Report 99-038

The Fan of an Experimental Design

M. Caboara, G. Pistone
E. Riccomagno and H. P. Wynn

ISSN: 1389-2355

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
EURANDOM
Riccomagno@eurandom.nl

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

1 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. 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.

2 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 $d = \{x^{(1)}, \ldots, x^{(n)}\}$ ($x^{(i)}$ all distinct) and a so-called monomial ordering τ, a unique reduced G-basis results and gives a unique set of monomial terms. This set is renamed $Est_{d,\tau}$. Linear combinations of such terms over a suitable coefficient space give identifiable linear models. 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.

The elements of $Est_{d,\tau}$, of which the estimable models 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 $Est_{d,\tau}$. That is to say that $Est_{d,\tau}$ is an *order ideal.* 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. Marginal functionality is the statistical notion corresponding to order ideal. Sometimes models with an order ideal property are called hierarchical models (see Rogantin, 1999).

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 $g_1(x), \ldots, g_r(x)$ the polynomials forming the reduced G-basis with respect to a given term-ordering τ for the design d, 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 $g_j(x)$ with respect to τ (see Section 2). The polynomial r is a linear combination of elements in Est above and is identifiable by the design d.

Since $g_j(x^{(i)}) = 0$ for all $j = 1, \ldots, r$ 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.

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. 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 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 $\{(\pm 1, \pm 1)\}$ there is only one saturated estimable model subject to (D). It is

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

with $Est_{d,\tau} = \{1, x_1, x_2, x_1 x_2\}$. In this case every monomial ordering τ gives the same saturated model.

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

The set of all designs with n-distinct points in m-dimensions can be decomposed into a finite number of non-intersecting classes. 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 of 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.

3 Basic Algebra

In this section we summarize the basic theory. We refer to Pistone and Wynn (1996), Holliday, Pistone, Riccomagno and Wynn (1999) 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 $d = \{x^{(1)}, \ldots, x^{(n)}\} \subset \mathbb{Q}^m$ 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 $\mathbb{Q}$ or $\mathbb{R}$. 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 Ideal(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 (1996). Given a term-ordering τ one can calculate the (unique) reduced G-basis, $G_{d,\tau}$ of the design ideal, Ideal(d). A set of polynomials is a G-basis for a polynomial ideal I and with respect to the term ordering τ if

$$\text{Ideal}(Lt_\tau(g) : g \in G) = \text{Ideal}(Lt_\tau(f) : f \in I)$$

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 Ideal$A$ indicates the ideal generated by the set of polynomials A.

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

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

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

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

and the terms in the remainder precedes the leading terms of the G-basis elements in the ordering τ. That is $Lt_\tau(r) <_\tau Lt_\tau(g_j)$ for all $j = 1, \dots, r$. 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 $k[x]/\text{Ideal}(d)$ as k-vector space. The following is an important but unstated fact within experimental design (see Pistone and Wynn, 1996). Namely *the dimension as a vector-space of $k[x]/Ideal(d)$ 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 Ideal(d), a vector-space basis of the remainder set $k[x]/\text{Ideal}(d)$ is calculated as the set of terms not divisible by any leading term in G_τ. It follows that $k[x]/\text{Ideal}(d)$ 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 $k[x]/\text{Ideal}(d)$ 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 $Est_{d,\tau}$.

Definition 1 *Given a design d and a term ordering τ, the set of monomials $Est_{d,\tau}$ is the standard vector space basis of the quotient space $k[x]/Ideal(d)$. It is computed as the set of monomials not divisible by the leading terms of the τ-Gröbner basis of Ideal(d). When clear by the context, one or both of the suffices in $Est_{\tau,d}$ are suppressed. Sometimes we write $Est_\tau(d)$ or $Est(d)$.*

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

$$\begin{aligned} Id_{d,\tau} \, : \, k[x] &\longrightarrow k[x]/\text{Ideal}(d) \\ f &\longmapsto Rem(f, G) \end{aligned}$$

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 $Rem(f, G_{d,\tau}) = Rem(g, G_{d,\tau})$. That is f and g are in the same equivalence class of the quotient space.

Note at this point that the G-basis G carries all the information about the design.

4 The Design-Est relationship

Theorem 1 *Let d_1 and d_2 be two designs such that $d_1 \subseteq d_2$. Let τ be a term ordering and $Est_\tau(d_1)$ and $Est_\tau(d_2)$ be the estimable set for d_1 and d_2 respectively. Then*

$$Est_\tau(d_1) \subseteq Est_\tau(d_2)$$

Proof. Let Ideal(d_i) be the design ideal for d_i ($i = 1, 2$) and $\{Lt_\tau(\text{Ideal}(d_i))\}$ the set of leading terms of Ideal(d_i) with respect to τ. The following relationships prove the theorem

$$\begin{aligned} d_1 \subseteq d_2 &\iff \text{Ideal}(d_1) \supseteq \text{Ideal}(d_2) \\ &\implies \{Lt_\tau(\text{Ideal}(d_1))\} \supseteq \{Lt_\tau(\text{Ideal}(d_2))\} \\ &\iff Est_\tau(d_1) \subseteq Est_\tau(d_2) \end{aligned}$$

Note that the last step uses the fact that for a design d, $Est_\tau(d)$ is the complementary set of $\{Lt_\tau(\text{Ideal}(d_i))\}$ equivalently of the set of leading terms of the Gröbner basis of Ideal(d_i). The second implication follows from the definition of $\{Lt_\tau(\text{Ideal}(d_i))\}$. ■

Theorem 1 implies that is we add points one by one to a design so we add terms to Est. This can be turned into an algorithm for computing the successive terms of Est which is statistical in flavour.

