The Fan of an Experimental Design

Massimo Caboara
Dipartimento di Matematica
Università di Pisa
caboara@lancelot.dima.unige.it

Giovanni Pistone
Dipartimento di Matematica
Politecnico di Torino
pistone@calvino.polito.it

Eva Riccomagno
Dept of Statistics
University of Warwick
emr@stats.warwick.ac.uk

Henry P. Wynn
Dept of Statistics
University of Warwick
hpw@stats.warwick.ac.uk

Summary

This paper continues work on the application of algebraic geometry to the design of experiments initiated in Pistone and Wynn (1996). It extends the theory of confounding to study the *fan* of design which, except for a monomial divisibility condition, is the full range of estimable polynomial models for a particular design. This gives a fuller understanding of confounding/aliasing and leads to the concepts of maximal fan and minimal fan designs.

Some key words: Computer algebra; Design and analysis of experiments; Identifiability; Fan of an ideal.

1. Introduction

In Pistone and Wynn (1996) two of the current authors introduce algebraic geometry ideas into experimental design and show how the theory of Gröbner bases (G-bases) can be used to find a saturated estimable (identifiable) linear polynomial model for a given design. That paper shows how, given a design $d = \left\{x^{(1)}, \ldots, x^{(n)}\right\}$ ($x^{(i)}$ all distinct) and given a so-called monomial ordering τ on the monomial terms of the form $x_1^{\alpha_1} \ldots x_m^{\alpha_m}$, a unique reduced G-basis results and gives a unique set of monomial terms for a linear model. This set is renamed $Est_{d,\tau}$. The number of such terms, namely the size of $Est_{d,\tau}$, is always equal to the sample size n of distinct design points and hence $Est_{d,\tau}$ is saturated in the usual statistical sense.

Section 2 of this paper will revisit in summary form the basic theory. This arises from considering an experimental design as an algebraic variety, namely the solution of a set of algebraic equations. G-bases are special choices of such equations. If $x = (x_1, \ldots, x_m)$ are the independent variables, $f(x)$ any (multivariate) polynomial model and $g_1(x), \ldots, g_r(x)$ are the polynomials forming the G-basis then

$$f(x) = \sum_{j=1}^{r} s_j(x) g_j(x) + r(x)$$

where $r(x)$ is *unique* and of lower order than $g_j(x)$ (see Section 2).

Since $g_j(x^{(i)}) = 0$ for all j and all design points $x^{(i)}$ we have

$$f(x^{(i)}) = r(x^{(i)}) \qquad (i = 1, \ldots, n)$$

When observations, $Y_1, \ldots, Y_n$ are taken (without error) so that for the response $Y_i = f(x^{(i)})$ then $r(x)$ is a polynomial interpolatory of the $(x^{(i)}, Y_i)$ which is unique given the G-basis. The elements of $Est_{d,\tau}$, of which the estimable models $r(x)$ are linear combinations, satisfy a divisibility condition (D) which is that if a term $x^\alpha = x_1^{\alpha_1} \ldots x_m^{\alpha_m}$ is in $Est_{d,\tau}$ then every term which divides x^α is also in $Est_{d,\tau}$. For example if $x_1^2 x_2$ is in $Est_{d,\tau}$ then so are x_1, x_2, x_1x_2, x_1^2 and the constant term, which we denote by 1.

A rough introduction to the general theory of confounding here is to say that two polynomial models $f_1(x)$ and $f_2(x)$ are aliased relative to a design and a G-basis (that is a certain term-ordering) if they have the same remainder $r(x)$ with respect to the G-basis. Thus, the theory of confounding in Pistone and Wynn (1996) is essentially relative to the choice of monomial ordering and hence G-basis.

This paper covers the description of the set of *all* saturated models $Est_{d,\tau}$ identifiable with a given design d as we range over all monomial orderings τ. This set of models is called here a *fan* after seminal work by Mora and Robbiano (1988). The paper will emphasise designs with a maximal fan that is designs for which there is a maximal number of saturated estimable models (subject to the divisibility condition (D)) and designs with a minimal fan. Designs with both kinds of fans always exist (see Sections 6 and 7).

We clarify this with a simple example. Consider designs with 4 points in two factors, x_1 and x_2. For the classical 2^2 factorial design $\{(\pm 1, \pm 1)\}$ the saturated estimable model subject to (D) is

$$\mu = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2$$

with $Est_{d,\tau} = \{1, x_1, x_2, x_1x_2\}$. In this case every monomial ordering τ gives the same saturated model. A set of polynomials which is a G-basis with respect to every monomial orderings is called a *universal G-basis.* For example the set $T = \{x_1^2 - 1, x_2^2 - 1\}$ is a universal G-basis for the above design. Notice that the universal G-basis T also has the property that the leading terms of its elements (that is the highest terms in each element of T with respect to the given term-ordering, see Section 2) are the same with respect to all term-orderings. A universal G-basis satisfying the above property is called here a *total* G-basis. G-bases for minimal fan designs are total G-bases. In particular the full-factorial is a minimal fan design and T is its total and reduced G-basis.

Now consider the design $\left\{(-1,-1), (-\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, -\frac{1}{2}), (1,1)\right\}$. One can check that it separately identifies the following five saturated models $\{1, x_1, x_1^2, x_1^3\}$, $\{1, x_1, x_1^2, x_2\}$, $\{1, x_1, x_2, x_1x_2\}$, $\{1, x_1, x_2, x_2^2\}$ and $\{1, x_2, x_2^2, x_2^3\}$. Moreover there are no other saturated models (subject to

(D)) identifiable by a 4-point design. In fact the collection of these five models is a maximal fan and the models are called the *leaves* of the fan.

Although in this paper we do not study the analysis of such models we may raise the issue of whether a design which allows us to estimate more saturated models is better than one which allows only one. In particular it should be especially interesting to consider screening experiments. If we first run a saturated experiment, and then re-use the same data fitting a model with less factors, different leaves of the fan could lead to quite different submodels.

The set of all designs with n-distinct points in m-dimensions can be decomposed into a finite number of non-intersecting (possibly empty) classes. Indeed two designs belong to the same class if and only if they have the same fan. For example for 4 point designs in 2 dimensions there are $31 = (2^5 - 1)$ such possible fans. Table 1 gives the classification all possible fans for three points in two dimensions. The last design has a maximal fan, in that every model with three terms subject to (D) is estimable. An interpretation is that such a design is in *general position* in an algebraic sense. In Sections 3 and 4 we summarize the algebraic theory for fans of ideals.

It seems to be a major challenge to try to determine a design for each possible fan of n-term leaves in m-dimensions. This could be called the problem of generalized confounding which then becomes a problem of algebraic geometry in general. At present the authors are able to compute the fan of a particular design using G-basis methods (Section 4) or simply computing determinants (in the manner of Section 7). As yet they have no comprehensive method of classifying all patterns which give the same fan. Our main results are included in the last three sections. The completeness of the algebraic method for identifiability is proved for models satisfying the (D) condition in Section 5.

