Algebraic techniques in design of experiments

Hugo Maruri, Henry Wynn
H.Maruri-Aguilar@lse.ac.uk

London School of Economics

SAMSI, 27th October 2008

Abstract

Techniques from algebraic geometry were introduced in design of experiments by Pistone and Wynn (1996). The techniques they introduced identify models of a low degree and provide a natural extension to basic concepts such as confounding and fractions of designs.

I intend to describe the basic ideas of the algebraic method, specially work relating term orderings with geometrical results. I will also present some of our recent work on aberration which leads to Nyquist-type lower bounds for models identified with algebra.

Order of the talk

1. Algebraic analysis of experiments
   a Generalised confounding
   b Examples
2. Weighted orders and zonotopes
3. The fan of a design
   a A column selection algorithm
   b Hilbert zonotope
4. Linear aberration
   a Examples
   b Bounds for minimal aberration
   c Comparison of designs
5. References

1. Algebraic analysis of experiments

Rings, polynomial division (Cox et al., 1996 [5])

- $\mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_d]$ 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 $I(\mathcal{D})$

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

.

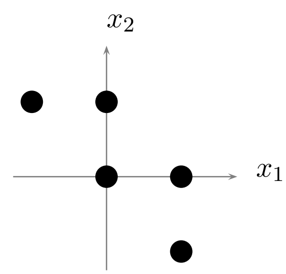
Quotient rings (Cox et al., 1996 [5])

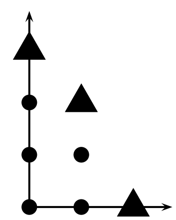
- $\mathbb{R}[\mathcal{D}]$ is the collection of polynomial functions $\phi : \mathcal{D} \mapsto \mathbb{R}$.
- The elements of $\mathbb{R}[\mathcal{D}]$ are in one to one correspondence with equivalence classes of polynomials modulo $I(\mathcal{D})$ and we have an isomorphism $\mathbb{R}[\mathcal{D}]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

Generalised confounding (Pistone and Wynn 1996 [13])

- 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 $\{1, x_1, x_2, x_3, x_1 x_3, x_2 x_3\}$.
- Confounding by normal form: NF$(x_1 x_2 x_3) = x_1 - x_2 + x_3$.
- Technique applicable to essentially, any design whose points have continuous factors: LH, RSM, optimal,...

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

2. Weighted orders and zonotopes

Term orders, weighted orders

- The requirement of a term order (total order in all of T^d) can be relaxed without losing generality. Use instead weighted order.

Def. w-**order** Let $B \subseteq T^d, B \neq \emptyset$, let $w \in \mathbb{Z}^d_{\geq 0}, w \neq \mathbf{0}$. For $x^\alpha, x^\beta \in B$ we say $x^\alpha \succeq_w x^\beta$ if $w \cdot \alpha \geq w \cdot \beta$.

- If $B = T^d$ then for any w, $\succeq_w$ is a partial order, i.e. there are ties among monomials.
- Even for finite B, it is easy to find w such that $\succeq_w$ is only partial ordering. e.g. $B = \{x_1, x_2\}$ take $w = (1, 1)$.
- But careful selection of w wrt B allows to use w-orders.

Theo. [10] Let $B \subset T^d$ be finite; let w be such that w is not orthogonal to any of $\{\alpha - \beta : x^\alpha, x^\beta \in B\}$. Then w-order is total ordering $\succ_w$ in B.

