

3. Decomposition of the GWLP and applications

In this section we show some formulae to compute the GWLP of a fraction. The first proposition shows that the elements of the GWLP of a fraction depend only on the mutual position of its design points.

Proposition 3.1. Given a fraction $F$ of size $n$,

(2)

$$n^2 A_j(F) = \sum_{f \in F} \sum_{g \in F} \sum_{A = \{a_1, \ldots, a_j\} \subseteq \{1, \ldots, m\}} S_{(f,g)}^{(f,g)} \cdots S_{(f,g)}^{(f,g)}$$

where

$$S_i^{(f,g)} = \begin{cases} -1 & \text{if } f_i \neq g_i \\ s_i - 1 & \text{if } f_i = g_i \quad i = 1, \ldots, m \end{cases}$$

and $j = 1, \ldots, m$. 
Proof. Given a fraction $\mathcal{F}$, let $R(\zeta) = \sum_{\alpha \in L} c_\alpha X^\alpha(\zeta)$, $\zeta \in D$, $c_\alpha \in \mathbb{C}$ be its counting function. From Eq. (I), the $j$-th term of the GWLP of $\mathcal{F}$ is

$$A_j(\mathcal{F}) = \sum_{|\alpha|=j} \left( \frac{\|c_\alpha\|_2}{c_0} \right)^2 = j = 0, \ldots, m.$$  

The counting function $R$ can be written as the sum of the counting functions of the points of $\mathcal{F}$. Let $R^{(f)}$ be the indicator function of a point $f \in D$. We have

$$R = \sum_{f \in \mathcal{F}} R^{(f)}.$$

We observe that $\mathcal{F}$ can be a multiset, i.e. it can exist $\zeta \in \mathcal{F}$ such that $R(\zeta) > 1$. The corresponding term in Eq. (3) will be

$$\frac{R^{(\zeta)} + \ldots + R^{(\zeta)}}{R(\zeta) \text{ times}}.$$  

We denote by $c_\alpha^{(f)}$ the coefficients of the indicator function of the point $f$, $R^{(f)} = c_\alpha^{(f)} X^\alpha$. Let us consider the case $m = 1$, $D = \{\omega_1^{(s_1)}, \ldots, \omega_{s_1-1}^{(s_1)}\}$. It follows that

$$D = \{\omega_f^{(s_1)} : f = 0, 1, \ldots, s_1 - 1\}$$

and a generic point $f \in D$ can be written, with a small abuse of notation, as $\omega_f^{(s_1)}$. It is not difficult to show that

$$R^{(f)}(\zeta) = \frac{1}{s_1} (1 + f s_1^{-1} \zeta + f s_1^{-2} \zeta^2 + \ldots + f \zeta(s_1-1))$$

that is

$$c_\alpha^{(f)} = \frac{1}{s_1} f s_1^{-\alpha} = \frac{1}{s_1} (\omega_f^{(s_1)})^{s_1-\alpha} = \frac{\omega_1^{(s_1)}}{s_1} \text{ for } \alpha = 0, 1, \ldots, s_1 - 1.$$  

The generalization to the case $m > 1$ is straightforward. Given $f = (f_1, \ldots, f_m) \equiv (\omega_{f_1}^{(s_1)}, \ldots, \omega_{f_m}^{(s_m)}) \in D$ we get

$$c_\alpha^{(f)} = \frac{1}{\#D} \omega_1^{(s_1)} \cdot \ldots \cdot \omega_{s_m}^{(s_m)}, \quad \alpha \in L.$$  

We can write

$$c_0^2 A_j(\mathcal{F}) = \sum_{|\alpha|=j} \|c_\alpha\|^2 = \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} \sum_{|\alpha|=j} c_\alpha^{(f)} c_\alpha^{(g)}, \quad j = 0, \ldots, m$$

and we obtain

$$c_\alpha^{(f)} c_\alpha^{(g)} = \frac{1}{\#D^2} \omega_{s_1}^{(s_1)} \cdot \ldots \cdot \omega_{s_m}^{(s_m)}.$$  

