OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

Rainer Schwabe

Simplify Designs:
Reduction Principles Revisited

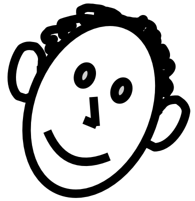
Simplify Designs:
Reduction Principles Revisited

joint work with

Fritjof Freise — Dortmund

Osama Idais
Eric Nyarko
Maryna Prus
Martin Radloff
Frank Röttger
Dennis Schmidt
Marius Schmidt
— Magdeburg

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG



CIRM ?

Outline

1 Initial Considerations

2 Majorization

3 Invariance and Equivariance

4 Constructions for Multiple Factors

5 Final Considerations

1 Initial Considerations

- all design problems
  seem to be solvable
  by numerical optimization

$$\xi^* = \arg\max_{\xi}\{\ldots\}$$

- but some may be substantially reduced
  - » multi-factorial models: $\mathbf{x} = (x_1, \ldots, x_K)$
  - » monotonicity and symmetry

Questions

- When can the set of optimal design points be reduced to a tractable (finite) set ?

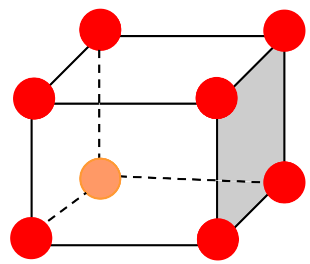
$$\mathbf{x}^* \in X^* \subset X$$

- When can symmetries be utilized ?

$$g(\xi^*) = \xi^*$$

- When can multi-factor optimal designs be constructed from marginal models ?

$$\xi^* = h(\xi_1^*, \ldots, \xi_K^*)$$



- … and how ?

Design

- aim: choose settings $\mathbf{x}$ to maximize information

- approximate design

$$\xi = \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_m \\ \xi(\mathbf{x}_1) & \cdots & \xi(\mathbf{x}_m) \end{pmatrix}$$

← settings

← proportions

- information matrix

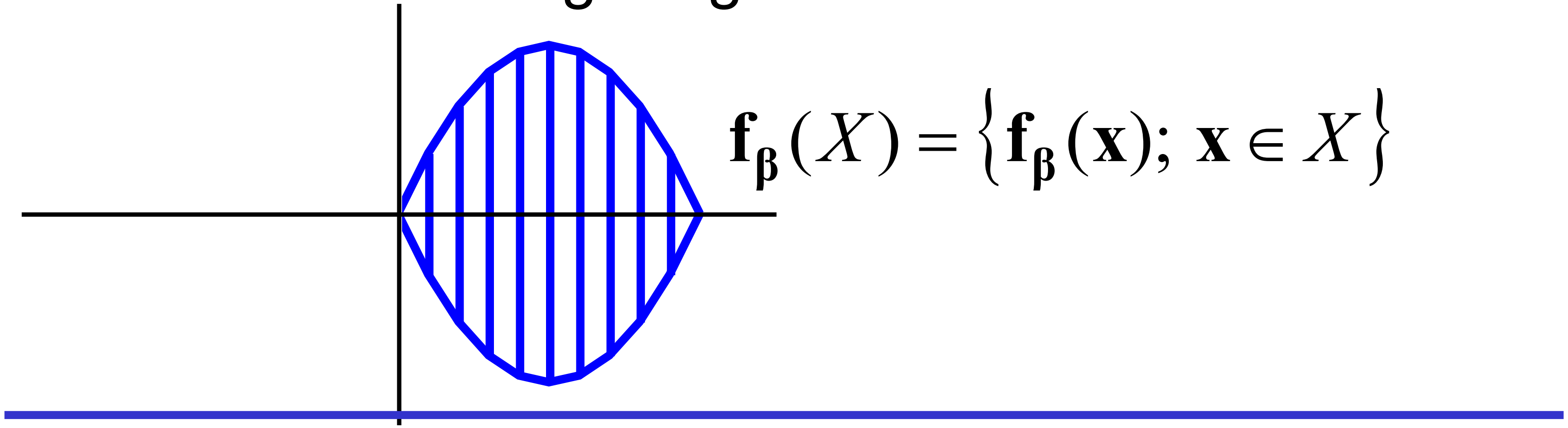
$$\mathbf{M}(\xi;\boldsymbol{\beta}) = \sum_{j=1}^{m} \xi(\mathbf{x}_j)\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}_j)\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}_j)^{\mathsf{T}}$$

