Generalised aberration for space filling designs

Hugo Maruri and Henry Wynn

Department of Statistics, London School of Economics

MUCM-SAMSI joint meeting April 2007

Generalised aberration for space filling designs

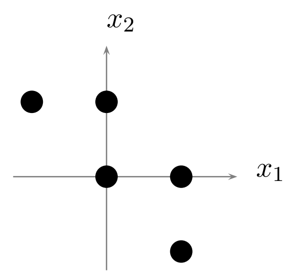
Abstract

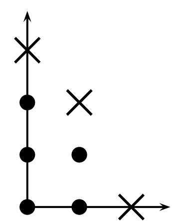
The aberration of a design measures its capability to identify models of a lower (weighted) degree. In a precise algebraic sense, this measure is related to the spread of the design points, and although it is defined for polynomial models, it is not restricted to them. We present recent work and aberration results for latin hypercube designs and Sobol sequences.

Algebraic design of experiments (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 Gröbner basis $G_\tau \subset I(\mathcal{D})$

Design $\mathcal{D}$





$\tau : x_1 > x_2$

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

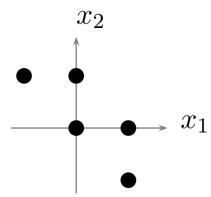
$$G_\tau = \{\underline{x_1^2} + 2x_1x_2 + x_2^2 - x_1 - x_2, \underline{x_2^3} - x_2, \underline{x_1x_2^2} - x_1x_2 - x_2^2 + x_2\}$$

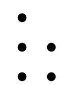
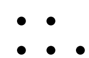
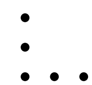
- Hierarchical polynomial model: it belongs to $\mathcal{C}_{d,n}$
- Saturated regression model = exact polynomial interpolator
- Link with aliasing/confounding $f(x) = g(x), x \in \mathcal{D}$

The fan of a design

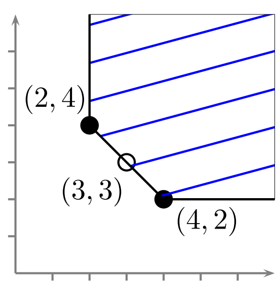
- As we scan over all possible term orders, we obtain the algebraic fan of $\mathcal{D}$ (Caboara, 1997; Maruri, 2007)
- Not all identifiable hierarchical models belong to the algebraic fan, i.e. $\emptyset \subset A \subseteq S \subseteq \mathcal{C}_{d,n}$

$\mathcal{D}$



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

- The models in A correspond to the vertexes of the state polytope $\mathcal{S}(I)$ (Bayer, 1988), e.g. we add up the exponent vectors for $L = \{1, x_1, x_2, x_1x_2, x_2^2\}$, $\bar{\alpha}_L = \sum_L \alpha = (2,4)$.



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

Linear aberration

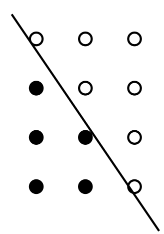
- Taking the motivation from the concept of aberration, we want to fill out lower degrees before higher:

$$\phi(L,c) = \sum c_i \bar{\alpha}_{L_i} = \langle c, \bar{\alpha}_L \rangle, c \geq 0, c \neq 0$$

Theorem: Any algebraic model minimises some $\phi(L,c)$.

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

- **Generic** designs minimise $\phi(L,c)$ over $\mathcal{C}_{d,n}$ and all vectors c.
- For generic designs, algebraic models are corner cut models (Onn, 1999)

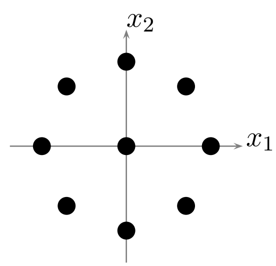
Corner cut model  $\{1, x_1, x_2, x_1x_2\}$ is not corner cut