Let us consider the $m$-th term of the GWLP of $\mathcal{F}$. We get

$$c_0^2 A_m(\mathcal{F}) = \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} \sum_{|\alpha|=m} c_\alpha^{(f)} c_\alpha^{(g)} =$$

$$= \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} \frac{1}{\#D^2} \omega_{s_1}^{(s_1)} \cdot \ldots \cdot \omega_{s_m}^{(s_m)}.$$  

$$\sum_{|\alpha|=m}$$

$$= \frac{1}{\#D^2} \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} \frac{1}{\omega_{s_1}^{(s_1)} \cdot \ldots \cdot \omega_{s_m}^{(s_m)}}$$

and

$$= \frac{1}{\#D^2} \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} \frac{1}{\omega_{s_1}^{(s_1)} \cdot \ldots \cdot \omega_{s_m}^{(s_m)}}.$$  

The $j$-th term of the GWLP of $\mathcal{F}$ is

$$A_j(\mathcal{F}) = \sum_{|\alpha|=j} \left( \frac{\|c_\alpha\|_2}{c_0} \right)^2 = j = 0, \ldots, m.$$
We observe that, for \( i = 1, \ldots, m \) we have
\[
\sum_{\alpha_i = 1}^{s_i - 1} \omega_{\alpha_i}^{(s_i)} (g_i - f_i) = \begin{cases} 
\sum_{\alpha_i = 0}^{s_i - 1} \omega_{\alpha_i}^{(s_i)} (g_i - f_i) & \text{if } f_i \neq g_i \\
\sum_{\alpha_i = 1}^{s_i - 1} \omega_{\alpha_i}^{(s_i)} & \text{if } f_i = g_i.
\end{cases}
\]
It follows
\[
c_0^2 A_m (\mathcal{F}) = \frac{1}{\# D^2} \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} (S_1^{(f,g)} \cdots S_m^{(f,g)})
\]
where
\[
S_i^{(f,g)} = \begin{cases} 
-1 & \text{if } f_i \neq g_i, \\
0 & \text{if } f_i = g_i.
\end{cases}
\]
For \( j < m \) it is sufficient to apply the formula above for all the subsets of size \( j \) of the set \( \{1, \ldots, m\} \)
\[
c_0^2 A_j (\mathcal{F}) = \frac{1}{\# D^2} \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} \sum_{\substack{A = \{a_1, \ldots, a_j\} \\ A \subseteq \{1, \ldots, m\}}} (S_1^{(f,g)} \cdots S_m^{(f,g)}).
\]
To complete the proof it is enough to observe that \( c_0 = \frac{1}{\# \mathcal{F}} \).

As a consequence, the GWLP of a singleton is a constant depending only on \( s_1, \ldots, s_m \), as stated below.

**Corollary 3.2.** Given a design point \( f \) in a factorial design \( \mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_m \), the \( j \)-th element of the GWLP of \( f \) is
\[
A_j (f) = \sum_{\substack{A = \{a_1, \ldots, a_j\} \\ A \subseteq \{1, \ldots, m\}}} (s_{a_1} - 1) \cdots (s_{a_j} - 1),
\]
where \( s_1, \ldots, s_m \) are the number of levels of \( \mathcal{D}_1, \ldots, \mathcal{D}_m \) respectively.

The formula in Eq. (7) becomes very simple for symmetric designs. Indeed, when \( s_1 = \ldots = s_m = s \), we have
\[
A_j (f) = \binom{m}{j} (s - 1)^j.
\]

Now we show some formulae to decompose the GWLP of a fraction. The proof of the first result can be found in [6]. The subsequent results exploit the results in Prop. 3.1 and Cor. 3.2 and will be useful for our purpose, i.e., to choose the best design points to be removed from a given fraction.