  » with regression functions $\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x})$ $\left[= \partial\mu(\mathbf{x})/\partial\boldsymbol{\beta}\right]$

  » for linear models $\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$

2 Majorization

- effective design points $\mathbf{x}^* \in X^* \subset X$
  - » induced design region



$$\mathbf{f}_{\boldsymbol{\beta}}(X) = \left\{ \mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x});\ \mathbf{x} \in X \right\}$$

- » X^* set of $\mathbf{x}^*$ such that

  $\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}^*)$ extremal point of $\mathbf{f}_{\boldsymbol{\beta}}(X)$

Ehrenberg 195?

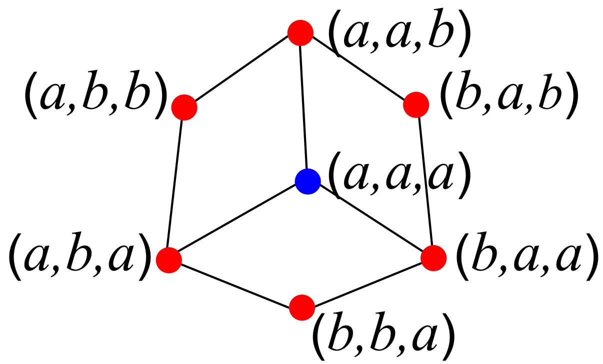
$$\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}) = \sum w_j \mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}_j^*)$$

Example

- Gamma model

$$\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) / (\mathbf{f}(\mathbf{x})^{\mathsf{T}} \boldsymbol{\beta})$$

  - 3 factor on $X=[a,b]^3$ (multilinear), without intercept
    - induced design region $\mathbf{f}_{\boldsymbol{\beta}}(X)$



    - reduced design region

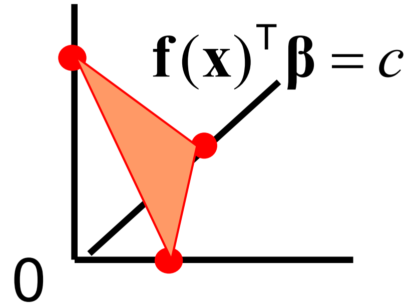
O. Idais/Gaffke/Sch. 2018

$$X^* = \{(a,a,b),(a,b,a),(b,a,a),\\ (a,b,b),(b,a,b),(b,b,a)\}$$

Generalized Linear Models and alike

- weighted regression function

$$\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}) = q\left(\mathbf{f}(\mathbf{x})^{\top}\boldsymbol{\beta}\right)\mathbf{f}(\mathbf{x})$$

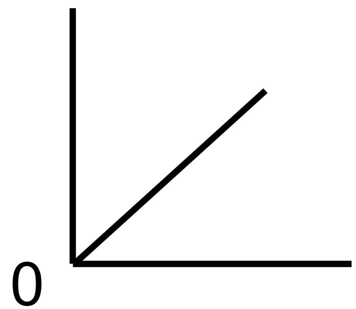


- » with weight q as a function of the linear predictor

$$\mathbf{f}(\mathbf{x})^{\top}\boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \ldots + \beta_K x_K$$