Linear aberration and algebraic models

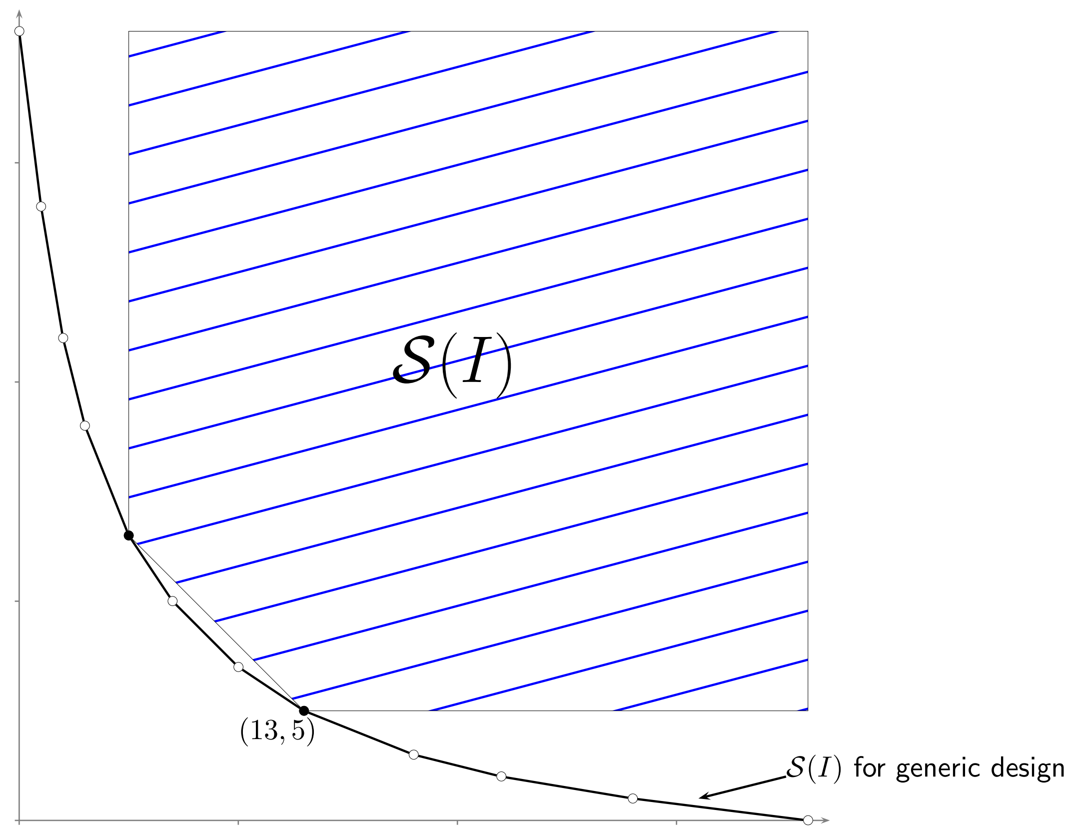
- The state polytope summarises information about linear aberration, i.e. its vertexes correspond to models that minimise $\phi(L, c)$ over the set of identifiable hierarchical models S.
- The vertexes of $\mathcal{S}(I)$ correspond to algebraic models A.
- The (minimum) aberration of designs can be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

