The algebraic method in experimental design: Betti numbers and Alexander duality

Hugo Maruri-Aguilar, Eduardo Sáenz de Cabezón, Henry Wynn*

Queen Mary University of London, University of Rioja, London School of Economics

DEMA, Newton Institute, 2 September, 2011

The algebraic approach

Identifications of models from a given design can be achieved by computational commutative algebra techniques (Pistone and Wynn, 1996), with the design $\mathcal{D}$ considered as an algebraic variety (i.e. a solution of a system of polynomial equations).

- Models are identified
- Confounding relations between factors are generalized

This approach has been worked in the context of industrial experimentation (Halliday et al., 1996), mixture designs (Maruri et al., 2006).

The collection of algebraic models (Caboara et al., 1997) has been identified to be of low average degree (Bernstein et al., 2010). This is a characterisation of the centroid of the model. Recent work characterises instead the border of the model (Maruri et al., 2010).

Rings, polynomial division (Cox et al., 1996)

- $\mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_d]$ the polynomial ring.
- The ideal generated by a finite set of points $\mathcal{D} \subset \mathbb{R}^d$ is $I(\mathcal{D}) = \{f \in \mathbb{R}[x] : f(x) = 0, x \in \mathcal{D}\} \subset \mathbb{R}[x]$.
- A term order τ is a total ordering in monomials in $T^d = \{x^\alpha : \alpha \in \mathbb{Z}^d_{\geq 0}\}$, compatible with monomial simplification: i) $x^\alpha \succ 1, \alpha \neq \mathbf{0}$, ii) $x^\alpha \succ x^\beta \Rightarrow x^{\alpha+\gamma} \succ x^{\beta+\gamma}$ for $x^\alpha, x^\beta, x^\gamma \in T^d$.
- A Gröbner basis G_τ is a finite subset of $I(\mathcal{D})$ such that $\langle \mathsf{LT}(g) : g \in G_\tau \rangle = \langle \mathsf{LT}(f) : f \in I(\mathcal{D}) \rangle$.
- For any $f \in \mathbb{R}[x]$, unique remainder r in division of f by $I(\mathcal{D})$

$$f = \sum_{g \in G_\tau} gh + r \tag{1}$$

.

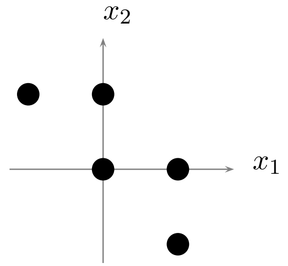
Quotient rings (Cox et al., 1996)

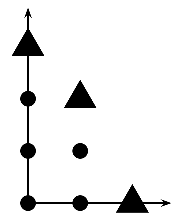
- $\mathbb{R}[\mathcal{D}]$ is the collection of polynomial functions $\phi : \mathcal{D} \mapsto \mathbb{R}$.
- The elements of $\mathbb{R}[\mathcal{D}]$ are in one to one correspondence with equivalence classes of polynomials modulo $I(\mathcal{D})$ and we have an isomorphism $\mathbb{R}[\mathcal{D}] \sim \mathbb{R}[x]/I(\mathcal{D})$.
- A basis for $\mathbb{R}[x]/I(\mathcal{D})$ is given by those monomials that cannot be divided by any of LT(g) for $g \in G_{\mathcal{T}}$.
- The reminder in Eq. (1) is known as the normal form of f (modulo $I(\mathcal{D})$), i.e. NF$(f) = r$.

Generalised confounding (Pistone and Wynn, 1996)

- Design $\mathcal{D}$, n points, d factors.
- Study the $\mathcal{D}$ through the design ideal $I(\mathcal{D}) \subset \mathbb{R}[x]$.
- The support for a model is given by those monomials not divisible by the leading terms of the RGröbner basis $G_\tau \subset I(\mathcal{D})$.

Design $\mathcal{D}$





term order $\tau : x_1 \succ x_2$

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

$$G_\tau = \{\underline{x_1^2} + 2x_1x_2 + x_2^2 - x_1 - x_2, \underline{x_2^3} - x_2, \underline{x_1x_2^2} - x_1x_2 - x_2^2 + x_2\}$$

- Exact polynomial interpolator = saturated regression model.
- Hierarchical polynomial model: staircases.
- Link with aliasing/confounding $f(x) = g(x), x \in \mathcal{D}$.

Examples

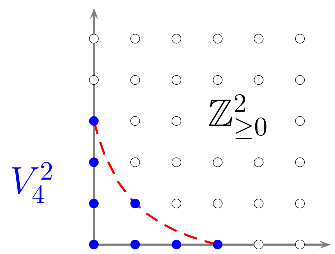
- Factorial design 2^d with levels ± 1. For any term ordering, its design ideal $I(\mathcal{D})$ has Gröbner basis $G_\tau = \{x_i^2 - 1, i = 1, \ldots, d\}$ and identifies the model $\{1, x_1\} \times \cdots \times \{1, x_d\}$
- Indicator function blends naturally to create the ideal of a design fraction, e.g. the indicator $(x_1 - x_2)(x_2 - x_3)$ removes the treatments $\pm(1, -1, 1)$ from the 2^3 design. The fraction $\mathcal{F}$ has six runs and for the standard term order in CoCoA, the model identified is $\{1, x_1, x_2, x_3, x_1x_3, x_2x_3\}$.
- Confounding by normal form: $\mathrm{NF}(x_1x_2x_3) = x_1 - x_2 + x_3$.
- Technique applicable to any design whose points have continuous factors: LH, RSM, optimal. Adaptable to other structures: block, row-column, …

…linear independence with a term order, but much more!

A column selection algorithm (Babson et al., 2003)

- Compute the design model matrix for the set of terms V_n^d.
- Using a term ordering $\succ_w$, order the columns of the matrix.
- Pick the first n columns which form a linearly independent set.