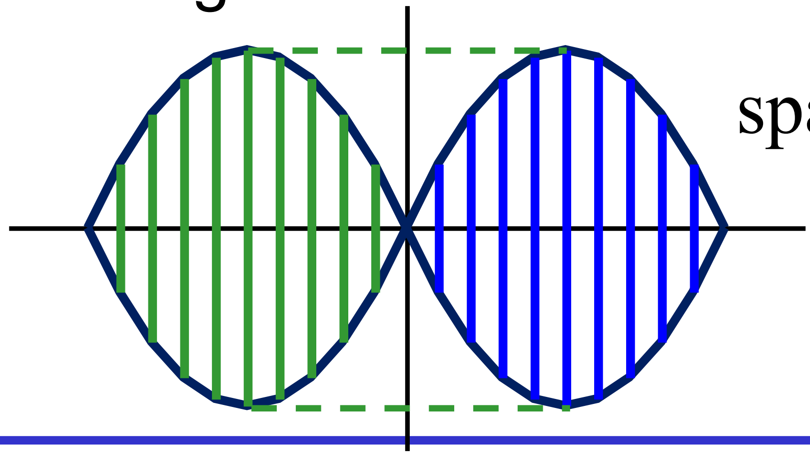
$$X = [0.\infty)^K$$

- reduced design region X^* on the edges



Majorization II

- effective design points $\mathbf{x}^* \in X^* \subset X$
  - Elfving set



$$\mathrm{span}\left(\mathbf{f}_{\boldsymbol{\beta}}(X) \cup -\mathbf{f}_{\boldsymbol{\beta}}(X)\right)$$

  - X^* set of $\mathbf{x}^*$ such that $\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}^*)$ extremal point of the Elfving set

Elfving 1952

$$\mathbf{M}(\mathbf{x};\boldsymbol{\beta}) = \sum w_j \mathbf{M}(\mathbf{x}_j^*;\boldsymbol{\beta})$$

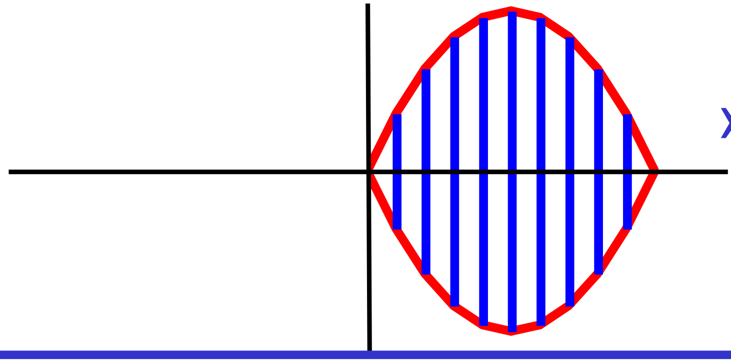
Example

Graßhoff/Großmann/Holling/Sch. 2001

- quadratic paired comparison

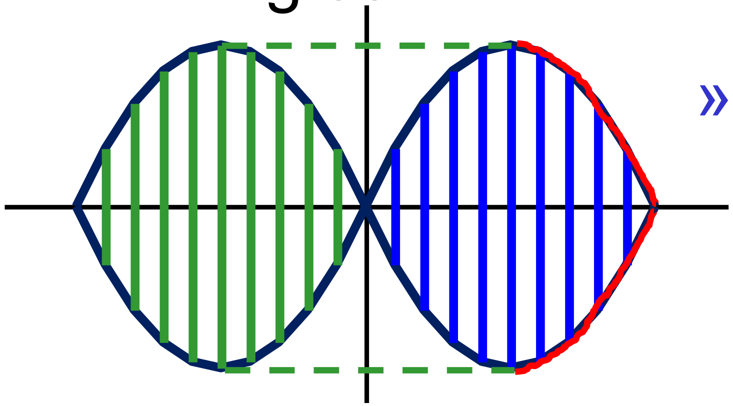
$$\mathbf{f}(\mathbf{x}) = (x_1 - x_2, x_1^2 - x_2^2)^T$$