Examples 1

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



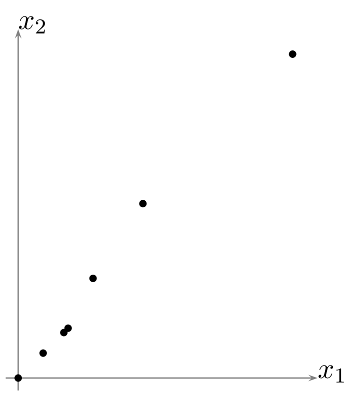
CCD



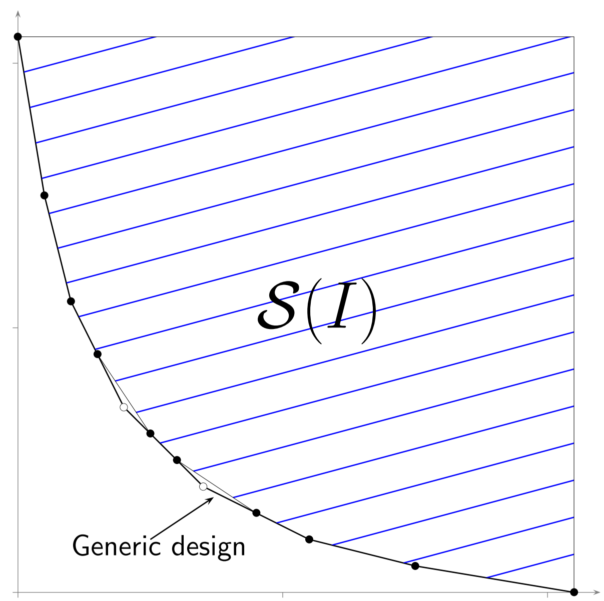
Algebraic $=\{1, x_1, x_1^2, x_1^3, x_1^4, x_2, x_1x_2, x_1^2x_2, x_2^2\}$ and its conjugate

Examples 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 $c \neq 0$.

Latin hypercube designs (McKay, 1979)

Construction of Latin hypercube (LH) designs

- Divide the range of each variable into n equal segments
- Choose a value in each segment (uniformly)
- Permute each coordinate randomly

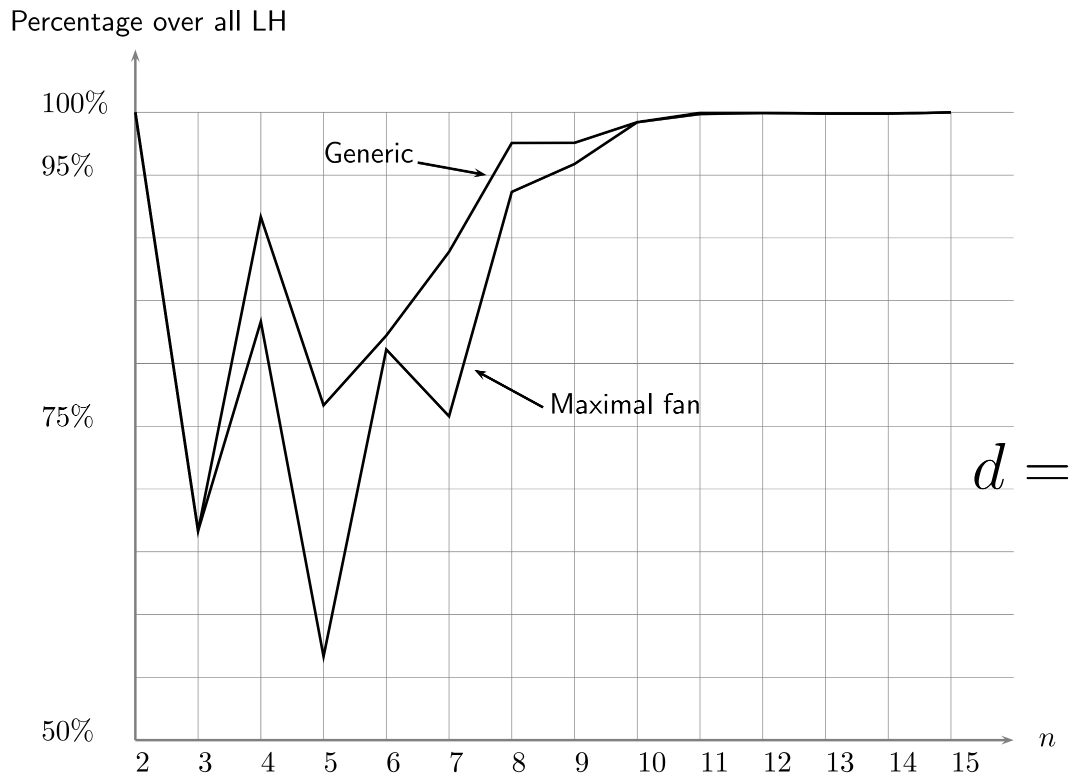
LH covers each coordinate evenly. Lower dimensional projections are also LH.