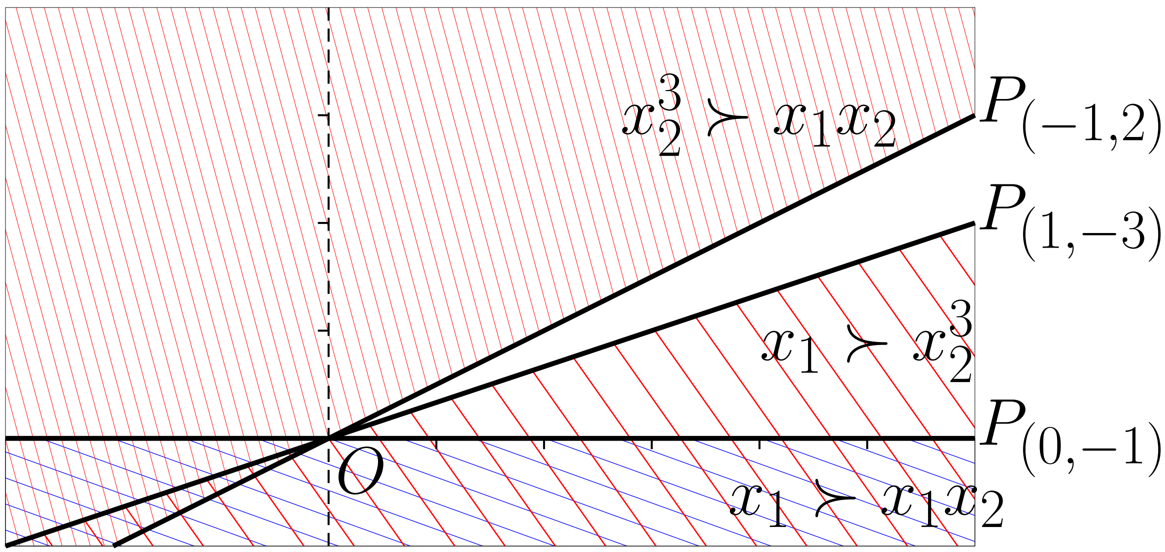
Equivalence classes of vectors [10]

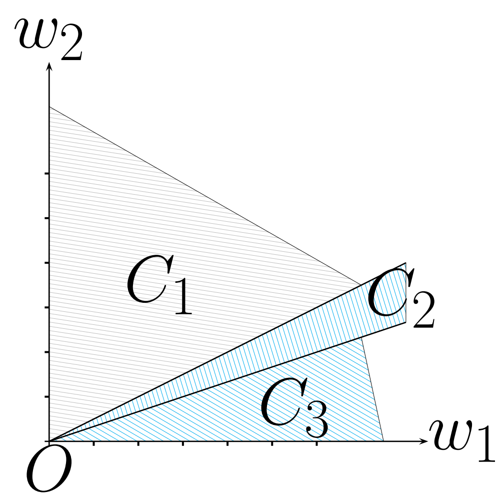
Def. Let $B \subset T^d$ be finite; let $\succ_{w_1}, \succ_{w_2}$ be total orderings in B. We say $w_1 \sim w_2$ when $x^\alpha \succ_{w_1} x^\beta$ iff $x^\alpha \succ_{w_2} x^\beta$ for all $x^\alpha, x^\beta \in B$.

For fixed B, the set of all weighing vectors that create $\succ_w$ is the equivalence class of w. It is a polyhedral cone (intersection of halfspaces) intersected with the positive integer lattice.

Theo. The central hyperplane arrangement constructed with all pairwise differences between exponents of monomials in B partitions the positive integer lattice into cones, the interior of which corresponds to equivalence classes $\sim$.

Example: $B = \{x_1, x_1x_2, x_2^3\}$





Equivalence classes and Minkowski sums [10]

Def. Let V be a finite set of vectors in $\mathbb{R}^d$. The **zonotope** of V is the Minkowski sum

$$Z(V) = \sum_{v \in V} [-v, v], \tag{2}$$

where $[-v, v]$ is the line segment between $-v$ and v.