  - induced design region: shaded area



  - reduced design region: boundary of the shaded area

  - Elfving set



  - reduced design region: boundary to the right of the shaded area

3 Invariance and Equivariance

Ford/Torsney/Wu 1992

- equivariance
  - transformation $g : X \to g(X)$

    such that $\mathbf{f}_{\boldsymbol{\beta}}(g(\mathbf{x})) = \mathbf{Q}\,\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x})$

    $\mathbf{f}_{\boldsymbol{\beta}}$ and g compatible

  - ξ^* optimal on X $\Rightarrow$ $g(\xi^*)$ optimal on g(X)

Giovagnoli/Pukelsheim/Wynn 1987, Sch. 1996

- invariance
  - group of transformations $g : X \to X$

    (permutations, sign change, rotations …)
  - invariant optimal designs

$$g(\xi^*) = \xi^*$$

Example

F. Freise/Holling/Sch. 2018

- multiple binary predictors

$$\mathbf{f}(\mathbf{x}) = (1, x_1, x_2, \ldots, x_K)^{\mathsf{T}} \qquad X \subset \{0,1\}^K$$

  - restricted design region

$$k_{\min} \le x_1 + x_2 + \ldots + x_K \le k_{\max}$$

  - optimal designs invariant (uniform) on orbits

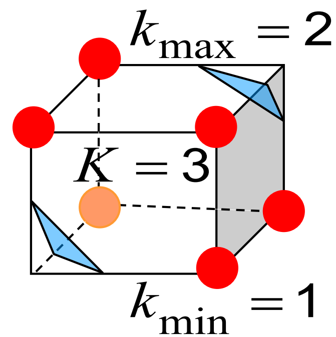
$$\{\mathbf{x}; x_1 + x_2 + \ldots + x_K = k\}$$

of equal number of active factors

  - $(K - \sqrt{K})/2 \le k_{\min} < k_{\max} \le (K + \sqrt{K})/2$
  - sensitivity function quadratic in k

optimal design uniform on the orbits

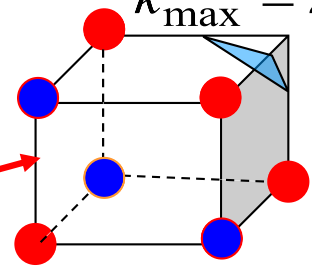
$k_{\min}$ and $k_{\max}$



Example (continued)

K	$\frac{K-\sqrt{K}}{K}$	$\frac{K+\sqrt{K}}{K}$	L	U	ℓ	w_L	w_U
2	0.29	1.71	0	2	1	0.2500	0.2500
3	0.63	2.37	0	2	~~1~~	0.2500	0.7500
			~~0~~	3	1	0.0000	0.2500
4	1.00	3.00	0	3	2	0.1667	0.3333
			0	4	2	0.1250	0.1250
			1	3	~~2~~	0.5000	0.5000
5	1.38	3.62	0	3	2	0.1667	0.8333
			0	4	2	0.0625	0.3125
			~~0~~	5	2	0.0000	0.1667
			1	4	2	0.1667	0.3333
6	1.78	4.22	0	4	3	0.1250	0.3750
			0	5	3	0.1000	0.1500
			0	6	3	0.0833	0.0833
			1	4	3	0.2500	0.5000
			1	5	3	0.1875	0.1875

3 orbits not unique



» sensitivity function constant in k
information equal to full factorial (without restriction)

Example

- paired comparisons (with interactions)

$$\mathbf{f}(\mathbf{x}) = (x_{1k} - x_{2k}, \ldots, x_{1k}x_{1\ell} - x_{2k}x_{2\ell}, \ldots)^{\mathsf{T}}$$

» optimal designs invariant (uniform) on orbits of equal comparison depth d

$$\{\mathbf{x}; \#\{k; x_{1k} \neq x_{2k}\} = d\}$$

» model without interactions

optimal design uniform on orbit K of full comparison depth

sensitivity linear in d

» model with second order interactions

sensitivity quadratic in d

D-optimal design uniform on orbit $d^* \approx K/2$ or two adjacent orbits $d^*, d^* + 1 \approx K/2$

Example (continued)