- Mixing LH with other criteria: Maximin, alphabetic optimality (e.g. D-optimality).

Theorem. The random LHS is generic with probability one.

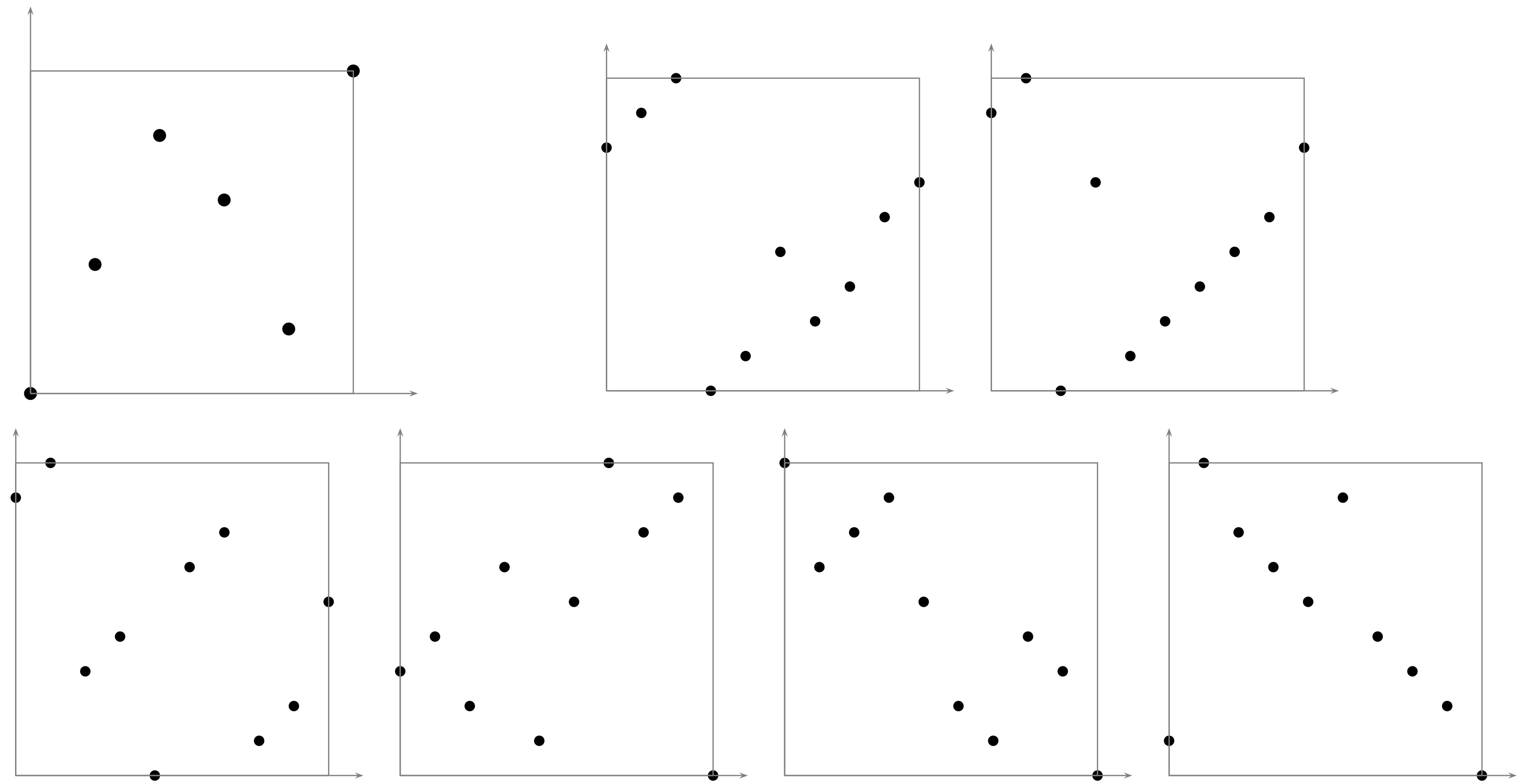
Minimal aberration and LH

Theorem. For $k, n \geq 2$, there always exist a generic non-random LHS design (proof not complete yet)



$d = 2$

Almost generic non-random LH designs ($d = 2, n = 6, 10$)

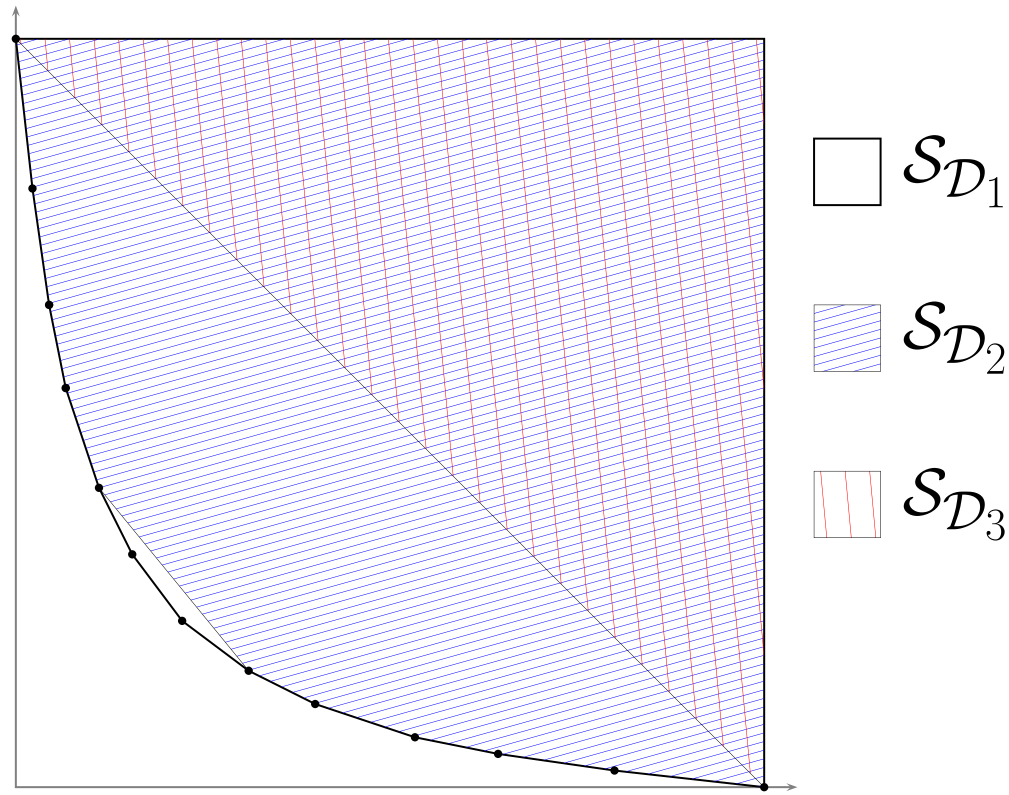


$\Rightarrow$ Identification of a model of total degree s is closely related to genericity for design size $n = \binom{d+s}{s}$.

LH rules

- There clearly exist a few bad cases of LH designs, e.g. when the points of the LH are along a line (e.g. next slide).
- However, by selecting a non-random latin hypercube (at random), the probability of it being generic increases rapidly with n.
- A necessary condition for LH to exist is given by the Shape Lemma (Cox, 2005). We note that the polynomials forming the GBasis for the design ideal must be of the highest possible order (a specialization of the Shape Lemma).

