Ann Inst Stat Math (2010) 62:673–698
DOI 10.1007/s10463-010-0291-8

Minimal average degree aberration and the state polytope for experimental designs

Yael Berstein · Hugo Maruri-Aguilar · Shmuel Onn · Eva Riccomagno · Henry Wynn

Received: 15 October 2008 / Revised: 8 February 2010 / Published online: 31 March 2010

Abstract For a particular experimental design, there is interest in finding which polynomial models can be identified in the usual regression set up. The algebraic methods based on Gröbner bases provide a systematic way of doing this. The algebraic method does not, in general, produce all estimable models but it can be shown that it yields models which have minimal average degree in a well-defined sense and in both a weighted and unweighted version. This provides an alternative measure to that based on "aberration" and moreover is applicable to any experimental design. A simple algorithm is given and bounds are derived for the criteria, which may be used to give asymptotic Nyquist-like estimability rates as model and sample sizes increase.

Keywords Corner cut · Design ideal · Factorial design · Latin hypercube sampling · Linear aberration · State polytope

1 Introduction

It is of considerable value to represent an experimental design as the solution of a set of polynomial equations. In the terminology of algebraic geometry a design is a zero-

Y. Berstein · S. Onn
Israel Institute of Technology, 32000 Haifa, Israel

H. Maruri-Aguilar (✉)
School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, UK
e-mail: H.Maruri-Aguilar@qmul.ac.uk

E. Riccomagno
Dipartimento di Matematica, Università di Genova, 16146 Genova, Italy

H. Wynn
Department of Statistics, London School of Economics, London WC2A 2AE, UK

dimensional variety and the corresponding ideal comprising all polynomials which are zero on every design point is called an "ideal of points". Pistone and Wynn (1996) first used explicit methods from algebraic geometry and in particular introduced Gröbner bases into designs. Issues to do with identifiability of polynomial regression models, or interpolators, can be translated into problems about such varieties and ideals (see Pistone et al. 2001).

The purpose of this paper is to introduce the notion of linear aberration of a polynomial model. Linear aberration is defined only for polynomial models, which are used routinely in statistical literature. A polynomial model with low order terms has low aberration, thus engaging low aberration with the standard practice of preferring polynomial models with low order terms. The preference for models with low order terms has been acknowledged in recent papers, see Li et al. (2003) and Balakrishnan and Yang (2006), although they do not refer to linear aberration.

Let $\alpha = (\alpha_1, \ldots, \alpha_d)$ be a nonnegative d-dimensional integer multi-index. A monomial in the indeterminates $x_1, \ldots, x_d$ is the power product $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. A *model basis* is a collection of distinct monomials $\{x^\alpha, \alpha \in L\}$, where L is a finite set of multi-indices. By combining linearly monomials in L we form polynomials:

$$\eta_L(x) = \sum_{\alpha \in L} \theta_\alpha x^\alpha,$$

where θ_α are real coefficients. The polynomial $\eta_L(x)$ is a candidate for interpolation or statistical modelling.

This paper is concerned with the following concept.

Definition 1 Let L be a model basis with n elements and let $w = (w_1, \ldots, w_d)$ be a collection of non-negative weights with $\sum_{i=1}^{d} w_i = 1$. We define the weighted linear aberration of L as

$$A(w, L) = \frac{1}{n} \sum_{(\alpha_1,\ldots,\alpha_d)\in L} \sum_{i=1}^{d} w_i \alpha_i. \tag{1}$$

The weight vector in Definition 1 regulates preference of variables for inclusion in the model. For instance, if all the components of w are similar, then it shows even preference of variables. Preference of a variable or group of variables over the remaining variables occurs when the respective components of w are bigger than those for remaining variables.

We are interested in studying aberration for models identifiable by an experimental design and along this paper, we compare models and designs of the same size n.

Definition 2 An experimental design D, of sample size $n = |D|$, is a set of points in $\mathbb{R}^d$.

We say that a model basis L with cardinality $|L| = n$ is identifiable by D if the design model matrix $X = [x^\alpha]_{x \in D, \alpha \in L}$ is invertible.

The term aberration is used to acknowledge the work on "minimum aberration" for regular fractional factorial designs of Wu and others (see Fries and Hunter 1980;

Wu and Wu 2002). For fractional factorial designs, the notion of estimation capacity is related to the ability of a design to identify models of low degree (see Cheng and Mukerjee 1998; Chen and Cheng 2004). We do not make a direct mathematical comparison with that work but simply point to a common motivation.

In Sect. 2, we review the basic ideas on algebraic identifiability. The search for identifiable models is driven by a divisibility condition, which makes the search problem tractable. We then introduce the *state polytope*, whose vertices correspond to the models identified using the algebra. In Sect. 3 we study aberration. The basic ideas on aberration are closely linked with the algebraic work on corner cut models and state polytopes in Onn and Sturmfels (1999). We are specially interested in obtaining minimal values for aberration for which we establish upper and lower bounds. An approximate approach to minimal aberration is discussed. In Sect. 4, we discuss various examples. In Sect. 5, we discuss possible extensions of the theory and, by example, a connection with the notion of aberration by Wu and others is discussed.

2 The G-basis method and the state polytope

The aberration $A(w, L)$ has remarkable connections with the algebraic method in experimental design introduced by Pistone and Wynn (1996) and developed in the monograph Pistone et al. (2001) and the joint work of Onn and Sturmfels (1999). In this section, we present the basic ideas on identifiability using algebraic techniques.

Let the set of all monomials in d indeterminates be $T^d = \{x^\alpha, \alpha \in \mathbb{Z}^d_{\geq 0}\}$, where $\mathbb{Z}_{\geq 0}$ is the set of non-negative integers and $\mathbb{Z}^d_{\geq 0}$ is the set of all vectors in d-dimensions and with entries in $\mathbb{Z}_{\geq 0}$. A polynomial is a finite linear combination of monomials in T^d with real coefficients. The set of all polynomials is denoted as $\mathbb{R}[x_1, \dots, x_d]$. It has the structure of a ring with the usual operations of sum and product of polynomials.

A term ordering $\succ$ on $\mathbb{R}[x_1, \dots, x_d]$ is a total ordering on T^d such that (1) $x^\alpha \succ 1$ for all $x^\alpha \in T^d, \alpha \neq (0, \dots, 0)$ and (2) for all $x^\alpha, x^\beta, x^\gamma \in T^d$ if $x^\alpha \succ x^\beta$ then $x^\alpha x^\gamma \succ x^\beta x^\gamma$. The leading term of a polynomial is the largest term with non-zero coefficient with respect to $\succ$. For a polynomial $f \in \mathbb{R}[x_1, \dots, x_d]$, we write its leading term as $\mathrm{LT}_\succ(f)$.

A partial order on T^d is defined by a vector $w \in \mathbb{R}^d_{\geq 0}$ as $x^\alpha \succeq_w x^\beta$ if $w^T\alpha \geq w^T\beta$, where w^T is the transposed vector of w, and $x^\alpha, x^\beta \in T^d$. Under some conditions on w (see Babson et al. 2003; Cox et al. 1997) this defines a term order. Given a term order $\succ$, there are w such that $x^\alpha \succ x^\beta$ if and only if $x^\alpha \succeq_w x^\beta$.

A design D, considered as a zero-dimensional variety gives rise to a *design ideal*, $I(D)$, which is the set of all polynomials which have zeros at all the points of D. We have that $I(D) \subset \mathbb{R}[x_1, \dots, x_d]$. The polynomial ideal I is generated by the set of polynomials $G = \{g_1, \dots, g_s\}$ if $I = \{\sum_{i=1}^s f_i g_i : f_i \in \mathbb{R}[x_1, \dots, x_d]\}$ and we write $I = \langle g_1, \dots, g_s \rangle$.

An important set of generators for the design ideal is the Gröbner basis. Gröbner bases were introduced by Buchberger (1966) and they have become a powerful computational tool in many fields (Cox et al. 1997, 2005). A Gröbner basis of $I(D)$ with respect to a term order $\succ$ is a finite subset $G_\succ(D) \subset I(D)$ such that $\langle \mathrm{LT}_\succ(g) : g \in G_\succ(D) \rangle = \langle \mathrm{LT}_\succ(f) : f \in I(D) \rangle$. The computation of Gröbner bases is implemented in standard computer programs such as CoCoA, Singular or Maple (see CoCoA Team 2007; Greuel et al. 2005; Monagan et al. 2005).

Two polynomials f and g in $\mathbb{R}[x_1, \ldots, x_d]$ are equivalent with respect to $I(D)$ if the following equivalent conditions hold:

(1) $f - g \in I(D)$
(2) $f(x) = g(x)$ for all $x \in D$

Given a term ordering $\succ$, the quotient ring $\mathbb{R}[x_1, \ldots, x_d]/I(D)$ has a unique $\mathbb{R}$-vector space basis given by the monomials in T^d that cannot be divided by the leading terms of the polynomials in $G_\succ(D)$ for $I(D)$. The monomial basis so obtained, or equivalently, the set of its exponents $L = L(D, \succ)$, has a *staircase* (also *echelon*, *order ideal*) property: for $\alpha \in L$, if $\beta \leq \alpha$ componentwise, then $\beta \in L$. Equivalently we say that for any $x^\alpha \in L$, if x^β divides x^α then $x^\beta \in L$. We call bases which have a staircase structure *staircase* models. The dimension of $\mathbb{R}[x_1, \ldots, x_d]/I(D)$ as $\mathbb{R}$-vector space is n, see Pistone and Wynn (1996), i.e. the number of points in D and of multi-indices in L is n.

