Towards an Equivalence Theorem for Computer Simulation Experiments?

Werner G. Müller

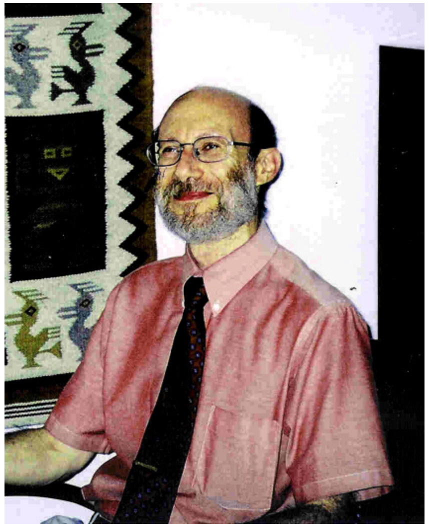
Department of Applied Statistics (IFAS)

Based on joint work with *M.Stehlík* and *Luc Pronzato (started a week ago)*

I take on challenge #1 from David's list:

„create designs that are tied to

our methods of analysis“

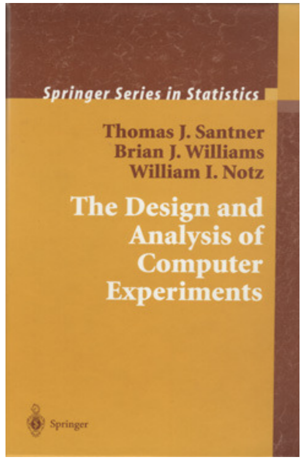


D. Steinberg (2009), ENBIS-EMSE workshop

The Setup

Random field: $$y(z) = \eta\big(x(z), \beta\big) + \varepsilon(z)$$

with $$E\big[\varepsilon(z)\varepsilon(z')\big] = c(z, z'; \theta) = c(d, \theta).$$



Two purposes: **prediction** or **estimation**

Universal Kriging: using the EBLUP and the corresponding GLS-estimator.

Alternative: Full ML or REML of (β, θ) and insert above.

Optimal Designs (for estimation)

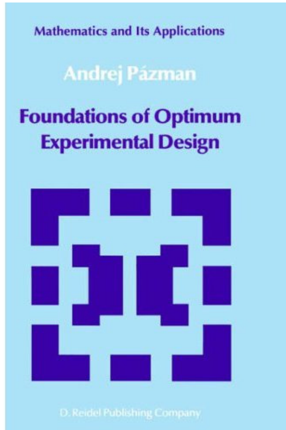
Classical: select the inputs (and weights)

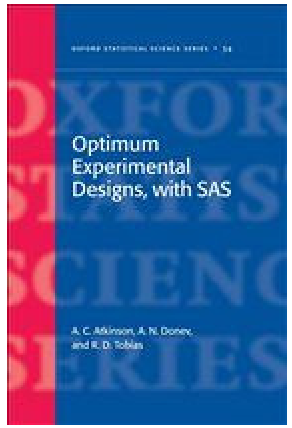
$$\xi_N = \begin{Bmatrix} p_1, p_2, \cdots, p_n \\ z_1, z_2, \cdots, z_n \end{Bmatrix}$$

such that a prespecified criterion

$$\max_{z_i, p_i} \Phi\left[M\left(\xi_N \right) \right]$$

is optimized.





Well developed theory for standard (uncorrelated) regression based on Kiefer's (1959) concept of design measures.

Three Practical Cases:

Case 1: We are interested only in the trend parameters β and consider θ as known or a nuisance.

Case 2: We are interested only in the covariance parameters θ (sometimes we set $\beta = 0$).

Case 3: We are interested in both sets of parameters equally.