Theorem 2 *Let τ be a term ordering. Let d_1 be a design, P a design point distinct from d_1 and $d_2 = d_1 \cup P$. Then $Est(d_2) = Est(d_1) \cup x^\beta$ where x^β is*

1. *one of the leading terms of the Gröbner basis of Ideal(d_1) with respect to τ*
2. *the smallest such term with respect to τ for which the design matrix of $Est_\tau(d_2)$ is non-singular.*

Proof. Property 1 holds because the order ideal property of $Est(d_2)$ must be preserved. Now consider Property 2 and let $Est(d_2) = Est(d_1) \cup x^\gamma$ and proceed by contradiction. Thus let β be defined as in the theorem and $\gamma \neq \beta$, $x^\beta <_\tau x^\gamma$. Now x^β remains a leading term of some Gröbner basis element $g(x)$ of d_2 which we can write

$$g(x) = \theta_\beta x^\beta + \sum_{\alpha \in L \cup \gamma} \theta_\alpha x^\alpha$$

where $Est(d_1) = \{x^\alpha : \alpha \in L\}$. But then since $x^\beta <_\tau x^\gamma$ we must have $\theta_\gamma = 0$. But since $g(x) = 0$ on d_2 and $Est(d_1) \cup x^\beta$ is invertible over d_2 all the coefficients of $g(x)$ must be zero, which is a contradiction. ∎

A graded monomial ordering τ is one for which, in addition to the basic definition, $\sum_{i=1}^m \alpha_i < \sum_{i=1}^m \beta_i$ implies $x^\alpha <_\tau x^\beta$; `tdeg` and `deglex` are the common examples (see Cox, Little and O'Shea, 1996). We show that for a fixed design d and any graded monomial orderings τ, $Est_\tau(d)$ has the same number of terms of a fixed degree.

Theorem 3 *Let d be a design and τ any graded ordering then the number of terms in $Est_\tau(d)$ of a given order s is a function $h(s)$ not depending (otherwise) on the ordering.*

Proof. This makes use of the idea of a Hilbert function $H_I(s)$ of an ideal I. The following equivalent computation of $H_I(s)$ is found in Proposition 3, Section 9.3 of Cox, Little and O'Shea (1996): (i) for all $s \geq 0$, $H_I(s)$ is the number of monomials not in I of total degree less or equal to s. Specialising to $I_\tau(d)$ we see from the definition of Est that the Hilbert function of $I_\tau(D)$, $H_{\tau,d}(s)$ is the number of terms of $Est_\tau(d)$ of degree less or equal to s. But proposition 4 of the same section says that $H_I(s)$ is the same for all graded orderings. Setting $h_I(s) = H_I(s) - H_I(s-1)$ we are done. ∎

4.1 Buchberger-Möller algorithm for design ideals

Theorem 2 leads to a sequential algorithm for finding $Est_{d,\tau}$. If $d_n = \{x^{(1)}, \ldots, x^{(n)}\}$ is the current design we can inspect the design matrix X_{n+1} obtained by testing the addition of the point $x^{(n+1)}$ and candidate Est member x^β. The algorithm is easily understood in tableau from which represents X_{n+1}^t. At each iteration a new column, for x^{n+1} and row for x^β are added. Row reduction can be used to test the rank of X_{n+1}^t.

Such a tableau representation aids the implementation of an algorithm to compute the Gröbner basis of design ideals based on linear manipulation of matrices was introduced by Buchberger and Möller (1982). Abbott, Bigatti, Kreuzer and Robbiano (1999) represent it and extend it to projective spaces.

The working object here is a matrix M whose columns represents design points and the rows represent monomials, the transpose of the design matrix in statistics. The idea is to perform a "row by row" LU decomposition of $M = LUR$ where L is a square unit lower triangular matrix, U is a square upper triangular matrix and R is the unique reduced echelon form of M and to keep track of the various passages. This will be clear with an example.