Example 1 Consider the design $D = \{(0, 0), (1, 0), (0, 1), (-1, 1), (1, -1)\}$ and its design ideal $I(D)$. For a term ordering in which $x_1 \succ x_2$, consider the set of polynomials $G = \{x_1^2 + 2x_1x_2 + x_2^2 - x_1 - x_2, x_2^3 - x_2, x_1x_2^2 - x_1x_2 - x_2^2 + x_2\} \subset I(D)$. The monomial ideal generated by the leading terms of G, $\langle x_1^2, x_2^3, x_1x_2^2\rangle$, equals the ideal of leading terms $\langle \mathrm{LT}_\succ(f) : f \in I(D)\rangle$, i.e. G is a Gröbner basis for $I(D)$. The monomial basis is given by the following monomials $1, x_1, x_2, x_1x_2, x_2^2$ which are not divisible by leading terms of G, and we have its exponent set $L = \{(0, 0), (1, 0), (0, 1), (1, 1), (0, 2)\}$.

For a given basis of the quotient ring with exponents in L and a set of real values (data) Y_x, $x \in D$, there exists a unique interpolator $\eta_L(x)$ such that $Y_x = \eta_L(x)$, $x \in D$. Other non-saturated statistical sub-models can be constructed from subsets of L (see Holliday et al. 1999; Peixoto 1987).

Definition 3 The algebraic fan of D is $\mathcal{L}_a(D) = \{L(D, \succ)$, where $\succ$ is a term ordering in $\mathbb{R}[x_1, \ldots, x_d]\}$. This is the collection of staircases $L(D, \succ)$ arising from a fixed design D by varying all monomial orderings.

The algebraic fan of a design was proposed by Caboara et al. (1997), constructing upon the algebraic fan of an ideal of Mora and Robbiano (1988). Babson et al. (2003) proposed a polynomial time algorithm to compute $\mathcal{L}_a(D)$. They compute an efficient set of weight vectors and perform a change of basis which stems from the so-called FGLM algorithm (see Faugère et al. 1993). In Sect. 3.1, an algorithm is presented to identify a model in the algebraic fan using a weight vector.

It is important to note that not all staircase models identified by D are in $\mathcal{L}_a(D)$. The set of all identifiable staircase models for a design D is denoted as $\mathcal{L}_s(D)$. In fact the algebraic fan is small relative to $\mathcal{L}_s(D)$, that is $\mathcal{L}_a(D) \subseteq \mathcal{L}_s(D)$, see Chapter 6 in the unpublished Ph.D. thesis by Maruri-Aguilar (2007) and Sect. 4 in Pistone et al. (2008).

We now establish the link between the algebraic fan of a design and the state polytope of the design ideal. For a model basis $L = \{\alpha_1, \ldots, \alpha_n\}$, $\alpha_i \in \mathbb{Z}^d_{\geq 0}$ define

$$\overline{\alpha}_L = \sum_{\alpha_i \in L} \alpha_i .$$

This vector appears in the definition of $A(w, L)$ and we can write $A(w, L) = (w^T \overline{\alpha}_L)/n$. The set all such vectors over $\mathcal{L}_a(D)$ gives the state polytope.

Definition 4 The state polytope $S(D)$ of a design D, or equivalently of the design ideal $I(D)$ is the convex hull

$$\mathcal{S}(D) := \text{conv}\left(\{\overline{\alpha}_L : L \text{ is a staircase in } \mathcal{L}_a(D)\}\right) .$$

The following theorem (Sturmfels 1996, Chap. 2) summarizes the connection between the state polytope and the set of models $\mathcal{L}_a(D)$, i.e. the relation between a design and its algebraic fan.

Theorem 1 *Let D be a design and let $S(D)$ be its state polytope. Then the set of vertices of the state polytope of D is in one to one correspondence with the algebraic fan of D.*

The state polytope does not only contain information concerning models in the algebraic fan of a design, but it also provides information about the term ordering vectors needed to construct it. We recall that a d-dimensional polytope is a bounded subset of $\mathbb{R}^d$, which corresponds to the solutions of a system of linear inequalities. The *normal cone* of a face of a polytope is the relatively open cone of those vectors in $\mathbb{R}^d$ uniquely minimised over the face of the polytope. The *normal fan* of a polytope is the collection of all the normal cones of the polytope.

Two ordering vectors w and w' are said to be equivalent (modulo $I(D)$) if $L(D, \succ_w) = L(D, \succ_{w'})$. The normal fan of the state polytope partitions $\mathbb{R}^d_{\geq 0}$ into equivalence classes of ordering vectors (see Babson et al. 2003; Fukuda et al. 2007; Sturmfels 1996). Indeed every vertex of $S(D)$ corresponds to a model in $\mathcal{L}_a(D)$. Moreover, the interior of the normal cone of a vertex in $S(D)$ contains those vectors w which correspond to the same equivalence class.

We motivate Theorem 2 below with a simple example. The black dots in Fig. 1 give a 5-point design in two-dimensions, D. They also give the set of exponents L obtained for any term ordering, indeed the size of the algebraic fan of D is one. The triangles represent the exponents of the leading terms of the Gröbner basis: (2, 0), (1, 2), (0, 3). The line separates the model exponents, L, from these leading terms. This is an example of a *corner cut* model. Note that equivalently the line separates L from its complement in $\mathbb{Z}^2_{\geq 0}$.

Definition 5 A model L, of size $|L| = n$, is said to be a corner cut model if there is a $(d-1)$-dimensional hyperplane separating L from its complement $\mathbb{Z}^d_{\geq 0} \setminus L$.

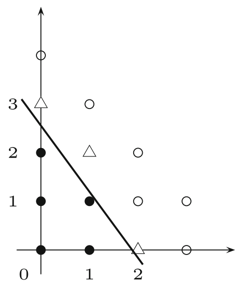


Fig. 1 Corner cut and separating hyperplane

Not all staircases are corner cuts, for example $L = \{(0,0), (1,0), (0,1), (1,1)\}$ is a staircase that cannot be separated by a hyperplane from its complement in $\mathbb{Z}^2_{\geq 0}$.

The set of exponents of a corner cut model is referred to as a corner cut staircase or simply, as a corner cut. Corner cuts were introduced by Onn and Sturmfels (1999). A generating function for the number of bi-dimensional corner cuts is given in Corteel et al. (1999), while the order of the cardinality of the set of corner cuts is proven bounded by $(n \log n)^{d-1}$ in Wagner (2002). A special class of designs is composed with those designs that identify all corner cut models of a given size.

Definition 6 A design $D \subset \mathbb{R}^d$ comprised of n distinct points is said to be generic if all corner cut models of size $n = |D|$ are identifiable.

A special polytope is constructed with the exponents for corner cut models. It will be used to compute the algebraic fan of generic designs.

Definition 7 The corner cut polytope is $CC(n,d) := \text{conv}(\{\overline{\alpha}_L : L$ is a corner cut staircase in d-dimensions and of size $n\})$.

For a discussion on the properties of bi-dimensional corner cut polytopes see the paper by Müller (2003). The algebraic fan of generic designs corresponds to the set of corner cut models, as stated in the following theorem.

Theorem 2 (Onn and Sturmfels 1999) *Let $D \subset \mathbb{R}^d$ be a generic design with n points. Then*

(1) *$S(D) = CC(n,d)$ and*
(2) *the algebraic fan of D is the set of corner cut models in d dimensions and with n elements.*

We remark that the corner cut polytope is an invariant object for the class of all the ideals generated by generic designs with the same sample size n and number of factors d and all generic designs have the same state polytope.

3 Minimal linear aberration

An important feature of the state polytope is that its vertices are automatically “lower” vertices in the sense of convexity. State polytopes relate directly to models with minimal linear aberration. In Sect. 3.1, an algorithm to compute a models of minimal aberration is presented.

Theorem 3 *Given a design $D \subset \mathbb{R}^d$ with n distinct points and a weight vector $w \in \mathbb{R}^d_{>0}$, there is at least one vertex $\alpha^* \in S(D)$ which minimises $A(w, L)$ over all identifiable staircase models $\mathcal{L}_s(D)$, that is*

$$\frac{1}{n}(w^T\alpha^*) = A(w, L^*) = \min_{L\in\mathcal{L}_s(D)} A(w, L)$$

for all L^* such that $\overline{\alpha}_{L^*} = \alpha^*$. Moreover, given a vertex of $S(D)$, there is at least one $w^* \in \mathbb{R}^d_{>0}$ such that this vertex (model) minimizes $A(w, L)$, that is,

$$A(w^*, \overline{L}) = \min_{w\in\mathbb{R}^d_{>0}} A(w, L)$$

for $\bar{L}$ such that $\overline{\alpha}_{\overline{L}} = \overline{\alpha}_L$.