D-optimal designs for estimating trend and covariance parameters

For the full parameter set the information matrix is

$$E\begin{Bmatrix} -\frac{\partial lnL(\beta,\theta)}{\partial\beta\partial\beta^T} & -\frac{\partial lnL(\beta,\theta)}{\partial\beta\partial\theta^T} \\ -\frac{\partial lnL(\beta,\theta)}{\partial\theta\partial\beta^T} & -\frac{\partial lnL(\beta,\theta)}{\partial\theta\partial\theta^T} \end{Bmatrix} = \begin{pmatrix} M_\beta(\xi;\theta,\beta) & 0 \\ 0 & M_\theta(\xi;\theta) \end{pmatrix}.$$

Use the (weighted) product of the respective determinants as an optimum design criterion (Müller and Stehlík, 2009):

$$\bar{\Phi}'[M_\beta, M_\theta] = |M_\beta(\xi)|^\alpha \cdot |M_\theta(\xi)|^{1-\alpha}$$

Xia G., Miranda M.L. and Gelfand A.E. (2006) suggest to use the trace.

Compound Designs

Single purpose criterion is inefficient, thus construct weighted averages

$$\bar{\Phi}[\xi \mid \alpha] = \alpha\Phi[M(\xi)] + (1-\alpha)\Phi'[M'(\xi)].$$

were introduced by Läuter (1976), related to constrained designs:

$$\xi^* = \arg\max_{\xi\in\Xi} \Phi[M(\xi)] \quad \text{s.t. } \Phi'[M'(\xi)] > \kappa(\alpha),$$

(cf. Cook and Wong, 1994); sometimes standardized (Mcgree et al., 2008):

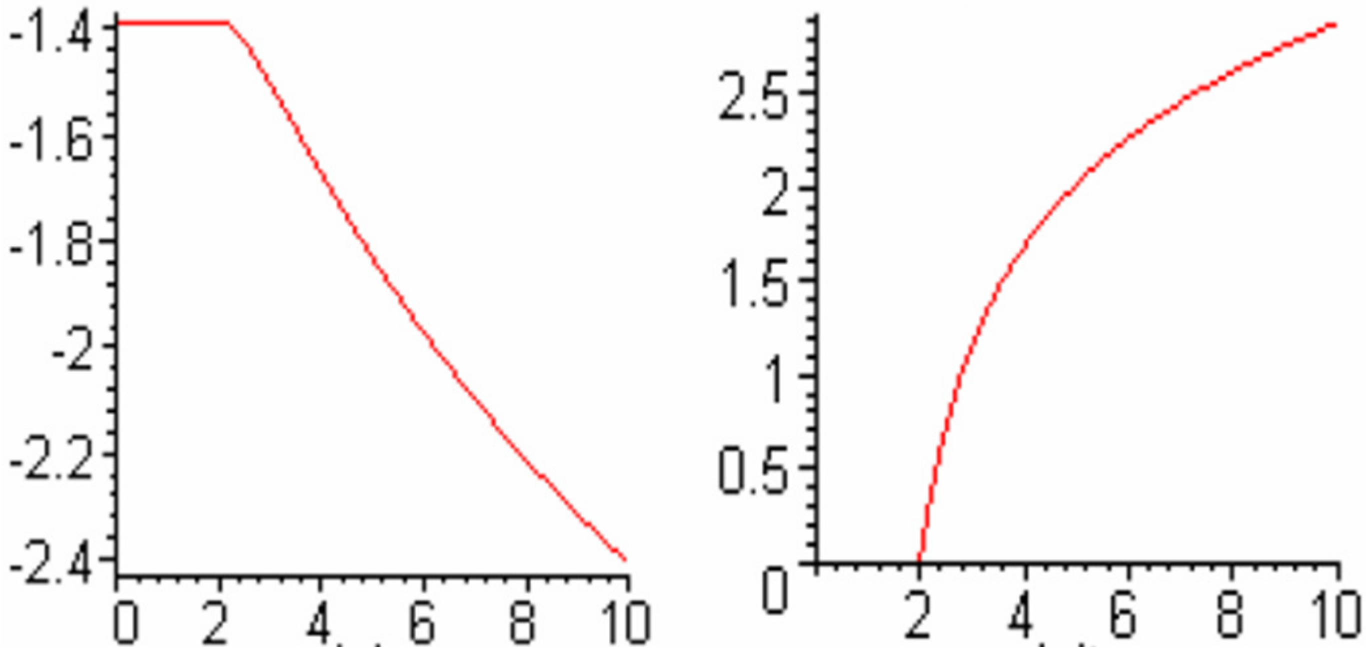
$$\bar{\Phi}[\xi \mid \alpha] = \alpha\Phi[M(\xi)] / \Phi[M(\xi^*)] + (1-\alpha)\Phi'[M'(\xi)] / \Phi'[M'(\xi^*)].$$