2. Basic Algebra

In this section we summarize the basic algebraic theory of identifiability. We refer to Pistone and Wynn (1996), Holliday, Pistone, Riccomagno and Wynn (1996) and Caboara and Riccomagno (1997).

Let $\mathbb{Q}$ and $\mathbb{R}$ represent the rational numbers and the real numbers respectively. The algebraic theory of identifiability assumes a finite set of distinct points with rational coordinates, that is a single replicate design $d = \left\{x^{(1)}, \ldots, x^{(n)}\right\} \subset \mathbb{Q}^m$ or $\mathbb{R}^m$, and a term-ordering τ on the terms in $k[x] = k[x_1, \ldots, x_m]$, the ring of all polynomials in m indeterminates with coefficients in k. In our case k is $\mathbb{Q}^m$ or $\mathbb{R}^m$. The terms or monomial terms of $k[x]$ are the elements of $k[x]$ of the form $x^\alpha = x_1^{\alpha_1} \ldots x_m^{\alpha_m}$ where $\alpha = (\alpha_1, \ldots, \alpha_m)$ is a vector of non-negative integers.

Let d be a design. The set of all polynomials whose zeros include the design points is an ideal of $k[x]$, it is denoted by $I(d)$ and is called the design ideal associated to d.

A term-ordering τ is a totally ordering relation on the monomials satisfying the following conditions

(i) if $x^\alpha <_\tau x^\beta$ then $x^{\alpha+\gamma} <_\tau x^{\beta+\gamma}$ for all non-negative integer vectors α, β, γ, that is τ is compatible with the division and multiplication of monomials.

(ii) τ is a well-ordering, that is any set of terms has a smallest element with respect to τ.

For examples of term-orderings we refer the reader to Cox, Little and O-Shea (1992) and Pistone and Wynn (1996). Given a term-ordering τ one can calculate the (unique) reduced G-basis, $G_{d,\tau}$ of the design ideal, $I(d)$. A set of polynomials is a G-basis for a polynomial ideal I and with respect to a term ordering τ if

$$< Lt_\tau(g) : g \in G >=< Lt_\tau(f) : f \in I >$$

where in general $Lt_\tau(q)$ is the leading term of the polynomial q, that is the highest term in q with respect to τ and $< A >$ indicates the ideal generated by the set A.

Since the G-basis, $G_{d,\tau} = \{g_1, \ldots, g_r\}$ is a basis of the ideal $I(d)$ every element $f \in I(d)$ can be decomposed in a non-unique way as

$$f(x) = \sum_{j=1}^{r} g_j(x) s_j(x) \qquad \text{for some } s_j(x) \in k[x] \text{ for all } j = 1, \ldots, r.$$

Moreover, and this is the main feature of G-bases, for any polynomial f in $k[x]$ there exists a unique polynomial r in $k[x]$, called the remainder, such that

$$f(x) = \sum_{j=1}^{r} g_j(x) s_j(x) + r(x) \qquad \text{for some } s_j(x) \in k[x] \text{ for all } j = 1, \ldots, r$$

and the terms in the remainder precedes the leading terms of the G-basis in the ordering τ. That is $Lt_\tau(r) <_\tau Lt_\tau(g_j)$ for all $j = 1, \ldots, r$. A shortened notation for the remainder of f with respect to the G-basis, G (and the term-ordering τ) is $Rem(f, G)$.

The set of all remainders is in one-to-one correspondence with the quotient ideal $k[x]/I(d)$. The following is an important but unstated fact within experimental design (see Cox, Little and O'Shea, 1992 and Pistone and Wynn, 1996). Namely *the dimension as a vector-space of $k[x]/I(d)$ equals the number of design points regardless of the term-ordering in which the calculations are done.*

Once we have the G-basis, G_τ of the design ideal $I(d)$, a vector-space basis of the remainder set $k[x]/I(d)$ is calculated as the set of terms not divisible by any leading term in G_τ. It follows that $k[x]/I(d)$ is the set of all models (subject to the (D) condition) identifiable by the design d with respect to the ordering τ. In particular the elements of a vector space basis of $k[x]/I(d)$ give the terms of a saturated model identifiable using d. This is the set $Est_{d,\tau}$ and the remainder $Rem(f, G)$ is a k-linear combination of elements of $Est_{d,\tau}$.

Note that $Est_{d,\tau}$ is an *order ideal* where E is an order ideal if (i) E is a finite set of monomials and (ii) if $x^\alpha \in E$ and x^β divides x^α then $x^\beta \in E$. In particular (ii) is the (D) condition which is the key condition for models in this paper.

All of the above is summarised in the following function, $Id_{d,\tau}$ that associates (through the division operation among polynomials) an estimable model satisfying the (D) condition with a model f

$$\begin{array}{rcl} Id_{d,\tau} \; : \; k[x] & \longrightarrow & k[x]/I(d) \\ f & \longmapsto & Rem(f, G) \end{array}$$

¿From this formulation we can infer other important facts. For example the polynomial model $f \in k[x]$ is aliased/confounded with the model g with respect to the design d and with respect to the term-ordering τ if and only if $Rem(f, G_{d,\tau}) = Rem(g, G_{d,\tau})$.

Note at this point that the G-basis G carries all the information about the design d and the term-ordering τ gives a total ordering on the variables $x_1, \ldots, x_m$.

3. Fan theory

We present the theory of a fan of a polynomial ideal first in a rather algebraic way. References are Mora and Robbiano (1988) and Sturmfels (1995).

Given a G-basis, G of the ideal I with respect to a term-ordering τ the monomial ideal generated by the leading terms of G is called the *initial ideal*

$$Init_\tau(G) =< Lt_\tau(g) : g \in G >$$

Notice that by the definition of a G-basis the following holds

$$Init_\tau(G) =< Lt_\tau(g) : g \in I >$$

The set of all monomials not divisible by any of the $Lt(g)$, $g \in G$, that is the monomials not in $Init_\tau(G)$ is an order ideal.