Proof First, for given w we minimise $w^T\overline{\alpha}_L$ for $L \in \mathcal{L}_a(D)$, which is a finite set, see Mora and Robbiano (1988). The $\overline{\alpha}_L$ for $L \in \mathcal{L}_a(D)$ are vertices of $S(D)$ by definition. Furthermore, because we restrict L to the algebraic fan of D, vectors $\overline{\alpha}_L$ can only be aligned when they are vertexes of a facet of $S(D)$, i.e. they cannot be interior points, see (Sturmfels 1996, Chap. 2). For the second claim, it is sufficient to take a vector w_L in the interior of a normal cone for $\overline{\alpha}_L$. By definition, $A(w, L)$ is minimised for vectors on the interior of the normal cone. □

Theorem 4 follows directly from Theorem 3.

Theorem 4 *For every weight vector w there is a design $D \subset \mathbb{R}^d$ which minimizes $A(w, L)$, among all designs with sample size n and identifiable staircases.*

This is stated compactly as:

$$A^*(w, n) = \min_{D:|D|=n} \min_{L\in\mathcal{L}_a(D)} A(w, L)$$

is achieved for a generic design, i.e. there exists a design (generic design) that achieves this minima at every vertex of its state polytope. In other words, if a design is generic then automatically its algebraic fan contains models of minimal aberration.

3.1 Computation of the minimal aberration model

The model minimizing linear aberration can be found by a *greedy algorithm*. Let D be a design; let w be a fixed weight vector in $\mathbb{R}^d_{>0}$ and let Γ be the following set of potential exponents

$$\Gamma := \left\{\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}^d_{\geq 0} : \prod_{i=1}^{d}(\alpha_i + 1) \leq n\right\}.$$

The set Γ contains all staircase models with n terms, see Babson et al. (2003). Now define the *weight* of $\alpha \in \Gamma$ to be $\omega(\alpha) := \frac{1}{n}\sum_{i=1}^{d} w_i\alpha_i = (w^T\alpha)/n$. Order the vectors in Γ by their weight $\omega(\cdot)$ in increasing order, that is, index them as $\alpha^1, \ldots, \alpha^{|\Gamma|}$ such that $\omega(\alpha^1) \leq \cdots \leq \omega(\alpha^{|\Gamma|})$, where $|\Gamma|$ is the cardinality of Γ. Then the set $L \subseteq \Gamma$ with the first n terms of Γ which are identifiable by D has minimum aberration.

The model basis L is constructed by the following procedure, which can be seen as a sequential method for constructing the design-model matrix X: initialize $L := \emptyset$; while $|L| < n$, find α^i of smallest index with respect to $\omega(\cdot)$ such that the column vectors x^α, $\alpha \in L \cup \{\alpha^i\}$, $x \in D$, are linearly independent; update $L := L \cup \{\alpha^i\}$ and repeat until $|L| = n$. We have the following theorem.

Theorem 5 *Let $D \subset \mathbb{R}^d$ be a design; let w be a fixed weight vector with positive entries and let L be the model basis constructed by the greedy algorithm. Then L belongs to the algebraic fan of the design.*

Example 2 Consider the design $D = \{(0, 0), (1, 0), (0, 1), (-1, 1)\}$ and the weight vector $w = (4, 1)$. The set of potential exponents, Γ contains 8 elements, which are sorted out using the weight function $\omega(\cdot)$ as

$$\begin{array}{llllllllll} \Gamma & = \{(0,0), & (0,1), & (0,2), & (0,3), & (1,0), & (1,1), & (2,0), & (3,0)\} \\ n\omega(\cdot) = & \quad 0 & 1 & 2 & 3 & 4 & 5 & 8 & 12 \end{array}$$

The first 4 elements in Γ such that their design columns are linearly independent are $L = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Thus the set L of minimal linear aberration corresponds to the model with terms $\{1, x_1, x_2, x_1x_2\}$.

3.2 Examples

We can compare different designs using aberration as long as they have the same number of factors d and the number of points n. For a design D, the state polyhedron of D is obtained by (Minkowski) addition of $\mathbb{R}^d_{\geq 0}$ to the state polytope $S(D)$ (see Babson et al. 2003). The state polyhedron yields the same information as the state polytope. Indeed the normal fan of the (negative) state polyhedron yields automatically the first orthant (see Fukuda et al. 2007).

Example 3 Consider a central composite design (CCD by Box and Wilson 1951) with two factors, one observation at the origin and axial distance $\alpha = \sqrt{2}$. The CCD has nine runs and its algebraic fan contains exactly two models, namely

$$\left\{1, x_1, x_1^2, x_1^3, x_1^4, x_2, x_1x_2, x_1^2x_2, x_2^2\right\} \tag{2}$$

together with the model obtained by permuting the roles of x_1 and x_2. Let L_1 be the set of exponents of the model support in Eq. (2). Clearly, $\overline{\alpha}_{L_1} = (13, 5)$ and the state polytope for the design ideal of the CCD is conv$(\{(13, 5), (5, 13)\})$, see left graph of Fig. 2. Now consider a generic design with the same number of runs as the CCD. In Corteel et al. (1999) and Onn and Sturmfels (1999), it is shown that there are 12

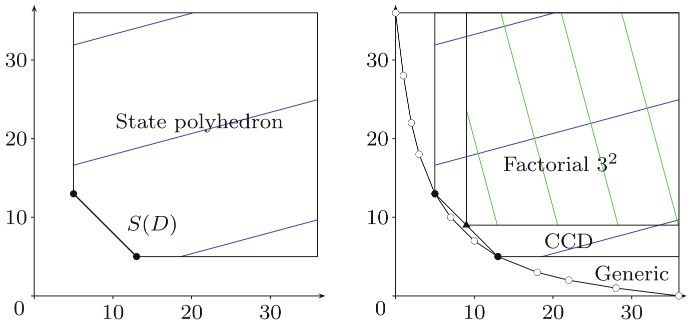
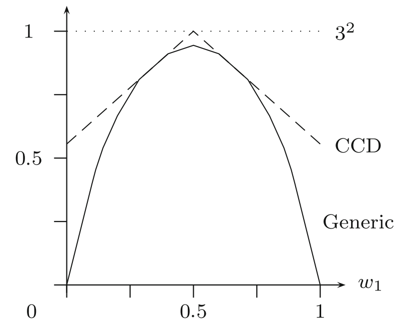


Fig. 2 The left graph depicts $S(D)$ and the state polyhedron for the CCD of Example 3. The right graph shows state polyhedra for the three designs of Example 3. The empty dots correspond to vertexes/models identified by the generic design only, while the triangle is for the sole model in the algebraic fan of the 3^2 design

Fig. 3 Minimal aberration for three designs in two factors and nine runs, see Example 3



corner cut models for $d = 2$ and $n = 9$. By Theorem 2, the algebraic fan of the generic design contains all the 12 corner cut models, including those in the algebraic fan of the CCD. We consider also a full factorial design 3^2, which identifies only the model with support $\{1, x_1, x_1^2\} \otimes \{1, x_2, x_2^2\}$, where $\otimes$ is the Kronecker product. Its state polytope is the point (9, 9). In the right graph of Fig. 2, we depict the state polyhedra for the three designs and in Fig. 3, we plot $\min_{L \in \mathcal{L}_a(D)} A(w, L)$ for $w = (w_1, w_2) \in [0, 1]^2$ and $w_1 + w_2 = 1$. For the CCD, this is

$$\begin{cases} \left((w_1, 1 - w_1)(13, 5)^T\right)/9 = (8w_1 + 5)/9 & \text{if } w_1 \leq 1/2 \\ \left((w_1, 1 - w_1)(5, 13)^T\right)/9 = (-8w_1 + 13)/9 & \text{if } w_1 > 1/2 \end{cases}$$

For the generic design, the aberration curve is a piecewise linear function with 12 segments. Finally, the aberration for the design 3^2 is constant for all weights. As expected, the aberration takes its minimum value for the generic design, over all possible weights.

Example 4 Consider the design $D = \{(0, 0), (1, 1), (2, 2), (3, 4), (5, 7), (11, 13), (\alpha, \beta)\}$, where $(\alpha, \beta) \approx (1.82997, 1.82448)$ is the only real solution of a system

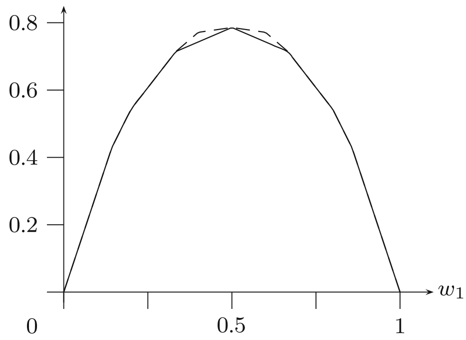


Fig. 4 Minimal aberration for G (*solid line*) and D (*dashed line*), see Example 4

of polynomial equations (see Onn and Sturmfels 1999, p. 47). The algebraic fan of the above design has ten models and its state polytope is

$$\text{conv}(\{(21, 0), (15, 1), (11, 2), (9, 3), (6, 5), (5, 6), (3, 9), (2, 11), (1, 15), (0, 21)\}).$$

Now consider a generic design G with the same number of runs and factors. The algebraic fan of G is the set of corner cut models which for seven points in two-dimensions has eight elements, see Corteel et al. (1999) and Onn and Sturmfels (1999) and thus its state polytope is the corner cut polytope:

$$CC(7, 2) = \text{conv}(\{(21, 0), (15, 1), (11, 2), (7, 4), (4, 7), (2, 11), (1, 15), (0, 21)\}).$$

In Fig. 4, we graph the aberration for both designs as a function of w_1. Although the size of the algebraic fan of D is bigger than that for a generic design, the weighted aberration takes minimal value for the generic design for all possible weight vectors $(w_1, 1 - w_1)$.