7 (9) Issues (surveyed in Müller & Stehlík, 2009)

1. Nonconvexity
2. Asymptotic unidentifiability (M_θ)
3. Nonreplicability
4. Non-additivity
5. Smit's paradox
6. Näther's paradox
7. Impact of dependence on information (M&S paradox)
8. Choice of dependence structure
9. Singular designs (the role of the nugget effect)

The Impact of Dependence

information from D-optimal design:



when θ is estimated or not estimated respectively.
(Müller & Stehlík, 2004)

Design for prediction (EK-optimality)

Criterion often based on kriging variance, e.g.

$$\min_{\xi} \max_{z} E[(\hat{y}(z|\xi) - y(z))^2]$$

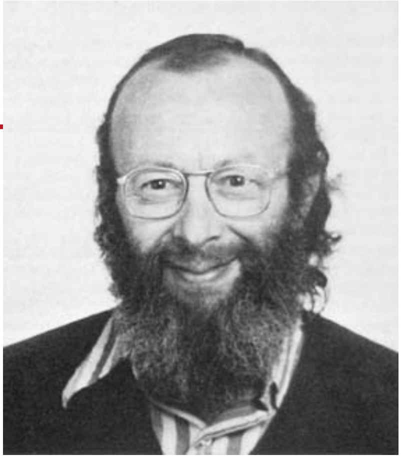
Additional uncertainty from estimation of θ is taken into account by Zhu (2002) and Zimmerman (2006):

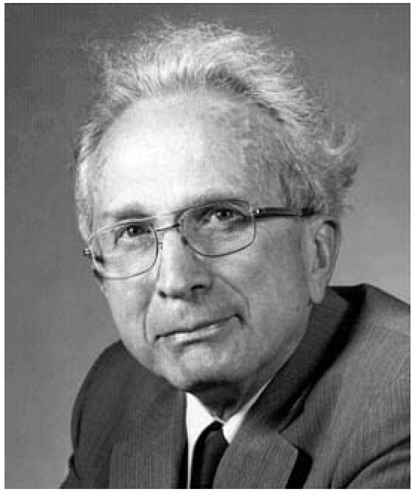
$$\min_{\xi} \max_{z} \left\{ \mathrm{Var}[\hat{y}(z)] + \mathrm{tr}\left\{ M_{\theta}^{'-1} \mathrm{Var}[\partial \hat{y}(z)/\partial \theta] \right\} \right\}$$

Abt (1999) and Zhu and Stein (2007) supplement this by

$$\left(\frac{\partial \mathrm{Var}[\hat{y}(z)]}{\partial \theta} \right)^T M_{\theta}^{'-1} \left(\frac{\partial \mathrm{Var}[\hat{y}(z)]}{\partial \theta} \right).$$

Recall the Kiefer-Wolfowitz Equivalence Theorem (1960)





(Case 1 with uncorrelated errors)

D-criterion: $$\max_{\xi} \left| M_{\beta}(\xi) \right|$$

and G-criterion: $$\min_{\xi} \max_{z} \mathrm{Var}[\hat{y}(z) \mid \xi]$$

yield same (approximate) optimal designs.

Conjecture:

One can always find an α such that the compound design based upon

$\bar{\Phi}'[M_\beta, M_\theta]$ is (in some to be defined sense) close to designs

following from Zhu's EK-(empirical kriging)-optimality.

JOHANNES KEPLER UNIVERSITY LINZ | JKU

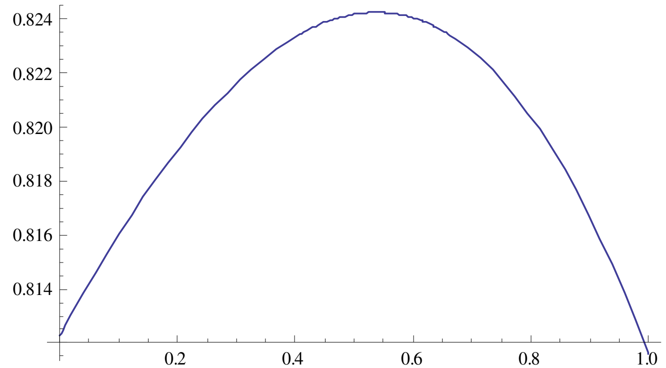
Example : Ornstein-Uhlenbeck process

constant trend $\eta(.) = \beta$

$$\operatorname{cov}(z, z') = \sigma^2 e^{-\frac{d}{\theta}}$$

Case 1 (Kiselak and Stehlík, 2007, Dette et al., 2007): uniform (space-filling) design is D-optimal!