1	x_1	x_2	x_1^2	x_1x_2	x_2^2	x_1^3	$\cdots$
1	0	0	0	0	0	0	
1	1	0	1	0	0	1	
1	0	1	0	0	1	0	
1	1	−1	1	−1	1	1	
1	−1	1	1	−1	1	−1	



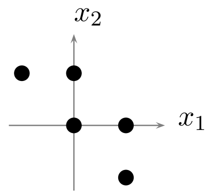
$V_n^d := \{x \in \mathbb{Z}^d_{\geq 0} : \prod_{i=1}^d (x_i + 1) \leq n\}$

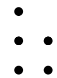
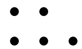
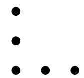
By row elimination, the methodology retrieves G_w for $I(\mathcal{D})$.

It is a variation of the FGLM algorithm for change of basis Faugere *et al.* (1993).

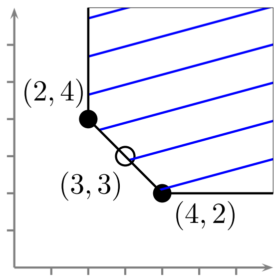
2. The fan of a design

- As we scan over all possible term orders, we obtain the algebraic fan of $\mathcal{D}$, see (Caboara *et al.*, 1997 and Maruri, 2007).
- Not all identifiable hierarchical models belong to the algebraic fan, i.e. $\emptyset \subset A \subseteq S \subseteq \mathcal{C}_{d,n}$.

$\mathcal{D}$ 

$A = \{$  ,  $\}$, $S \setminus A = \{$  $\}$

- The models in A correspond to the vertexes of the state polyhedron $\mathcal{S}(I)$, e.g. we add up the exponent vectors for $L = \{1, x_1, x_2, x_1x_2, x_2^2\}$, $\bar{\alpha}_L = \sum_L \alpha = (2, 4)$.



$$\mathcal{S}(I) = \text{conv}(\bar{\alpha}_L : L \in A) + \mathbb{R}^d_{\geq 0}$$

3. Linear aberration

Linear aberration of model L

- Taking the motivation from the concept of aberration, we want to fill out lower degrees before higher:

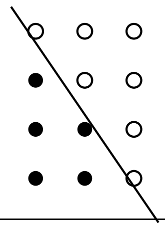
$$A(w, L) = \frac{1}{n} \sum w_i \bar{\alpha}_{L_i}$$

$w_i \geq 0$, $\sum w_i = 1$.

Theo. There exists w such that $A(w, L)$ is minimised by an algebraic model.

Proof. Use LP arguments for the lower boundary of $\mathcal{S}(I)$.

- **Generic** designs minimise $A(w, L)$ over $\mathcal{C}_{d,n}$ and all vectors w.
- For generic designs, algebraic models are corner cut models [**??**]

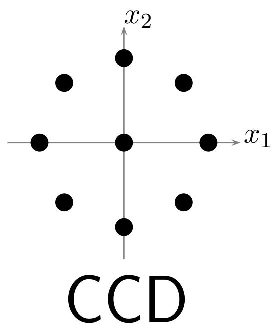
Corner cut model  $\{1, x_1, x_2, x_1x_2\}$ is not corner cut

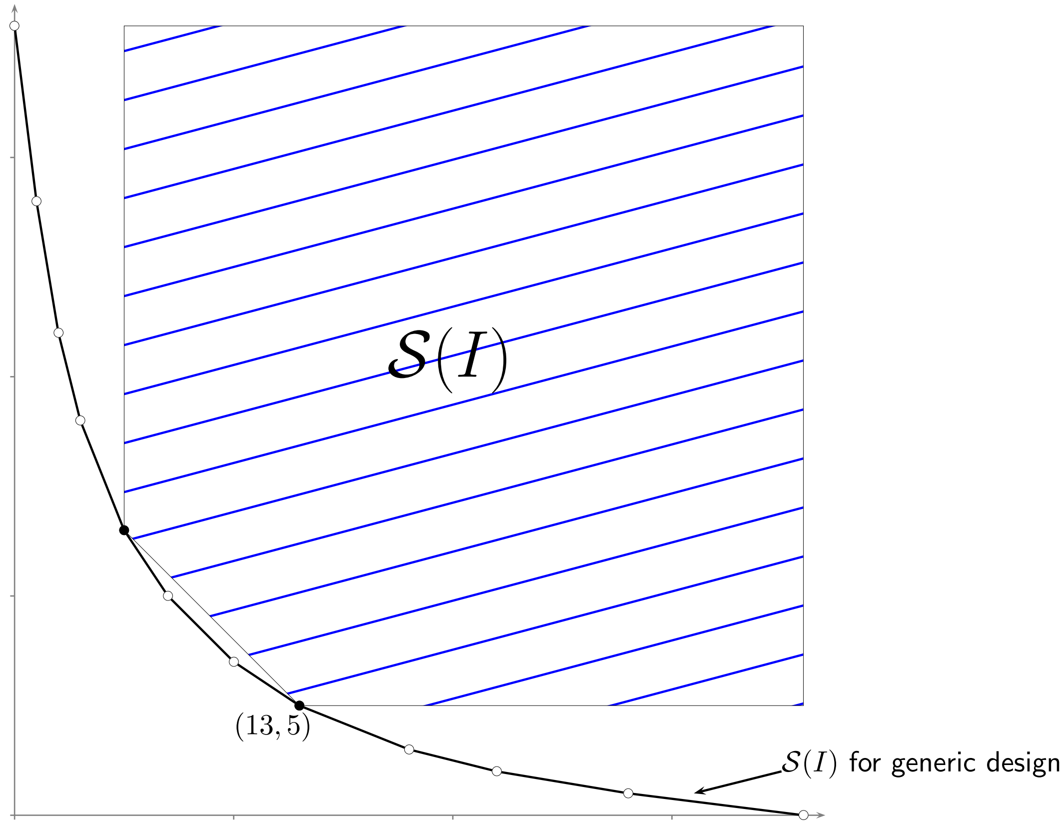
Linear aberration and algebraic models

- The state polytope summarises information about linear aberration, i.e. its vertexes correspond to models that minimise $A(w, L)$ over the set of identifiable hierarchical models S.
- The vertexes of $\mathcal{S}(I)$ correspond to algebraic models A.
- The (minimum) aberration of designs can be compared through their state polytopes.
- However, there may be non-algebraic models on the lower boundary (and thus minimising $A(w, L)$ for some w) or in the interior of $\mathcal{S}(I)$.

