STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

On the optimal designs for the prediction of Ornstein–Uhlenbeck sheets

Sándor Baran [a], Kinga Sikolya [a], Milan Stehlík [b],*

[a] Faculty of Informatics, University of Debrecen, Kassai út 26, H-4028 Debrecen, Hungary
[b] Institut für Angewandte Statistik, Johannes Kepler University in Linz, Altenberger Straße 69, A-4040 Linz a. D., Austria

ARTICLE INFO

ABSTRACT

Computer simulations are often used to replace physical experiments for exploring the complex relationships between input and output variables. We study the optimal design problem for the prediction of a stationary Ornstein–Uhlenbeck sheet on a monotonic set with respect to the integrated mean square prediction error criterion and the entropy criterion. We show that there is a substantial difference between the shapes of optimal designs for Ornstein–Uhlenbeck processes and sheets. In particular, we show that the optimal prediction based on the integrated mean square prediction error does not necessarily lead to space-filling designs.

1. Introduction

A common problem in spatial statistics is deciding how to choose a set of sample locations in order to predict a random process in an optimal way. In the present paper we study the problem of optimal design for the prediction of an Ornstein–Uhlenbeck (OU) sheet on a monotonic set with respect to the integrated mean square prediction error (IMSPE) criterion and the entropy criterion.

We consider the stationary process

$$Y(s, t) = \theta + \varepsilon(s, t) \tag{1.1}$$

with the design points taken from a compact design space $\mathcal{X} = [a_1, b_1] \times [a_2, b_2]$, where $a_1 < b_1$ and $a_2 < b_2$ and $\varepsilon(s, t)$, $s, t \in \mathbb{R}$, is a stationary OU sheet, that is a zero-mean Gaussian process with covariance structure

$$\mathsf{E}\,\varepsilon(s_1, t_1)\varepsilon(s_2, t_2) = \frac{\widetilde{\sigma}^2}{4\alpha\beta} \exp\left(-\alpha|s_1 - s_2| - \beta|t_1 - t_2|\right), \tag{1.2}$$

where $\alpha > 0$, $\beta > 0$, $\widetilde{\sigma} > 0$. We remark that $\varepsilon(s, t)$ can also be represented as

$$\varepsilon(s, t) = \frac{\widetilde{\sigma}}{2\sqrt{\alpha\beta}} e^{-\alpha s - \beta t} \mathcal{W}\left(e^{2\alpha s}, e^{2\beta t}\right), \tag{1.2}$$

where $\mathcal{W}(s, t)$, $s, t \in \mathbb{R}$, is a standard Brownian sheet (Baran et al., 2003; Baran and Sikolya, 2012).

---

* Corresponding author. Tel.: +43 732 2468 6806; fax: +43 732 2468 9846.
E-mail address: Milan.Stehlik@jku.at (M. Stehlík).

In order to apply the usual notation of spatial modelling (Kiseľák and Stehlík, 2008) we introduce $\sigma := \widetilde{\sigma}/(2\sqrt{\alpha\beta})$ and instead of (1.2) we investigate

$$\mathsf{E}\,\varepsilon(s_1, t_1)\varepsilon(s_2, t_2) = \sigma^2 \exp\left(-\alpha|s_1 - s_2| - \beta|t_1 - t_2|\right), \tag{1.3}$$

where σ is considered as a nuisance parameter.

Further, we require that the shifted OU sheet (1.1) is measured at the points constituting a *monotonic set*. A monotonic set can be defined in an arbitrary Hilbert space H, with real or complex scalars. For $x, y \in H$, we denote by $\langle x, y \rangle$ the real part of the inner product. A set $E \subset H \times H$ is called monotonic (see Minty (1962, 1963)) provided that for all $(x_1, y_1), (x_2, y_2) \in E$ we have $\langle x_1 - x_2, y_1 - y_2 \rangle \geq 0$. A practical example of such a set is provided by the measurements on isotherms of a stationary temperature field, with several applications in thermal slab modelling (see e.g. Babiak et al. (2005) or Koizumi and Jin (2012)). In this paper we consider the following version of a monotonic set:

Condition D. The design points $\{(s_1, t_1), (s_2, t_2), \ldots, (s_n, t_n)\} \subset \mathcal{X}, n \in \mathbb{N}$, *are not overlapping (observations without repetitions); moreover,* $0 < s_1 < s_2 < \cdots < s_n$ *and* $0 < t_1 < t_2 < \cdots < t_n$ *hold.*

For the shifted OU sheet (1.1) observed at points satisfying *Condition D*, Baran and Stehlík (submitted for publication) derived the exact form of the Fisher information (matrices) corresponding to the trend parameter θ and to covariance parameters $r = (\alpha, \beta)$, and the authors investigated the problem of D-optimal designs for the estimation of these quantities.

We derive the exact form of the IMSPE recently used in several papers (see e.g. Crary (2002) or Sacks et al. (1989)), and show that in contrast to the case for the OU process on a compact interval investigated by Baldi Antognini and Zagoraiou (2010), the equidistant design is usually not optimal. This is an important observation, and also consistent with the results of Pronzato and Müller (2012), since statistical folklore is that space-filling designs have superior properties when it comes to the prediction of emulator functions. We also investigate the properties of the optimal design with respect to the entropy criterion (Baldi Antognini and Zagoraiou, 2010; Shewry and Wynn, 1987).