Case 2 (Müller and Stehlik, 2009, Zagoraiou and Baldi-Antognini, 2009): D-optimal design collapses!



Case 3: regulatory version.

Exchange algorithms (Fedorov/Wynn, 1972)

consist in a simple exchange of points from the two sets S_{ξ_s} and $\bar{X}_s \setminus S_{\xi_s}$ at every iteration, namely

$$\xi_{s+1} = \left\{ \xi_s \setminus \left\{ x_s^-, \frac{1}{n_s} \right\} \right\} \cup \left\{ x_s^+, \frac{1}{n_s} \right\},$$

where

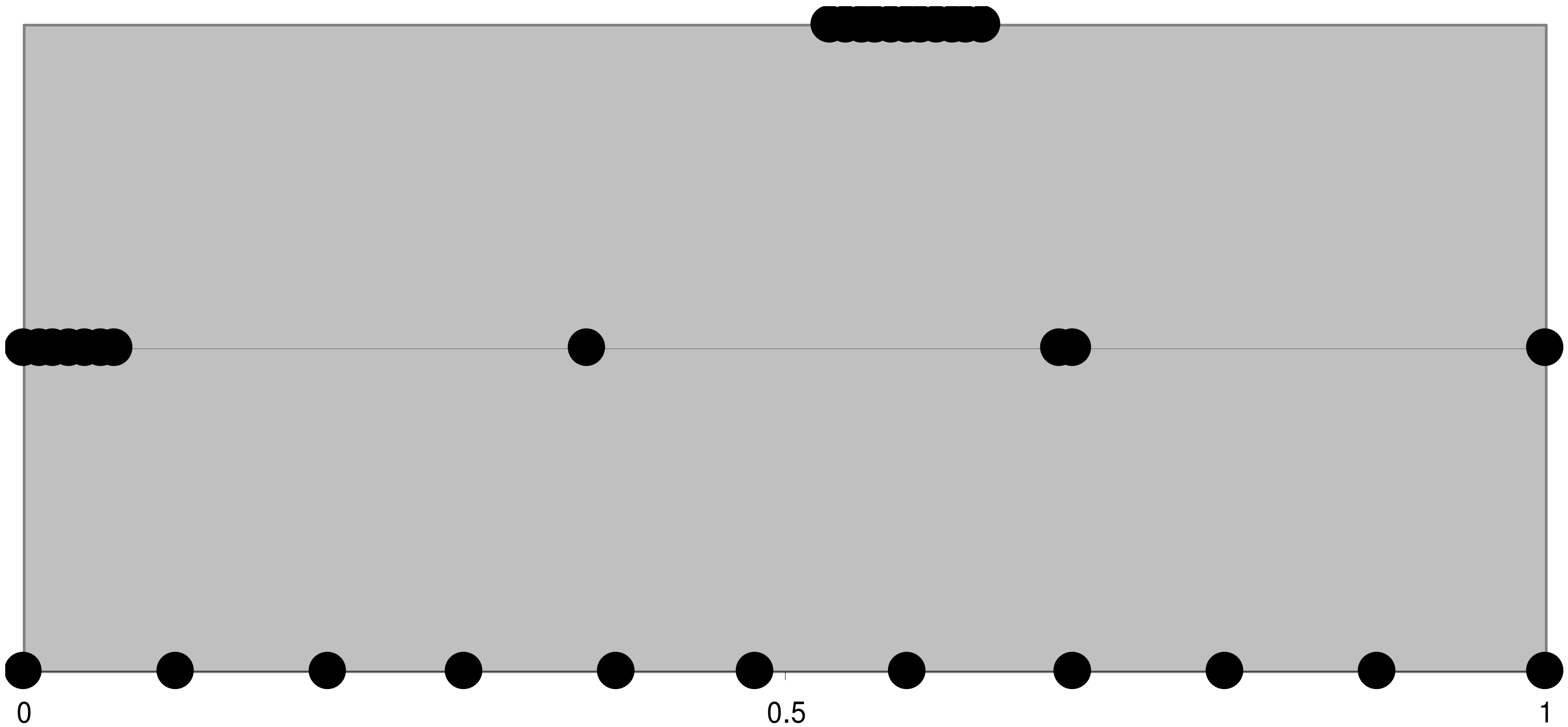
$$x_s^+ = \arg \max_{x \in \bar{X}_s \setminus S_{\xi_s}} \phi(x, \xi_s) \quad \text{and} \quad x_s^- = \arg \min_{x \in S_{\xi_s}} \phi(x, \xi_s).$$