Example aberration 1

Central composite design (CCD, Box, 1957) with $d = 2$, $n = 9$ and axial distance $= \sqrt{2}$

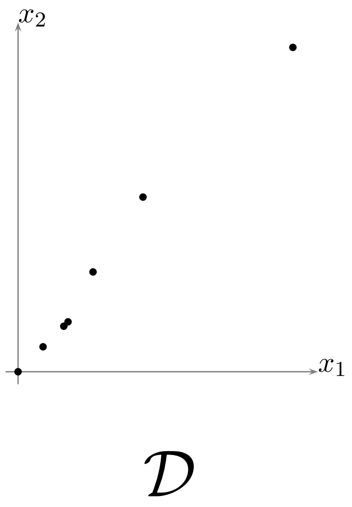


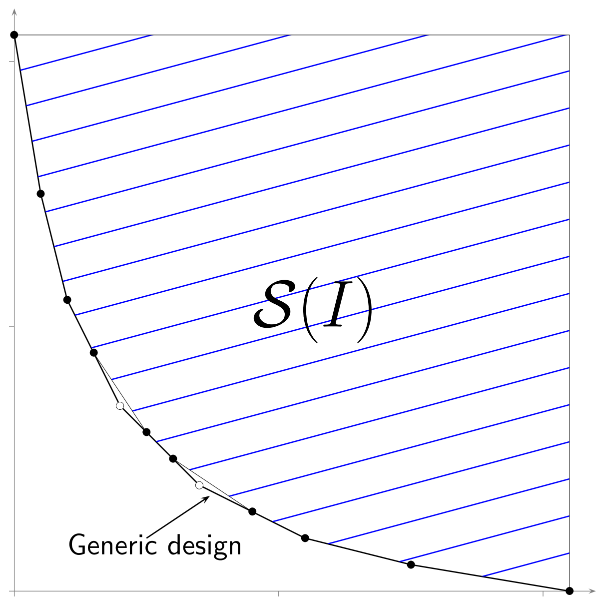


Algebraic $=\{1, x_1, x_1^2, x_1^3, x_1^4, x_2, x_1x_2, x_1^2x_2, x_2^2\}$ and its conjugate

Example aberration 2

Consider $\mathcal{D} = \{(0,0),(1,1),(2,2)(3,4),(5,7),(11,13),(\alpha,\beta)\}$, with $(\alpha,\beta) \approx (1.82997, 1.82448)$ (Onn, 1999)





$\Rightarrow$ The set of algebraic models can be larger in size than the set of corner cut models. However, corner cut models are always of lowest possible degree over all vectors $c \neq 0$.

Minimal aberration (Bernstein et al. (2010) [??,??])

L a model support, $w > 0$ vector of weights, $\sum w_i = 1$

Compute the aberration: $A(w, L) = \frac{1}{n} \sum w_i \bar{\alpha}_{L_i}$

Minimal aberration II (Bernstein et al. (2010) [??,??])

For fixed w, define

$$A^* = \min_{L \in \mathcal{L}} A(w, L).$$

This value is achieved for algebraic models in a generic design.

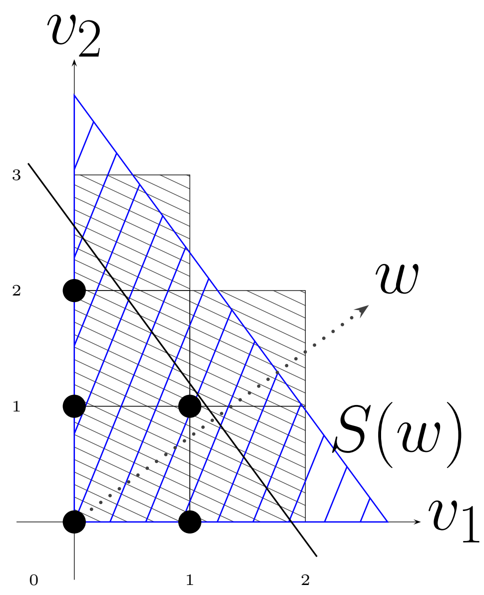
Theorem. In a generic design, minimal aberration A^* obbeys the following bounds:

$$A^+ - 1 \leq A^* \leq A^+ + 1$$

with $A^+ = (nd!)^{\frac{1}{d}} \frac{d}{d+1} g(w)$ and $g(w) = \sqrt[n]{w_1 \cdots w_d}$

Proof.

Define equivalent simplex $S(w)$ ($\int S = n$) and lower and upper cells $\underline{Q}$ and $\bar{Q}$.

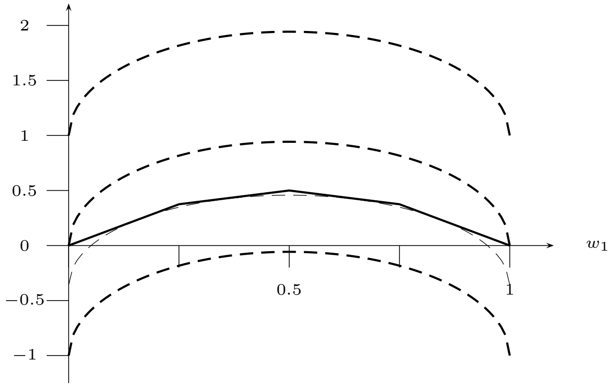


Aberration over S is $E(w^T X)$ with $X \sim U(S)$. The bounds follow from the following inequalities:

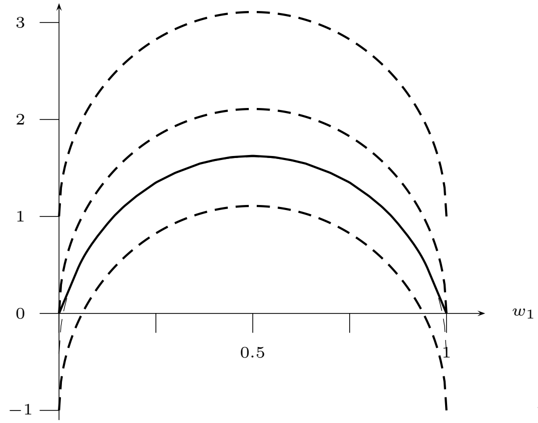
$$A(w, \underline{S}(w)) \leq A(w, \underline{Q}) \leq A(w, S(w)) \leq A(w, \overline{Q}),$$

and noting that $A(w, \underline{S}(w)) = A(w, S(w)) - 1$. We denote A^+ for $A(w, S(w))$.