Observe that by condition *Condition D*, the results presented here are not applicable for space-filling designs. In the general case the main problem is the difficulty of finding the inverse of the covariance matrix of the observations. The authors have some preliminary results for regular grid designs, where the covariance matrix is the Kronecker product of two simple covariance matrices corresponding to the two spatial dimensions, and this special form allows finding a reasonable form of the inverse and calculating the objective function for the IMSPE criterion. However, the presentation of these results is beyond the space limitations of this work.

2. Optimal design with respect to the IMSPE criterion

Suppose we observe our process $Y(s, t)$ at the design points $\{(s_1, t_1), (s_2, t_2), \ldots, (s_n, t_n)\}$ satisfying *Condition D*. The main aim of the kriging technique consists in predicting the output of the simulator on the experimental region, and for any untried location $x = (x_1, x_2) \in \mathcal{X}$ the estimation procedure is focused on the best linear unbiased estimator of $Y(x)$ given by $\widehat{Y}(x) = \widehat{\theta} + \mathcal{Q}^\top(x)C^{-1}(n, r)(\mathbf{Y} - \mathbf{1}_n\widehat{\theta})$, where $\mathbf{Y} = (Y(s_1, t_1), \ldots, Y(s_n, t_n))^\top$ is the vector of observations, $\mathbf{1}_n$ is the column vector with entries 1 of length n, $C(n, r)$ is the covariance matrix of $\mathbf{Y}$, $\mathcal{Q}(x)$ is the vector of correlations between $Y(x)$ and $\mathbf{Y}$ defined by $\mathcal{Q}(x) = (\varrho(x, s_1, t_1), \ldots, \varrho(x, s_n, t_n))^\top$ with $\varrho(x, s_i, t_i) := \exp\left(-\alpha|x_1 - s_i| - \beta|x_2 - t_i|\right)$ and $\widehat{\theta}$ is the generalized least squares estimator of θ, that is $\widehat{\theta} = \left(\mathbf{1}_n^\top C^{-1}(n, r)\mathbf{1}_n\right)^{-1} \mathbf{1}_n^\top C^{-1}(n, r)\mathbf{Y}$. Usually, correlation parameters α, β are unknown and will be estimated by the maximum likelihood method. Thus, the kriging predictor is obtained by substituting the maximum likelihood estimators (MLE) $(\widehat{\alpha}, \widehat{\beta})$ for (α, β) and in such a case $\widehat{Y}(x)$ is called the MLE–empirical best linear unbiased predictor (Santner et al., 2003).

Thus natural criteria will minimize suitable functionals of the mean squared prediction error (MSPE) given by

$$\text{MSPE}\big(\widehat{Y}(x)\big) := \sigma^2 \left[1 - \big(1,\ \mathcal{Q}^\top(x)\big) \begin{bmatrix} 0 & \vdots & \mathbf{1}_n^\top \\ \cdots & \cdots & \cdots \\ \mathbf{1}_n & \vdots & C(n, r) \end{bmatrix}^{-1} \big(1,\ \mathcal{Q}^\top(x)\big)^\top \right]. \tag{2.1}$$

Since the prediction accuracy is often related to the entire prediction region $\mathcal{X}$, the design criterion IMSPE is given by

$$\text{IMSPE}\big(\widehat{Y}\big) := \sigma^{-2} \int_{\mathcal{X}} \text{MSPE}\big(\widehat{Y}(x)\big)\,\mathrm{d}x.$$

Without loss of generality, we can assume that the design space $\mathcal{X} = [0, 1]^2$. Therefore, as extrapolative prediction is not advisable in kriging, we can set $t_1 = s_1 = 0$ and $t_n = s_n = 1$.

Theorem 2.1. *In our set-up,*

$$\text{MSPE}\big(\widehat{Y}(x)\big) = \sigma^2 \left[1 - \varrho(x, s_n, t_n)^2 - \sum_{i=1}^{n-1} \frac{(\varrho(x, s_i, t_i) - \varrho(x, s_{i+1}, t_{i+1})q_i)^2}{1 - q_i^2} \right.$$