A finite set of points $x^{(1)}, \ldots, x^{(n)}$ in m dimension and a term-ordering τ are assumed. The monomials x^α are ordered with respect to τ, let us say $1 = x^{\alpha_1}, x^{\alpha_1}, \ldots, x^{\alpha_1}, \ldots$. Then the matrix M is built row by row. The first row is the evaluation of 1 in $x^{(1)}, \ldots, x^{(n)}$ respectively. The second row is the evaluation of x^{α_1} in $x^{(1)}, \ldots, x^{(n)}$. Next the second row is reduced with respect to the first one by a linear combination

$$\left((x^{(1)})^{\alpha_2}, \ldots, (x^{(n)})^{\alpha_2}\right) - a_1 \left((x^{(1)})^{\alpha_1}, \ldots, (x^{(n)})^{\alpha_1}\right)$$

Then construct the third row by evaluating x^{α_3} in $x^{(1)},\dots,x^{(n)}$. and reduce it with respect to the previous two. At the k step one has

$$\left((x^{(1)})^{\alpha_k},\dots,(x^{(n)})^{\alpha_k}\right)-\sum_{i=1}^{k-1}a_i\left((x^{(1)})^{\alpha_i},\dots,(x^{(n)})^{\alpha_i}\right)$$

If the resulting vector is zero, then the polynomial $x^{(k)}-\sum_{i=1}^{k-1}a_ix^{(i)}$ is an element of the Gröbner basis. If it is non zero, then consider the next monomial which is not divisible by any of the $x^{(i)}$ for $i\leq k$. The algorithm stops when all the remaining monomials are to be avoided, as we are considering design ideals.

The reductions performed transform M as

$$M=LUR$$

where R the $n\times n$ upper part is the identity matrix and the remaining rows are all zeros. The identity part encodes the indicator functions of the points, and the zero part the Gröbner basis. This is clarified by an example.

Consider the design $P_1=(0,5,7),P_2=(3,0,2),P_3=(4,1,7)$ and the $\mathtt{tdeg}(z<y<x)$ term-ordering. The first two rows of M are

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1\\ z & 7 & 2 & 7\end{array}$$

which can be reduced to $(7,2,7)-7(1,1,1)$ to give

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1\\ 1-7z & 0 & -5 & 0\end{array}$$

Next

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1\\ z & 7 & 2 & 7\\ y & & 0 & 1\end{array}$$

with reduction

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1\\ z-7 & 0 & -5 & 0\\ y-5-(z-7) & 0 & 0 & -4\end{array}$$

Thus $y-5-(z-7)$ is the indicator function of P_3.

Next

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1\\ z & 7 & 2 & 7\\ y & 5 & 0 & 1\\ x & 0 & 3 & 4\end{array}$$

the last row reduces to $(0,0,0)$ by the transformation $\frac{-2}{10}z-\frac{11}{15}+\frac{1}{3}(x+y)$. This is an element of the sought Gröbner basis, with leading term x. No multiple of x will be further considered.

Next

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1 \\ z & 7 & 2 & 7 \\ y & 5 & 0 & 1 \\ x & 0 & 3 & 4 \\ z^2 & 49 & 4 & 49 \end{array}$$

which is reduced to $(0,0,0)$ by the transformation $\frac{1}{7}z^2 - 3z + 14$. This is an element of the sought Gröbner basis, with leading term z^2. No multiple of z^2 will be further considered.

Next

$$\begin{array}{c|ccc} 1 & 1 & 1 & 1 \\ z & 7 & 2 & 7 \\ y & 5 & 0 & 1 \\ x & 0 & 3 & 4 \\ z^2 & 49 & 4 & 49 \\ yz & 35 & 0 & 7 \end{array}$$

which is reduced to $(0,0,0)$ by the transformation $yz - 7y$. Thus no multiple of yz will be further considered.

Next

$$M := \begin{array}{c|ccc} 1 & 1 & 1 & 1 \\ z & 7 & 2 & 7 \\ y & 5 & 0 & 1 \\ x & 0 & 3 & 4 \\ z^2 & 49 & 4 & 49 \\ yz & 35 & 0 & 7 \\ y^2 & 25 & 0 & 1 \end{array}$$

which is reduced to $(0,0,0)$ by the transformation $y^2 - 6y + z - 2$. Thus no multiple of y^2 will be further considered.

All the remaining monomials are multiples of y^2, yz, z^2, x. Thus the algorithm terminates and the Gröbner basis is $y^2-6y+z-2, yz-7y, \frac{1}{7}z^2-3z+14, \frac{-2}{10}z-\frac{11}{15}+\frac{1}{3}(x+y)$. The indicator functions are $Sep(P_3) = y - z + 2, Sep(P_2) = z - 7, Sep(P_1) = x$. The LUR decomposition of M is

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 7 & 2 & 7 \\ 5 & 0 & 1 \\ 0 & 3 & 4 \\ 49 & 4 & 49 \\ 35 & 0 & 7 \\ 25 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 1 & 0 & 0 & 0 & 0 & 0 \\ 5 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -3/5 & -1 & 1 & 0 & 0 & 0 \\ 49 & 9 & 0 & 0 & 1 & 0 & 0 \\ 35 & 7 & 7 & 0 & 0 & 1 & 0 \\ 25 & 5 & 6 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

5 The fan of a design

We shall need the theory of a fan of a polynomial ideal first in a rather algebraic way, see Mora and Robbiano (1988) and Sturmfels (1995). Given a G-basis, G of the polynomial ideal I with respect to a term-ordering τ the monomial ideal generated by the leading terms of G is called the *initial ideal*

$$Init_\tau(G) = \text{Ideal}(Lt_\tau(g) : g \in G)$$

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

$$Init_\tau(G) = \text{Ideal}(Lt_\tau(g) : g \in I)$$

and thus we also write $Init_\tau(I)$. 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

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

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 $I = \text{Ideal}(d)$ 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.