The following is proved for example in Sturmfels (1995): *every ideal $I \subset k[x]$ has only finitely many distinct initial ideals, equivalently order ideals.* This allows us to define an equivalence relation splitting the infinite set of term-orderings into a finite number of classes, as mentioned in Section 1. Two orderings, τ_1 and τ_2 are equivalent with respect to an ideal I (and we shall say with respect to a design d) if and only if they have the same initial ideal

$$Init_{\tau_1}(I) =< Lt_{\tau_1}(g) : g \in G_{\tau_1} >=< Lt_{\tau_2}(g) : g \in G_{\tau_2} >= Init_{\tau_2}(I)$$

where G_{τ_j} is the G-basis of I with respect to τ_j, $j = 1, 2$. The partition so induced on the set of term-orderings is called the *fan* of the ideal I, in symbols $\mathcal{F}(I)$ or $\mathcal{F}(d)$ when $I = I(d)$ for some design d. Each one of these equivalence classes is called a *leaf*. In particular leaves are characterised by initial ideals, that is τ_1 and τ_2 belong to the same leaf, L if and only if $Init_{\tau_1}(I) = Init_{\tau_2}(I)$. Moreover to each leaf L of the fan one can associated an order ideal E_L namely the set of terms which are not divisible by any of the elements in the initial ideal corresponding to the leaf. In the case of interest in this paper, which is when I is a design ideal $I(d)$, E_L is finite and it is precisely $Est_{d,\tau}$ for all $\tau \in L$.

As an example the design $d = \{(0,0),(1,1),(3,2)\}$ is maximal fan since it estimates all the two-dimensional models estimable by a three point design which are $\{1, x_1, x_1^2\}$, $\{1, x_1, x_2\}$ and $\{1, x_2, x_2^2\}$. A term-ordering representative for the leaf labeled $\{1, x_1, x_1^2\}$ is the lexicographic order with $x_2 > x_1$; for $\{1, x_1, x_2\}$ it is the `tdeg` ordering and for $\{1, x_2, x_2^2\}$ the lexicographic order with $x_1 > x_2$. We show the algorithm that returns the fan (constrained fan) of an ideal in a next section.

4. Computing the fan

Ideally one would like to input all the information available on the term-ordering before starting the computation, that is to define a pre-ordering on the variables instead of a term-ordering. The computer algebra package CoCoA, which we mainly use, allows us to define a pre-ordering.

The algorithm to calculate fans of ideals receives as input a basis of the design ideal. At each step it chooses the possible leading terms compatible with the known ordering information, applies the important S-polynomial test to check whether a set of polynomials is a G-basis with respect to a given monomial ordering (see Cox, Little and O'Shea, 1992, Section 2.6 and below) and keeps iterating to create new leaves of the fan. When the S-polynomial test is positive over one leaf it returns the G-basis associated with that leaf and the conditions which the term-ordering s of that leaf must satisfy. This algorithm was first introduced in Mora and Robbiano (1988). The usual improvements to the Buchberger algorithm for reduced G-bases can be applied.

We recall that the S-polynomial of two polynomials f and g is defined as

$$Spoly(f,g) = \frac{LCM(Lt(f),Lt(g))}{Lt(f)LC(f)}f - \frac{LCM(Lt(f),Lt(g))}{Lt(g)LC(g)}g$$

where LC is the coefficient of the leading term and LCM is the least common multiple. The S-polynomial test says that a set G is a G-basis with respect to τ if and only if $Rem(Spoly(f,g),G) = 0$ for all $f,g \in G$.

Let us show the details with an example. Consider the design $d = \{(0,0),(1,2),(2,1)\}$ from Table 1 and impose the condition $x_1 > x_2$ on the term ordering. The design d is the set of solution of the following system of polynomial equations

$$\begin{aligned} f &= x_2^3 - 3x_2^2 + 2x_2 \\ g &= x_1 + 3/2x_2^2 - 7/2x_2 \end{aligned}$$

The possible leading terms of g (compatible with $x_1 > x_2$) are x_1 and x_2^2, and for f we have only x_2^3. We create two leaves in the fan $\mathcal{F}(d)$ characterised by the conditions $x_1 > x_2^2$ and $x_2^2 > x_1$ respectively. The S-polynomials are

$$\begin{aligned} Spoly(f,g) &= -3x_2^2x_1 + 2x_1x_2 - 3/2x_2^5 + 7/2x_1^4 && \text{for } x_1 > x_2^2 \\ Spoly(f,g) &= -\tfrac{2}{3}x_2^2 + 2x_2 - \tfrac{2}{3}x_2x_1 && \text{for } x_2^2 > x_1 \end{aligned}$$

Their remainders with respect to f and g are

$$\begin{aligned} p &= Rem(Spoly(f,g),\{f,g\}) = 0 && \text{for } x_1 > x_2^2 \\ h &= Rem(Spoly(f,g),\{f,g\}) = -\tfrac{2}{3}x_1x_2 + \tfrac{4}{9}x_1 + \tfrac{4}{9}x_2 && \text{for } x_2^2 > x_1 \end{aligned}$$

Since $p = 0$, by the S-polynomial test we have that for all the orderings such that $x_1 > x_2^2$ the set $\{f,g\}$ is a (reduced) G-basis which gives $\{1, x_2, x_2^2\}$ as the estimable set.

We have to continue the calculation for the orderings such that $x_2^2 > x_1$. The new generating set is $\{f, g, h\}$ and the only possible leading term of h is x_1x_2. Thus

$$\begin{aligned} Spoly(f,h) &= -7/3x_1x_2^2 + 2x_1x_2 + 2/3x_2^3 \\ Spoly(g,h) &= \frac{2}{3}(x_1^2 + x_2^2) - \frac{5}{3}x_1x_2 \end{aligned}$$

and

$$\begin{aligned} l &= Rem(Spoly(f,h), \{f,g,h\}) &= -14/9x_1^2 + 98/27x_1 - 28/27x_2 \\ m &= Rem(Spoly(g,h), \{f,g,h\}) &= \tfrac{2}{3}x_1^2 - \tfrac{14}{9}x_1 + \tfrac{4}{9}x_2 \end{aligned}$$

Because of the prior condition $x_1 > x_2$ on the ordering the only possible leading term of l and g is x_1^2. The S-polynomial test shows that for the term-orderings such that $x_2^2 > x_1$ and $x_1^2 > x_2$ the set $\{f, g, h, l, m\}$ is a G-basis. The estimable set is $\{1, x_1, x_2\}$. In conclusion the fan of the design d with the constrained $x_1 > x_2$ is $\{\{1, x_2, x_2^2\}, \{1, x_1, x_2\}\}$.

If no condition on the ordering is imposed the above algorithm returns the fan of the ideal given as input.

5. Interpolation

For a particular design $d = \left\{x^{(1)}, \ldots, x^{(n)}\right\}$ let E_L be the order ideal corresponding to a particular leaf L of the fan of d and let p_j for $j = 1, \ldots, n$ be the elements of E_L, thus

$$E_L = \{p_1(x), \ldots, p_n(x)\}$$

Then the usual design matrix $X(E_L, d)$ for the model is

$$X(E_L, d) = \left\{p_j(x^{(i)})\right\}_{i,j=1}^{n}$$

Since E_L is estimable the matrix $X(E_L, d)$ is invertible and equivalently $\det(X(E_L, d)) \neq 0$. Now the maximal set of leaves of dimension n subject to the (D) condition is well defined and finite. For $m = 2$ dimensions each such model can be mapped into a partition of n where the models (order ideals E_L) can be represented by solid dots on an integer grid. For example for $m = 2$, $n = 5$ the pattern

$$\begin{array}{ccc} \cdot & \cdot & \\ \cdot & \cdot & \cdot \end{array}$$

corresponding to $5 = 2 + 2 + 1$ gives the model $1, x_1, x_1^2, x_2, x_1x_2$. There are 7 models hence the fan of a 5-point design in 2-dimension will have at most 7 leaves. This easily generalises to $m > 2$ using a generalised partition of integers.

