Working with Henry Wynn: from Fourier to Gröbner

Eva Riccomagno

Department of Mathematics
Università di Genova
eva.riccomagno@unige.it

mODa14 15 June, 2026

1990-1995
Engineering Design and Quality Center
City University, London

1994
Workshop at Wendisch-Riez
Germany

Low discrepancy sequences for Fourier regression

Optimum orthogonal fractions for Fourier regression models based on integer lattice designs, as an alternatives to product designs

For $(x_1, \ldots, x_d) \in [0,1)^d$ and $A_l = \{\alpha \in \{-1,1\}^l : \alpha_1 = 1\}$
$F(d; m_1, \ldots, m_d; M)$ denotes the complete M-factor interaction model

$$\mathrm{E}(Y(x_1, \ldots, x_d)) = \theta_0$$
$$+\sqrt{2} \sum_{l=1}^{M} \sum_{k_1<k_l} \sum_{r_1=1}^{m_{k_1}} \cdots \sum_{r_l=1}^{m_{k_l}} \sum_{\beta \in A_l} \sin\left(2\pi\left(\beta_1 r_1 x_{k_1} + \cdots + \beta_l r_l x_{k_;}\right)\right) \theta_{k_1 \ldots k_l, r_1 \ldots r_l, \beta}$$
$$+\sqrt{2} \sum_{l=1}^{M} \sum_{k_1<k_l} \sum_{r_1=1}^{m_{k_1}} \cdots \sum_{r_l=1}^{m_{k_l}} \sum_{\beta \in A_l} \cos\left(2\pi\left(\beta_1 r_1 x_{k_1} + \cdots + \beta_l r_l x_{k_;}\right)\right) \phi_{k_1 \ldots k_l, r_1 \ldots r_l, \beta}$$

Let $g = (g_1, \ldots, g_d)$ be an integer valued vector with positive entries, $g_1 = 1$, such that the greatest common divisor of its components and the sample size n is one. Define the grid

$$L_{g,n} = \left\{ \left(\frac{jg_1 \mod n}{n}, \ldots, \frac{jg_d \mod n}{n} \right) \ : \ j = 0, \ldots, n-1 \right\}$$

Orthogonality conditions are expressed in terms of the dual lattice

$$L^{-} = \left\{ h \in \mathbb{R}^d : h^\top x \in \mathbb{Z} \text{ for all } x \in L_{g,n} \right\}$$

$$h^\top x = k \text{ with } k \in \mathbb{Z}$$

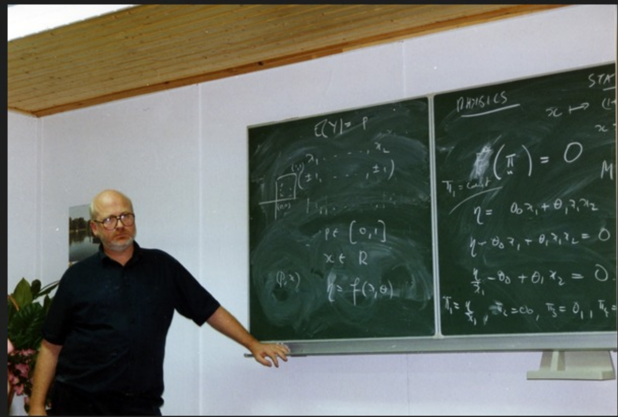
$$\frac{1}{n} \sum_{x \in L \cap [0,1)^d} e^{2\pi i h^\top x} = \begin{cases} 1 & \text{if } h \in L^{-} \\ 0 & \text{otherwise} \end{cases}$$

The power law gives necessary, sufficient, computable and algebraic conditions on g, n and $m_1, \ldots, m_d$ so that the uniform design on $L_{g,n}$ is orthogonal for $F(d; m_1, \ldots, m_d; M)$

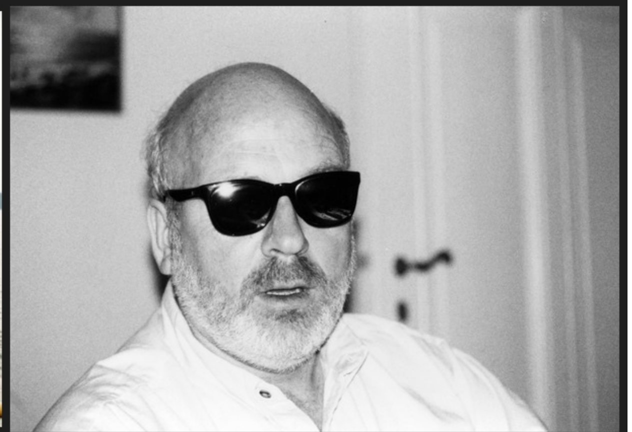
THEOREM 2. *In the M-factor interaction model $F(d; m_1, \ldots, m_d; M)$, the n-point lattice design generated by $(g_1, \ldots, g_d)$ is D-optimum for the parameters up to the S-factor interactions ($S \leq M$) if and only if the members in the first $S+1$ rows of the array*