We specialise these ideas to the present context. Thus when I is a design ideal Ideal(d), E_L is finite and it is precisely $Est_{d,\tau}$ for all $\tau \in L$. Consider as an example

$$d = \{(-1,-1), (-1/2, 1/2), (1/2, -1/2), (1, 1)\}$$

With respect to the $\texttt{tdeg}(x_1 > x_2)$ term-ordering the G-basis is

$$\left\{-3x_1 + 8x_2^3 - 5x_2, x_1^2 - x_2^2, -5x_2^2 + 2 + 3x_1x_2\right\}$$

the corresponding initial ideal is $Init(\text{Ideal}(d)) = \left\{x_1^2, x_2^3, x_1x_2\right\}$ and the corresponding leaf is $Est = \left\{1, x_1, x_2, x_2^2\right\}$. These two sets are represented in Figure 1 with the symbols $\circ$ and $\bullet$ respectively

6 Computing the fan

Ideally one would like to input all the information available on the term-ordering before starting the computation. Such information are generally not enough to determine a term-ordering but only a pre-ordering on the variables, sometimes not even that. Some computer algebra packages allow the user to define a pre-ordering.

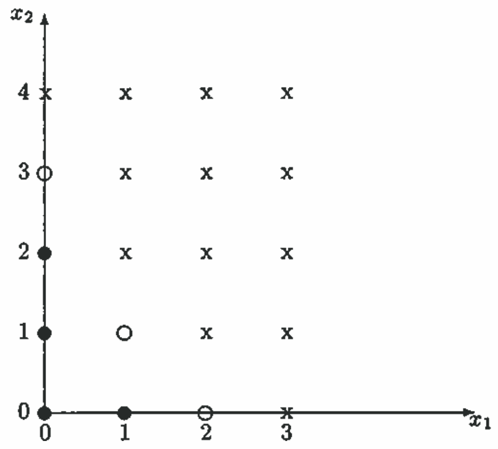


Figure 1: An example of order ideal and initial ideal

The algorithm to compute the fan of ideals receives as input a basis of the design ideal and a pre-ordering, if it is known. At each step it chooses the possible leading terms compatible with the known ordering information, applies the important S-polynomial test (see Cox, Little and O'Shea, 1992, Section 2.6 and below) to check whether a set of polynomials is a G-basis (with respect to the set of term-orderings satisfying the given condition) and creates 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-orderings 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.

Given a term-ordering τ, the S-polynomial of the two polynomials f and g is defined as

$$\text{S-poly}(f,g) = \frac{LCM(Lt(f),Lt(g))}{Lt(f)LC(f)}f - \frac{LCM(Lt(f),Lt(g))}{Lt(g)LC(g)}g$$

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

$$\begin{cases} \text{S-poly}(f,g) &= -3x_2^2x_1 + 2x_1x_2 - 3/2x_2^5 + 7/2x_1^4 = 0 \quad \text{for } x_1 > x_2^2 \\ \text{S-poly}(f,g) &= -\frac{2}{3}x_2^2 + 2x_2 - \frac{2}{3}x_2x_1 = 0 \quad \text{for } x_2^2 > x_1 \end{cases}$$

Their remainders with respect to f and g are

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

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

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

and

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

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.

Theorem 3 has implications for the nature of the sub-fan consisting of all leaves L_τ for graded ordering τ. We use the term *graded fan* for this sub-fan. It says, simply, that every such leaf has the same number of terms of degree s for s positive integer. With a slight abuse of notation we might write the number as $h_d(s)$, where d is the design. It is useful also to think of growing the design sequentially using the algorithmic version of Theorem 2. As we add points to the design for any graded ordering we jump to a higher degree of Est element at the same time.

7 An example: star composite design

Theorem 4 *Let d be the star composite design with central point in m dimensions. To fix notation, assume that the central point is $0 = (0, \ldots, 0)$, the levels of the 2^m full factorial part are $\pm 1\}$ and that the arms are at levels ± 2. Then the fan of d has m leaves. One leaf is (with respect to any term-ordering such that $x_m < x_i$ for all $i = 1, \ldots, m-1$)) is*

$$
\begin{array}{lll}
L = \{ & 1, & \\
& x_i^2, & \\
& x_i x_1^2, & \textit{(for all } i = 1, \ldots, m) \\
& x_1^4, & \\
& \prod_{i \in I} x_i, & \textit{(for all } I \textit{ with } N \textit{ elements and } I \subset \{1, \ldots, d\} \\
& & \textit{and } N = 1, \ldots, m) \qquad \}
\end{array}
$$

The other leaves are obtained by permutation of the variables.