Version: Cook and Nachtsheim, 1980
Survey: Royle, 2002

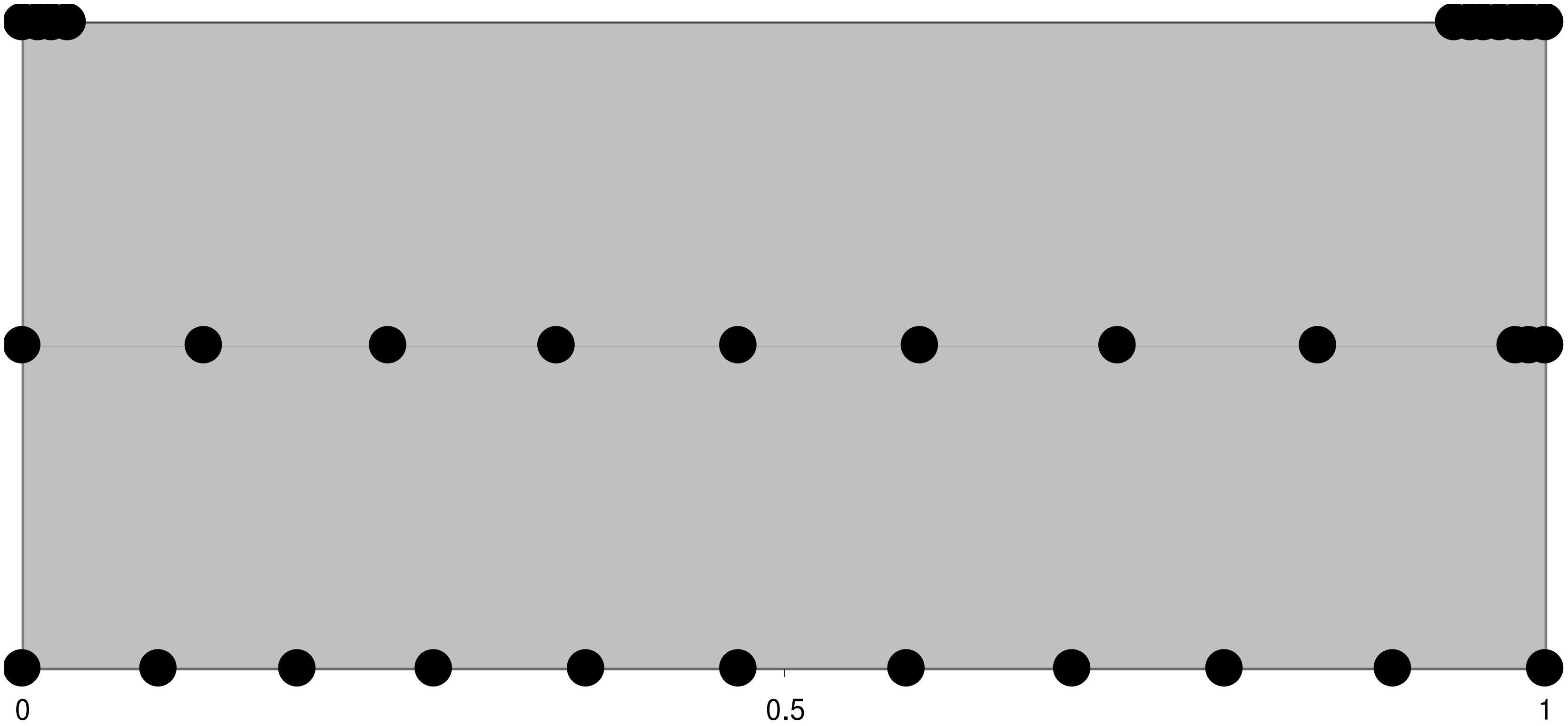
Suggested Variant: Hybrid with Simulated Annealing