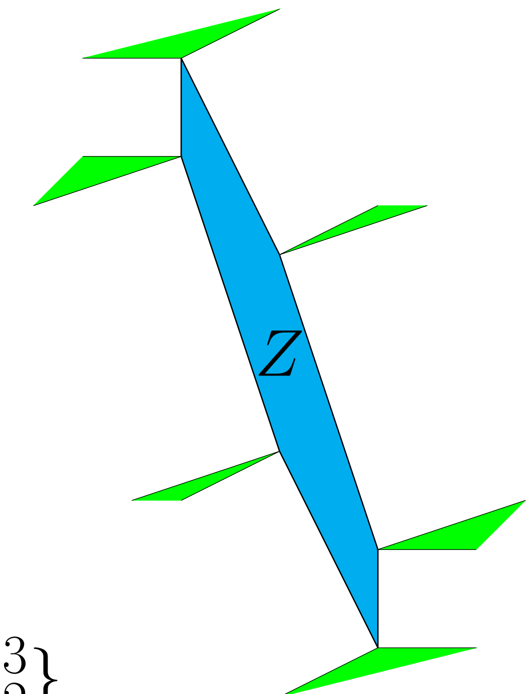
Theo. Let $B \subset T^d$ be finite, let $D = D(B) = \{\alpha - \beta : x^\alpha, x^\beta \in B\}$ and let $Z(D)$ be the zonotope of D. Then the restricted *normal fan* of $Z(D)$ partitions the positive integer lattice into the cones $\sim$.

Universal set of weighing vectors $W_+(B)$

For $B \subset T^d$ finite, there is a one to one correspondence between

Equivalence classes of ordering vectors $\leftrightarrow$ Cones in the first orthant of the normal fan of the zonotope of $D(B)$.

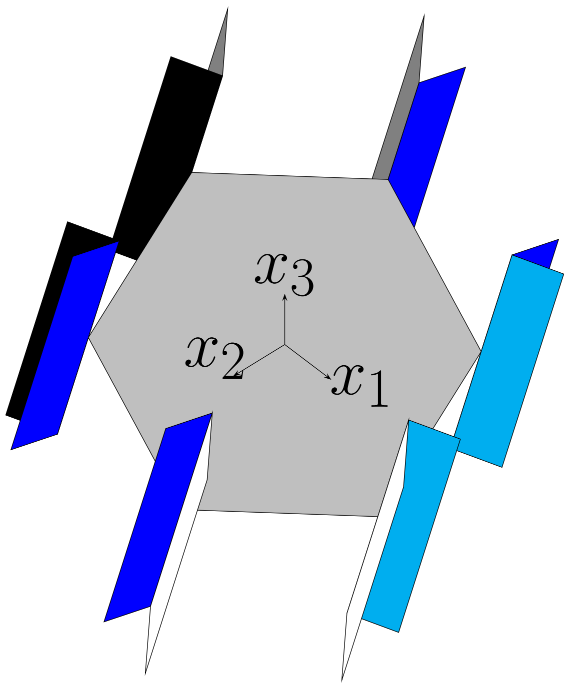
Set of representatives: $W_+(B)$



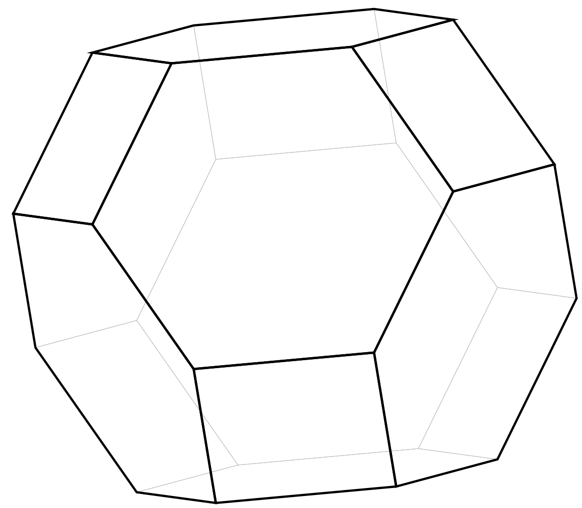
$$B = \{x_1, x_1x_2, x_2^3\}$$

Permutahedron

When $B = \{x_1, \ldots, x_d\}$ then the associated zonotope is the permutahedron Π_{d-1}.