Proof. First notice that the design d and the model L have the same number of elements. Briefly the computation goes as follows: for d, $2^d + 2d + 1$ and for L, $1 + d + d - 1 + \sum_{k=1}^{d} \binom{d}{k} = 2d + 2^d - 1$.

To prove that L is identifiable by d simply run the algebraic procedure for identifiability with respect to any term-ordering for which $x_1 < x_i$ for example $\mathtt{tdeg}(x_1 < \ldots < x_m)$. From the symmetry of d infer that all the models obtained from L by permuting the factors are identifiable. Thus the fan of d includes $m!$ leaves at least.

We are left to prove that there is no other leaf in the fan. A set of equations interpolating the design points is

$$
\begin{array}{ll}
x_1^5 - 5x_1^3 + 4x_1 & \\
x_i x_1 (x_1^2 - 1) & i = 2, \ldots, m \\
x_i x_j (x_1^2 - 1) & i \neq j, i, j = 1, \ldots, m \\
x_1^3 + 3x_1 x_i^2 - 4x_1 & i = 2, \ldots, m \\
x_i^3 + 3x_1^2 x_i - 4x_i & i = 2, \ldots, m \\
x_j (x_i^2 - x_1^2) & i \neq j, i, j = 2, \ldots, m
\end{array}
$$

Let us compute the fan of Ideal(d). By symmetry again we can assume $x_1 < x_i$. Under such assumption each polynomial above has only one possible leading term,

$$x_1^5, x_i x_1^3, x_i x_j x_1^2, x_1 x_i^2, x_i^3, x_j x_i^2$$

respectively. One can check that the equations above form a Gröbner basis using the S-polynomial test or running the algebraic procedure as above. The computation is here omitted. That ends the proof. ∎

8 Interpolation and Statistical fan

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

$$E_L = \{x^{\alpha_1}, \ldots, x^{\alpha_n}\}$$

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

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

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

$$\begin{array}{lll} \circ & \circ & \\ \circ & \circ & \circ \end{array}$$

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. In more than two dimension not much is know on the set of m dimensional order ideals with n terms. Some bounds are know on the cardinality of such sets (see e.g. Bhatia, Prasad and Arora, 1997) but the study of such sets is still an open problem in combinatorics.

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.

For a given number of factors m, $\mathcal{E}(d)$ 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 $E = \{1, x_1, x_1^2, x_2, x_2^2\}$ 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 $k[x]/\text{Ideal}(d)$. For details see Pistone and Wynn (1996), Section 4. Theorem 5 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

$$E(x) = (x^{\alpha_1}, \ldots, x^{\alpha_n})^t$$

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

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

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 $x^\alpha \in \text{Ideal}(Lt_k) \subset \text{Ideal}(Lt_j : j = 1, \ldots, J) = Init$. 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

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

■

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 $Init$

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

Theorem 5 *It there exists a term-ordering τ such that $Lt_j(x)$ is the leading term of $g_j(x)$ for all $j = 1, \ldots, J$, then the set $\{g_1, \ldots, g_J\}$ is the reduced Gröbner bases of Ideal(d) with respect to τ. That is $E \in \mathcal{F}(d)$.*

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 Ideal($g_j(x)$) is the design ideal, Ideal(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 $p(x) \in$ Ideal($g_j(x)$).

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

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

and by evaluating it at the design points. Since S-poly$(g_l, g_k) \in$ Ideal($g_j(x)$), 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. ∎

For the previous example 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

$$\begin{aligned} x_1x_2 &= -x_1^2 + x_2^2/2 + x_1 + x_2/2 \\ x_1^3 &= x_1 \\ x_2^3 &= x_2 \end{aligned}$$

The condition in Theorem 5 is not met since there does not exist a term-ordering such that x_1x_2 is leading term of the first polynomial. Indeed it should simultaneously be $x_1x_2 > x_1^2$ and $x_1x_2 > x_2^2$, that is $x_2 > x_1$ and $x_1 > x_2$ which is not possible in a total ordering.

Theorem 2 leads to a simple updating formula for interpolators. We change to the notation d_n to indicate a n point design and d_{n+1} to denote the same design with one more point.