- Make the best exchange between a point from sets S_{ξ_s} and a randomly chosen point from $\bar{X}_s \setminus S_{\xi_s}$ at every iteration
- If there is no improvement, give more weight to points farer from the selected and draw anew.
- Perhaps use a stochastic acceptance operator (decreasing temperature) to improve performance.

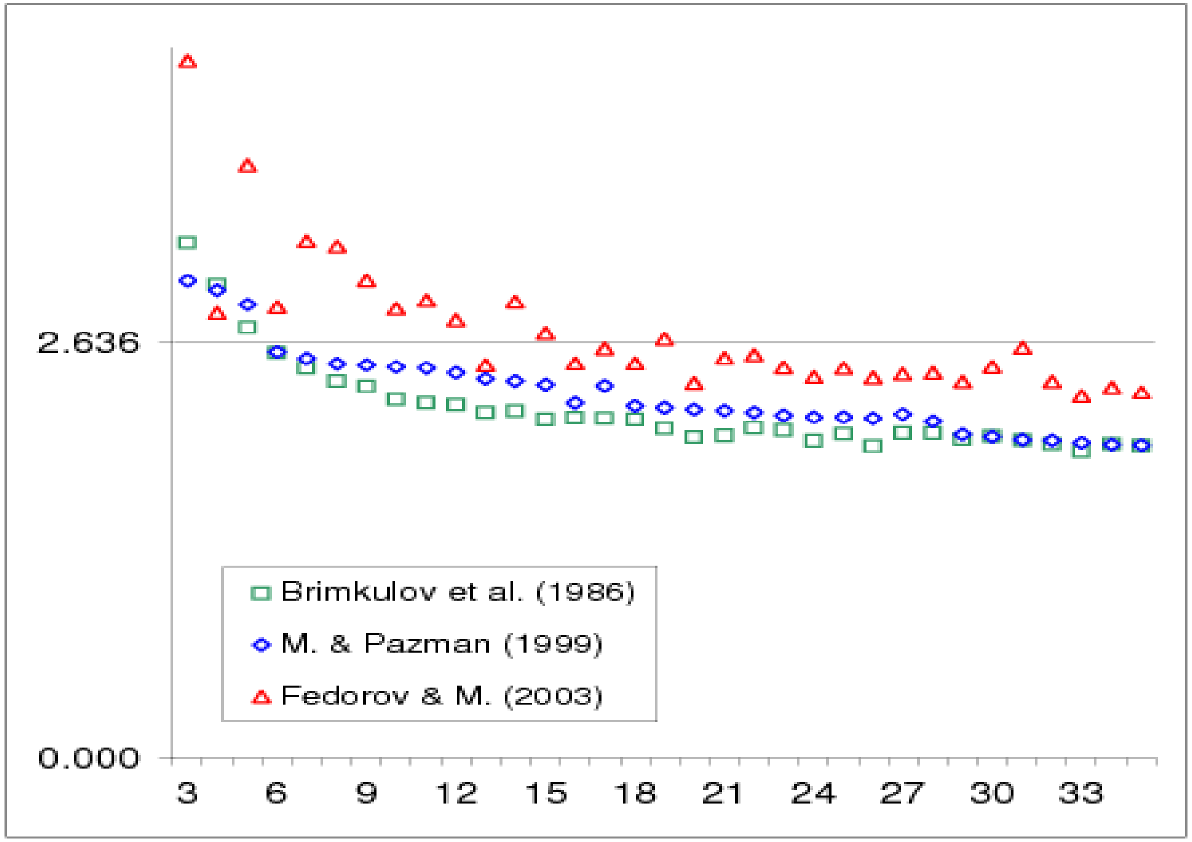
Case 3 for $\theta = 1$ and varying $\alpha = 0, 0.7, 1$