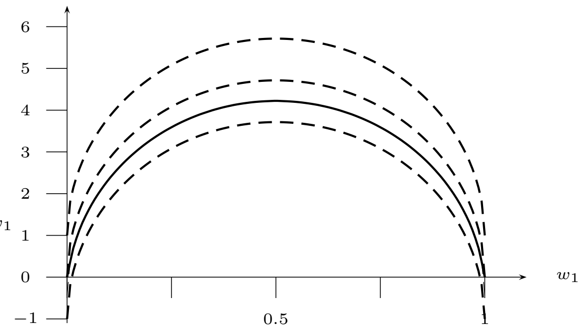
Example: minimal aberration $d = 2$



$n = 4$



$n = 20$



$n = 100$

The graphs include A^*, A^+ and $A^+ \pm 1$ of Theorem; also approximate $\tilde{A}$ is shown.

4. The border of the model

For a polynomial ideal I, the *ideal of leading terms* $LT(I)$ is the monomial ideal generated by the leading terms of polynomials in I. In what remains of the talk, I denotes a monomial ideal.

The Hilbert function $H_{R/I}$ counts the number of monomials not in I, for each degree. E.g. for $I =< x^3, y^2 >$ we have $H_{R/I} = 1, 2, 2, 1$.

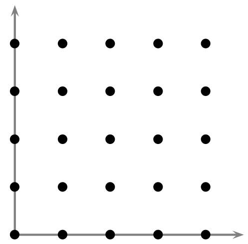
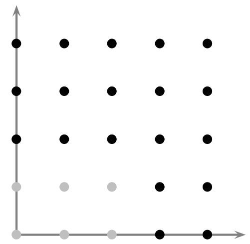
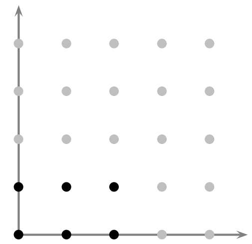
The generating function for those terms is the *multigraded Hilbert series* $HS_{R/I}$. In the current example $HS_{R/I} = 1 + x + y + x^2 + xy + x^2y$.

We can however compute the Hilbert function and the Hilbert series for terms in I, and we have the following equality for HS:

$$\sum_{\alpha \geq 0} x^\alpha = \sum_{\alpha \in I} x^\alpha + \sum_{\alpha \notin I} x^\alpha$$

$$\frac{1}{\prod_{i=1}^{d}(1 - x_i)} = HS_I + HS_{R/I}$$

The table of Betti numbers describes the composition of (numerators of) sums, i.e. entry (i, j) contains number of terms of degree $i + j$.

 $=$  $\bigcup$ 

$$\frac{1}{(1-x)(1-y)} = \frac{y^2+x^3-x^3y^2}{(1-x)(1-y)} + (1 + x + y + x^2 + xy + x^2y)$$

	0
0:	1
Tot:	1

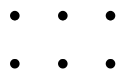
	0	1
2:	1	-
3:	1	-
4:	-	1
Tot:	2	1

	0	1	2
0:	1	-	-
1:	-	1	-
2:	-	1	-
3:	-	-	1
Tot:	1	2	1

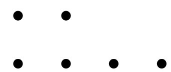
$$\frac{1}{(1-s)^2} = \frac{s^2+s^3-s^5}{(1-s)^2} + (1 + 2s + 2s^2 + s^3)$$

CoCoA code for computing Hilbert function H, Hilbert series HS and Betti table.

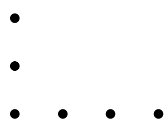
```
Use T::=Q[x,y];
I:=Ideal(x^3,y^2);
I;
Hilbert(T/I);
HilbertSeries(T/I);
BettiDiagram(I);
BettiDiagram(T/I);
```



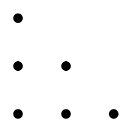
	0	1
2:	1	-
3:	1	-
4:	-	1
Tot:	2	1



	0	1
2:	1	-
3:	1	1
4:	1	1
Tot:	3	2



	0	1
2:	1	-
3:	1	1
4:	1	1
Tot:	3	2



	0	1
3:	4	3
Tot:	4	3

We can only compare tables of Betti numbers for ideals that have the same Hilbert function. In such case, the following theorem guarantees existence of an ideal (called *lex segment ideal*) that attains maximal Betti numbers.

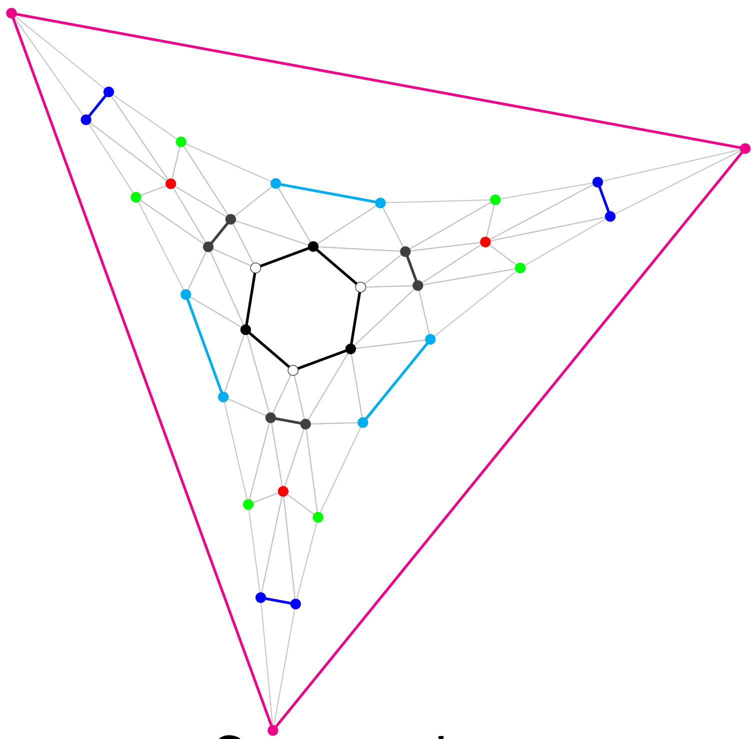
Theorem[Bigatti-Hulett] Let $I \subseteq R$ and L be the lex ideal such that $H_{R/I} = H_{R/L}$. Then $\beta_{i,j}(R/L) \geq \beta_{i,j}(R/I)$ for all i, j.