It will be shown in Section 7 that there is always a design of sample size n with which to estimate a model with n terms subject to the (D) condition.

Let $\mathcal{E}(d)$ be the set of models satisfying the (D)-condition and with n terms, where n is the size of the design d, and such that their design matrices at d are invertible. We say that the elements of $\mathcal{E}(d)$ are identifiable in a statistical sense. Let $\mathcal{F}(d)$ be the fan of the design d calculated as in Section 4. Elements of $\mathcal{F}(d)$ are algebraically identifiable. By Pistone and Wynn (1996) we have that algebraic identifiability implies statistical identifiability, that is $\mathcal{F}(d) \subseteq \mathcal{E}(d)$ and Caboara and Robbiano (1997) show with a counterexample that the inclusion may be strict: the model $E = \{1, x, x^2, y, y^2\}$ is statistically but not algebraically identifiable by the design $d = \{(0,0), (0,-1), (1,0), (1,1), (-1,1)\}$. However notice that the k-vector space generated by any model E in $\mathcal{E}(d)$ is isomorphic to the quotient $k[x]/I(d)$. For details see Pistone and Wynn (1996), Section 4. Theorem 1 below shows that subject to an additional condition to avoid designs and models in $\mathcal{E}(d) \setminus \mathcal{F}(d)$, there is a correspondence between interpolation and algebraic identifiability.

Let d be a n-point design and E an element of $\mathcal{E}(d)$. With an abuse of notation we list the terms of the saturated estimable model in a vector as follows

$$E(x) = (p_1(x), \ldots, p_n(x))^t$$

Suppose that the usual $n \times n$ design matrix

$$X = \left\{p_j(x^{(i)})\right\}_{i,j=1}^{n}$$

is invertible. We want to construct the initial ideal leading to E.

First we observe that given a term-ordering every polynomial $f \in k[x]$ can be decomposed as a leading term $Lt(f,x) = Lt(f)$ and a *tail* $t(f,x) = Lt(f) - f$ in such a way that $f(x) = Lt(f,x) - t(f,x)$. Let G be a reduced G-basis. Then for all $h \in G$ none of the terms in $t(h,x)$ is divisible by any $Lt(g,x)$ for all $g \in G$. In other words for all $j = 1, \ldots, J$ there exist a vector of length n with scalar entries, Θ_j such that the tail t_j is a linear combination of elements in $E(x)$

$$t_j(x) = E(x)^t \Theta_j$$

where J is the number of elements in G.

Next we observe that the complementary set of $E(x)$ in the set of all monomial terms in the variables x is a monomial ideal and thus by the Dickson's Lemma (see Little, Cox, O'Shea, 1992) we can construct a unique minimal finite basis of monomials of such a set. Let us denote such a basis by $Init = \{Lt_j(x)\}_{j=1}^{J}$. By construction the elements of $E(x)$ are those monomials not divisible by any of the $Lt_j(x)$, for $j = 1, \ldots, J$. Indeed let x^α be an element of $E(x)$. By definition $x^\alpha \notin Init$. Let us suppose that x^α is divisible by one of the Lt_k for a k in $\{1, \ldots, J\}$. Thus there exists a monomial x^β such that $x^\alpha = x^\beta Lt_k$, that is $x^\alpha \in < Lt_k > \subset < Lt_j : j = 1, \ldots, J >= Init$. This is a contradiction and we are done.

Then we construct polynomials $t_j(x)$ which interpolate each of the terms in $Init$ using the model based on $E(x)$ at the design d, that is to say solve the following J linear systems of equations with respect to Θ_j

$$\begin{cases} Lt_j(x^{(1)}) = E(x^{(1)})^t \Theta_j = X\Theta_j \\ \vdots \\ Lt_j(x^{(n)}) = E(x^{(n)})^t \Theta_j = X\Theta_j \end{cases}$$

Thus the t_j are uniquely determined because of the invertibility of X. Then define

$$g_j(x) = Lt_j(x) - t_j(x) \qquad j = 1, \ldots, J \tag{1}$$

Now we are ready to state Theorem 1.

Theorem 1 *Let d be a design and E an element of $\mathcal{E}(d)$. Let $Init$ be constructed as above and let us suppose that there exists a term-ordering τ such that $Lt_j(x)$ is the leading term of $g_j(x)$ with respect to τ for all $j = 1, \ldots, J$. Then $E \in \mathcal{F}(d)$.*

Proof. The existence of τ follows by the fact that the hypothesis in the theorem defines the leading terms of the $g_j(x)$'s. That hypothesis is essential to avoid situations similar to the counterexample of Caboara and Robbiano (1997). We show that the ideal generated by the $g_j(x)$'s namely $< g_j(x) >$ is the design ideal, $I(d)$. Certainly by construction the design ideal includes the ideal generated by the g_j's. Conversely let p be a polynomial in the design ideal and expand it in the g_j's by the division algorithm using the τ in the statement of the theorem:

$$p(x) = \sum_{j=1}^{J} s_j(x) g_j(x) + r(x)$$

Since $p(x)$ belongs to the design ideal and $g_j(x^{(i)}) = 0$ at all design points $x^{(i)}$ $(i = 1, \ldots, n)$ and for all $j = 1, \ldots, J$ we have

$$p(x^{(i)}) = r(x^{(i)}) = 0$$

Now the division algorithm always yields a remainder $r(x)$ every monomial of which is dominated by the leading terms of the $g_j(x)$, in this case the $Lt_j(x)$. But by the assumption in the theorem the monomials must be from $E(x)$. But the design matrix for $E(x)$ at the design d is invertible and thus $r(x) = 0$ identically. This implies that $p(x) \in < g_j(x) >$.