Corollary 1 *Following the use of Est for interpolation let $p_n(x)$ be the interpolator of values $\left\{(x^{(i)}, Y_i)\right\}_{i=1}^{n}$ based on the design $d_n = \left\{x^{(1)}, \dots, x^{(n)}\right\}$ and $Est_\tau(d_n)$ for some monomial ordering. Let $d_{n+1} = d_n \cup x^{(n+1)}$ where $x^{(n+1)}$ is distinct from d_n. Let $Est_\tau(d_{n+1}) = Est_\tau(d_n) \cup x^\beta$, and in Theorem 2 and let $g_n(x)$ be the element of the Gröbner basis element of $I(d_n)$ which has x^β as leading term. Let $p_{n+1}(x)$ be the interpolator of $\left\{x^{(i)}, Y_i\right\}_{i=1}^{n}$ then*

$$p_{n+1}(x) = p_n(x) + (Y_{n+1} - p_n(x))\,\frac{g_n(x)}{g_n(x_{n+1})}$$

Proof. Since $g_n(x) = 0$ on d_n, $p_{n+1}(x^{(i)}) = p_n(x^{(i)}) = Y_i$ $(i = 1, \dots, n0)$. But at $x^{(n+1)}$, $p_{n+1}(x^{(n+1)}) = Y_{n+1}$ provided that $g_n(x_{n+1}) \neq 0$. But this cannot happen because then $g_n(x) = 0$ on d_{n+1} and the fact that $Est_\tau(d_{n+1}) = x^\beta \cup Est(d_n)$ is non-singular on d_{n+1} would force $g_n(x) \equiv 0$, similarly to the proof of Theorem 2. ∎

9 Minimal fan designs

Definition 2 *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

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

has the single leaf

$$\left\{ \begin{array}{cccc} x_2^3, & x_2^3x_1, & x_2^3x_1^2, & x_2^3x_1^3, \\ x_2^2, & x_2^2x_1, & x_2^2x_1^2, & x_2^2x_1^3, \\ x_2 & ,x_2x_1, & x_2x_1^2, & x_2x_1^3, \\ 1, & x_1, & x_1^2, & x_1^3 \end{array} \right\}$$

The following fundamental class of designs generalises this remark.

Definition 3 *A design $d \subset Z_+^m$ is called a generalised echelon design if for any design point $(d_1, \dots, d_m)$ all points of the form $(y_1, \dots, y_m)$ with $0 \leq abs(y_j) \leq abs(d_j)$, for all $j = 1, \dots, m$ belong to the design d, where $abs(x)$ is the absolute value of x.*

Robbiano and Rogantin (1998) prove that an echelon design is a minimal fan designs. The associated (reduced) Gröbner basis (the same with respect to any term ordering) consists of "distractions" of its leading terms. Let x^α be a leading term then its distraction is the polynomial

$$\prod_{i=1}^{\alpha_1}(x_1 - a_{1,i}) \cdots \prod_{i=1}^{\alpha_1}(x_m - a_{m,i})$$

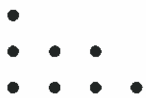
where $a_{i,j}$ are coordinates of the design points.

Another interesting example of minimal fan designs is echelon designs.

Definition 4 *A design $d \subset Z_+^m$ is called an echelon design if for any design point $(d_1, \dots, d_m)$ all points of the form $(y_1, \dots, y_m)$ with $0 \leq y_j \leq d_j$, for all $j = 1, \dots, m$ belong to the design d.*

For example consider the design

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



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

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

Echelon designs are examples of generalised echelon designs. The fan of an echelon design consists of a single *echelon* leaf whose elements are $x_1^{d_1} \dots x_m^{d_m}$ for all $(d_1, \dots, d_m)$ in the echelon design. Thus the design and the model have the same patter.

Definition 5 *Let N be a positive integer. A N-mixture design is the variety defined by*

$$\begin{array}{rcll} \prod_{h=0}^{N}(x_i - h) & = & 0 & \textit{for } i = 1, \dots, m \\ \sum_{i=1}^{d} x_i & = & N & \end{array}$$

Note that one of the equations $\prod_{h=0}^{N}(x_i - h)$ is superfluous and for example we can parametrise with respect to the m-factor.

The projection on any factor of a mixture designs is an echelon design. In particular, with respect to any term-ordering for which $x_d > x_i$ for all i the corresponding leaf is

$$\left\{ x_1^{\alpha_1} \dots x_{m-1}^{\alpha_{m-1}} : \sum_{i=1}^{m-1} \alpha_i \leq N \right\}$$

It follows that the fan of a mixture design has as many leaves as there are factors. And one moves between leaves by substituting $x_j = N - \sum_{i \neq j} x_i$.

9.1 Echelon designs and Newton finite difference formulae

We now give an alternative proof in m dimensions more statistical in style of the minimality of the fan of an echelon design. The same argument applies to generalised echelon designs. For an integer $r \geq 1$ define the univariate polynomial

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

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

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

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

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

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:

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

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

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

In the one-dimensional case interpolation using the univariate polynomials $p(r,x) = x(x-1)\dots(x-r+1)$ leads to Newton's divided difference formula. Thus from the structure of Z and Z^{-1} we have that the parameters are simply the divided differences. For example

$$\begin{aligned} \Phi_0 &= Y[z_0] = Y(z_0) \\ \Phi_1 &= Y[z_0, z_1] = \frac{Y[z_1] - Y[z_0]}{z_1 - z_0} \\ &\vdots \\ \Phi_n &= Y[z_0, \dots, z_n] = \frac{Y[z_1, \dots, z_n] - Y[z_0 - z_{n-1}]}{z_n - z_0} \end{aligned}$$

in the case $z_i = i$ $(i = 1, \dots, n-1)$.