We want to describe the (borders of) models in the algebraic fan of a design, and we first study generic designs. For generic designs, the models in the algebraic fan are corner cuts.

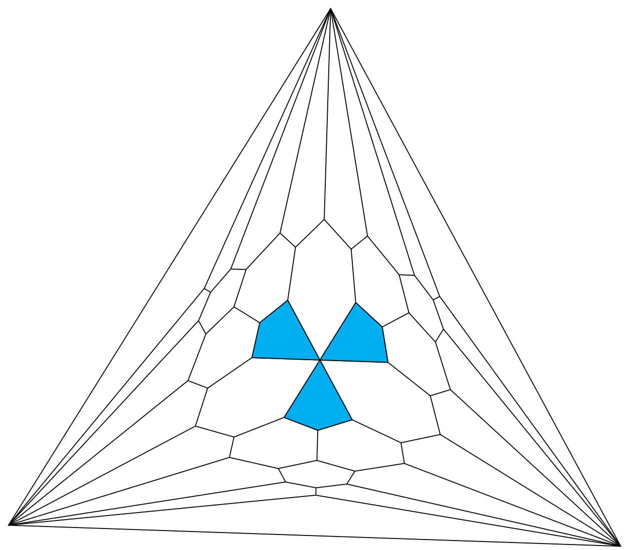
In two dimensions, the relation between ideals generated by corner cuts staircases and lex-segment ideals is one-to-one. In other words, the models in the algebraic fan are precisely those that maximise Betti numbers (Maruri *et al.*, 2011).

For more than two dimensions, the relationship between lex-segment ideals and ideals which are the complement of corner cut staircases and is not necessarily one-to-one. For some cases, the ideal of a corner cut model may attain maximal Betti number despite not being a lex-segment ideal, while in other cases it may not attain maximal Betti numbers.

Example. Generic design $n = 7, d = 3$. Fan with 36 models, of which 3 are not lex segment, yet they still attain maximal Betti numbers.

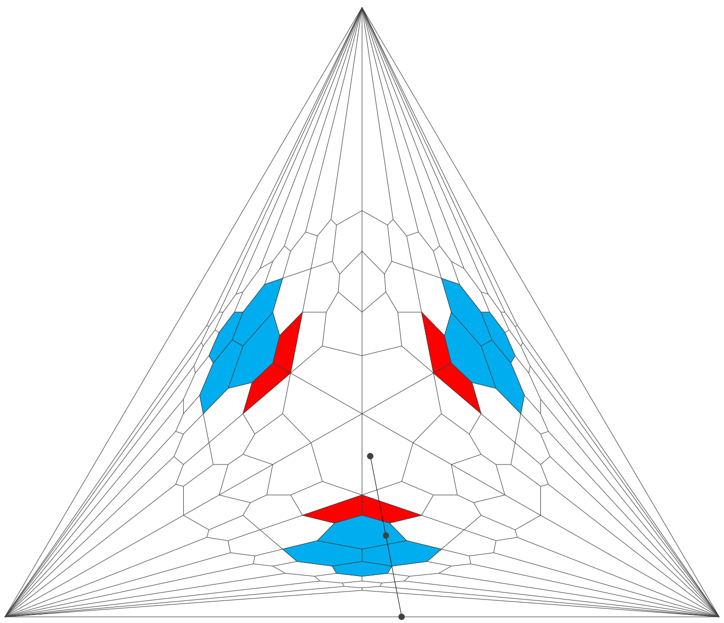


State polytope



Fan (dual of SP)

We give some (neccesary) conditions for identification of corner cut models whose ideals are lex segment ideals. The construction builds a trajectory $w = (C - \gamma^{d-1}, C - \gamma^{d-2}, \ldots, C - 1)$ in the dual of the state polytope. The initial and final points of the trajectory are lex-segment ideals.



Trajectory ($\gamma = 4$) and the normal fan of the corner cut polytope, $d = 3, n = 12$.

5. Betti numbers, the squarefree case

Design whose points are fractions of factorial design with two levels 2^d have important role in experimentation.

Fractions can be selected to satisfy orthogonality conditions and thus not only economy of runs is achieved, but also independent estimation.

Models are (hierarchical) squarefreemodels and thus they can be seen as simplicial complexes.

For a simplicial complex (model) Δ, we construct the *Stanley Reisner* ideal I_Δ, which will be used to analyze the complexity of the border of Δ. Those ideals will be compared against (square free) lex segment ideals.

5. Betti numbers: An example with Plackett-Burman designs

- Small fractions of 2^d with d factors and $n = d + 1$ runs.
- Designs constructed by circular shifts of a generator, available for $d = 7, 11, 15, 19, 23, \ldots$
- PB designs possess a complicated aliasing table, but they have an orthogonal design-model matrix for the linear model with all factors

$$E(y) = \beta_0 + \sum_{i=1}^{d} \beta_i x_i \tag{2}$$

- PB designs are a popular choice for screening in a first stage of experimentation.

We study their algebraic fan and describe the structure of models with the aid of Betti numbers.

PB8

Consider a Plackett-Burman (PB8) design with eight runs, seven factors a, b, c, d, e, f, g and generator $+--+-++$.

a	b	c	d	e	f	g
1	-1	-1	1	-1	1	1
1	1	-1	-1	1	-1	1
1	1	1	-1	-1	1	-1
-1	1	1	1	-1	-1	1
1	-1	1	1	1	-1	-1
-1	1	-1	1	1	1	-1
-1	-1	1	-1	1	1	1
-1	-1	-1	-1	-1	-1	-1

PB8 (cont.)

The design PB8 has 218 models in its fan, which are summarized in Table 2, where representatives of six equivalence classes (up to permutation of factors) are shown.