$$\left. + \left(1 + \sum_{i=1}^{n-1} \frac{1 - q_i}{1 + q_i} \right)^{-1} \left(1 - \varrho(x, s_n, t_n) - \sum_{i=1}^{n-1} \frac{\varrho(x, s_i, t_i) - \varrho(x, s_{i+1}, t_{i+1})q_i}{1 + q_i} \right)^2 \right],$$
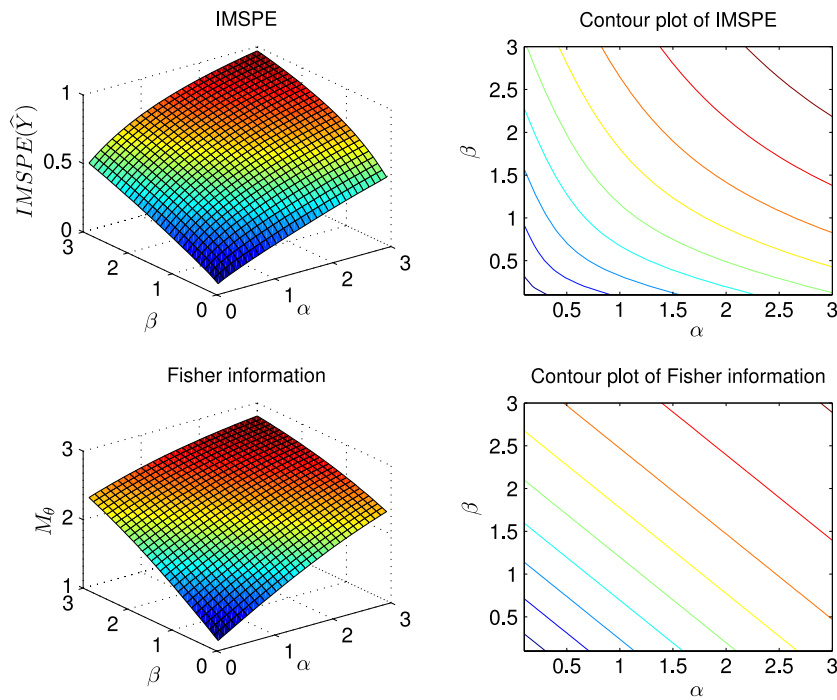
Fig. 1. IMSPE corresponding to the optimal design for the prediction and Fisher information on θ as functions of correlation parameters (α, β) for $n = 3$.

where $q_i := \exp(-\alpha d_i - \beta \delta_i)$ with $d_i := s_{i+1} - s_i$ and $\delta_i := t_{i+1} - t_i$, $i = 1, 2, \ldots, n - 1$. Further,

$$\text{IMSPE}\left(\widehat{Y}\right) = 1 - A_n + \left(1 + \sum_{i=1}^{n-1} \frac{1 - q_i}{1 + q_i}\right)^{-1} B_n, \tag{2.2}$$

where

$$A_n = R_{1,1} + \sum_{i=1}^{n-1} \frac{R_{i,i} q_i^2 - 2R_{i+1,i} q_i + R_{i+1,i+1}}{1 - q_i^2}, \tag{2.3}$$

$$\begin{aligned}
B_n = {} & 1 - 2\omega_n + R_{n,n} + \sum_{i=1}^{n-1}\left(\frac{R_{i,i} - 2R_{i+1,i} q_i + R_{i+1,i+1} q_i^2}{(1 + q_i)^2} + 2\frac{(R_{n,i} - \omega_i) - (R_{n,i+1} - \omega_{i+1})q_i}{1 + q_i}\right) \\
& + 2\sum_{i=2}^{n-1}\sum_{j=1}^{i-1} \frac{R_{i,j} - R_{i+1,j} q_i - R_{i,j+1} q_j + R_{i+1,j+1} q_i q_j}{(1 + q_i)(1 + q_j)},
\end{aligned} \tag{2.4}$$

with

$$\omega_i = \frac{1}{\alpha\beta}\left[2 - e^{-\alpha(d_1 + \cdots + d_{i-1})} - e^{-\alpha(d_i + \cdots + d_{n-1})}\right]\left[2 - e^{-\beta(\delta_1 + \cdots + \delta_{i-1})} - e^{-\beta(\delta_i + \cdots + \delta_{n-1})}\right], \tag{2.5}$$

$$\begin{aligned}
R_{i,j} = {} & \left[\frac{1}{2\alpha}\left(2e^{-\alpha(d_j + \cdots + d_{i-1})} - e^{-\alpha(2d_1 + \cdots + 2d_{j-1} + d_j + \cdots + d_{i-1})} - e^{-\alpha(d_j + \cdots + d_{i-1} + 2d_i + \cdots + 2d_{n-1})}\right) \right.\\
& \left. + (d_j + \cdots + d_{i-1})e^{-\alpha(d_j + \cdots + d_{i-1})}\right] \\
& \times \left[\frac{1}{2\beta}\left(2e^{-\beta(\delta_j + \cdots + \delta_{i-1})} - e^{-\beta(2\delta_1 + \cdots + 2\delta_{j-1} + \delta_j + \cdots + \delta_{i-1})} - e^{-\beta(\delta_j + \cdots + \delta_{i-1} + 2\delta_i + \cdots + 2\delta_{n-1})}\right) \right.\\
& \left. + (\delta_j + \cdots + \delta_{i-1})e^{-\beta(\delta_j + \cdots + \delta_{i-1})}\right], \quad i, j \in \mathbb{N}, \ 1 \le j \le i \le n.
\end{aligned} \tag{2.6}$$