Moreover, the fact that each parameter in the general case arise from the product design/model with corner at the corresponding site means that there is a generalisation of the

Newton formula in this case. Consider again $\Phi_{(2,1)}$ for the above example. Then the product design with corner at $(2,1)$ is $\{(0,0),(1,0),(2,0),(0,1),(1,1),(2,1)\}$ which is

$$\{0,1,2\} \otimes \{0,1\}.$$

The Z-matrix for this design and the model

$$\{1, x_1, x_1^2, x_2, x_1x_2, x_1^2x_2\} = \{1, x_1, x_1^2\} \otimes \{1, x_2\}$$

is

$$Z_{(2,1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 2 & 2 & 1 & 2 & 2 \end{bmatrix}$$

But

$$Z_{(2,1)} = Z_1 \otimes Z_2$$

where $Z_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $Z_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 2 \end{bmatrix}$ are the Z-matrices for the one-dimensional models $\{1, x\}$ and $\{1, x, x(x-1)\}$ respectively. Moreover

$$\begin{aligned} Z_{(2,1)}^{-1} &= Z_1^{-1} \otimes Z_2^{-1} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & -1 & \frac{1}{2} & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 1 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix} \end{aligned}$$

The general formula, which is easily established, is that for a general monomial model term $x^\alpha = x_1^{\alpha_1} \dots x_m^{\alpha_m}$

$$Z_\alpha^{-1} = \bigotimes_{i=1}^{m} Z_{\alpha_i},$$

with obvious notation. This, then, leads to a natural generalisation of Newton's formula to m dimension for echelon designs. Thus, for example,

$$\Phi_{(2,1)} = Y(x_1, x_2)[0,1,2]_1, [0,1]_2$$

where $[\ ]_j$ means differencing in the j-th dimension. In general, again with obvious notation, for x^α the parameter is

$$\Phi_\alpha = Y(x)[0,1,\dots,\alpha_1]_1[0,1,\dots,\alpha_2]_2 \dots [0,1,\dots,\alpha_m]_m$$

All the above extends to non equally spaced grids with distinct levels. A fuller development if given in Riccomagno and Wynn (1999).

10 Maximal fan designs

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

Theorem 6 *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 $d = \{x^{(i)} : i = 1, \dots, n\}$ 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 $\{q_1, \dots, q_m\}$ be the first m prime numbers $\{1, 2, \dots\}$. Then define $d = \{x^{(i)}\}_{i=1}^{n}$ where

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

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 $\{0,1,2,3,4,5\}^2$.

Consider $m = n = 3$. Then the following simple argument shows that no maximal fan design exists on the integer grid $\{0,1,2\}^3$. The full fan is this case consists of the six models:

$\{1, x_1, x_1^2\}$, $\{1, x_2, x_2^2\}$, $\{1, x_3, x_3^2\}$, $\{1, x_1, x_2\}$, $\{1, x_1, x_3\}$ and $\{1, x_2, x_3\}$. For a maximal fan design to exist every one of the two dimensional projections would need to be maximal fan designs for the relevant two variables. That is an interpolating set of polynomials for a maximal fan design is of the type $g_1(x_1), x_i - g_i(x_1)$ where $i = 2, \dots, m$, the degree of the univariate polynomial g_1 is n, the sample size and the value of g_i at the sample points are all distinct, that is $g_i(x^{(j)}) \neq g_i(x^{(k)})$ for all pairs j, k of design points. In algebraic terminology we say that we are in the Shape Lemma structure (see Cohen, Cuypers, Sterk, 1999).

It is clear from this example that equally spaced grids may not be the appropriate support and that more haphazard space-filling configurations are suitable, for example the Latin hypercube sampled designs used in computer experiments or a special constructed sequence in m dimensions as used in numerical integration. The use of prime numbers in *(ii)* above and in the construction of such sequences is a good omen for such a construction. Alternatively one may make a conjecture that for fixed m a maximal fan design exists on the n^m grid for n sufficiently large. Further work on this is in progress.

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

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

where $\mathcal{G}$ is the set of all models subject to (D), in m factors 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

$$\prod_{E \in \mathcal{G}} \left[\det\left(X(E,d)\right)^{\alpha_E}\right] \qquad (\alpha_E > 0) \tag{3}$$

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 2 gives designs maximal with respect to (3) on integer grids for sample sizes $n = 3, \dots, 7$.

It should be noted 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 models identifiable in a statistical sense and models identifiable in an algebraic sense, namely they belong to the same equivalence class in the quotient space and one can move between them by the division, *Rem* operator which acts linearly on the coefficients. 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.