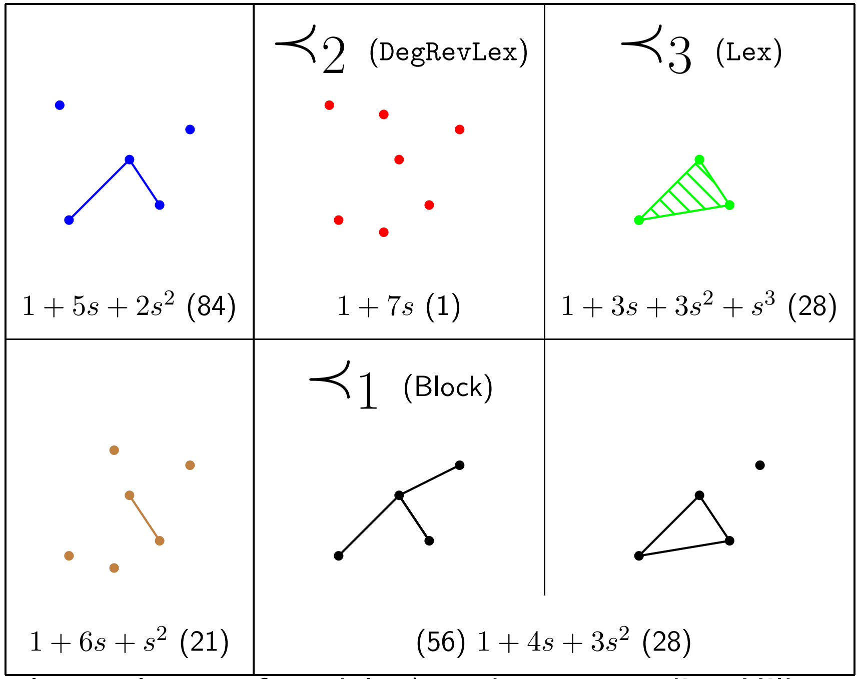
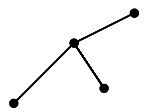


Table 2: Equivalence classes of models Δ and corresponding Hilbert Series for PB8.

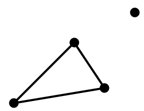
PB8: Comparing models

Betti table | Model

	0	1	2	3	4	5	6	7
0:	1	3	3	1	-	-	-	-
1:	-	7	29	48	40	17	3	-
2:	-	-	6	26	45	39	17	3
Tot:	1	10	38	75	85	56	20	3

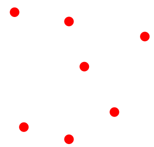


	0	1	2	3	4	5	6	7
0:	1	3	3	1	-	-	-	-
1:	-	7	30	52	47	24	7	1
2:	-	1	10	33	52	43	18	3
Tot:	1	11	43	86	99	67	25	4



Disconnected model $\prec_2$ DegRevLex:

	0	1	2	3	4	5	6
0:	1	-	-	-	-	-	-
1:	-	21	70	105	84	35	6
Tot:	1	21	70	105	84	35	6



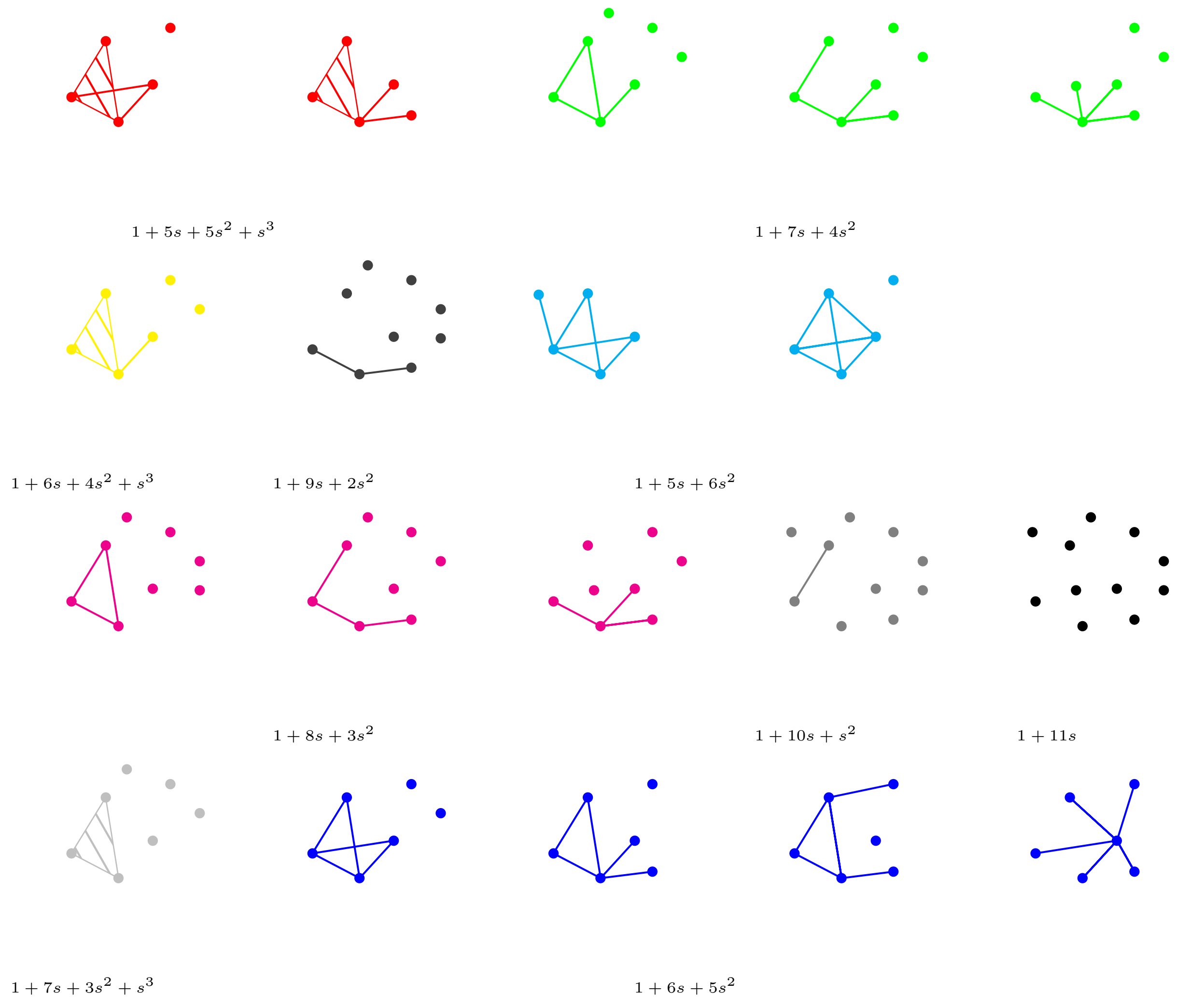
PB12

This design has 12 runs in 11 factors and generator
$++-+++---+-$.