Comparison of space filling designs: Sobol, maximin LH and worst case LH for $d = 2, n = 10$



State polyhedrons for maximin LHS design ($\mathcal{D}_1$), a Sobol sequence ($\mathcal{D}_2$) and a "worst case" LHS ($\mathcal{D}_3$).

Future work

- Using other kernels
- Genericity in kriging: ?
- Link between different (concepts) criteria
    - Maximin (special criteria)
    - Discrepancy (number theoretic)
    - Genericity (algebraic)
    - Optimality (statistical)
- Estimating the covariance kernel

Model Based Optimal Design (Non-Sequential Approach)

(MUCM) PhD student: Noha Youssef
Supervisor: Henry Wynn

Department of Statistics, London School of Economics

MUCM-SAMSI joint meeting April 2007

Maximum entropy sampling (MES: Wynn, 1987)

- Partition: $Y_S = [Y_s : Y_{S\setminus s}]$
- $\mathsf{Ent}(Y_S) = \mathsf{Ent}(Y_s) + \mathsf{E}_{Y_s}\mathsf{Ent}(Y_{S\setminus s}|Y_s)$
- Since $\mathsf{Ent}(Y_S)$ is fixed, an experiment that maximises the entropy of the (marginal/prior) distribution of Y_s will minimise the preposterior entropy
- In the Gaussian case MES: maximise ϕ where $\phi = \log|\mathsf{cov}(Y_s)|$

Branch and Bound for MES (Ko, 1995)

- Branch and Bound (B&B) algorithm
  - The algorithm searches over a finite set of candidate points in a grid.
  - At every step, a branch is done and bounds for ϕ are computed: $LB \leq \phi \leq UB$. Then based on the bounds (LB, UB) computed, one of the following decisions is taken: add a problem to the live set of problems, kill a problem or terminate the algorithm.
- Future work
  - Extending the use of B&B for different kernels
  - Trying to use the B&B in sequential construction of designs

Karhunen-Loeve Expansion and its Applications

PhD Candidate: Limin Wang
Supervisor: Henry Wynn

Department of Statistics, London School of Economics

MUCM-SAMSI joint meeting April 2007

Karhunen-Loeve (K-L) Expansion

- Karhunen-Loeve (K-L) Expansion, Mercer's Theorem and Reproducing Kernel Hilbert Space (Karhunen, 1947)
- Fredholm Equation

$$\int K(s,t)\phi(s)ds = \lambda\phi(t) \tag{1}$$

- Examples for K-L Expansion: Brownian Motion, Brownian Bridge, O-U process, Integrated Brownian Motion, &c
- Numerical Methods: Integration based method and Galerkin method

K-L Expansion: Applications

- Maximum Entropy Sampling (MES, Wynn 1987) for Continuous Gaussian Processes
- D-Optimal Design and MES
  Numerical Approximation: using Fourier bases and Haar wavelets
- Conditional K-L expansions
- Conditional Mean Square Error for finite expansions
- Kriging for finite expansions
- Multivariate K-L expansions

References

1. Generalised aberration
   - Caboara *et al*. (1997). Res. Rep. 33, Univ. of Warwick.
   - Cox *et al*. (2005). *Using algebraic geometry*.
   - Maruri (2007). Ph.D. Thesis, Univ. of Warwick.
   - McKay *et al*. (1979). *Technometrics* 21, 239-245.
   - Onn and Sturmfels (1999). *Adv. Appl. Math.* 23,29-48.
   - Pistone and Wynn (1996). *Biometrika* 83, 653-666.
2. Model based optimal design
   - Ko *et al*. (1995). *Oper. Res.* 43, 684-691.
   - Shewry and Wynn (1987). *J. Appl. Stat.* 14, 165-170.
3. Karhunen-Loeve expansions
   - Karhunen (1947). *Ann. Acad. Sci. Finn.* A.I. 37