Example 2.2. Consider a three-point design, that is $n = 3$, $s_1 = t_1 = 0$, $s_2 := d$, $t_2 := \delta$, $s_3 = t_3 = 1$. For this particular design we obviously have $d_1 = d$, $d_2 = 1 - d$, $\delta_1 = \delta$, $\delta_2 = 1 - \delta$. By analysing the partial derivatives of IMSPE $\left(\widehat{Y}\right)$ one can easily show that the minimum of IMSPE is reached at $d = 1/2$, $\delta = 1/2$, so for $n = 3$ the equidistant design is optimal. In Fig. 1, IMSPE $\left(\widehat{Y}\right)$ corresponding to the optimal design and the Fisher information on the shift parameter θ are plotted as functions of correlation parameters (α, β). Having a look at the figures for IMSPE and M_θ one can observe that there is
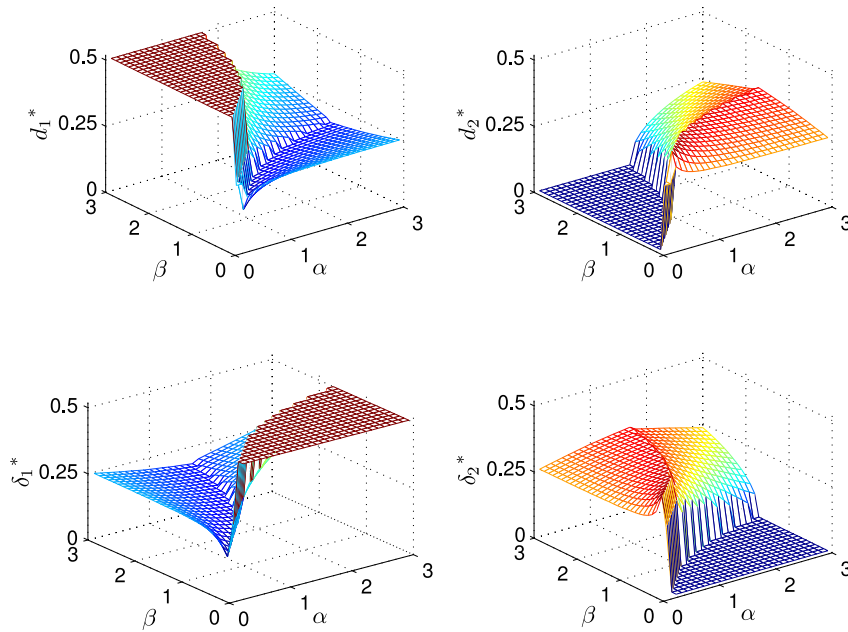
Fig. 2. The values of d_1^*, d_2^* and δ_1^*, δ_2^* giving the optimum of IMSPE as functions of correlation parameters (α, β) for $n = 5$.

a substantial difference between the variations of the estimation and the prediction criterion with respect to correlation parameters (α, β). The equilateral lines for M_θ differ considerably from equilateral curves for IMSPE, which shows the different level of sensitivity of these criteria to the misspecification of correlation parameters. For more information on the sensitivity of optimal design criteria with respect to parameter misspecification, see e.g. Stehlík (2004).

Theorem 2.3. *IMSPE given by* (2.2) *is symmetric in the sense that the interchange of pairs of distances* (d_i, δ_i) *and* (d_{n-i}, δ_{n-i}), $i = 1, 2, \ldots, n - 1$, *does not change the value of* IMSPE $\left(\widehat{Y}\right)$.

Theorem 2.3 gives us some idea of the behaviour of the optimal design with respect to the IMSPE criterion.

Corollary 2.4. *The optimal design with respect to the IMSPE criterion is symmetric, that is if the optimum of* IMSPE $\left(\widehat{Y}\right)$ *is reached at* $(d_1^*, d_2^*, \ldots, d_{n-1}^*, \delta_1^*, \delta_2^*, \ldots, \delta_{n-1}^*)$ *then* $d_i^* = d_{n-i}^*$ *and* $\delta_i^* = \delta_{n-i}^*$, $i = 1, 2, \ldots, n$.

Example 2.5. Consider a five-point design, that is $n = 5$ and $s_1 = t_1 = 0$, $s_5 = t_5 = 1$. By Corollary 2.4, for the optimal design we have $d_1^* = d_4^*$, $d_2^* = d_3^*$ and $\delta_1^* = \delta_4^*$, $\delta_2^* = \delta_3^*$. In Fig. 2 the values of d_1^*, d_2^* and δ_1^*, δ_2^* are plotted as functions of correlation parameters α and β both varying between 0 and 3. We remark that by symmetry, $d_1^* + d_2^* = \delta_1^* + \delta_2^* = 1/2$.

3. The optimal information gain for a shifted OU sheet

Another approach to optimal design is to find locations which maximize the amount of information obtained. Following the ideas of Shewry and Wynn (1987), one has to maximize the entropy Ent(**Y**) of the observations corresponding to the chosen design, which in the Gaussian case form an n-dimensional normal vector with covariance matrix $\sigma^2 C(n, r)$, that is

$$\mathrm{Ent}(\mathbf{Y}) = \frac{n}{2}\left(1 + \ln(2\pi\sigma^2)\right) + \frac{1}{2}\ln\det C(n, r).$$

Theorem 3.1. *Under the conditions of* Theorem 2.1, *entropy* Ent(**Y**) *has the form*

$$\mathrm{Ent}(\mathbf{Y}) = \frac{n}{2}\left(1 + \ln(2\pi\sigma^2)\right) + \frac{1}{2}\sum_{i=1}^{n-1}\ln\left(1 - q_i^2\right). \tag{3.1}$$

For any sample size, the equidistant design when values $\alpha d_i + \beta \delta_i$, $i = 1, 2, \ldots, n - 1$ *(and in this way* q_i), *are constant is optimal with respect to the entropy criterion.*

Remark 3.2. Observe that the equispaced monotonic design $d_1 = d_2 = \cdots d_n$, $\delta_1 = \delta_2 = \cdots = \delta_n$ clearly optimizes the entropy criterion.