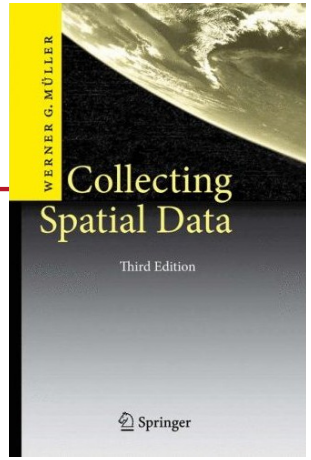
Case 3 for $\alpha=0.9$ and varying $\theta=0.1,1,10$



Efficiency Comparison



References (www.ifas.jku.at)



Müller, W.G., "Collecting Spatial Data. ", 3rd revised and extended edition, Springer Verlag, Heidelberg, 2007.

Müller, W.G. and Stehlík, M., “An Example of D-optimal Designs in the Case of Correlated Errors”, in J. Antoch (Ed.), COMPSTAT2004 Proceedings in Computational Statistics, Springer, 1519-1526, 2004.

Müller, W.G. and Stehlík, M., “D-Optimal Spatial Designs for Estimating Trend and Covariance Parameters”, in Statistics for Spatio-Temporal Modelling, Proceedings of the 4th International Workshop on Statistics for Spatio-Temporal Modelling (METMA 5), Mateu, Porcu, Zocchi (Eds.), 69-77, 2008.

Müller, W.G. and Stehlík, M., “Issues in the Optimal Design of Computer Simulation Experiments” in *Applied Stochastic Models in Business and Industry*, 25(2), 153-177, 2009.

Müller, W.G. and Stehlík, M., “Compound Optimal Spatial Design”, accepted for *Environmetrics*.

Stehlík, M. and Müller, W.G., “Fisher Information in the Design of Computer Simulation Experiments” in *Journal of Physics: Conference Series, 2008.*

ISBIS: The International Society for Business and Industrial Statistics

ISBIS is a new international society and one of the newest Sections of the International Statistical Institute (ISI) founded in April 2005.

http://www.isbis.org/

Soliciting proposals for invited sessions at ISI Dublin 2011 !

Newsletter of the
International Society for Business and Industrial Statistics
A Section of the International Statistical Institute (http://isi.cbs.nl/)

ISBIS News
Volume 1 – Issue 1
November 2008

Editor:
Werner G. Müller
Department of Applied Statistics,
Johannes-Kepler-University of Linz,
Altenbergerstraße 69,
A-4040 Linz, Austria,
Tel. +43-732-2468-5880, E-mail:
werner.mueller@jku.at

Introduction to the first issue of ISBIS News

Welcome to the first issue of ISBIS News, a bi-monthly bulletin of the Society's current activities. It has the sorts of usual Newsletter aims: to provide timely information about

- General ISBIS activities
- Upcoming conferences

know about an interesting conference? Have you seen an interesting new book? Do you want to broadcast your view about something in a Letter to the Editor? Please email us.

2. We need your help! The newsletter sections listed above won't write themselves. If you'd like to coordinate news under one of these headings, please let us know. For example, would you

Problem #1: Non-additivity of the Information Matrix

Leads to unseparability of information contributions through design measures!

$$M(\xi_N) = \frac{1}{N} \sum_{z} \sum_{z'} X(z) \left[C^{-1}(\xi_N) \right]_{z,z'} X^T(z')$$

Remedy: e.g. interpretation of design measures as amount of noise suppression (Pázman & M., 1998, M.+P., 2003)

Problem #2: Use of Fisher Information Matrix

If covariance parameters θ are included in the estimation (cases 2 & 3), the FI matrix contains a block

$$M'(\xi,\theta)\}_{ij} = \frac{1}{2} tr \left\{ C^{-1}(\xi,\theta) \frac{\partial C(\xi,\theta)}{\partial \theta_i} C^{-1}(\xi,\theta) \frac{\partial C(\xi,\theta)}{\partial \theta_j} \right\}$$

Then its interpretation as being inversely proportional to asymptotic covariance matrix of parameters fails (Abt & Welch, 1998).

Remedy: small normal error theory by Pázman (2007).