Finally we show that the set $G = \{g_j(x) : j = 1, \ldots, J\}$ is a (reduced) G-basis for the design ideal. We use the S-polynomial test. Consider a generic S-polynomial and proceed as above by expanding it on G

$$Spoly(g_l, g_k) = \sum_{j=1}^{J} s_j(x) g_j(x) + r(x)$$

and by evaluating it at the design points. Since $Spoly(g_l, g_k) \in < g_j(x) >$, it must be zero at the design points leading to $r(x^{(i)}) = 0$ for all design points. But again since $r(x)$ is a linear combination of elements in $E(x)$ which is estimable we must have $r(x) = 0$ identically. Notice that by construction $\{g_j(x) : j = 1, \ldots, J\}$ is reduced. ∎

The following example clarifies the three steps of the proof. Consider the two-dimensional design $d = \{(0,0), (1,0), (0,1), (2,1)\}$ and the estimable model $E = \{1, x_1, x_2, x_1^2\}$. We check estimability simply by checking that the design matrix

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 4 \end{pmatrix}$$

is invertible. The set of leading terms giving E is $Init = \{x_1^3, x_1x_2, x_2^2\} = \{Lt_1(x), Lt_2(x), Lt_3(x)\}$. Note that the condition in Theorem 1 is satisfied. We have the interpolators of the elements of $Init$

$$\begin{aligned} t_1(x) &= 3x_1^2 - 2x_1 \\ t_2(x) &= x_1^2 - x_1 \\ t_3(x) &= x_2 \end{aligned}$$

Thus the G-basis is

$$\begin{aligned} g_1(x) &= x_1^3 - 3x_1^2 + 2x_1 \\ g_2(x) &= x_1x_2 - x_1^2 + x_1 \\ g_3(x) &= x_2^2 - x_2 \end{aligned}$$

The leading term of g_2 must be x_1x_2 and thus we require that $x_1x_2 > x_1^2$ which implies that the term-orderings such that $x_2 > x_1$ belong to the leaf of $E(x)$.

For the counterexample mentioned above the set of interpolating polynomials is as follows

$$\begin{aligned} xy &= -x^2 + y^2/2 + x + y/2 \\ x^3 &= x \\ y^3 &= y \end{aligned}$$

The condition in Theorem 1 is not met since there does not exist a term-ordering such that xy is leading term of the first polynomial. Indeed it should simultaneously be $xy > x^2$ and $xy > y^2$, that is $y > x$ and $x > y$ which is not possible in a total ordering.

6. Minimal fan designs

A minimal fan design is defined as a design whose fan has only one leaf. A special case of such designs are the full factorial, or product, designs. For example the fan of the design in $\mathbb{R}^2$ $\{0,1,2,3\} \times \{0,1,2\}$ which has as representation

$$\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

has the single leaf

$$\begin{array}{lcccc} \{ & x_2^3, & x_2^3x_1, & x_2^3x_1^2, & x_2^3x_1^3, \\ & x_2^2, & x_2^2x_1, & x_2^2x_1^2, & x_2^2x_1^3, \\ & x_2 & , x_2x_1, & x_2x_1^2, & x_2x_1^3, \\ & 1, & x_1, & x_1^2, & x_1^3 \quad \} \end{array}$$

The following fundamental class of designs generalises this remark.

Definition 1 *A design $d \subset Z_+^m$ is called an echelon design if for any design point $(d_1, \ldots, d_m)$ all points of the form $(y_1, \ldots, y_m)$ with $0 \leq y_j \leq d_j$, for all $j = 1, \ldots, m$ belong to the design d.*

In this section we prove that the fan of an echelon design consists of a single *echelon* leaf whose elements are $x_1^{d_1} \ldots x_m^{d_m}$ for all $(d_1, \ldots, d_m)$ in the echelon design. First we show it in two dimensions. Consider the design d

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

$$\begin{array}{cccc} \cdot & & & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

A (non reduced) G-basis for the design ideal is

$$\begin{cases} x_2(x_2-1)(x_2-2) \\ x_1x_2(x_2-1) \\ x_1(x_1-1)x_2(x_2-1) \\ x_1(x_1-1)(x_1-2)x_2 \\ x_1(x_1-1)(x_1-2)(x_1-3) \end{cases}$$

The leading terms are x_2^3, $x_1x_2^2$, $x_1^2x_2^2$, $x_1^3x_2$ and x_1^4 with respect to any term-ordering. The saturated set of estimable terms is as follows

$$\begin{array}{cccc} x_2^2 & & & \\ x_2 & x_1x_2 & x_1^2x_2 & \\ 1 & x_1 & x_1^2 & x_1^3 \end{array}$$

Notice that the structure of the design mimics the structure of the estimable model. Let now d be a generic echelon design in two dimensions. It is the union of columns of points of the form

$$\begin{array}{ll} (0,h): & h = 0, \ldots, k_0 \\ (1,h): & h = 0, \ldots, k_1 \\ \vdots & \\ (l,h): & h = 0, \ldots, k_l \end{array}$$

where $k_0 \geq k_1 \geq \ldots \geq k_l$. Next we show that the following polynomials form a G-basis for the design ideal

$$\begin{cases} p_0(x_2) & = & \prod_{j=0}^{k_0}(x_2 - j) \\ p_1(x_1, x_2) & = & x_1 \prod_{j=0}^{k_1}(x_2 - j) \\ p_2(x_1, x_2) & = & x_1(x_1-1) \prod_{j=0}^{k_2}(x_2 - j) \\ \vdots & & \\ p_l(x_1, x_2) & = & \prod_{j=0}^{l-1}(x_1 - j) \prod_{j=0}^{k_l}(x_2 - j) \\ p_{l+1}(x_1) & = & \prod_{j=0}^{l}(x_1 - j) \end{cases} \quad (2)$$

First they clearly interpolate the design points. Next apply the S-polynomial test. Without loss of generality choose a, b such that $a, b \in \{0, \ldots, l+1\}$ and $a < b$, thus $k_a \geq k_b$. Then by definition

$$p_a(x_1, x_2) = A_{x_1} A_{x_2} \qquad \text{and} \qquad p_b(x_1, x_2) = B_{x_1} B_{x_2}$$

where

$$A_{x_1} = \prod_{j=0}^{a-1} (x_1 - j) \qquad \text{and} \qquad A_{x_2} = \prod_{j=0}^{k_a} (x_2 - j),$$
$$B_{x_1} = \prod_{j=0}^{b-1} (x_1 - j) \qquad \text{and} \qquad B_{x_2} = \prod_{j=0}^{k_b} (x_2 - j)$$

and $A_{x_1} = 1$ for $a = 0$ and $B_{x_2} = 1$ for $b = 0$. Note that A_{x_1} divides B_{x_1} and B_{x_2} divides A_{x_2}. ¿From Lemma 1 below it follows that

$$Spoly\,(p_a(x_1, x_2), p_b(x_1, x_2)) = A_{x_1} B_{x_2} Spoly\left(\frac{A_{x_2}}{B_{x_2}}, \frac{B_{x_1}}{A_{x_1}}\right)$$

and by Proposition 4 in Cox, Little and O'Shea (1992) on coprime polynomials we have that

$$Rem\left(Spoly\left(\frac{A_{x_2}}{B_{x_2}}, \frac{B_{x_1}}{A_{x_1}}\right), \left\{\frac{A_{x_2}}{B_{x_2}}, \frac{B_{x_1}}{A_{x_1}}\right\}\right) = 0$$

This implies that the S-polynomial test on p_a and p_b is positive since

$$\begin{aligned} Spoly(p_a, p_b) &= A_{x_1} B_{x_2} Spoly\left(\frac{A_{x_2}}{B_{x_2}}, \frac{B_{x_1}}{A_{x_1}}\right) \\ &= A_{x_1} B_{x_2} \left[s_1 \frac{A_{x_2}}{B_{x_2}} + s_2 \frac{B_{x_1}}{A_{x_1}}\right] \qquad \text{for some polynomial } s_1, s_2 \\ &= A_{x_1} s_1 A_{x_2} + s_2 B_{x_2} B_{x_1} \\ &= s_1 p_a + s_2 p_b \end{aligned}$$

As p_a, p_b are any two polynomials of the interpolating basis (2) we conclude that it is a G-basis.