Example 5 The aberration of some sets of multi-indices does not depend on w. For instance, consider the following sets in two dimensions

$$\begin{aligned} L_n &= \{(i, i) : i = 0, \ldots, n-1\} \\ M_n &= \{(i, j) : i, j = 0, \ldots, n-1\} \\ N_n &= \{(i, j) : 0 \leq i + j \leq n\} \end{aligned}$$

for which the aberrations are $A(w, L_n) = (n-1)/2$, $A(w, M_n) = (n-1)/2$ and $A(w, N_n) = n/3$. To properly compare the above aberrations, the sets L, M, N must have the same size. In Table 1, values m, n and $p = m^2$ are presented such that $\#L_p = \#M_m = \#N_n$ for m up to 8,000. As sample size grows, the aberration of the triangular set N_n remains smaller than for the square set M_m.

3.3 Bounds for the aberration

Although the minimal value of the aberration $A^*(w, n)$, depends on the weight vector $w = (w_1, \ldots, w_d)$, we can carry out a special normalisation which leads to bounds for

Table 1 Aberration for sets of multi-indices L_n, M_n and N_n

m	n	$A(w, L_p)$	$A(w, M_m)$	$A(w, N_n)$
1	0	0	0	0
6	8	17.5	2.5	2.6
35	49	612.0	17.0	16.3
204	288	20807.5	101.5	96.0
1,189	1,981	7.0×10^5	594.0	560.3
6,930	9,800	2.4×10^7	3,464.5	3,266.6
40,391	57,121	8.1×10^8	20,195.0	19,040.3

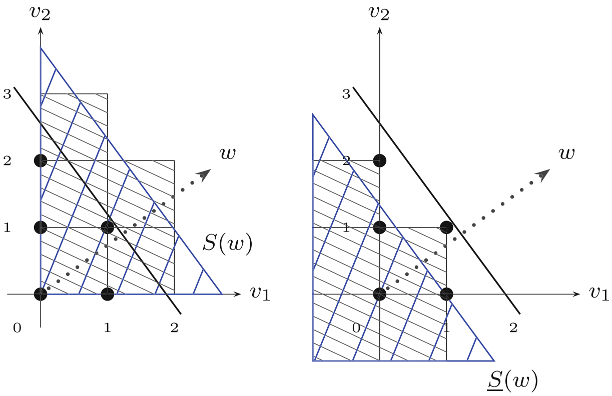


Fig. 5 Bidimensional corner cut together with upper (*left diagram*) and lower cells (*right diagram*) $\overline{Q}$ and $\underline{Q}$. In both diagrams the vector w, a separating hyperplane and equivalent simplexes $S(w)$ and $\underline{S}(w)$ were added

the minimal aberration. These bounds depend only on a simple function of the weights, surprisingly the geometric mean. Our construction is based upon the expected value of auxiliary random variables which are suitably constructed.

For the rest of this section, let $D \subset \mathbb{R}^d$ be a generic design with n points. Let w be a fixed weight vector with positive elements and let L be the corner cut model identified by w. We recall that $|L| = n$.

For an integer multindex α define its *upper cell* as the unit cube with lower vertex at α

$$\overline{c}(\alpha) = \{v \in \mathbb{R}^d : \alpha_i \le v_i \le \alpha_i + 1\}$$

and similarly the *lower cell* of α is

$$\underline{c}(\alpha) = \{v \in \mathbb{R}^d : \alpha_i - 1 \le v_i \le \alpha_i\}.$$

Define $\underline{Q} = \cup_{\alpha \in L} \underline{c}(\alpha)$ and $\overline{Q} = \cup_{\alpha \in L} \overline{c}(\alpha)$. See Fig. 5 for a depiction of lower and upper cells with L a corner cut.

Clearly, the volume of $\overline{Q}$ and of $\underline{Q}$ equals n, that is the cardinality of L. We now create a simplex $S(w) \subset \mathbb{R}^d$ which is directed by the vector w and has volume n. We call this simplex and the subset of the first orthant below it the *equivalent simplex*, which is formally $S(w) = \{v \in \mathbb{R}^d_{\geq 0} : \sum_{i=1}^d v_i w_i \leq c\}$. The volume of $S(w)$ is determined up to the constant $c > 0$. We find the value of this constant by setting the total volume of the equivalent simplex equal to n:

$$n = \frac{c^d}{d! \prod_{i=1}^d w_i},$$

giving

$$c = (nd!)^{\frac{1}{d}} g(w), \tag{3}$$

where

$$g(w) = \left(\prod_{i=1}^d w_i \right)^{\frac{1}{d}}$$

is the geometric mean of the components of the weight vector w. We call $H(w)$ an hyperplane, orthogonal to w, which limits the equivalent simplex, that is $H(w) = \{v \in \mathbb{R}^d_{\geq 0} : \sum_{i=1}^d v_i w_i = c\}$.

The expected value of a random variable with uniform support over $S(w)$ will be used now to compute bounds for aberration. We can compute a notional value of A, the linear aberration as the expectation $A(w, S(w)) = \mathrm{E}(\sum w_i X_i)$, for the random vector $(X_1, \ldots, X_d)$ with uniform distribution over $S(w)$. Thus for the equivalent simplex we have that

$$A(w, S(w)) = \frac{1}{n} \frac{d}{(d+1)!} \frac{c^{d+1}}{\prod_{i=1}^d w_i} = (nd!)^{\frac{1}{d}} \frac{d}{d+1} g(w), \tag{4}$$

after substituting Eq. (3) in $A(w, S(w))$.

We observe that the region $\underline{Q}$ is obtained from $\overline{Q}$ by a negative shift $(-1, \ldots, -1)$. As before, we consider a random vector with joint uniform distribution over $\underline{Q}$. We then use the expected value of $\sum w_i X_i$ as the aberration $A(w, \underline{Q})$. Analogously, we define $A(w, \overline{Q})$ and we have

$$A(w, \underline{Q}) = A(w, \overline{Q}) - 1.$$

Similarly, we can create a region $\underline{S}(w)$ by the same downward shift, and we have

$$A(w, \underline{S}(w)) = A(w, S(w)) - 1.$$

As D is generic and thus L is a corner cut there exist cutting hyperplanes separating L from its complement in $\mathbb{Z}^d_{\geq 0}$. Moreover if w is in the interior of the normal cone of the corner cut polytope, then we can select a cutting hyperplane H which is orthogonal to w and thus parallel to $H(w)$ (see Onn and Sturmfels 1999).

Example 6 Consider $D = \{(0, 0), (1, 2), (2, 1)\}$, which is a generic design with $d = 2, n = 3$. Take $L = \{(0, 0), (1, 0), (2, 0)\}$ and $w = (1, 2)$. The weight vector w is not in the interior of a normal cone of the corner cut polytope $CC(2, 3) = \mathrm{conv}(\{(3, 0), (1, 1), (0, 3)\})$. Indeed w is on the boundary of the normal cone separating L from the corner cut model $\{(0, 0), (1, 0), (0, 1)\}$, i.e. none of the hyperplanes perpendicular to w is a cutting hyperplane for L.

By a simple argument the simplex S_H with faces $x_i = 0,\ (i = 1, \ldots, d)$ and H lies wholly within the upper quadrant region $\overline{Q}$ because otherwise, the cutting hyperplane hypothesis for H would be violated and thus S_H has volume less than n. Recall that the equivalent simplex $S(w)$ has volume n.

There is one additional argument that leads to our first inequality. Since the region $\overline{Q}$ and the equivalent simplex $S(w)$ have the same volume n, it must be that $\overline{Q}$ protrudes beyond $S(w)$. Equivalently we may move mass from $\overline{Q}$ inside $S(w)$. As this mass occurs orthogonally to w, we claim that this movement diminishes the aberration, thus

$$A(w, S(w)) \leq A(w, \overline{Q}).$$

This property is also inherited by the downward shifted version, and we have $A(w, \underline{S}(w)) \leq A(w, \underline{Q})$. The same argument, based on $\underline{Q}$ being below $S(w)$ shows the middle inequality in the following sequence:

$$A(w, \underline{S}(w)) \leq A(w, \underline{Q}) \leq A(w, S(w)) \leq A(w, \overline{Q}).$$

By Theorem 4, as the design is generic and L is the model identified by w, clearly we have

$$A(w, \underline{Q}) \leq A^*(w, n) \leq A(w, \overline{Q}).$$

By comparing continuous mass to point masses of the model L, we see that $A(w, \overline{Q}) \leq A(w, S(w)) + 1$. Collecting the above inequalities, we have our result.