Π_2



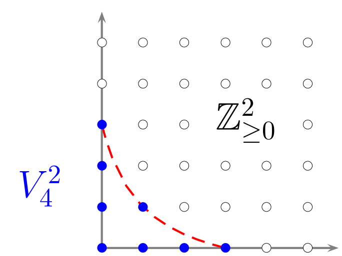
Π_3

3. The fan of a design

A column selection algorithm (Babson et al. 2003 [1])

- 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	



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

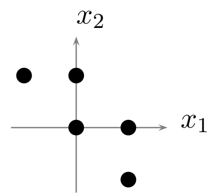
By row elimination, the methodology retrieves G_w for $I(\mathcal{D})$.

It is a variation of the FGLM algorithm for change of basis [6].

The fan of a design

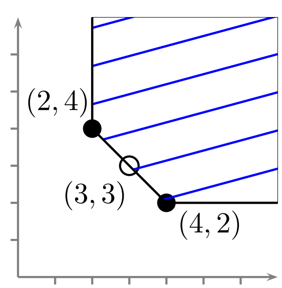
- As we scan over all possible term orders, we obtain the algebraic fan of $\mathcal{D}$, see [3][10].
- Not all identifiable hierarchical models belong to the algebraic fan, i.e. $\emptyset \subset A \subseteq S \subseteq \mathcal{C}_{d,n}$.

$\mathcal{D}$



$$A = \{ \ \vdots\!: \ , \ ::. \ \}, S \setminus A = \{ \ \vdots.. \ \}$$

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



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

Constructing the fan

One to one correspondence between:

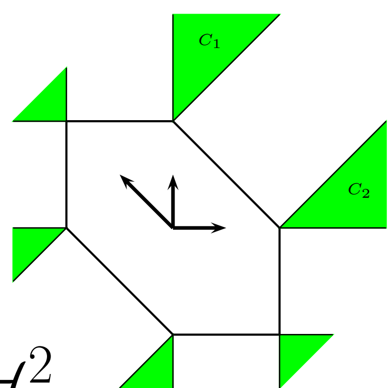
Models retrieved by algebra $\leftrightarrow$ vertexes of state polytope

1. Reverse search and change of basis. Particular construction for every ideal $I(\mathcal{D})$. Software available (Gfan).
2. Hilbert zonotope [12,1] $\mathcal{H}_n^d$. Universal construction for any $I(\mathcal{D})$, depends only on d, n.
   Idea behind $\mathcal{H}_n^d$: to generate the universal set of w-ordering vectors it is sufficient to order the set $V_n^d = \{$union of all staircases with n elements in d undeterminates$\}$.
   $\Rightarrow$ run the column selection algorithm with a universal set of w-orders.

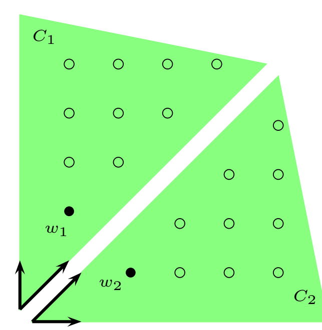
Hilbert zonotope

Hilbert schemes are algebraic varieties that parameterize families of ideals in polynomial rings i.e.Hilb$_2^2$ consists of all $I \subset R[x]$ for which $R[x]/I$ has dimension 2 as $\mathbb{R}$-vector space.

Theo. [1]: $\mathcal{H}_n^d$ is a refinement of the state polytope of every member of Hilb$_n^d$.