In the following proposition we consider the union of \( k \) fractions, \( k \geq 2 \). Let us consider fractions \( \mathcal{F}_1, \ldots, \mathcal{F}_k \) with \( n_1, \ldots, n_k \) design points, respectively. Let us denote by \( R^{(i)} = \sum_{\alpha \in L} c^{(i)}_\alpha X^\alpha \) the counting function of \( \mathcal{F}_i \), \( i = 1, \ldots, k \). When we consider the union \( \mathcal{F} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_k \) of size \( n = n_1 + \cdots + n_k \), the counting function of \( \mathcal{F} \) is clearly \( R = \sum_{i=1}^k R^{(i)} \).

**Proposition 3.3.** With the notation above, let us consider fractions \( \mathcal{F}_1, \ldots, \mathcal{F}_k \) with \( n_1, \ldots, n_k \) design points, respectively, and their union \( \mathcal{F} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_k \). The \( j \)-th element of the GWLP of \( \mathcal{F} \) is
\[
A_j (\mathcal{F}) = \sum_{i=1}^k \frac{n_i}{n^2} A_j (\mathcal{F}_i) + \frac{\# \mathcal{D}}{n^2} \sum_{i_1 \neq i_2} \sum_{\alpha \in L} c^{(i_1)}_\alpha c^{(i_2)}_\alpha, \quad j = 0, \ldots, m.
\]
Remark 1. The term $\sum_{|\alpha|=j} c_{\alpha}^{(f_1)} c_{\alpha}^{(f_2)}$ in Eq. (9) can be viewed as a kind of covariance between the coefficients of order $j$ of the two counting functions $R^{(f_1)}$ and $R^{(f_2)}$.

We consider now two special cases of Prop. 3.3. In the first proposition we decompose a fraction as the union of singletons, while in the second one we explicitly write the formula for the GWLP of a fraction with one removed run.

**Proposition 3.4.** Let $\mathcal{F} = \{f_1, \ldots, f_n\}$ be a fraction with $n$ runs, and let $R^{(f_i)} = \sum_{\alpha} c_{\alpha}^{(f_i)} X^\alpha$ be the indicator function of the point $f_i$. Then

$$A_j(\mathcal{F}) = \frac{1}{n^2} \sum_{i=1}^n A_j(f_i) + \frac{(\#D)^2}{n^2} \sum_{i_1 \neq i_2} \sum_{|\alpha|=j} c_{\alpha}^{(f_1)} c_{\alpha}^{(f_2)} j = 0, \ldots, m.$$  \hspace{1cm} (10)

Now, let us take a fraction $\mathcal{F}$ and a design point $f \in \mathcal{F}$. We denote with $\mathcal{F}_j$ the fraction with $n-1$ runs obtained by removing $f$ from $\mathcal{F}$.

**Proposition 3.5.** Let $\mathcal{F} = \{f_1, \ldots, f_n\}$ be a fraction with $n$ runs, and let $R^{(f_i)} = \sum_{\alpha} c_{\alpha}^{(f_i)} X^\alpha$ be the indicator function of the point $f_i$. Then

$$A_j(\mathcal{F}) = \left(\frac{n-1}{n}\right)^2 A_j(\mathcal{F}_j) + \frac{1}{n^2} A_j(f) + \frac{(\#D)^2}{n^2} \sum_{g \in \mathcal{F}, g \neq f} \sum_{|\alpha|=j} c_{\alpha}^{(f)} c_{\alpha}^{(g)}$$  \hspace{1cm} (11)

for $j = 0, \ldots, m$.

Notice that in the Equations (10) and (11) the term involving a singleton is constant.

Before the use of the previous results in actual computations, some remarks are in order.

**Remark 2.** The last summand in Equations (9), (10), and (11) is independent on the choice of orthonormal contrasts, because it is the difference of GWLPs, see [16], page 1069. Therefore, the choice of the complex coding is due merely to computational reasons.

**Remark 3.** From Eq. (7), we note that the elements $A_j(\mathcal{F})$ of the GWLP of a fraction $\mathcal{F}$ depends only on the mutual position of the runs, and in this sense the elements $A_j(\mathcal{F})$ of the GWLP of a fraction with two runs can be viewed as a kind of distance of the design points.