Remark 3.3. If the design space $\mathcal{X}$ is the unit square then the optimal value of the entropy equals

$$\frac{n}{2}\left(1 + \ln(2\pi\sigma^2)\right) + \frac{n-1}{2}\ln\left(1 - \exp\left(-(\alpha + \beta)/(n - 1)\right)\right) \to -\infty \quad \text{as } n \to \infty.$$

Table 1
IMSPE values corresponding to the optimal and to the equispaced monotonic design and relative efficiency of the equispaced monotonic design.

n		$\alpha = 0.5, \beta = 0.8$	$\alpha = 1, \beta = 1$	$\alpha = 1, \beta = 10$	$\alpha = 2.5, \beta = 1.5$	$\alpha = 3, \beta = 3$
4	Optimal	0.2602	0.4008	0.9266	0.6549	0.8487
	Equispaced	0.2693	0.4010	0.9326	0.6598	0.8493
	Rel. eff. (%)	96.62	99.95	99.36	99.26	99.93
5	Optimal	0.2309	0.3699	0.8290	0.5981	0.7866
	Equispaced	0.2473	0.3700	0.8409	0.6065	0.7873
	Rel. eff. (%)	93.37	99.97	98.58	98.62	99.91
6	Optimal	0.2130	0.3529	0.7593	0.5640	0.7502
	Equispaced	0.2351	0.3530	0.7766	0.5763	0.7509
	Rel. eff. (%)	90.60	99.97	97.77	97.87	99.91
7	Optimal	0.2007	0.3423	0.7066	0.5241	0.7269
	Equispaced	0.2274	0.3424	0.7288	0.5571	0.7275
	Rel. eff. (%)	88.26	99.97	96.95	94.08	99.92
8	Optimal	0.1692	0.3351	0.6655	0.5007	0.7111
	Equispaced	0.2222	0.3352	0.6918	0.5441	0.7115
	Rel. eff. (%)	76.15	99.97	96.20	92.02	99.94
9	Optimal	0.1620	0.3300	0.6325	0.4858	0.6997
	Equispaced	0.2184	0.3301	0.6626	0.5348	0.7001
	Rel. eff. (%)	74.18	99.97	95.46	90.84	99.94
10	Optimal	0.1570	0.3262	0.6057	0.4756	0.6912
	Equispaced	0.2155	0.3262	0.6390	0.5278	0.6915
	Rel. eff. (%)	72.87	99.98	94.79	90.11	99.95

Table 2
Optimal four-point designs with respect to the IMSPE criterion.

	$\alpha = \beta = 0.5$	$\alpha = \beta = 1$	$\alpha = \beta = 1.5$
$d_1 = \delta_1$	0.3372	0.3401	0.3421
$d_2 = \delta_2$	0.3256	0.3199	0.3158
$d_3 = \delta_3$	0.3372	0.3401	0.3421

4. Numerical experiments

4.1. Comparison of the IMSPE criterion and the entropy criterion

In order to compare the performances of the two criteria we compare the optimal values of $\mathrm{IMSPE}(\widehat{Y})$ calculated using the `fmincon` function of Matlab to its values corresponding to the equispaced monotonic design which is optimal for the entropy criterion. In Table 1 the values of IMSPE are given for both designs together with the relative efficiency of the equispaced monotonic design with respect to the optimal one for various sample sizes and combinations of parameters (α, β). Unfortunately, larger sample sizes cause technical problems in optimization, since for n design points one has to find numerically the constrained minimum of functions with $2n - 2$ parameters.

Observe that for symmetric models when $\alpha = \beta$, the efficiency of the equispaced monotonic design is nearly 100%, while for different covariance parameters it reduces significantly. However, even in this special case the optimal designs with respect to the IMSPE and entropy criteria do not coincide. As an example, consider Table 2 where distances d_i and δ_i corresponding to the optimal four-point design are given for three different values of $\alpha = \beta$. Obviously, in all three cases we have $d_1 + \delta_1 \neq d_2 + \delta_2$, so according to Theorem 3.1 these designs cannot be optimal with respect to the entropy criterion.

4.2. IMSPE on a monotonic set and on a regular grid

In Theorem 2.1, the exact form of IMSPE is derived only for designs satisfying *Condition D*. However, one might ask what is the relative efficiency of the optimal value of IMSPE on monotonic sets containing $n = m^2$ design points compared to the IMSPE of a regular grid with the same number of points. Table 3 gives the optimal values of IMSPE on monotonic sets, IMSPE values for regular designs and the relative efficiencies of the optimal IMSPE values on monotonic sets for different sample sizes and combinations of parameters (α, β). Observe, that for four points the optimal monotonic design gives much better IMSPE values than the regular grid and for $\alpha = \beta = 3$ the relative efficiency is slightly above 100% even in the case $n = 9$.