Theorem 6 *Let $D \subset \mathbb{R}^d$ be a generic design with n points; let $w \in \mathbb{R}^d$ be a vector of positive weights. Then the minimal aberration $A^*(w, n)$ satisfies the bounds*

$$A(w, S(w)) - 1 \leq A^*(w, n) \leq A(w, S(w)) + 1, \tag{5}$$

where $A(w, S(w))$ is computed in Eq. (4).

There are various kinds of asymptotic that this formula leads to. From the inequality between geometric and arithmetic mean we have $g(w) \leq \frac{1}{d}$. This suggests the condition $\lim_{d\mapsto\infty} g(w) = \beta/d$ for some constant $0 \leq \beta \leq 1$. Now for $w_i = (1 + \delta_i)/d$,

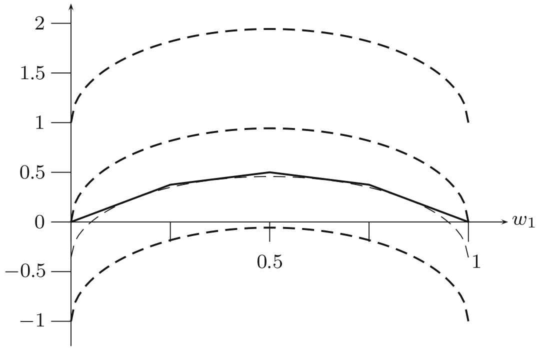


Fig. 6 Minimal aberration $A^*(w, n)$ (*solid line*) for a generic design with $d = 2, n = 4$; bounds $A(w, S(w))$ and $A(w, S(w)) \pm 1$ of Theorem 6 (*dashed lines*). We also show approximate aberration $\tilde{A}$ using Theorem 7 (*thin dashed line*)

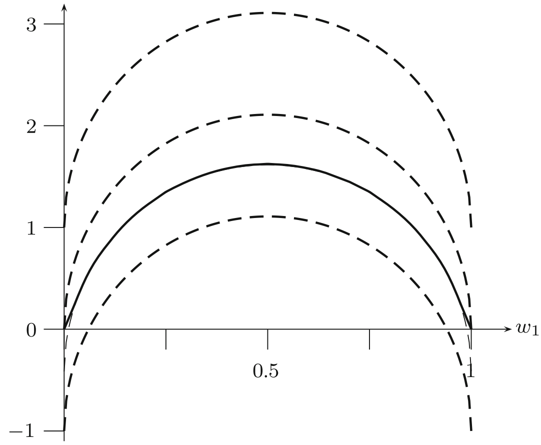


Fig. 7 Minimal aberration $A^*(w, n)$ (*solid line*) for a generic design with $d = 2, n = 20$; bounds $A(w, S(w))$ and $A(w, S(w)) \pm 1$ and (*dashed lines*) of Theorem 6. The figure also shows approximate aberration $\tilde{A}$ of Theorem 7 (*thin dashed line*) which almost overlaps the *solid line*

with $\sum \delta_i = 0$, and assuming convergence of $\sum \delta_i^2$ and $n = k^d, k > 0$, we use Stirling's approximation to obtain

$$\lim_{d \mapsto \infty} A^*(w, n) = \frac{k\beta}{e}.$$

Such limits may be considered as asymptotic identifiability rates, analogous to the more familiar Nyquist rates in Fourier analysis.

Example 7 For small d and n the bounds of Eq. (5) are rather coarse. Figure 6 shows the bounds $A(w, S(w)) \pm 1$ of Theorem 6 together with the minimal aberration $A^*(w, n)$, plotted as function of w_1 for $d = 2$ and $n = 4$. Notice that, as function of w, the minimal aberration $A^*(w, n)$ is a piece-wise linear graph (this is a general fact, consequence of Definition 1), each segment corresponding to a different vertex (different corner cut) of the corner cut polytope. Figures 7 and 8 give the bounds and minimal aberration for $n = 20$ and $n = 100$. In Figs. 6, 7 and 8, we also added a curve for the approximate aberration which is presented in Theorem 7 below.

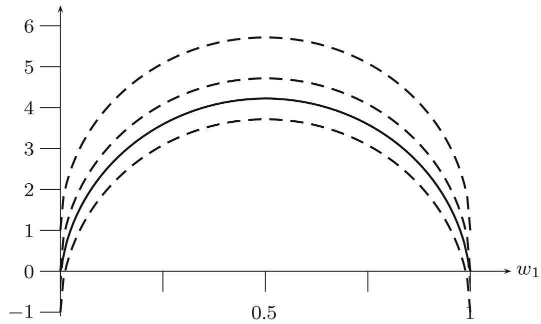


Fig. 8 Minimal aberration $A^*(w, n)$ (*solid line*) for a generic design with $d = 2, n = 100$; bounds $A(w, S(w))$ and $A(w, S(s)) \pm 1$ (*dashed lines*). The approximate aberration $\tilde{A}$ of Eq. (8) (*thin dashed line*) is also plotted, but is undistinguishable from the minimal aberration

3.4 Approximated state polytope for generic designs

Note that as w changes the hyperplanes $H(w)$ are tangent to the surface defined by

$$\prod_{i=1}^{d} x_i = c^d = \frac{nd!}{d^d}$$

and the (normalised) centroids of the equivalent simplices lie on the surface defined by

$$\prod_{i=1}^{d} x_i = b^+ = \frac{nd!}{(d+1)^d}. \tag{6}$$

We can solve an equivalent optimisation problem to the computations of $A(w, S(w))$ in terms of the tangent surfaces: for all centroids lying above or on the surface of Eq. (6), the minimum value of $A(w, S(w))$ is achieved at the centroid of the tangent.

In the above argument, we are essentially using the surface inEq. (6) to approximate the lower border of the state polytope for a generic design, i.e. the lower border of the corner cut polytope. In order to improve the bounds given in Theorem 6, it seems natural simply to take a surface defined by

$$\prod_{i=1}^{d} (x_i + a) = b \tag{7}$$

with fixed a, b. In Theorem 6, we have set $a = \pm 1$ and $b = b^+$ of Eq. (6). In Appendix B we discuss an approach to select the values a, b to obtain a good approximation of the corner cut polytope. The following theorem estimates minimal aberration for generic designs using the approximation of Eq. (7). The proof is given in Appendix A.

Theorem 7 *Let $D \subset \mathbb{R}^d$ be a generic design with n points; let $w = (w_1, \ldots, w_d)$ be a fixed positive weight vector with $\sum_{i=1}^{d} w_i = 1$. Let the state polytope of $I(D)$ be approximated by Eq. (7). Then the value*

$$\tilde{A}(w) = db^{1/d} g(w) - a \tag{8}$$

is an approximation of $A^*(w, n)$.

We recall that $g(w)$ is the geometrical mean of the components in w. Figures 6, 7 and 8 give examples ($d = 2$ factors, $n = 4, 20, 100$) of the minimal aberration $\tilde{A}(w)$ in Theorem 7. The values a, b for each case were selected using the technique in Appendix B.

4 Examples

In this section, we discuss through extended examples other possible uses of the ideas on generic designs and aberration. In Sect. 4.1, we explore and conjecture the existence of generic designs over Latin hypercubes for all factors and sample sizes. In Sect. 4.2, we compare fractional factorial designs through their state polytopes.

4.1 Latin hypercube design

Latin hypercube designs (LH) were first proposed by McKay et al. (1979) in the context of computer experiments. Latin hypercubes are designs with reasonable space filling properties and good projections in lower dimensions.

Theorem 4 relates minimal aberration to generic designs, i.e. if the design is generic, then it identifies models of lower weighted degree (and minimal aberration) for any weight vector w. In what follows we study LH using Definition 6 of generic designs.

The construction of a Latin hypercube design can be summarised as follows.

1. Divide the range of each factor into n equal segments.
2. Select a value in each segment using a random uniform distribution, or any other continuous distribution.
3. Randomly permute the list for each factor.

By Theorem 30 in Pistone et al. (2001), a Latin hypercube design constructed as above is generic with probability one.

We now consider a special type of LH designs. This type is constructed by selecting a fixed value in every segment in Step 2. For instance, we could select the minimum, maximum or the midpoint value for every segment. We show by example that for this type of LH designs, the probability of being generic is close, but generally not equal to one.

There are a few obvious cases of LH designs which are not generic, for example when the points of the design lie on a line. We have performed exhaustive search for a few cases of LH in two-dimensions. Our search points out to the existence of generic LH for different values of d, n. In fact for the values we tried the proportion of generic LH tends clearly to one. See Figs. 9 and 10 for a depiction of the results, where we additionally plot the proportion of *maximal fan* designs among LH, i.e. LH designs that identify all possible staircase models for given d, n. We have the following conjecture for the existence of generic LHS for any value of d, n.