**Remark 4.** Another consequence of Equations (7) and (10) is that the terms

$$\sum_{|\alpha|=j} c_{\alpha}^{(f_1)} c_{\alpha}^{(f_2)}$$

in the last summand of Eq. (10) can be used also for all sub-fractions.

The results above suggest to introduce a sequence of matrices based on Eq. (2) to easily compute the GWLP of a fraction and its sub-fractions.

**Definition 3.6.** Given a fraction $\mathcal{F}$ with $n$ runs of a full factorial design with $m$ factors, define for each $j = 1, \ldots, m$ the $n \times n$ matrix $W_j$ with generic element

$$W_j(f_1, f_2) = (\#D)^2 \sum_{|\alpha|=j} c_{\alpha}^{(f_1)} c_{\alpha}^{(f_2)} = \sum_{A = \{a_1, \ldots, a_j\}} \{a_j \in \{1, \ldots, m\}\} \sum_{\alpha \subseteq A} c_{\alpha}^{(f_1, f_2)}$$  \hspace{1cm} (12)
It follows from Proposition 3.1 that the \( j \)-th aberration can be written as
\[
A_j(\mathcal{F}) = \frac{1}{n^2} \sum_{f \in \mathcal{F}} \sum_{g \in \mathcal{F}} W_j(f, g), \quad j = 1, \ldots, m.
\]

It also follows that given a point \( f \in \mathcal{F} \)
\[
(13) \quad n^2A_j(\mathcal{F}) = (n - 1)^2A_j(\mathcal{F}_f) + w_{j,f}
\]
where
\[
w_{j,f} = \sum_{c=1}^{n} W_j(f, c) + \sum_{r=1}^{n} W_j(r, f) - W_j(f, f).
\]

Exploiting the formula in Eq. (13) it is easy to use \( w_{j,f} \) in order to choose the best point to be removed, i.e., the run which is candidate to be the last run of the OA. In the next section several examples are illustrated to show how this procedure works. Notice that in principle it would be easy to define a step-by-step algorithm removing one run at a time. In fact, the new matrices \( W_j \) for the sub-fraction with \( n - 1 \) runs can be obtained simply by deleting the row and column pertaining to the removed run, and therefore such matrices allows us to compute the GWLPs of fractions with two removed runs, and so on. However, as discussed in Sect. 4.1, such a procedure is in general not hierarchical.

4. Examples

In this section we study the effect on the GWLP of the removal of one, two or three points from an OA of strength \( t \). We consider both symmetric and mixed level OAs and we do not restrict the number of levels to be prime or prime power. Most of the examples here are chosen from the OA catalogue in [5].

4.1. \( OA(12, 2^5, t = 2) \). We consider an OA with 12 runs, five 2-level factors and strength 2. Writing the runs as columns and the factors as rows, the fraction \( \mathcal{F} \) is

\[
\mathcal{F} = \begin{pmatrix}
  f_1 & f_2 & f_3 & f_4 & f_5 & f_6 & f_7 & f_8 & f_9 & f_{10} & f_{11} & f_{12} \\
  1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
  1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
  1 & 1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
  1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
  1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\
  1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
\end{pmatrix}
\]

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

We remove each of the twelve points from \( \mathcal{F} \) and we compute the corresponding GWLPs. The results are reported in Table 1. We observe that, according to the results of Sect. 3, \( A_1(\mathcal{F}_f) = 5/(12 - 1)^2 = 0.041 \) and \( A_2(\mathcal{F}_f) = 10/(12 - 1)^2 = 0.083 \). It is worth noting that there are two different GWLPs. More specifically there are 10 fractions \( \mathcal{F}_f \) with \( A_3(\mathcal{F}_f) = 1.140 \) and 2 fractions \( \mathcal{F}_f \) with \( A_3(\mathcal{F}_f) = 1.405 \).