5. Conclusion

We derive the exact form of the IMSPE for an OU sheet on a monotonic set. We show that the optimal design for the prediction based on IMSPE may differ substantially from the equidistant one. This is in contrast both to the optimal design

Table 3
IMSPE values corresponding to the optimal monotonic and to the regular grid design and relative efficiency of the optimal monotonic design.

n		$\alpha = 0.5, \beta = 0.8$	$\alpha = 1, \beta = 1$	$\alpha = 1, \beta = 10$	$\alpha = 2.5, \beta = 1.5$	$\alpha = 3, \beta = 3$
4	Monotonic	0.2602	0.4008	0.9266	0.6549	0.8487
	Regular	0.3580	0.5389	1.1449	0.8370	1.0094
	Rel. eff. (%)	137.59	134.46	123.56	127.81	118.93
9	Monotonic	0.1620	0.3300	0.6325	0.4858	0.6997
	Regular	0.1527	0.3018	0.5762	0.4850	0.7011
	Rel. eff. (%)	94.26	91.45	91.10	99.84	100.20

for estimation on a monotonic set (Baran and Stehlík, submitted for publication) and to the optimal design for the prediction of the OU process on a compact interval investigated by Baldi Antognini and Zagoraiou (2010). We also investigate the properties of the optimal design with respect to the entropy criterion, where constraining sample points from a rectangle to a monotonic set, as expected, decreases the entropy of the Gaussian field. Simulations illustrate selected cases of optimal designs for small number of sampling locations. Since the above discussed designs depend on the values of the correlation parameters, the optimal designs obtained are only locally optimal. We briefly study the dependence of the designs obtained on these parameters, too. Such knowledge may be crucial for an experimenter for increasing the efficiency of a design in a practical set-up.

Acknowledgements

Appendix

A.1. Proof of Theorem 2.1

To shorten our formulae, in what follows instead of $\varrho(x, s_i, t_i)$ we are using simply ϱ_i, $i = 1, 2, \ldots, n$. Consider first MSPE $\left(\widehat{Y}(x)\right)$ given by (2.1). Short matrix algebraic calculations show that

$$
\begin{bmatrix} 0 & \mathbf{1}_n^\top \\ \mathbf{1}_n & C(n, r) \end{bmatrix}^{-1} = \begin{bmatrix} 0 & \mathbf{0}_n^\top \\ \mathbf{0}_n & C^{-1}(n, r) \end{bmatrix} - \left(\mathbf{1}_n^\top C^{-1}(n, r)\mathbf{1}_n\right)^{-1} \begin{bmatrix} 1 & -\left(C^{-1}(n, r)\mathbf{1}_n\right)^\top \\ -C^{-1}(n, r)\mathbf{1}_n & C^{-1}(n, r)\mathbf{1}_n\left(C^{-1}(n, r)\mathbf{1}_n\right)^\top \end{bmatrix},
$$

and according to the results of Baran and Stehlík (submitted for publication) we have $\mathbf{1}_n^\top C^{-1}(n, r)\mathbf{1}_n = 1 + \sum_{i=1}^{n-1} \frac{1-q_i}{1+q_i}$. We remark that $\mathbf{1}_n^\top C^{-1}(n, r)\mathbf{1}_n$ is the Fisher information on θ based on $\mathbf{Y}$. In this way, we obtain

$$
\text{MSPE}\left(\widehat{Y}(x)\right) = \sigma^2 \left[1 - \mathcal{Q}^\top(x)C^{-1}(n, r)\mathcal{Q}(x) + \left(\mathbf{1}_n^\top C^{-1}(n, r)\mathbf{1}_n\right)^{-1}\left(1 - \mathcal{Q}^\top(x)C^{-1}(n, r)\mathbf{1}_n\right)^2 \right]
$$

$$
= \sigma^2 \left[1 + \frac{\varrho_1^2 - 2\varrho_1\varrho_2 q_1}{q_1^2 - 1} + \frac{\varrho_n^2}{q_{n-1}^2 - 1} - \sum_{i=2}^{n-1}\left(\frac{2\varrho_i\varrho_{i+1}q_i}{q_i^2 - 1} + \frac{\varrho_i^2(1 - q_i^2 q_{i-1}^2)}{(q_i^2 - 1)(q_{i-1}^2 - 1)} \right) \right.
$$

$$
- \left(1 + \sum_{i=1}^{n-1} \frac{1-q_i}{1+q_i}\right)^{-1}
$$

$$
\left. \times \left(1 + \frac{\varrho_1 - (\varrho_1 + \varrho_2)q_1}{q_1^2 - 1} + \frac{\varrho_n}{q_{n-1}^2 - 1} - \sum_{i=2}^{n-1}\left(\frac{(\varrho_i + \varrho_{i+1})q_i}{q_i^2 - 1} + \frac{\varrho_i(1 - q_i^2 q_{i-1}^2)}{(q_i^2 - 1)(q_{i-1}^2 - 1)} \right)\right)^2 \right]
$$

$$
= \sigma^2 \left[1 - \varrho_n^2 - \sum_{i=1}^{n-1} \frac{(\varrho_i - \varrho_{i+1}q_i)^2}{1 - q_i^2} + \left(1 + \sum_{i=1}^{n-1} \frac{1-q_i}{1+q_i}\right)^{-1}\left(1 - \varrho_n - \sum_{i=1}^{n-1} \frac{\varrho_i - \varrho_{i+1}q_i}{1 + q_i}\right)^2 \right].
$$