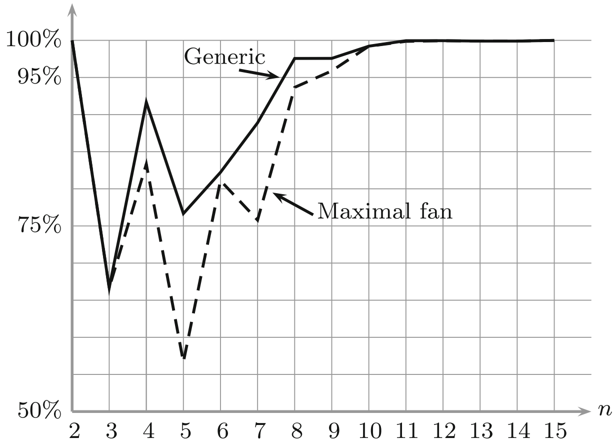


Fig. 9 Percentage of generic LHS designs for $d = 2$ and $n \leq 15$

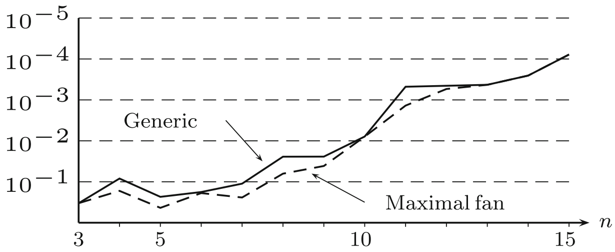


Fig. 10 Minus logarithm of the percentage of non generic LHS designs for $d = 2$ and $n \leq 15$

Conjecture 1 For every $d \geq 2$ and $n \geq 2$ there exists at least one generic LH design, constructed by setting a fixed value for every one of the n segments in the above procedure.

Experimentally we observed that when the sample size is $n = \binom{k+1}{d}$ for $k \geq 1$, the genericity of a LH design is closely linked to the identification of a model of total degree $k - 1$. For example for $k = 4$, $d = 2$, $n = 10$ there are 10! LH of which 99% are generic. Of the remaining 1% which are not generic only 6 designs (up to reflection and rotation), which are given in Fig. 11, identify the cubic model with exponent set

$$L = \{(0, 0), (1, 0), (0, 1), (2, 0), (1, 1), (0, 2), (3, 0), (2, 1), (1, 2), (0, 3)\}.$$

4.2 Orthogonal fractions

In this section, we consider some of the techniques of this paper for the class of fractional factorial designs with two levels. We first explore the relation between state

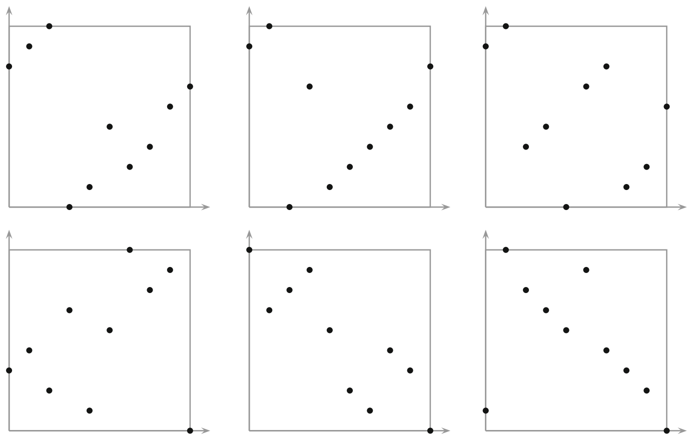


Fig. 11 LH on $[0, 1]^2$ for $d = 2$, $n = 10$ which are not generic and identify L

polyhedron and then later propose a tool to compare the identification capability of designs.

In Examples 3 and 4, we observed that in general, nesting of state polyhedra for two designs does not imply any easy relation between the algebraic fan of the designs. If instead we restrict to the family of designs with two levels then there is a clear relation between such nesting and algebraic fans. We have the following Lemma from Chapter 6 in the Ph.D. thesis by Maruri-Aguilar (2007).

Lemma 1 *Let F_1 and F_2 be two fractional factorial designs with two levels and let S_1 and S_2 be their corresponding state polyhedra of $I(F_1)$, $I(F_2)$. Then the nesting of state polyhedra $S_1 \subset S_2$ implies nesting of algebraic fans $\mathcal{L}_a(F_1) \subset \mathcal{L}_a(F_2)$.*

The following example is based upon Lemma 1 and presents an interesting relation between resolution and identifiability. That is, bigger resolution points to more models in the algebraic fan.

Example 8 Let F_1 and F_2 be the 2^{4-1}_{IV} and 2^{4-1}_{III} fractional factorial designs with eight runs in four factors and respective generators $x_1x_2x_3x_4 - 1 = 0$ and $x_1x_2x_3 - 1 = 0$. The subindices III, IV refer to the *resolution* of the fraction, see Box and Hunter (1961a,b). Their corresponding state polyhedra are nested, i.e. $S(F_2) \subset S(F_1)$ and by direct computation, we confirm that the algebraic fans are also nested. The algebraic fan $\mathcal{L}_a(F_2)$ has four models, while $\mathcal{L}_a(F_1)$ includes 12 elements.

For fractional factorial designs, the estimation of interactions in a design was related to the resolution of the design through the property termed *hidden projection* (see Evangelaras and Koukouvinos 2006; Wang and Wu 1995). We conjecture the nesting of algebraic fans of two designs 2^{k-p} with different resolution. However, exploiting this nesting property of fans to compare designs using *aberration* might need additional considerations.

Example 9 Let F_1, F_2 be the fractions 2^{7-2}_{IV} given by generators $x_6 - x_1x_2x_3 = 0, x_7 - x_2x_3x_4 = 0$ and $x_6 - x_1x_2x_3x_4 = 0, x_7 - x_1x_2x_3x_5 = 0$, respectively.

Although both fractions have the same resolution, the fraction F_2 corresponds to a *minimum aberration design* using the definition of Fries and Hunter (1980). The state polyhedron $S(F_1)$ has 133 vertices while $S(F_2)$ has 1,708. There is no nesting of the state polyhedra and $\mathcal{L}_a(F_1) \cap \mathcal{L}_a(F_2) \neq \emptyset$.

A proposal to compare two designs D_1, D_2 of the same size through their state polytopes is to map the vertices of the state polytopes $S(D_1), S(D_2)$ with a function $f : \mathbb{R}^d \to \mathbb{R}$. In this way the state polytopes of D_1 and D_2 are compared by the univariate projections of their vertices. We propose a weighted sum of the vertex coordinates

$$f(v_1, \ldots, v_d) = \sum_{i=1}^{d} w_i v_i, \tag{9}$$

with positive weights $w_i > 0$. We use $w_i = 1$ for $i = 1, \ldots, d$ and thus Eq. (9) allows for direct comparison of designs based on the distribution of total degrees for models in the algebraic fan.

Example 10 (Continuation of Example 9) We transform the vertices of the state polytopes for F_1 and F_2 using Eq. (9). In Table 2 in Appendix B, we summarize the results for each fraction as the distribution of absolute and relative frequencies. Clearly, the fraction F_2 with minimum aberration for generators identifies models with a smaller total degree than that for F_1 and in that sense it has smaller linear aberration. See Fig. 12 for a histogram of the relative frequencies for F_1 and F_2.

5 Discussion

5.1 Generalised concave aberration

This paper is partly concerned with a problem of linear programming, i.e. optimising a linear function $f : \mathbb{R}^d \to \mathbb{R}$ over a convex polytope. We now discuss extensions of our work using other types of aberration. When we consider concave aberration criteria, some of our results still hold.

Consider any concave function $f : \mathbb{R}^d \to \mathbb{R}$. Now, given a model L, define its aberration by

$$A(f, L) := f\left(\sum_{\alpha \in L} \alpha_1, \ldots, \sum_{\alpha \in L} \alpha_d\right).$$

The linear aberration of Definition 1 is the special case where f is the following linear (hence concave) function,

$$\begin{aligned} f : \mathbb{R}^d &\longrightarrow \mathbb{R} \\ x = (x_1, \ldots, x_d) &\mapsto \frac{1}{n} \sum_{i=1}^{d} w_i x_i. \end{aligned}$$

Table 2 Absolute (AF) and relative (RF) frequencies of total degrees for models identified by fractions F_1 and F_2 of Example 10 and F_3 of Example 11

Total degree	AF F_1	AF F_2	AF F_3	RF F_1	RF F_2	RF F_3
58	–	–	2,290	–	–	0.84
59	–	–	5,437	–	–	1.99
60	–	–	15,036	–	–	5.51
61	–	8	34,574	–	0.47	12.66
62	–	52	55,025	–	3.04	20.15
63	–	108	57,848	–	6.32	21.18
64	–	124	47,851	–	7.26	17.52
65	–	220	28,511	–	12.88	10.44
66	–	268	13,928	–	15.7	5.1
67	–	204	6,837	–	11.94	2.5
68	72	340	3,378	54.55	19.91	1.24
69	–	60	1,596	–	3.51	0.58
70	–	136	567	–	7.96	0.21
71	–	8	140	–	0.47	0.05
72	48	144	33	36.36	8.43	0.01
73	–	–	12	–	–	0.00
74	–	20	5	–	1.17	0.00
80	12	16	–	9.09	0.94	–
Total	132	1,708	273,069	100.00	100.00	100.00

– zero

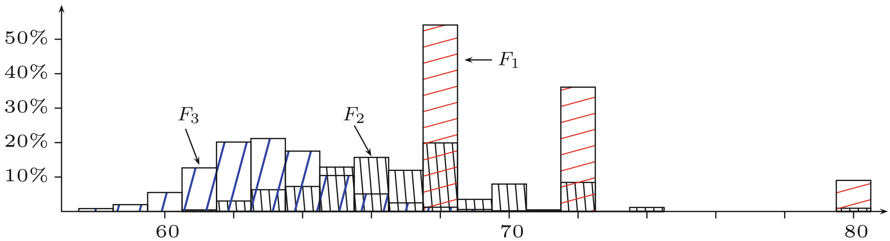


Fig. 12 Histograms of relative frequencies for fractions F_1 and F_2, see Example 10. We added F_3 of Example 11

Since we only appealed to convexity, Theorem 3 is valid when we replace $A(w, L)$ by the more general form $A(f, L)$. That is to say, the set of lower vertices of the state polytope (corresponding to models in the algebraic fan) contains the solution to minimising any concave aberration function. This can be understood as minimisation over a matroid, which was studied further in Berstein et al. (2008). A further development is to consider aberration $A(w, S(w))$ with respect to other distributions rather than the uniform.