The symmetric matrix \( W_3 \) (as defined in Sect. 3) is written in the columns labeled \( f_1, \ldots, f_{12} \) of Table 2. The last column of Table 2 reports the value of \( w_{3,f}, f \in \mathcal{F} \). It follows from Eq. (13) that if we want to choose a single point \( f \) to be removed in a way that \( A_j(\mathcal{F}_f) \) is as small as possible we must select one of the points for which \( w_{3,f} \) is as large as possible. In this case, for minimizing \( A_3(\mathcal{F}_f) \) we must select \( f \notin \{f_3, f_{10}\} \). These results are confirmed by the values of \( A_3(\mathcal{F}_f) \) in Table 1.
Table 1. GWLPs of the fractions with one removed run for the OA in Sect. 4.1

| point | $A_1(F_f)$ | $A_2(F_f)$ | $A_3(F_f)$ | $A_4(F_f)$ | $A_5(F_f)$ |
|-------|-----------|-----------|-----------|-----------|-----------|
| $f_1$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_2$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_3$ | 0.041     | 0.083     | 1.405     | 0.372     | 0.008     |
| $f_4$ | 0.041     | 0.083     | 1.405     | 0.372     | 0.008     |
| $f_5$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_6$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_7$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_8$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_9$ | 0.041     | 0.083     | 1.14      | 0.636     | 0.008     |
| $f_{10}$ | 0.041  | 0.083 | 1.405 | 0.372 | 0.008 |
| $f_{11}$ | 0.041 | 0.083 | 1.14 | 0.636 | 0.008 |
| $f_{12}$ | 0.041 | 0.083 | 1.14 | 0.636 | 0.008 |

*a*this column specifies the removed run.

Table 2. The $W_3$ matrix of $OA(12, 2^5, t = 2)$ for the OA in Sect. 4.1

| $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ | $f_8$ | $f_9$ | $f_{10}$ | $f_{11}$ | $f_{12}$ | $w_{3,f}$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------|----------|----------|-----------|
| 10    | -2    | -2    | 2     | 2     | -2    | 2     | 2     | 2     | 2       | 2        | 2        | 22        |
| -2    | 10    | 2     | 2     | -2    | 2     | -2    | 2     | -2    | -10     | 2        | 2        | -10       |
| -2    | 2     | 10    | -2    | -2    | 2     | 2     | 2     | 2     | 2       | 2        | 2        | 22        |
| 2     | -2    | 2     | 10    | -2    | 2     | 2     | 2     | -2    | 2       | 2        | -2       | 22        |
| 2     | 2     | -2    | -2    | 10    | 2     | 2     | -2    | 2     | 2       | 2        | 2        | 22        |
| 2     | 2     | -2    | 2     | 10    | 2     | 2     | -2    | -2    | 2       | 2        | 2        | 22        |
| -2    | 2     | -2    | 2     | 2     | 2     | 10    | -2    | 2     | 2       | 2        | -2       | 22        |
| 2     | 2     | -2    | 2     | -2    | 2     | -2    | 10    | 2     | 2       | 2        | -2       | 22        |
| 2     | -2    | -10   | 2     | -2    | 2     | -2    | 2     | 10    | -2      | -2       | -2       | -2        |
| 2     | 2     | -2    | -2    | 2     | -2    | 2     | -2    | 2     | 10      | -2      | -2       | 22        |
| 2     | 2     | 2     | -2    | -2    | 2     | -2    | -2    | 2     | 10      | -2      | -2       | 22        |

A simple sequential strategy can be defined. Once a run has been removed, the new $W_j$ matrix is obtained by simply deleting the row and the column corresponding to the removed point. Then the second point to be removed could be chosen by computing the new value of $w_{j,f}$ based on the new $W_j$ matrix. The problem is that, in general, this strategy does not lead to an optimal selection of the pair of points to be removed. In the case under study it is possible to verify that if we remove $f_1$ in the first step than the best possible choices for the second point to be removed would be $f_6$ or $f_9$ for which the aberration $A_1(F_{f_1,f_6}) = A_1(F_{f_1,f_9}) = 0.04$. But if we select as the pair of points to be removed \{f_3, f_{10}\} we obtain $A_1(F_{f_3,f_{10}}) = 0$. It follows that to have an optimal strategy the number of points to be removed must be fixed in advance.