Lemma 1 *Let f, A, B be polynomials. Then $Spoly(fA, fB) = fSpoly(A, B)$.*

Proof. The following facts are easily seen: (i) $Lt(AB) = Lt(A)Lt(B)$ for all A, B polynomials, (ii) $LCM(x^\gamma x^\alpha, x^\gamma x^\beta) = x^\gamma LCM(x^\alpha, x^\beta)$ for all $x^\gamma, x^\alpha, x^\beta$ monomials. It follows that

$$LCM(Lt(fA), Lt(fB)) = Lt(f) LCM(Lt(A), Lt(B))$$

Now

$$\begin{aligned} Spoly(fA, fB) &= \frac{LCM(Lt(fA), Lt(fB))}{Lt(fA)} fA - \frac{LCM(Lt(fA), Lt(fB))}{Lt(fB)} fB \\ &= \frac{LCM(Lt(A), Lt(B))}{Lt(A)} fA - \frac{LCM(Lt(A), Lt(B))}{Lt(B)} fB \\ &= fSpoly(A, B) \qquad \blacksquare \end{aligned}$$

The polynomials defined in (2) form a total G-basis whose leading terms with respect to any term-ordering are

$$x_2^{k_0}, x_1 x_2^{k_1}, \ldots, x_1^l x_2^{k_l}, x_1^{l+1}$$

giving as saturated estimable model the following model

$$\begin{array}{lcl} x_2^h & : & h = 0, \ldots, k_0 \\ x_1 x_2^h & : & h = 0, \ldots, k_1 \\ \vdots & & \\ x_1^l x_2^h & : & h = 0, \ldots, k_l \end{array}$$

which follows exactly the structure of the echelon design. Since the leading terms are unique (total G-basis) this model is unique (using the algebraic procedure of identifiability by Pistone and Wynn, 1997) and thus the fan of the echelon design has just this one leaf. This proof is easily extended to more than two dimensions. However the proof can be traced in the work of Robbiano and Rogantin (1997).

We now give an alternative proof in m dimensions more statistical in style. For an integer $r \geq 1$ define the univariate polynomial

$$p(r, z) = z(z-1)\cdots(z-r+1)$$

and for $x = (x_1, \ldots, x_m)$ and an integer vector $\beta = (\beta_1, \ldots, \beta_m)$ define

$$P(\beta, x) = \prod_{j=1}^{m} p(\beta_j, x_j) = \prod_{j=1}^{m} \prod_{k=0}^{\beta_j - 1} (x_j - k)$$

Note first that the echelon design (and corresponding model) is defined via a unique set of leading terms (by the Dickson's lemma). These terms are defined by certain integer vectors

$$\alpha^{(1)}, \ldots, \alpha^{(K)}$$

where no $x^{\alpha^{(i)}}$ divides an $x^{\alpha^{(j)}}$ for all $i \neq j$ and $i, j = 1, \ldots, K$. Note that the corresponding echelon design is all points in $\mathbb{Z}_+^m$ not dominated by $\alpha^{(1)}, \ldots, \alpha^{(K)}$. For the above example the leading terms are given by the crosses

$$\begin{array}{ccccc} \times & & & & \\ . & \times & & & \\ . & . & . & \times & \\ . & . & . & . & \times \end{array}$$

namely the points

$$(4,0), (3,1), (1,2), (0,3)$$

The corresponding leading terms are

$$x_1^4, x_1^3 x_2, x_1 x_2^2, x_2^3$$

We first show that the X-matrix for the echelon design and corresponding model, $X(E,d)$ is invertible. First list the design and the model in the same order in such a way that the monomial term $x^{\alpha(j)}$ of the model and the design point $\alpha(j)$ of the echelon design occupies the same position in the order. Next reparametrise replacing monomial $x^{\alpha^{(j)}}$ by the polynomials $P(\alpha^{(j)},x)$ themselves. The mapping from the functional class $x^{\alpha^{(j)}}$ to $P(\alpha^{(j)},x)$ is invertible and linear. For example for the model above we have

$$\begin{bmatrix} 1 \\ x_1 \\ x_1(x_1-1) \\ x_1(x_1-1)(x_1-2) \\ x_2 \\ x_1x_2 \\ x_1(x_1-1)x_2 \\ x_2(x_2-1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \\ x_1^3 \\ x_2 \\ x_1x_2 \\ x_1^2x_2 \\ x_2^2 \end{bmatrix}$$

The invertibility follows immediately from the lower triangular form of the transformation matrix, Q. If Z is the X-matrix for the $\{P(\alpha^{(j)},x)\}$ and $X = X(E,d)$ then

$$Z = XQ^t$$

Now from the structure of the echelon design Z is also invertible and lower triangular. For the example

$$Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 6 & 6 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 2 & 2 & 0 \\ 1 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \end{bmatrix}$$

Now in general $\det(Q) = 1$ and

$$\begin{aligned} \det(X) = \det(Z) &= \prod_{j=1}^{K} P(\alpha^{(j)},x) \\ &= \prod_{j=1}^{K}\prod_{i=1}^{m} \alpha_i^{(j)}! > 0 \end{aligned}$$

For the above example $\det(X) = 48$. It is straightforward to show that the X matrix for any other model (of size N satisfying the D-condition) is singular. This then shows that in the statistical sense the fan of an echelon design has a single leaf. But from the discussion before Theorem 1 it must also be single leaf in the algebraic sense.