Table 3 Design F_3 of Example 11

x_1	x_2	x_3	x_4	x_5	x_6	x_7
+	+	+	+	−	−	+
+	−	+	−	−	+	+
+	−	+	+	−	+	−
+	+	+	−	+	+	−
+	+	−	−	−	−	+
+	−	+	+	−	−	+
+	−	−	−	+	+	+
+	−	−	+	−	−	+
−	+	+	−	+	−	−
+	−	−	+	−	+	−
+	−	+	−	+	−	−
−	+	+	+	−	−	+
−	+	+	+	+	−	−
+	−	−	+	+	+	−
−	−	−	−	−	−	−
+	−	−	−	+	−	−
−	+	+	+	+	−	+
−	−	+	+	−	+	−
+	−	−	−	−	+	−
−	−	−	−	+	+	+
−	−	+	−	−	+	+
+	−	+	−	+	+	−
−	+	+	−	+	+	−
−	+	−	−	−	+	+
−	−	−	+	+	−	+
+	+	−	−	+	+	+
+	+	+	+	−	+	+
−	−	−	−	−	−	+
−	−	+	−	+	−	+
+	+	−	−	+	−	+
−	−	−	−	−	+	+
+	+	−	−	−	−	−

The signs $+$ and $-$ correspond to $+1$ and -1

5.2 Connection with aberration of Wu and others

In the statistical literature, the word aberration has been used to refer to properties of the generators for fractional factorial designs (see Chen and Hedayat 1998; Fries and Hunter 1980; Wu and Wu 2002). A topic of future research is to link minimal aberration of Definition 1 with the traditional measure based on generators for a fractional factorial design.

We conjecture that among the class of orthogonal fractions of 2^d designs there is some kind of correspondence between the minimal linear aberration of this paper and minimum generator aberration of Wu and others. If we select non-orthogonal fractions, the situation is more complex, as the next example shows.

Example 11 Let F_3 be the non-orthogonal fraction with size $n = 32$ of a 2^7 design given in Table 3 of Appendix B. We also consider the designs F_1 and F_2 of Examples 9 and 10. The three designs have the same size, but the design F_3 cannot be compared with F_1 or F_2 in traditional terms as it is not even orthogonal. However, we can compare the designs based in the distribution of degrees in their algebraic fans.

An interpolation as presented in Appendix B suggests that the minimum degree of models identified by a generic design with $n = 32, d = 7$ is $53.5 \approx 54$. This number is a lower bound for the total degree of models identified by designs F_1, F_2 and F_3. In other words, the set of total degrees for models in algebraic fan of F_1, F_2 and F_3 is lower bounded by 54, e.g. $54 \leq \min(\{\sum_{i=1}^{d} \bar{\alpha}_L : L \in \mathcal{L}_a(F_i)\})$ for $i = 1, 2, 3$.

Initial results show that

(i) the size of $\mathcal{L}_a(F_3)$ is much longer (it has around 3×10^5 models) than that for designs F_1 and F_2, see Table 2 in Appendix B;
(ii) the algebraic fans of F_1 and F_2 are not contained in the algebraic fan of F_3, and
(iii) the design F_3 identifies model of lower degree than F_1 or F_2 (indeed of total degree 58), and the bound 54 is verified.

The design F_3 has smaller minimal linear aberration than F_1 and F_2, see Fig. 12. We also note that the histogram for F_3 presents more symmetry than F_1 and F_2.

The authors appreciate that it would be very useful to relate the notations of aberration, both those in this paper and in work of other authors, to measures of complexity of models. That is to say low aberration implies low complexity in the same sense. There are more refined measures of complexity based on the topological structure of the monomial ideal $\langle \mathrm{LT}_{\succ}(f) : f \in I(D) \rangle$ (of Sect. 2), tackled in recent research by the authors. It is a challenging problem to relate this work to aberration.

Appendix A: Proof of Theorem 7

Proof We minimise $\sum_{i=1}^{d} w_i x_i$ over the first orthant, subject to the constraint $\prod_{i=1}^{d}(x_i + a) = b$. The change of coordinates $x_i' = x_i + a$ for $i = 1, \ldots, d$ turns the problem into minimisation of $\sum_{i=1}^{d} w_i x_i'$ subject to $\prod_{i=1}^{d} x_i' - b = 0$. The Lagrange multiplier $L(x', \lambda) = \sum_{i=1}^{d} w_i x_i' - \lambda(\prod_{i=1}^{d} x_i' - b)$ is constructed and the system of equations $\nabla L(x', \lambda) = 0$, $\frac{\partial L(x,\lambda)}{\partial \lambda} = 0$ is solved. The solution vector is $x^{*'} = (x_1^{*'}, \ldots, x_d^{*'})$ where

$$x_i^{*'} = b^{1/d} \frac{\prod_{i=1}^{d} w_i^{1/d}}{w_i}.$$

The convexity of the functions $\sum_{i=1}^{d} w_i x_i'$ and $\prod_{i=1}^{d} x_i' = b$ over the first orthant guarantees that $x^{*'}$ is indeed the minimum. The aberration at this minimum point is $\sum_{i=1}^{d} w_i x_i^{*'} = db^{1/d} g(w)$. The final value $\tilde{A}$ of Eq. (8) is achieved by substituting back $x_i^* = x_i^{*'} - a$. □

We remark that for a fixed w, x_i^* serves as an approximation to the centroid of the corresponding corner cut model and therefore $\tilde{A}$ is an approximation to $A^*(w, n)$. Although the approximate aberration $\tilde{A}$ does not depend on the actual corner cut identified by L, the minimal aberration $A^*(w, n)$ does depend on it. If L is the corner cut directed by w, the practical validity of the approximate aberration $\tilde{A}$ relies on x_i^* being close enough to $\frac{1}{n} \sum_{\alpha \in L} \alpha_i$. This closeness depends ultimately on a, b. See Appendix B for a proposal to compute a, b.

Appendix B: Computing values *a*, *b* for the approximate corner cut polytope

In Sect. 3.4, we proposed the continuous function of Eq. (7) to approximate the corner cut polytope (which is piecewise linear surface). In this section, we discuss on the selection of the values a, b so that the approximation is good enough. In general, the values a, b will depend on the number of dimensions d and number of points in the design n. However, for fixed d, the approximation will be coarse for small values of n.

For our approximation, we use the following properties of the corner cut polytope, which have been studied as well in Müller (2003) and Onn and Sturmfels (1999).

Lemma 2 *The corner cut polytope satisfies the following properties.*

(i) *The intersection of the corner cut polytope with the axes occurs at the point* $\binom{n}{2}$.
(ii) *When for* $k \geq 1$, *the sample size* n *satisfies*

$$n = \binom{k + d - 1}{d} \tag{10}$$

then the corner cut polytope is pointed.

Proof (i) The intersection is the the sum of exponents for any marginal model of the form $\{1, x_i, x_i^2, \ldots, x_i^{n-1}\}$. Therefore the intersection must occur at $\sum_{i=0}^{n-1} i = \binom{n}{2}$.
(ii) The corner cut polytope is pointed when the sample size is the same as the size of a model of total degree $k - 1$, that is, there are $\binom{d+1-j}{j}$ terms of degree j in the model where $j = 0, \ldots, k - 1$. Therefore the sample size must be $n = \sum_{j=0}^{k-1} \binom{d+1-j}{j} = \binom{k+d-1}{d}$.

□

Remark 1 When Eq. (10) is satisfied, the tip of the pointed corner cut polytope has coordinates $\alpha_L = \left(\binom{k+d-1}{d+1}, \ldots, \binom{k+d-1}{d+1}\right)$.