Further, according to the definition of the IMSPE criterion, we can write

$$
\text{IMSPE}\left(\widehat{Y}\right) = 1 - A_n + \left(1 + \sum_{i=1}^{n-1} \frac{1-q_i}{1+q_i}\right)^{-1} B_n,
$$

where

$$A_n := \int_{\mathcal{X}} \mathcal{Q}^\top(x) C^{-1}(n, r) \mathcal{Q}(x) dx = \mathrm{tr}\left[C^{-1}(n, r) \mathcal{R}\right],$$

$$B_n := 1 - 2\,\mathbf{1}_n^\top C^{-1}(n, r)\mathcal{W} + \mathbf{1}_n^\top C^{-1}(n, r)\mathcal{R}C^{-1}(n, r)\mathbf{1}_n,$$

with

$$\mathcal{W} = (\omega_1, \ldots, \omega_n)^\top := \int_{\mathcal{X}} \mathcal{Q}(x) dx \quad \text{and} \quad \mathcal{R} = \{R_{i,j}\} := \int_{\mathcal{X}} \mathcal{Q}(x)\mathcal{Q}^\top(x) dx.$$

Obviously,

$$\omega_i = \frac{1}{\alpha\beta}\left[2 - e^{-\alpha s_i} - e^{-\alpha(1-s_i)}\right]\left[2 - e^{-\beta t_i} - e^{-\beta(1-t_i)}\right]$$

$$= \frac{1}{\alpha\beta}\left[2 - e^{-\alpha(d_1+\cdots+d_{i-1})} - e^{-\alpha(d_i+\cdots+d_{n-1})}\right]\left[2 - e^{-\beta(\delta_1+\cdots+\delta_{i-1})} - e^{-\beta(\delta_i+\cdots+\delta_{n-1})}\right],$$

$$R_{i,j} = \left[\frac{1}{2\alpha}\left(2e^{-\alpha|s_i-s_j|} - e^{-\alpha(s_i+s_j)} - e^{-\alpha(2-s_i-s_j)}\right) + |s_i - s_j|e^{-\alpha|s_i-s_j|}\right]$$

$$\times \left[\frac{1}{2\beta}\left(2e^{-\beta|t_i-t_j|} - e^{-\beta(t_i+t_j)} - e^{-\beta(2-t_i-t_j)}\right) + |t_i - t_j|e^{-\beta|t_i-t_j|}\right]$$

$$= \left[\frac{1}{2\alpha}\left(2e^{-\alpha(d_{i\wedge j}+\cdots+d_{i\vee j-1})} - e^{-\alpha(2d_1+\cdots+2d_{i\wedge j-1}+d_{i\wedge j}+\cdots+d_{i\vee j-1})} - e^{-\alpha(d_{i\wedge j}+\cdots+d_{i\vee j-1}+2d_{i\vee j}+\cdots+2d_{n-1})}\right)\right.$$

$$\left. + (d_{i\wedge j}+\cdots+d_{i\vee j-1})e^{-\alpha(d_{i\wedge j}+\cdots+d_{i\vee j-1})}\right]$$

$$\times \left[\frac{1}{2\beta}\left(2e^{-\beta(\delta_{i\wedge j}+\cdots+\delta_{i\vee j-1})} - e^{-\beta(2\delta_1+\cdots+2\delta_{i\wedge j-1}+\delta_{i\wedge j}+\cdots+\delta_{i\vee j-1})} - e^{-\beta(\delta_{i\wedge j}+\cdots+\delta_{i\vee j-1}+2\delta_{i\vee j}+\cdots+2\delta_{n-1})}\right)\right.$$

$$\left. + (\delta_{i\wedge j}+\cdots+\delta_{i\vee j-1})e^{-\beta(\delta_{i\wedge j}+\cdots+\delta_{i\vee j-1})}\right],$$

where $i \wedge j := \min\{i, j\}$, $i \vee j := \max\{i, j\}$, $i, j \in \mathbb{N}$, and the empty sum is defined to be zero. In this way, we obtain

$$A_n = R_{n,n} + \sum_{i=1}^{n-1}\frac{R_{i,i} - 2R_{i+1,i}q_i + R_{i+1,i+1}q_i^2}{1 - q_i^2} = R_{1,1} + \sum_{i=1}^{n-1}\frac{R_{i,i}q_i^2 - 2R_{i+1,i}q_i + R_{i+1,i+1}}{1 - q_i^2},$$

$$B_n = 1 - 2\omega_n - 2\sum_{i=1}^{n-1}\frac{\omega_i - \omega_{i+1}q_i}{1 + q_i} + R_{n,n}$$

$$+ 2\sum_{i=1}^{n-1}\frac{R_{n,i} - R_{n,i+1}q_i}{1 + q_i} + \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\frac{R_{i,j} - R_{i+1,j}q_i - R_{i,j+1}q_j + R_{i+1,j+1}q_iq_j}{(1 + q_i)(1 + q_j)}$$

$$= 1 - 2\omega_n + R_{n,n} + \sum_{i=1}^{n-1}\left(\frac{R_{i,i} - 2R_{i+1,i}q_i + R_{i+1,i+1}q_i^2}{(1 + q_i)^2} + 2\frac{(R_{n,i} - \omega_i) - (R_{n,i+1} - \omega_{i+1})q_i}{1 + q_i}\right)$$

$$+ 2\sum_{i=2}^{n-1}\sum_{j=1}^{i-1}\frac{R_{i,j} - R_{i+1,j}q_i - R_{i,j+1}q_j + R_{i+1,j+1}q_iq_j}{(1 + q_i)(1 + q_j)}. \qquad \square$$