E. Nyarko/Sch. 2018

» model with second order interactions

K	d^*	w^*
3	1	0.750
4	2	0.857
5	2	0.833
6	3	0.732
7	3	0.697
8	(3)	0.609
9	4	0.577
10	4	0.538

sensitivity cubic in d

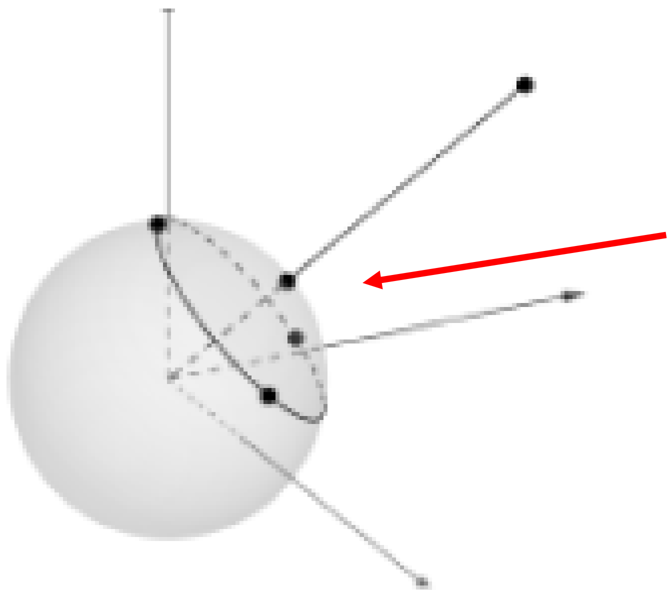
» D-optimal design uniform on two orbits K and d^* with weight w^* on d^*

Example

M. Radloff/Sch. 2018

- Poisson regression on a ball X

$$\mathbf{f}_{\boldsymbol{\beta}}(\mathbf{x}) = \exp(\mathbf{f}(\mathbf{x})^{\mathsf{T}}\boldsymbol{\beta}/2)\,\mathbf{f}(\mathbf{x})$$



$$\arg\max_{\boldsymbol{\beta}}(\mathbf{f}(\mathbf{x})^{\mathsf{T}}\boldsymbol{\beta})$$

» D-optimal design

on K+1 points on the surface

with equal weights

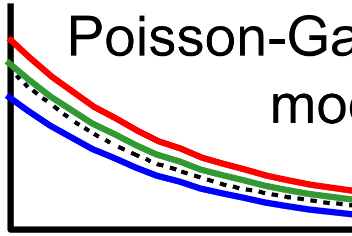
Example

- mixed models

Schmelter 2007, M. Schmidt/Sch. 2018

- » estimation

  in linear and generalized linear mixed models

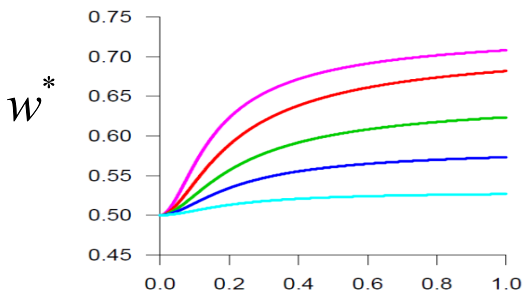


single-group designs are optimal

M. Prus/Benda/Sch. 2018

- » prediction

  invariance -> balancing



Example

F. Röttger/Kahle/Sch. 2018

- Bradley-Terry model
  - paired comparisons (4 alternatives)

  - optimality regions in the parameter space -> poster

4 Constructions for Multiple Factors

- additive linear models

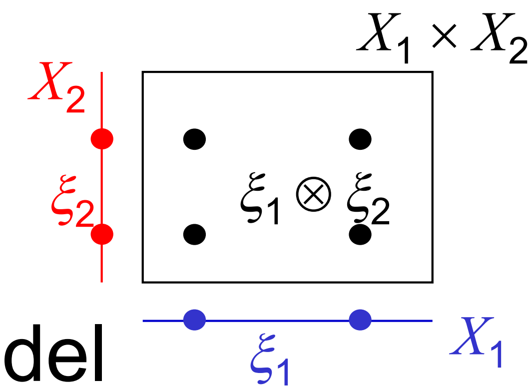
Cook and Thibodeau 1980, Sch./Wierich 1995, Sch. 1996