The structure of Z is of some interest. Let $<$ denote the partial order of the exponents vectors corresponding to divisibility. For the example we can draw the partial ordering

$$
\begin{array}{ccccccc}
(0,2) & & & & & & \\
\vee & & & & & & \\
(0,1) & < & (1,1) & < & (2,1) & & \\
\vee & & \vee & & \vee & & \\
(0,0) & < & (1,0) & < & (2,0) & < & (3,0)
\end{array}
$$

Then indexing Z by the $\alpha^{(j)}$'s

$$
Z(\alpha^{(i)}, \alpha^{(j)}) = \begin{cases} P(\alpha^{(i)}, x^{\alpha^{(j)}}) & \text{where } \alpha^{(i)} < \alpha^{(j)} \\ 0 & \text{otherwise} \end{cases}
$$

The theory of Möbius inversion (see for example Constantine, 1987, Chapter 9) can be invoked to show that the inverse Z^{-1} has the same structure. For the example

$$
Z^{-1} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & -1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
-\frac{1}{6} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{6} & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\
-\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \frac{1}{2} & -1 & \frac{1}{2} & 0 \\
\frac{1}{2} & 0 & 0 & 0 & -1 & 0 & 0 & \frac{1}{2}
\end{bmatrix}
$$

It is clear that the value of any parameter $\Phi_{\alpha(j)}$ (estimator in the statistical sense) in the interpolator based on the $\{P(\alpha^{(j)}, x)\}$ namely

$$
Y(x) = \sum_{j=1}^{K} \Phi_{\alpha(j)} P(\alpha^{(j)}, x)
$$

depends only on the values of $Y(x)$ at the special set of design points lower than $\alpha^{(j)}$ in the set of conditions:

$$
\left\{\beta^{(i)} : 1 \leq i \leq K,\ 0 \leq \beta^{(i)} \leq \alpha^{(j)}\right\}
$$

An interpretation is that each $\Phi_{\alpha(j)}$ depends only on the "product model" and design with corner at $\alpha^{(j)}$. Thus for example the point $(2,1)$ gives

$$
\begin{aligned}
\Phi_{(2,1)} &= -\frac{1}{2}Y((0,0)) + Y((1,0)) - \frac{1}{2}Y((2,0)) + \frac{1}{2}Y((0,1)) \\
&\quad -Y((1,1)) + \frac{1}{2}Y((2,1))
\end{aligned}
$$

7. Maximal fan designs

A n-point design is maximal fan in the statistical sense if all the models with n-terms satisfying the (D) condition are identifiable.

Theorem 2 *A maximal fan design with n distinct points in m dimensions always exists.*

Proof. We give two proofs.

(i) The condition $\det(X(E,d)) = 0$ defines a variety in the $n \times m$ space of all coordinates of $d = \left\{x^{(i)} : i = 1, \ldots, n\right\}$ which is of dimension less than $n \times m$. This follows from the linear independence of the monomials in any fan. Let $\mathcal{F}$ be the set of all models satisfying the (D) condition and with n terms. Then the set

$$\cup_{E \in \mathcal{F}} \{d : \det(X(E,d)) = 0\}$$

remains of dimension less then $n \times m$ since $\mathcal{F}$ is finite. Any design d whose coordinates do not lie on this variety (technically any point in the open set which is the union of the complement of the individual varieties $\det(X(E,d)) = 0$) will have all $\det(X(E,d)) \neq 0$. A statistical interpretation is that if d is chosen by any distribution which is continuous with respect to the Lebesgue measure then d will have a maximal fan with probability one.

(ii) The second proof is constructive. Let $\{q_1, \ldots, q_m\}$ be the first m prime numbers $\{1, 2, \ldots\}$. Then define $d = \left\{x^{(i)}\right\}_{i=1}^{n}$ where

$$x_j^{(i)} = q_i^{j-1} \qquad (j = 1, \ldots, m)$$

Then consider the second row ($i = 2$) of a typical $X(E,d)$. The elements of this row are distinct because each entry represents a distinct primes power decomposition. Now all other rows of $X(E,d)$ are distinct powers of this second row that is $X(E,d)$ is of Vandermonde type and therefore has non zero determinant. ∎

For the example $\{1, x_1, x_1^2, x_2, x_1x_2\}$ we have

$$\det(X(E,d)) = \det \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 3 & 6 \\ 1 & 4 & 16 & 9 & 36 \\ 1 & 8 & 64 & 27 & 216 \\ 1 & 16 & 256 & 81 & 1296 \end{pmatrix}$$

By an exhaustive search the authors found that in two dimensions there are 4 maximal fan designs with 3 points based on the integer grid $\{0,1,2\}^2$, specifically the design $\{(0,0),(1,2),(2,1)\}$ and the designs obtained by rotating it anti-clockwise by 90, 180 and 270 degrees. That there are 20 maximal fan designs with 4 points based on the integer grid $\{0,1,2,3\}^2$, 68 maximal fan designs with 5 points based on the integer grid $\{0,1,2,3,4\}^2$ and 584 maximal fan designs with 6 points based on the integer grid $\{0,1,2,3,4,5\}^2$. We conjecture that a maximal fan design on an integer grid of this kind exists for any n and m.

An alternative to seeking combinatorial type maximal fan designs is to appeal to the principals of optimal experimental design. For fixed sample size n one may seek to maximise through choice of d in some region

$$\prod_{E\in\mathcal{F}} \det\left(X^t(E,d)X(E,d)\right) = \prod_{E\in\mathcal{F}} \left[\det\left(X(E,d)\right)^2\right] \tag{3}$$

where $\mathcal{F}$ is the set of all models subject to (D) and with n terms. We call this fan-optimality (in this case fan D-optimality). Provided the design space for d is an open set in $\mathbb{R}^{n\times m}$ then such a design will always exist and be a maximal fan design. Optimal designs for such a weighted product of information matrices have a long history (see Atkinson and Cox, 1974 and Pukelsheim, 1993, Chapter 11). One can also weight different fan elements differently and maximise

$$\prod_{E\in\mathcal{F}} \left[\det\left(X(E,d)\right)^{\alpha_E}\right] \qquad (\alpha_E > 0) \tag{4}$$

Since the sample size is fixed in the present discussion it is not appropriate to consider the continuous optimal design theory (Kiefer-Wolfowitz, 1959) because that theory does not restrict the support of the design. Figure 1 gives designs maximal with respect to (4) on integer grids for sample sizes $n = 3, \ldots, 7$.

We would like to stress that we have considered maximal fan designs in a statistical sense. Let us rename minimal and maximal fan design in the statistical and algebraic sense by m_a, m_s, M_a and M_s respectively. We have $m_s \subseteq m_a \subseteq M_a \subseteq M_s$ and that echelon designs are both statistically and algebraically minimal fan. Recall however that there exists an isomorphism between any model identifiable in a statistical sense and any model identifiable in an algebraic sense. It is certainly true that some designs are maximal fan design in both the statistical and the algebraic sense but it remains a conjecture that such designs exist for all sample sizes and dimensions.