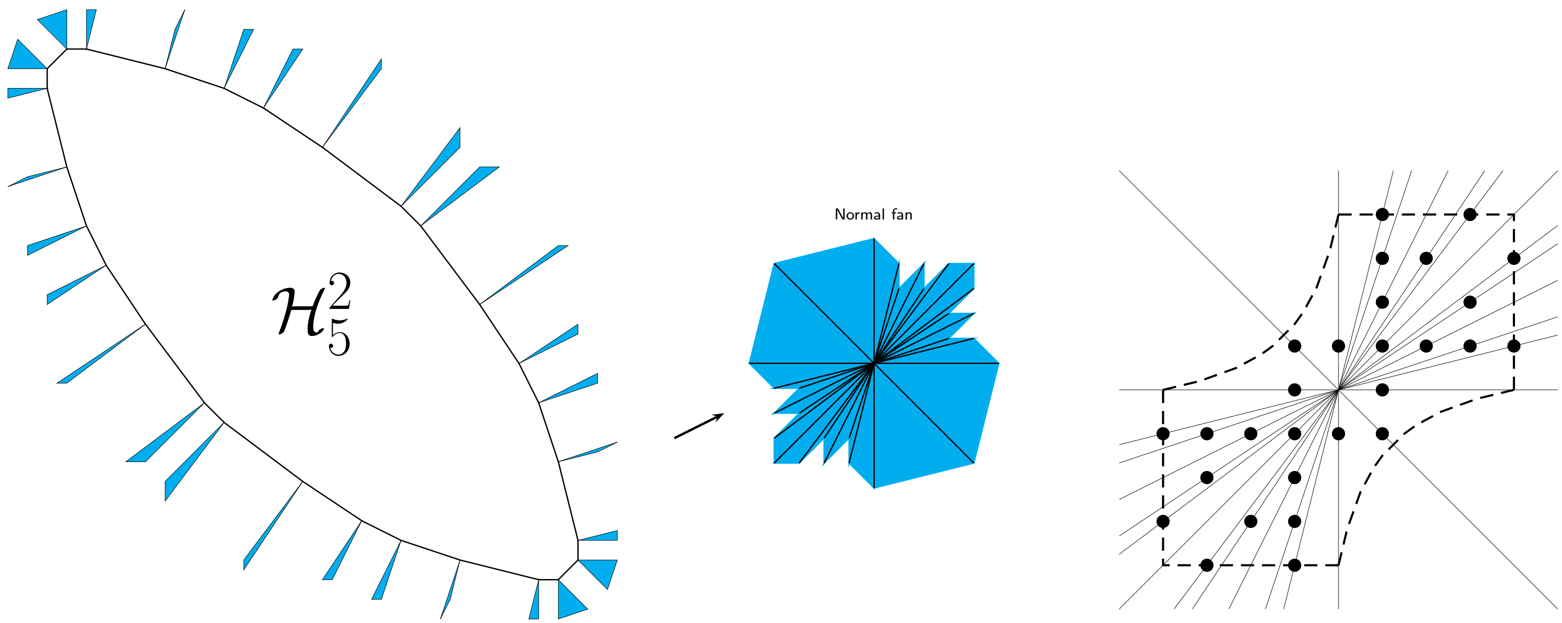


$\mathcal{H}_2^2$



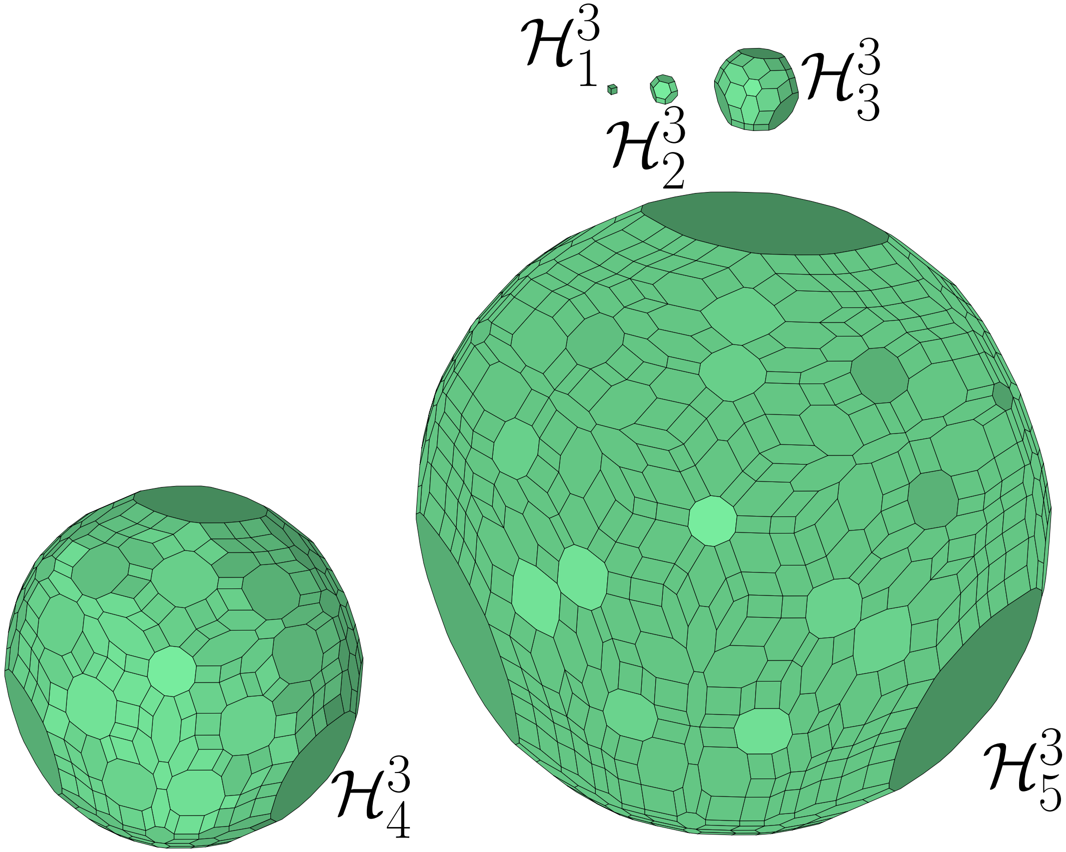
Hilbert zonotope, $d = 2$

Theo. [9] The cones in of $\mathcal{N}(\mathcal{H}_n^2) \bigcap \mathbb{R}^2_{\geq 0}$ are generated by the directions of $[0, n-1]^2$.



Hilbert zonotope, higher dimensions

For $d > 2$, the complexity of $\mathcal{N}(\mathcal{H}_n^k) \bigcap \mathbb{R}^d_{\geq 0}$ is not captured with the directions of $[0, n-1]^d$.



Any hints?

4. Linear aberration

Linear aberration

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

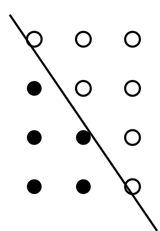
$$A(w, L) = \frac{1}{n} \sum w_i \bar{\alpha}_{L_i}$$

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

Theorem: Any algebraic model minimises some $A(w, L)$.

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 [12]

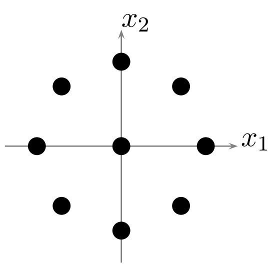
Corner cut model  $\{1, x_1, x_2, x_1 x_2\}$ is not corner cut