The algebraic fan of PB12 is very complex, showing a rich variety of simplicial models.

Despite its enormous size (around 3×10^5), models have been classified in nineteen classes (up to permutations of variables), which in turn share only ten distinct Hilbert Series.

PB12 (cont.)


1 + 5s + 5s^2 + s^3
1 + 7s + 4s^2
1 + 6s + 4s^2 + s^3
1 + 9s + 2s^2
1 + 5s + 6s^2
1 + 8s + 3s^2
1 + 10s + s^2
1 + 11s
1 + 7s + 3s^2 + s^3
1 + 6s + 5s^2

Comments

Betti numbers provide a description of the model border. Maximal Betti numbers are related to models of low degree (low aberration). However, differently to aberration, we can only compare models (ideals) that have the same Hilbert function.

For generic designs, we found that corner cut models fall into three cases:

a) when they are lex segment and thus they have maximal Betti numbers,

b) they are not lex segment yet they still have maximal Betti,

c) they are not lex segment nor with maximal Betti.

For fractions of 2^d, we have only found classes a) and c). Work in progress...

Main theorem on Alexander duality

Given a monomial ordering $\prec$ a full factorial design F and a fraction $D \subset F$, the bases of the quotient rings of D and the complementary design $\bar{D} = F \setminus D$, with respect to $\prec$, are Alexander dual (relative to F).

$$\bar{L}(D) = L(F) \setminus L(D).$$

Lemma

Let $\{g_i\}$ and $\{h_j\}$ be the G-bases for D and $\bar{D}$, respectively, with respect to $\prec$. Let the leading terms be $LT(g_i) = x^{\alpha^{(i)}},\ LT(h_j) = x^{\beta^{(j)}}$. Let the basis for F be

$$\{x^\gamma, \gamma \in \bigotimes_{i=1}^{d} \{0, \dots n_i - 1\}\}.$$

Then for all i, j

$$\alpha^{(i)} + \beta^{(j)} \in Z^+ \setminus L(F)$$

Proof of lemma

Since g_i is zero on D and h_j is zero on $\bar{D}$, $g_i h_j$ is zero on $D \cup \bar{D} = F$. It follows that $LT(g_i h_j)$ is in the leading term ideal of F which is all monomials $x^\delta, \delta \in Z^+ \setminus L(F)$. But by the properties of monomial orderings.

$$LT(g_i h_j) = LT(g_i) LT(h_j) = x^{\alpha^{(i)} + \beta^{(j)}}.$$

It follows that $\alpha^{(i)} + \beta^{(j)} \in Z^+ \setminus L(F)$.

Proof of theorem

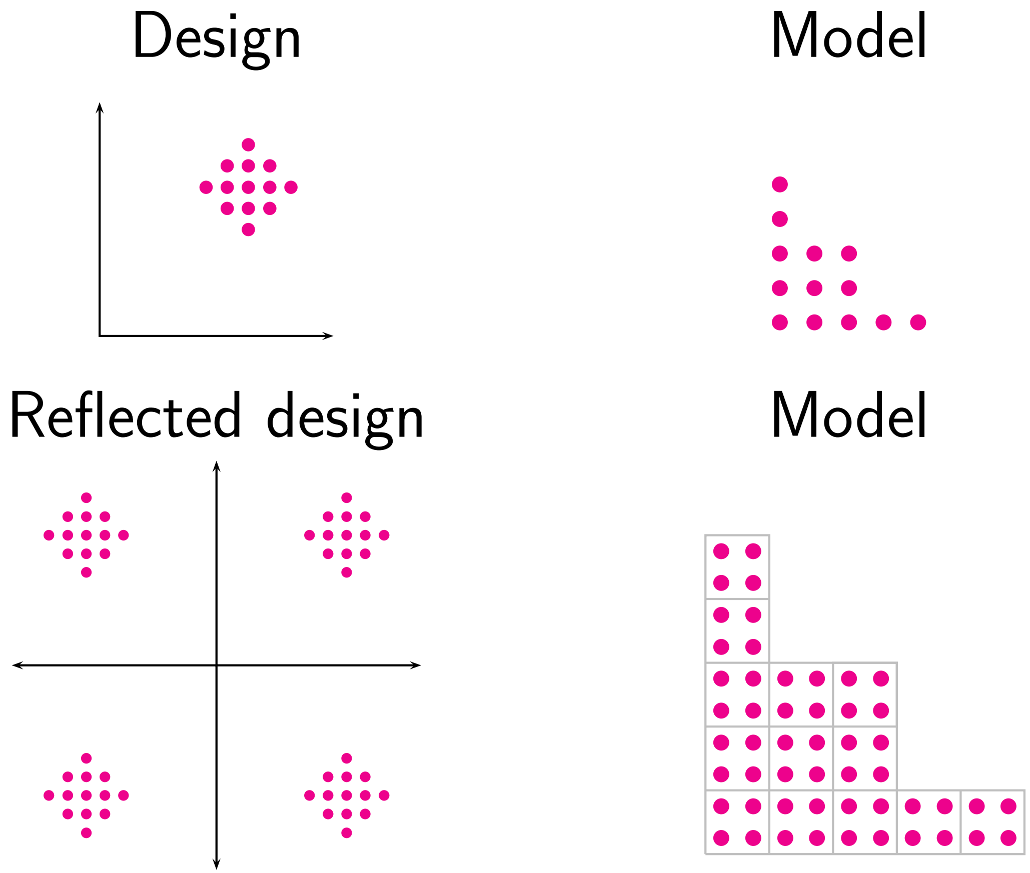
The proof is by contradiction. The cardinalities agree:

$$|L(\bar{D})^+| = |L(\bar{D})| = |L(F) - L(D)| = N - n.$$

Thus, if we suppose (1) is not true then there is a point $\gamma \in L(F)$, neither in $L(\bar{D})^+$, nor in $L(F)$. We can say that $\gamma \notin L(D)$ and (flipping back) $\gamma^* - \gamma \notin L(\bar{D})$. But then γ and $\gamma^* - \gamma$ must be in their respective leading terms ideals. Thus there exist an $\alpha^{(i)} \leq \gamma$ and a $\beta^{(j)} \leq \gamma^* - \gamma$. But then $\alpha^{(i)} + \beta^{(j)} \leq \gamma^*$ which is in $L(F)$, contradicting Lemma 1.