ξ_k^* D-optimal in the K marginal models

$$Y_k(x_k) = \mu_k + \mathbf{f}_k(x_k)^{\mathsf{T}} \boldsymbol{\beta}_k + \varepsilon$$

then $\xi^* = \xi_1^* \otimes \ldots \otimes \xi_K^*$

is D-optimal in the additive model



$$Y(x_1, \ldots, x_K) = \mu_0 + \mathbf{f}_1(x_1)^{\mathsf{T}} \boldsymbol{\beta}_1 + \ldots + \mathbf{f}_K(x_K)^{\mathsf{T}} \boldsymbol{\beta}_K + \varepsilon$$

GLM: Additive Predictors (multiple regression)

$$\mu(x_1, x_2, \ldots, x_K) = \mu_0 + x_1\beta_1 + \ldots + x_K\beta_K$$

» design regions $X_k = [0, \infty)$

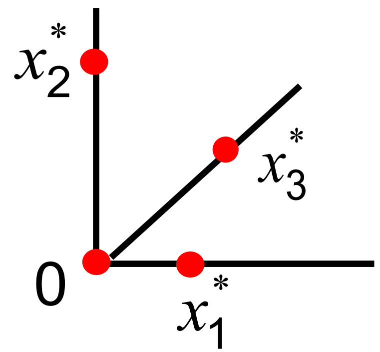
D. Schmidt/Sch. 2017

➢ ξ_k^* with support 0 and x_k^* D-optimal in the univariate models, then the design

$$\mathbf{x}_k^* = x_k^* \mathbf{e}_k, k = 1, \ldots, K, \quad \mathbf{x}_{K+1}^* = \mathbf{0}$$

with equal weights $1/(K+1)$

is D-optimal in the additive model

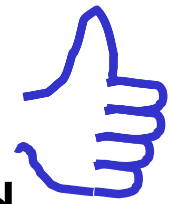


» also valid in GLMMs

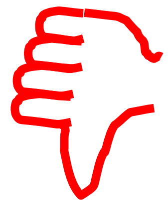
M. Schmidt/Sch. 2018

5 Final Considerations

- for high-dimensional models
  computational burden can be reduced

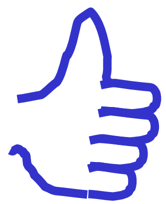


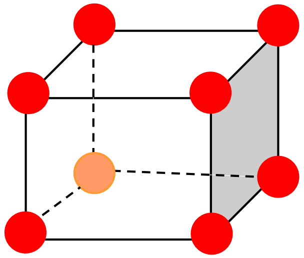
- invariant and product-type designs
  have a quite high number
  of design points



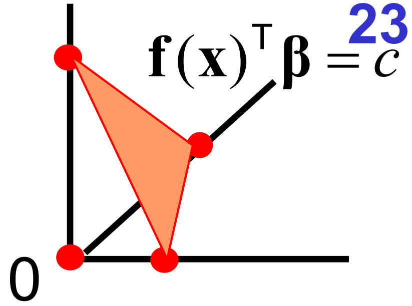
  - » potential reduction by fractional factorials
    or discretization

- but these designs may serve
  as a benchmark

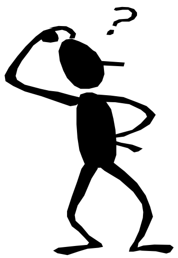




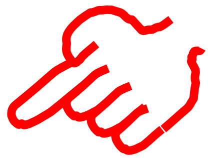
Simplify Designs:



Reduction Principles Revisited



message to go



"Utilize the structure of models!"