4.2. Plackett-Burman $OA(12, 2^{11}, t = 2)$. We consider the Plackett-Burman design with 12 runs and eleven two-level factors. It has strength 2, i.e., resolution
Table 3. One to three points removed - unique GWLPs for the Plackett-Burman design in Sect. 4.2

| p | N   | A_1 | A_2   | A_3 | ... | A_{10} | A_{11} |
|---|-----|-----|-------|-----|-----|--------|--------|
| 1 | 12  | 0.091 | 0.455 | 19.545 | ... | 0.091 | 1      |
| 2 | 66  | 0.2  | 1     | 21   | ... | 0.2    | 1      |
| 3 | 220 | 0.333 | 1.667 | 22.778 | ... | 0.333 | 1      |

p is the number of removed points, N is the number of fractions with the same GWLP.

Table 4. One to two points removed - unique GWLPs for the OA in Sect. 4.3

| p | N   | A_1  | A_2  | A_3  | A_4  |
|---|-----|------|------|------|------|
| 1 | 18  | 0.024 | 0.062 | 0.567 | 1.522 |
| 2 | 27  | 0.023 | 0.188 | 0.617 | 1.547 |
| 3 | 18  | 0.039 | 0.141 | 0.664 | 1.531 |
| 2 | 27  | 0.047 | 0.117 | 0.688 | 1.523 |
| 2 | 54  | 0.063 | 0.117 | 0.641 | 1.555 |
| 2 | 27  | 0.07  | 0.117 | 0.617 | 1.57  |

p is the number of removed points, N is the number of fractions with the same GWLP.

III. We point out that the removal of even a single point leads to a design where the number of parameters to be estimated is larger than the number of runs. We observe that any choice of one run to be removed leads to the same GWLP, which is reported in the first row of Table 3. Similarly any choice of two (three) points leads to the same GWLP which is reported in the second (third) row of Table 3.

4.3. OA(18, 2^{13}, t = 2). We consider an orthogonal array with 18 runs, one two-level and 3 three-level factors, and strength 2. In this case any choice of one point to be removed leads to the same GWLP, which is reported in the first row of Table 3. But different choices of pairs of points give different GWLPs. From Table 3 for example we observe that A_1 varies between 0.023 and 0.07.

4.4. OA(16, 2^44^2, t = 2). As the last example we consider an orthogonal array with 16 runs, four 2-level and two 4-level factors of strength 2. The results are very similar to those of the previous case. The choice of one single point to be removed does not affect the GWLP while the choice of different pairs of points can lead to different GWLPs as reported in Table 3.

5. Two-level designs

In the case of two-level designs the full factorial design is \( D = \{-1, 1\}^m \), the coefficients of the counting function are real numbers and therefore the aberrations in Eq. 1 are simply

\[
a_\alpha = \left( \frac{c_\alpha}{c_0} \right)^2
\]
Table 5. One to two points removed - unique GWLPs for the OA in Sect. 4.4

| p  | N  | A_1 | A_2 | A_3 | A_4 | A_5 | A_6 |
|----|----|-----|-----|-----|-----|-----|-----|
| 1  | 16 | 0.044 | 0.173 | 6.311 | 8.316 | 0.187 | 1.036 |
| 2  | 8  | 0.041 | 0.551 | 6.449 | 8.796 | 0.367 | 1.082 |
| 2  | 32 | 0.082 | 0.388 | 6.694 | 8.633 | 0.408 | 1.082 |
| 2  | 32 | 0.102 | 0.327 | 6.735 | 8.673 | 0.347 | 1.102 |
| 2  | 32 | 0.102 | 0.367 | 6.653 | 8.673 | 0.429 | 1.061 |
| 2  | 16 | 0.122 | 0.347 | 6.612 | 8.714 | 0.449 | 1.041 |

*p* is the number of removed points, *N* is the number of fractions with the same GWLP.

and all the computations yielding the elements of the GWLP involve only real numbers. Also the GWLP of a singleton assumes an easy form, since Eq. 48 reduces to

\[ A_j(f) = \binom{m}{j}. \]

But the main feature of the two-level case is that we can establish a recursive formula for the matrices \( W_j \).