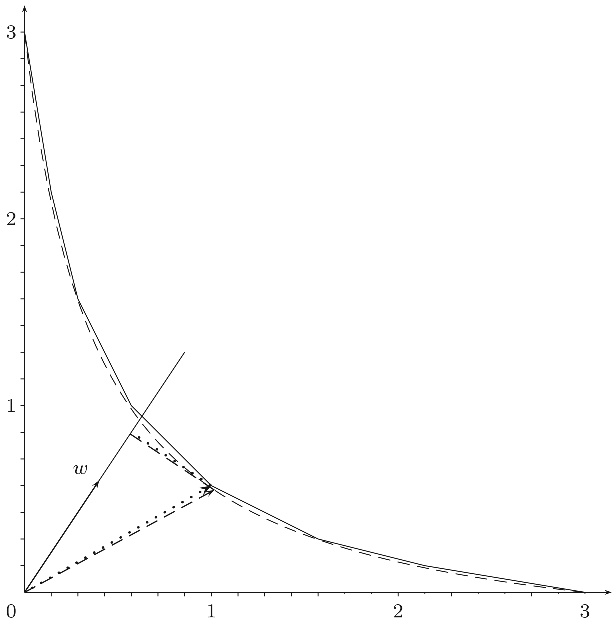


Fig. 13 Minimal aberration using the corner cut polytope. The corner cut polytope is the piecewise linear solid curve, while the approximation is the dashed curve. The minimal aberration is the projection over the direction of w of the vertex (*dotted line*), and an approximate value uses Eq. (7) (*dashed line*)

We propose to force Eq. (7) to satisfy the condition of Item 1 in Lemma 2 and pass through the tip point α_L for the model of total degree $k-1$. To summarize, when sample size satisfies Eq. (10) then a, b must satisfy the following equations:

$$b = a^{d-1}\left(\frac{n-1}{2} + a\right) \quad \text{and} \quad b = (s+a)^d,$$

where $s = \frac{1}{n}\binom{k+d-1}{d+1}$ is the scaled tip of the corner cut polytope. When design size, n, is not of the form $n = \binom{k+d-1}{d}$ for some $k \geq 1$, we propose to interpolate the value for s, the scaled tip of the polytope, that is to solve Eq. (10) for k and interpolate the corresponding tip with $\frac{1}{n}\binom{k+d-1}{d+1}$.

For two-dimensions ($d = 2$) by interpolation and solving the two conditions above we obtain the following formulæ for a, b in terms of n:

$$a = \frac{5 - 3\sqrt{1+8n} + 4n}{3(3 - 2\sqrt{1+8n} + 3n)}, \quad b = a\left(\frac{n-1}{2} + a\right).$$

See Fig. 13 for a depiction of the corner cut polytope and the approximate curve for $d = 2, n = 7$. The interpolation above is difficult to manipulate for $d > 2$ and we have to rely on approximations. The following formulæ are rough approximations for

a, b obtained by truncation of the binomial expansions

$$a \approx \left(\frac{2d!n}{(d+1)^d(n-1)}\right)^{\frac{1}{d-1}}, b = a^{d-1}\left(\frac{n-1}{2}+a\right) \approx \frac{d!n}{(d+1)^d}.$$

Acknowledgments The research of Shmuel Onn and Henry Wynn was partially supported by the Joan and Reginald Coleman-Cohen Exchange Program during a stay of Henry Wynn at the Technion-Israel Institute of Technology. Yael Berstein was supported by an Irwin and Joan Jacobs Scholarship and by a scholarship from the Graduate School of the Technion. Shmuel Onn was also supported by the ISF: Israel Science Foundation. Henry Wynn and Hugo Maruri-Aguilar were also supported by the Research Councils UK (RCUK) Basic Technology grant "Managing Uncertainty in Complex Models" (EP/D048893/1).

References

Babson, E., Onn, S., Thomas, R. (2003). The Hilbert zonotope and a polynomial time algorithm for universal Gröbner bases. *Advances in Applied Mathematics, 30*(3), 529–544.

Balakrishnan, N., Yang, P. (2006). Connections between the resolutions of general two-level factorial designs. *Annals of the Institute of Statistical Mathematics, 58*, 595–608.

Berstein, Y., Lee, J., Maruri-Aguilar, H., Onn, S., Riccomagno, E., Weismantel, R., Wynn, H. (2008). Nonlinear matroid optimization and experimental design. *SIAM Journal on Discrete Mathematics, 22*(3), 901–919.

Box, G., Hunter, J. (1961a). The 2^{k-p} fractional factorial designs. I. *Technometrics, 3*, 311–351.

Box, G., Hunter, J. (1961b). The 2^{k-p} fractional factorial designs. II. *Technometrics, 3*, 449–458.

Box, G., Wilson, K. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society Series B, 13*(1), 1–45.

Buchberger, B. (1966). On finding a vector space basis of the residue class ring modulo a zero dimensional polynomial ideal (in German), Ph.D. thesis. Department of Mathematics, University of Innsbruck.

Caboara, M., Pistone, G., Riccomagno, E., Wynn, H. (1997). The fan of an experimental design. SCU Research Report 33. Department of Statistics, University of Warwick.

Chen, H., Hedayat, A. S. (1998). Some recent advances in minimum aberration designs. In *New developments and applications in experimental design (Seattle, WA, 1997)*. IMS lecture notes monography series (Vol. 34, pp. 186–198). Hayward: Institute of Mathematical Statistics.

Chen, H. H., Cheng, C.-S. (2004). Aberration, estimation capacity and estimation index. *Statistica Sinica, 14*(1), 203–215.

Cheng, C.-S., Mukerjee, R. (1998). Regular fractional factorial designs with minimum aberration and maximum estimation capacity. *The Annals of Statistics, 26*(6), 2289–2300.

CoCoATeam. (2007). CoCoA: A system for doing computations in commutative algebra. Available at http://cocoa.dima.unige.it.

Corteel, S., Rémond, G., Schaeffer, G., Thomas, H. (1999). The number of plane corner cuts. *Advances in Applied Mathematics, 23*(1), 49–53.

Cox, D., Little, J., O'Shea, D. (1997). *Ideals, varieties, and algorithms* (2nd ed.). New York: Springer.

Cox, D., Little, J., O'Shea, D. (2005). *Using algebraic geometry*, *graduate texts in mathematics* (Vol. 185, 2nd ed.). New York: Springer.

Evangelaras, H., Koukouvinos, C. (2006). A comparison between the Gröbner bases approach and hidden projection properties in factorial designs. *Computational Statistics & Data Analysis, 50*(1), 77–88.

Faugère, J., Gianni, P., Lazard, D., Mora, T. (1993). Efficient computation of zero-dimensional Gröbner bases by change of ordering. *Journal of Symbolic Computation, 16*(4), 329–344.

Fries, A., Hunter, W. (1980). Minimum aberration 2^{k-p} designs. *Technometrics, 22*(4), 601–608.

Fukuda, K., Jensen, A. N., Thomas, R. R. (2007). Computing Gröbner fans. *Mathematics of Computation, 76*(260), 2189–2212 (electronic).

Greuel, G., Pfister, G., Schönemann, H. (2005). SINGULAR 3.0. A computer algebra system for polynomial computations. Centre for Computer Algebra, University of Kaiserslautern. http://www.singular.uni-kl.de.

Holliday, T., Pistone, G., Riccomagno, E., Wynn, H. (1999). The application of computational algebraic geometry to the analysis of designed experiments: A case study. *Computational Statistics, 14*(2), 213–231.

Li, W., Lin, D. K. J., Ye, K. Q. (2003). Optimal foldover plans for two-level nonregular orthogonal designs. *Technometrics, 45*(4), 347–351.

Maruri-Aguilar, H. (2007). Algebraic statistics in experimental design. Ph.D. thesis. Department of Statistics, University of Warwick.

McKay, M. D., Beckman, R. J., Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*(2), 239–245.

Monagan, M. B., Geddes, K. O., Heal, K. M., Labahn, G., Vorkoetter, S. M., McCarron, J., DeMarco, P. (2005). *Maple 10 programming guide*. Waterloo, ON: Maplesoft.

Mora, T., Robbiano, L. (1988). The Gröbner fan of an ideal. *Journal of Symbolic Computation, 6*(2–3), 183–208. (Computational aspects of commutative algebra).

Müller, I. (2003). Corner cuts and their polytopes. *Beiträge zur Algebra und Geometrie. Contributions to Algebra and Geometry, 44*(2), 323–333.

Onn, S., Sturmfels, B. (1999). Cutting corners. *Advances in Applied Mathematics, 23*(1), 29–48.

Peixoto, J. (1987). Hierarchical variable selection in polynomial regression models. *The American Statistician, 41*(4), 311–313.

Pistone, G., Riccomagno, E., Wynn, H. P. (2001). *Algebraic statistics*, *monographs on statistics and applied probability* (Vol. 89). Boca Raton: Chapman & Hall/CRC.

Pistone, G., Riccomagno, E., Rogantin, M. (2008). Algebraic statistics methods in DOE (with a contribution by Maruri-Aguilar, H.). In L. Pronzato, A. A. Zhigljavsky (Eds.), *In search for optimality in optimization and statistics*. Berlin: Springer.

Pistone, G., Wynn, H. (1996). Generalised confounding with Gröbner bases. *Biometrika, 83*(3), 653–666.

Sturmfels, B. (1996). *Gröbner bases and convex polytopes*. University lecture series (Vol. 8). Providence: American Mathematical Society.

Wagner, U. (2002). On the number of corner cuts. *Advances in Applied Mathematics, 29*(2), 152–161.

Wang, J. C., Wu, C.-F. J. (1995). A hidden projection property of Plackett–Burman and related designs. *Statistica Sinica, 5*(1), 235–250.

Wu, H., Wu, C. F. J. (2002). Clear two-factor interactions and minimum aberration. *The Annals of Statistics, 30*(5), 1496–1511.