ACKNOWLEDGMENTS

This work was first presented at the CoCoAV conference at Herstmonceux Castle, 3-6 of June, 1997 and has benefited from cordial interaction with the eminent algebraic geometry group in Genoa, Italy and particularly with L. Robbiano and T. Mora. Exchange of early versions of papers led to a correction to an early version of Theorem 1 (see the reference to counterexamples) and the understanding that at the same time the echelon design/model pair had been studied (under the name distractions) in the context of inversion problems by M.P. Rogantin and L. Robbiano. We are grateful for their generosity in bringing additional algebraic insight to our statistical approach to this rapidly developing area. This work was supported by the UK Engineering and Physical Sciences Research Council.

References

[1] Adams,W. W. and Loustaunau, P. (1994). *An Introduction to Gröbner Bases.* Graduate Studies in Mathematics, AMS.

[2] Atkinson, A.C. & Cox, D.R. (1974). Planning experiments for discriminating between models, J. Royal Stats. Soc. B 36:321–348 (with discussion).

[3] Caboara, M. & Riccomagno. E. (1997) An algebraic computational approach to the identifiability of Fourier Models. (Under revision).

[4] Caboara, M. & Robbiano, L. (1997) Families of Ideals in Statistics. ISSAC '97 Conference, Maui (Hawai) (to appear).

[5] Capani, A., Niesi, G. & Robbiano, L. (1995) CoCoA, *a system for doing Computations in Commutative Algebra*. Available via anonymous ftp from `lancelot.dima.unige.it`.

[6] Char, B., Geddes, K., Gonnet, G., Leong, B., Monogan, M & Watt, S. (1991) *MAPLE V Library Reference Manual*. Springer-Verlag, New York.

[7] Constantine, Gregory M. (1987). *Combinatorial theory and statistical design*. John Wiley & Sons, New York.

[8] Cox, D., Little, J. & O'Shea, D. (1992) *Ideal, Varieties, and Algorithms*. Springer-Verlag, New York.

[9] Fontana, R., Pistone, G. & Rogantin, M-P. (1997) Algebraic analysis and generation of two-level designs. Stats. Appl. 9, 1:15–29.

[10] Holliday, T., Pistone, G., Riccomagno, E. & Wynn, H.P. (1996) The application of computational geometry to the design and analysis of experiments: a case study. (Submitted).

[11] Kiefer, J. & Wolfowitz, J. (1959). Optimum design in regression problem. Ann. Math. Statist. 30:271–294.

[12] Mora, T. & Robbiano, L. (1988) The Gröbner fan of an ideal. J. Symbolic Computation 6, 183–208.

[13] Pistone, G. & Wynn, H.P. (1996). Generalised Confounding with Gröbner Bases. Biometrika 83,3:653–666.

[14] Pukelsheim, F. (1993). *Optimal Design of Experiments*. Wiley, New York.

[15] Robbiano, L. & Rogantin, M.P. (1997). *Factorial designs and distracted fractions*. Preprint Dipartimento di Matematica, U. Genova No. 344. Submitted to Proceedings of the International Conference "33 years of Gröbner basis".

[16] Sturmfels B. *Gröbner Bases and Convex polytopes*. Ed. AMS, Providence R I, 1995.

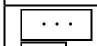
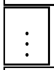
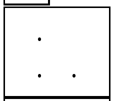
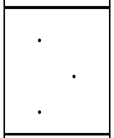
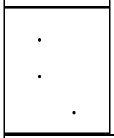
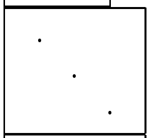
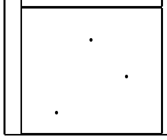
pattern	Example	Fan
	$\{(0,0),(1,0),(2,0)\}$	$\{1, x_1, x_1^2\}$
	$\{(0,0),(0,1),(0,2)\}$	$\{1, x_2, x_2^2\}$
	$\{(0,0),(0,1),(1,0)\}$	$\{1, x_1, x_2\}$
	$\{(0,0),(0,2),(1,1)\}$	$\{1, x_1, x_2\}$ and $\{1, x_2, x_2^2\}$
	$\{(0,0),(2,0),(1,1)\}$	$\{1, x_1, x_2\}$ and $\{1, x_1, x_1^2\}$
	$\{(0,0),(2,2),(1,1)\}$	$\{1, x_2, x_2^2\}$ and $\{1, x_1, x_1^2\}$
	$\{(0,0),(1,2),(2,1)\}$	$\{1, x_1, x_2\}$ and $\{1, x_1, x_1^2\}$ and $\{1, x_2, x_2^2\}$

Table 1: Three points designs and corresponding fans.

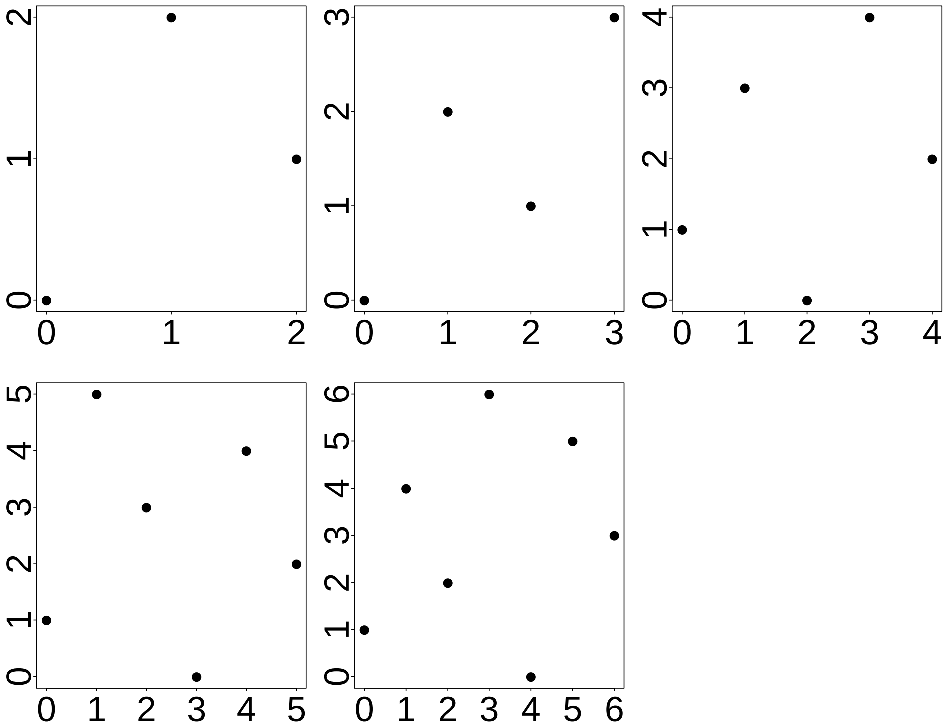


Figure 1: `Two dimensional maximal fan designs with` n `points based on the integer grid` $n \times n$ `(`$n = 3, \ldots, 7$`).`