**Proposition 5.1.** Let \( F \) be a fraction with \( n \) runs of a two-level full factorial design \( D = \{-1,1\}^m \), and denote by \( X \) the design matrix of \( F \). The sequence of matrices \( W_0, W_1, \ldots, W_m \) satisfy the recursive formula

\[
W_0 = J \quad (14)
\]

\[
W_1 = XX^t \quad (15)
\]

\[
W_j = \frac{1}{j!} (W_1 \ast W_{j-1} - (j-1)(m-j+2)W_{j-2}) \quad j = 2, \ldots, m \quad (16)
\]

where \( J \) is an \( n \times n \) matrix with all entries equal to 1 and \( \ast \) denotes the element-wise product of two matrices.

**Proof.** For \( W_0 \) the computation is trivial and for \( W_1 \) it is enough to observe that

\[ W_1(f_1, f_2) = \sum_i \omega_i(f_1)\omega_i(f_2). \]

To shorten the notation in the proof of the recursive formula, we write \( \varphi_i = \omega_i(f_1)\omega_i(f_2) \) and \( \varphi_{i_1 \ldots i_j} = \varphi_{i_1} \ldots \varphi_{i_j} \). We have:

\[
W_j(f_1, f_2) = \#D^2 \sum_{|\alpha|_0=j} c_{\alpha}(f_1)c_{\alpha}(f_2) = \frac{1}{j!} \sum_{\varphi_{i_1 \ldots i_j}} =
\]

\[
= \frac{1}{j!} \sum_{i_1} \varphi_{i_1} \left( \sum_{\varphi_{i_2 \ldots i_j} \text{ distinct} \atop i_2 = i_3 = \ldots = \ldots = i_{j-1}} \varphi_{i_2 \ldots i_j} - \sum_{\varphi_{i_2 \ldots i_j} \text{ distinct} \atop i_1 = i_2} \varphi_{i_2 \ldots i_j} - \ldots - \sum_{\varphi_{i_2 \ldots i_j} \text{ distinct} \atop i_1 = i_{j-1}} \varphi_{i_2 \ldots i_j} \right)
\]
The first sum in the round bracket is $W_{j-1}(f_1, f_2)$ while the last $(j - 1)$ terms are clearly equal. Thus,

$$= \frac{1}{j!} \left\{ W_1(f_1, f_2)W_{j-1}(f_1, f_2) - (j - 1) \sum_{i_1} \left( \varphi_{i_1}^2 \sum_{i_3, \ldots, i_j \text{ distinct}} \varphi_{i_3 \ldots i_j} \right) \right\} =$$

changing the order of the sums in the last term

$$= \frac{1}{j!} \left\{ W_1(f_1, f_2)W_{j-1}(f_1, f_2) - (j - 1) \sum_{i_3, \ldots, i_j \text{ distinct}} \varphi_{i_3 \ldots i_j} \left( \sum_{i_1} \varphi_{i_1}^2 \right) \right\} =$$

since $\varphi_{i_1}^2 = 1$ and $(m - j + 2)$ terms are summed up we obtain

$$= \frac{1}{j!} (W_1(f_1, f_2)W_{j-1}(f_1, f_2) - (j - 1)(m - j + 2)W_{j-2}(f_1, f_2)) .$$

\[\square\]