$$\begin{array}{lll}
\mathbf{0}, & & \\
r_k g_k, & r_k \in \mathcal{N}_k, & k = 1, \ldots, d, \\
r_k g_k + r_l g_l, & r_k \in \mathcal{N}_k \text{ and } r_l \in \mathcal{N}_l, & 1 \leq k < l \leq d, \\
\vdots & & \\
r_{k_1} g_{k_1} + \cdots + r_{k_m} g_{k_m}, & r_{k_1} \in \mathcal{N}_{k_1}, \ldots, r_{k_m} \in \mathcal{N}_{k_m}, & 1 \leq k_1 < \cdots < k_m \leq d, \\
\vdots & & \\
r_{k_1} g_{k_1} + \cdots + r_{k_M} g_{k_M}, & r_{k_1} \in \mathcal{N}_{k_1}, \ldots, r_{k_M} \in \mathcal{N}_{k_M}, & 1 \leq k_1 < \cdots < k_M \leq d,
\end{array}$$

where $\mathcal{N}_k = \{-m_k, \ldots, -1, 1, \ldots, m_k\}$, are distinct (in the cyclic group $\mathcal{G}_n$) and if they are, additionally (for $S < M$), different from those members in the last $M - S$ rows in $\mathcal{G}_n$.






det Σ

1998 R A Bates, ER, R Schwabe, H P Wynn *Lattices and Dual Lattices in Experimental Design for Fourier Models*. Computational Statistics and Data Analysis, 28:3, 283–296

R A Bates, ER, R Schwabe, H P Wynn *The use of lattices in the design of high-dimensional experiments*. New Developments and Applications in Experimental Design (N Flournoy, W F Rosenberger, W W Wong eds.) IMS Lecture Notes, 34, 26–35

1997 R Schwabe, ER, HPW *Lattice-based D-optimum designs for Fourier regression models*. The Annals of Statistics, 25:6, 2313–2327

1996 R A Bates, R J Buck, ER, H P Wynn *Experimental Design and Observation for Large Systems (with discussion)*. Journal of the Royal Statistical Society, B, 58:1, 77–94

1995 R A Bates, ER, R Schwabe, HPW *Lattices and Dual Lattices in Optimal Experimental Design for Fourier Models*. Proceedings, Workshop on Quasi-Monte Carlo Methods and their Applications, December 11-13, 1995 Hong Kong Baptist University (K-T Fang, F-J Hickernell eds.), 1–17

Algebraic Statistics

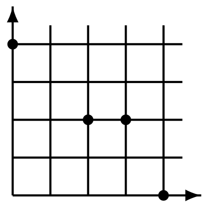
Pistone, Giovanni, and Henry P. Wynn. *Generalised confounding with Gröbner bases.* Biometrika (1996), 653-666

Polynomials and ratios of polynomials appear in statistics and probability under various forms, in model representations as well as in inferential procedures

Many problems of confounding and identifiability for polynomial and multidimensional polynomial models can be solved using methods of algebraic geometry aided by modern computational algebra packages

Pistone, Riccomagno, Wynn (2001). *Algebraic Statistics.* Chapman & Hall

From designs to polynomial ideals



$$\begin{array}{c}(x,y)\\(0,4)\\(2,2)\\(4,0)\\(3,2)\end{array} \qquad \begin{cases} g_1 := x(x-2)(x-4)(x-3) \\ g_2 := (y-4)(y-2)y \\ g_3 := (y-2)(x+y-4) \\ g_4 := (x-3)(x+y-4) \end{cases}$$

$$\text{Ideal}(\mathcal{D}) = \left\{ \sum_{i=1}^{4} f_i(x,y) g_i(x,y) : f_i \text{ are polynomials} \right\}$$

Ideal($\mathcal{D}$) is computable from the coordinates of the points, finitely generated, encodes information about the design structure and allows reframing statistical concepts in algebraic terminology

The space of functions over a design

The set of real polynomial functions over $\mathcal{D}$, the coordinate ring of $\mathcal{D}$, is isomorphic to a computable space of polynomials

$$\mathcal{L} = \mathbb{R}[\mathcal{D}] = \{f : \mathcal{D} \longrightarrow \mathbb{R}\} \sim_{\mathbb{R}} \mathbb{R}[x_1, \ldots, x_k]/\,\mathrm{Ideal}(\mathcal{D})$$

Computations are performed via Gröbner bases of Ideal($\mathcal{D}$) `GBasis`

They are finitely many Gröbner bases for a polynomial ideal `Fan`

A vector space basis of $\mathcal{L}$ can be easily derived from a Gröbner basis of Ideal($\mathcal{D}$) `QuotientBasis`


The Second CREST-SBM International Conference
Harmony of Gröbner Bases and the Modern Industrial Society

From this, recently coauthored by Henry

- identifiability and estimability
- aliasing, confounding, aberration
- algebraic vs statistical fans
- extensions to reliability (with G Giglio, E Sáenz-de-Cabezón, F Mohammadi, ...)
- orthogonal fractions and indicator methods (with the Genoa crowd)
- compositional data
- saturated and supersaturated designs, sparse polynomial prediction (with Hugo Maruri-Aguilar et al)
- toric models and conditional independence models (with G. Pistone, P Caines et al)
- circuits for randomisation (with F Rapallo, R Fontana, E Pesce et al)
- asymptotically efficient estimators (with K Kobayashi et al)