Reflections of an echelon design

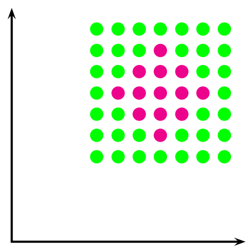
The model identified by reflections of an echelon design remains echelon and its model shares the same structure of that for the original design.



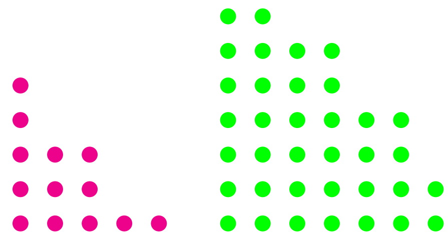
Reflections of an echelon design and Alexander duality

By Alexander duality, the complement of an echelon design remains so, thus its reflections lead to a model stemming from the Alexander dual.

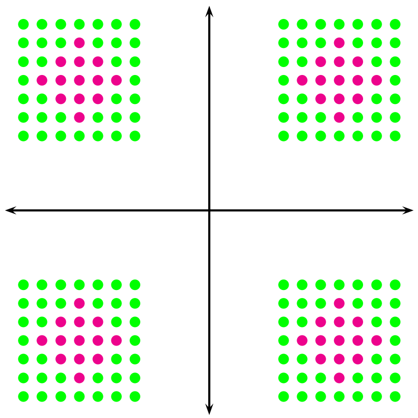
Design (and complement)



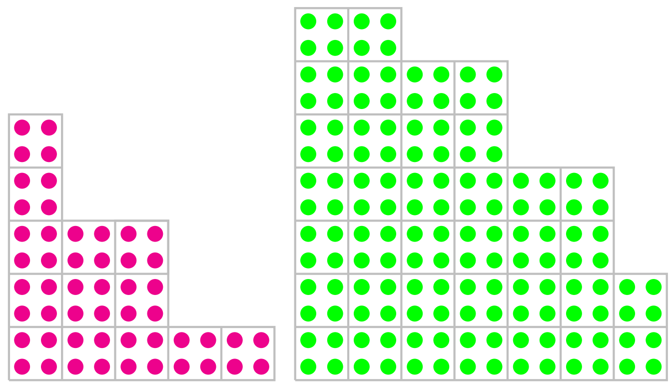
Model (and A. dual)



Reflected designs

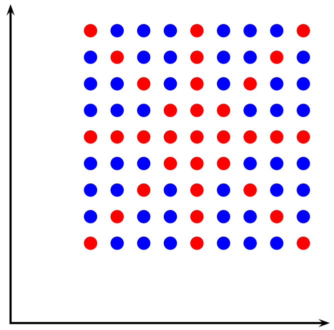


Model (and A. dual)

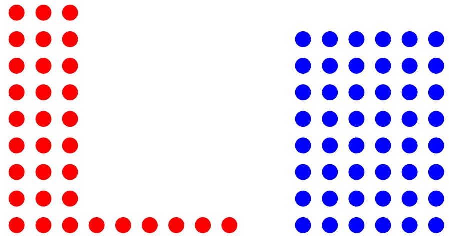


Reflections of a non-echelon design

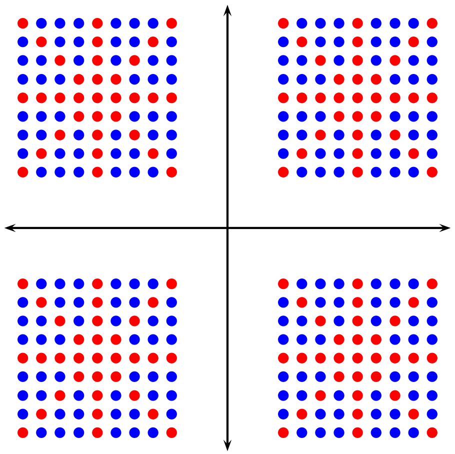
Design (and complement)



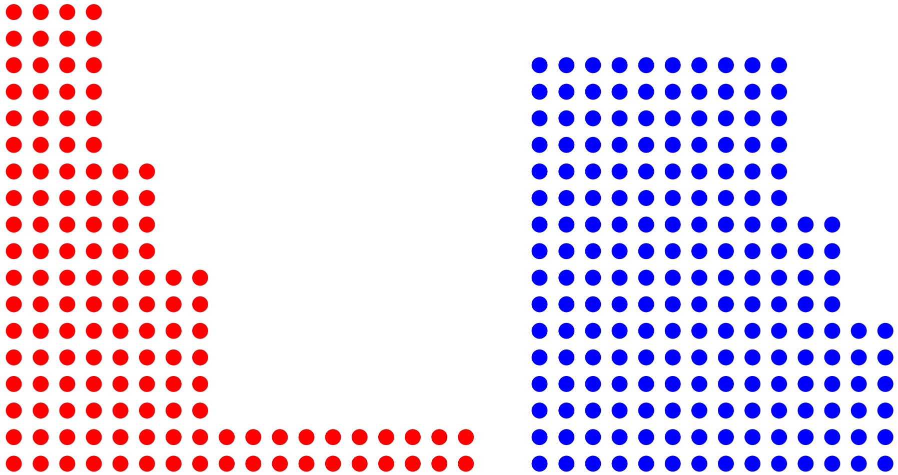
Model (and A. dual)



Reflected design



Model (and A. dual)



Although Alexander duality holds for models, the structure is not inhererited to the model for the reflected design.