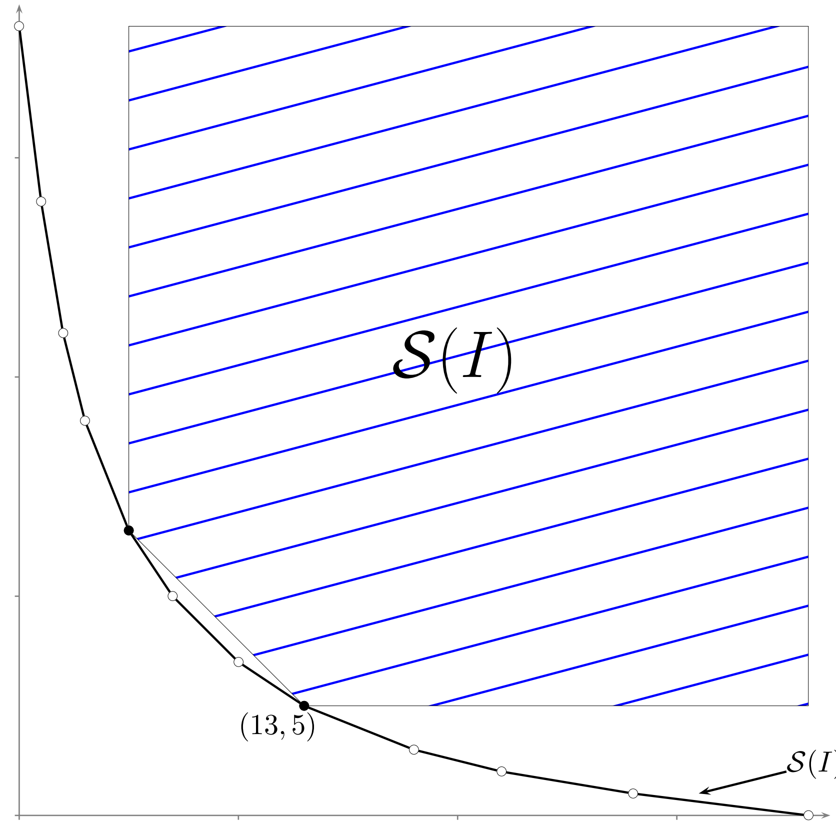
Linear aberration and algebraic models

- 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 $= \sqrt{2}$

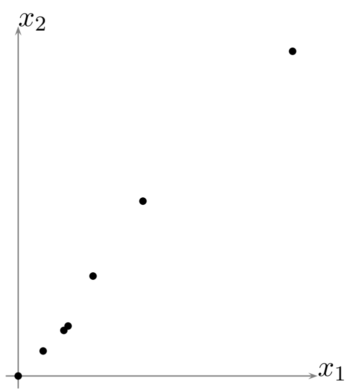




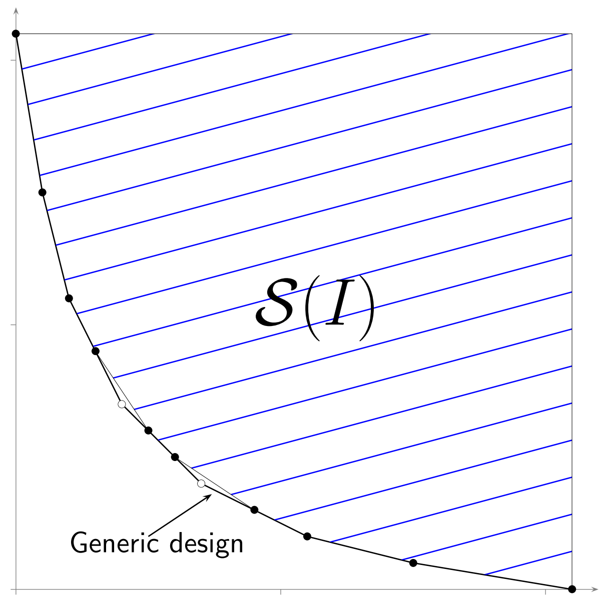
Algebraic $=\{1, x_1, x_1^2, x_1^3, x_1^4, x_2, x_1x_2, x_1^2x_2, x_2^2\}$ and its conjugate

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)



$\mathcal{D}$



$\Rightarrow$ 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. (2007) [2,8])

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

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

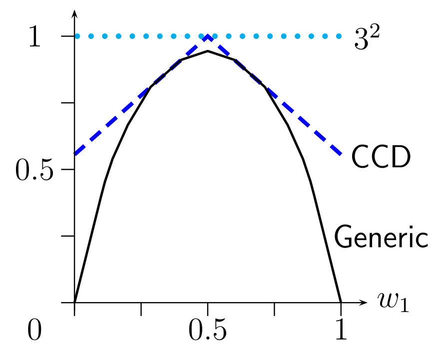


Figure 1: Minimal aberration for three designs in two factors and nine runs.

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

Bounds for A^*: $A^+ - 1 \leq A^* \leq A^+ + 1$; $A^+ = k \cdot g(w)$

Approximate minimal aberration: $\tilde{A} = db^{1/d} g(w) - a$; $\tilde{A} \to A^*$

Comparison of designs using A^*: generic LHS, fractional 2^d

Comparison through aberration

F_1, F_2 fractional factorial designs 2^{7-2} of resolution IV with F_2 of minimum *generator aberration* (Fries & Hunter, 1980). F_3 is a randomly selected non-orthogonal fraction.