6. Final remarks

The formulae introduced in this paper allow us to easily compute the GWLP of Orthogonal Arrays with removed runs. However, as discussed in Sect. 4.1, the choice of the best GWLP is not hierarchical when more than one point is removed. Future work on this topic could focus on the characterization of some classes of Orthogonal Arrays with constant GWLP over all sub-fractions, when removing 1, 2, 3, and possibly more than 3 runs. It would also be interesting to compare the GWLP of such Orthogonal Arrays with the $R^2$ and the canonical correlations as in [9], extending the analysis to designs which are different from Orthogonal Arrays. Finally, the connections between our approach based on the aberrations and the $D$-optimality criterion for some classical statistical models are worth exploring. Some efficient algorithms for finding such Orthogonal Arrays would be helpful, for example by exploiting the notion of mean aberration for mixed level Orthogonal Arrays introduced in [7].

References

[1] Neil A. Butler and Victorino M. Ramos. Optimal additions to and deletions from two-level orthogonal arrays. J. R. Stat. Soc. Ser. B. Stat. Methodol., 69(1):51–61, 2007.
[2] Stavros A. Chatzopoulos, Fotini Kolyva-Machera, and Kashinath Chatterjee. Optimality results on orthogonal arrays plus $p$ runs for $s^m$ factorial experiments. Metrika, 73(3):385–394, 2011.
[3] Shao-Wei Cheng and Kenny Q Ye. Geometric isomorphism and minimum aberration for factorial designs with quantitative factors. Ann. Statist., 32(5):2168–2185, 2004.
[4] Aloke Dey and Rahul Mukerjee. Fractional Factorial Plans. John Wiley & Sons, New York, 2009.
[5] Pieter Endebak and Eric Schoen. Complete series of non-isomorphic orthogonal arrays. http://pieterendeback.nl/oapage/, 2018. Accessed: 2018-07-31.
[6] Roberto Fontana and Fabio Rapallo. Unions of orthogonal arrays and their aberrations via Hilbert bases. Technical Report arXiv:1801.00501, 2018. Submitted.
[7] Roberto Fontana, Fabio Rapallo, and Maria Piera Roganti. Aberration in qualitative multi-level designs. J. Statist. Plann. Inference, 174:1–10, 2016.
[8] Arthur Fries and William G. Hunter. Minimum aberration $2^{k-p}$ designs. Technometrics, 22(4):601–608, 1980.
[9] Ulrike Grömping and Hongquan Xu. Generalized resolution for orthogonal arrays. *Ann. Statist.*, 42(3):918–939, 2014.

[10] A. S. Hedayat, Neil J. A. Sloane, and John Stufken. *Orthogonal Arrays: Theory and Applications*. Springer, New York, 2012.

[11] Rahul Mukerjee and C. F. Jeff Wu. *A Modern Theory of Factorial Design*. Springer, New York, 2007.

[12] Giovanni Pistone and Maria-Piera Rogantin. Indicator function and complex coding for mixed fractional factorial designs. *J. Statist. Plann. Inference*, 138(3):787–802, 2008.

[13] Deborah J. Street and Emily M. Bird. D-optimal orthogonal array minus t run designs. *J. Stat. Theory Practice*, 12(3):575–594, 2018.

[14] P.C. Wang and H.W. Jan. Designing two-level factorial experiments using Orthogonal Arrays when the run order is important. *The Statistician*, 44(2):379–388, 1995.

[15] Rafel Xampeny, Pere Grima, and Xavier Tort-Martorell. Which runs to skip in two-level factorial designs when not all can be performed. *Quality Engineering*, 2018.

[16] Hongquan Xu and C. F. Jeff Wu. Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Statist.*, 29(4):1066–1077, 2001.

DEPARTMENT DISMA, POLITECNICO DI TORINO, TORINO, ITALY
*E-mail address*: roberto.fontana@polito.it

DEPARTMENT DISIT, UNIVERSITY OF PIEMONTE ORIENTALE, ALESSANDRIA, ITALY
*E-mail address*: fabio.rapallo@uniupo.it