Acknowledgements

This work was first presented at the CoCoA V conference at Herstmonceux Castle, 3-6 of June, 1997 and has benefited from cordial interaction with the algebraic geometry group in Genoa, Italy. Exchange of early versions of papers led to a correction to an early version of Theorem 1 and acknowledgement of the introduction of the echelon designs under the name "distractions". Donka Taneva, as part of an EU "Tempus" grant, performed useful calculations on maximal fan designs. This work was to the largest extent supported by the UK Engineering and Physical Sciences Research Council.

References

[1] Abbott, J., Bigatti, A. Kreutzer, M and Robbiano, L (1999). Computing ideal of points. Journal of Symbolic Computation

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

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

[4] Bhatia, D.P., Prasad, M.A. and Arora, D (1997). Asymptotic results for the number of multidimensional partitions of an integer and directed compact lattice animals. *Journal of Physics A-Mathematical and General*, 30,7:2281-2285.

[5] Buchberger, B. and Möller, H. M. (1982). The construction of multivariate polynomials with preassigned zeros. In Calmet, J. editor, *Proceedings of the European Computer Algebra Conference (EUROCAM '82)*, vol. 144 of *Lecture Notes in Comp. Sci*, 24–31, Marseille, France. Springer.

[6] Caboara, M. & Riccomagno. E. (1998) An algebraic computational approach to the identifiability of Fourier Models. Journal of Symbolic Computation, 29,2:245–260.

[7] Caboara, M. & Robbiano, L. (1997) Families of Ideals in Statistics. *Proceedings of ISSAC '97, Külchlin ed.*, ACM, New York, 404–117.

[8] 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`.

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

[10] Cohen, A. M., Cuypers, H and Sterk, H. (Eds) (1999). *Some Tapas of Computer Algebra.* Springer-Verlag, Berlin.

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

[12] Cox, D., Little, J. & O'Shea, D. (1996) *Ideal, Varieties, and Algorithms.* Springer-Verlag, New York, Second edition.

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

[14] Holliday, T., Pistone, G., Riccomagno, E. & Wynn, H.P. (1999) The application of computational geometry to the design and analysis of experiments: a case study. Computational Statistics 14,2:213–231.

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

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

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

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

[19] Riccomagno, E. (1997). *Algebraic Identiafiability in Experimental Design and Related Fields.* PhD thesis, Department of Statistics, University of Warwick.

[20] Riccomagno, E. and Wynn, H.P. (1999). G-bases, order ideals and a generalised divided difference formula (submitted).

[21] 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".

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

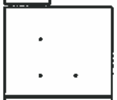
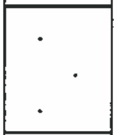
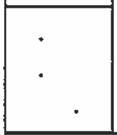
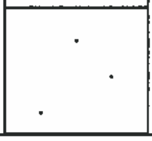
pattern	Example	Fan
	$\{(0,0),(1,0),(2,0)\}$	$\{1, x_1, x_1^2\}$
	$\{(0,0),(0,1),(0,2)\}$	$\{1, x_2, x_2^2\}$
	$\{(0,0),(0,1),(1,0)\}$	$\{1, x_1, x_2\}$
	$\{(0,0),(0,2),(1,1)\}$	$\{1, x_1, x_2\}$ and $\{1, x_2, x_2^2\}$
	$\{(0,0),(2,0),(1,1)\}$	$\{1, x_1, x_2\}$ and $\{1, x_1, x_1^2\}$
	$\{(0,0),(2,2),(1,1)\}$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

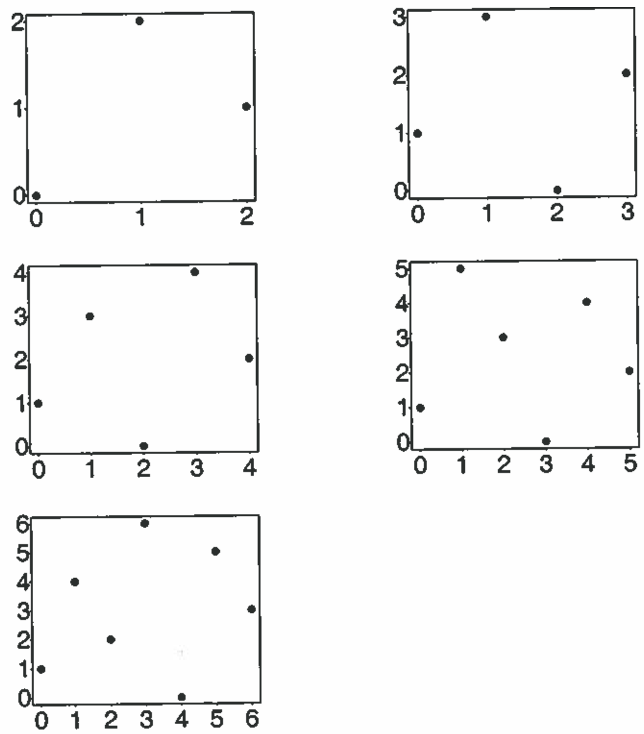


Figure 2: Two dimensional maximal fan designs with n points based on the integer grid $n \times n$ ($n = 3, \ldots, 7$).