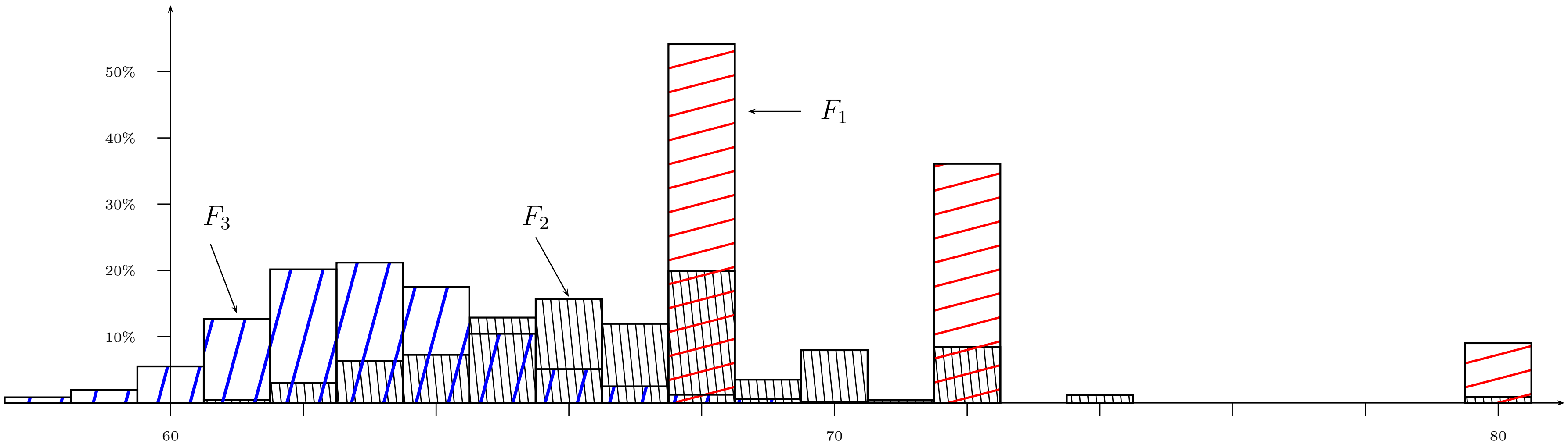


Figure 2: Histograms of relative frequencies of model degrees for fractions F_1, F_2 and F_3.

Work on progress and future work

- Universal w-orders. Extending/finding (zonotope) result to higher dimensions.
- Approximate computations? Computing algebraic fan is expensive even for, say $n = 50, d = 8$, need to retort to other methods, possibly approximate. Select w vectors at random to generate the algebraic fan.
- Aberration: Link linear aberration with traditional generator aberration and estimation capacity [4].

5. References

1. Babson, Onn, Thomas (2003). *Adv. Appl. Math.* **30**(3), 529-544.
2. Bernstein *et al.* (2007). *Minimal aberration* `arXiv:0808.3055`.
3. Caboara *et al.* (1997). Res. Rep. 33, Univ. of Warwick.
4. Cheng, Mukerjee (1998). *Ann. Statist.* **26**(6), 2289-2300.
5. Cox, Little, O'Shea (1996). *Ideals, varieties and algorithms*.
6. Faugere *et al.* (1993). *J. Symb. Comput.* **16**, 329-344.
7. Gritzmann, Sturmfels (1993). *SIAM JDM* **6**(2), 246-269.
8. Lee *et al.* (2008) *SIAM JDM* **22**(3), 901-919.
9. Maruri (2006). *RIMUT* **37**(1-2), 95-119.
10. Maruri (2007). Ph.D. Thesis, Univ. of Warwick.
11. Maruri, Notari, Riccomagno (2007). *Stat. Sin.* **17**, 1417-1440.

12. Onn, Sturmfels (1999). *Adv. Appl. Math.* **23**, 29-48.
13. Pistone and Wynn (1996). *Biometrika* **83**(3), 653-666.
14. Pistone *et al.* (2000). *Algebraic Statistics*. CRC.