A.2. Proof of Theorem 2.3

Let $\widetilde{d}_i := d_{n-i}$, $\widetilde{\delta}_i := \delta_{n-i}$, $\widetilde{q}_i := \exp\left(-\alpha\widetilde{d}_i - \beta\widetilde{\delta}_i\right)$, $i = 1, 2, \ldots, n - 1$, and denote by $\widetilde{\omega}_k$ and $\widetilde{R}_{k,\ell}$ the values calculated from (2.5) and (2.6), respectively, using distances $(\widetilde{d}_i, \widetilde{\delta}_i)$, $i = 1, 2, \ldots, n - 1$. Further, let q_i, ω_k and $R_{k,\ell}$ denote the corresponding quantities calculated from the original distances (d_i, δ_i), $i = 1, 2, \ldots, n - 1$. It is easy to see that

$$\widetilde{q}_i = q_{n-i}, \quad i = 1, 2, \ldots, n - 1, \quad \text{and} \quad \widetilde{\omega}_k = \omega_{n-k+1}, \qquad \widetilde{R}_{k,\ell} = R_{n-k+1,n-\ell+1}, \quad k, \ell = 1, 2, \ldots, n, \qquad (A.1)$$

and in this way we obtain

$$\sum_{i=1}^{n-1}\frac{1 - \widetilde{q}_i}{1 + \widetilde{q}_i} = \sum_{i=1}^{n-1}\frac{1 - q_i}{1 + q_i}.$$

Now, denote by $\widetilde{A}_n$ and $\widetilde{B}_n$ the values calculated from (2.3) and (2.4), respectively, using distances $(\widetilde{d}_i, \widetilde{\delta}_i)$, $i = 1, 2, \ldots, n-1$. Then with the help of (A.1), after straightforward but tedious calculations one can show that $\widetilde{A}_n = A_n$ and $\widetilde{B}_n = B_n$, which completes the proof. $\square$

A.3. Proof of Theorem 3.1

According to the results of Baran and Stehlík (submitted for publication),

$$
C(n, r) = \begin{bmatrix}
1 & q_1 & q_1 q_2 & q_1 q_2 q_3 & \cdots & \cdots & \prod_{i=1}^{n-1} q_i \\
q_1 & 1 & q_2 & q_2 q_3 & \cdots & \cdots & \prod_{i=2}^{n-1} q_i \\
q_1 q_2 & q_2 & 1 & q_3 & \cdots & \cdots & \prod_{i=3}^{n-1} q_i \\
q_1 q_2 q_3 & q_2 q_3 & q_3 & 1 & \cdots & \cdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & & \vdots \\
\vdots & \vdots & \vdots & \vdots & & \ddots & q_{n-1} \\
\prod_{i=1}^{n-1} q_i & \prod_{i=2}^{n-1} q_i & \prod_{i=3}^{n-1} q_i & \cdots & \cdots & q_{n-1} & 1
\end{bmatrix}.
$$

Hence, one can use the same type of decomposition as Baldi Antognini and Zagoraiou (2010, Lemma 3.1), that is

$$ C(n, r) = LDL^\top, $$

where L is a lower triangular matrix with entries 1 in its main diagonal and D is a diagonal matrix where in the diagonal we have $\left(1, 1 - q_1^2, 1 - q_2^2, \ldots, 1 - q_{n-1}^2\right)$. In this way

$$ \det C(n, r) = \det D = \prod_{i=1}^{n-1} (1 - q_i^2), $$

which proves (3.1).

In order to find the optimal design one has to find the constrained maximum of

$$ F(q_1, q_2, \ldots, q_{n-1}) := \sum_{i=1}^{n-1} \ln\left(1 - q_i^2\right) \quad \text{under condition} \quad \sum_{i=1}^{n-1} \ln q_i = -\alpha - \beta. $$

By analysing the first partial derivatives and the Hessian of the Lagrange function

$$ \Lambda(q_1, q_2, \ldots, q_{n-1}; \lambda) := \sum_{i=1}^{n-1} \ln\left(1 - q_i^2\right) + \lambda\left(\sum_{i=1}^{n-1} \ln q_i + \alpha + \beta\right) $$

one can easily see that the maximum is reached when $q_1 = q_2 = \cdots = q_{n-1}$, which completes the proof. $\square$

References

Babiak, J., Minárová, M., Petráš, D., 2005. Principles and calculations of temperature distribution in an active slab depending up various operation modes of TABS using FEM software. In: Proceeding in CLIMA Congress, Lausanne.

Baldi Antognini, A., Zagoraiou, M., 2010. Exact optimal designs for computer experiments via Kriging metamodelling. J. Statist. Plann. Inference 140, 2607–2617.

Baran, S., Pap, G., Zuijlen, M.V., 2003. Estimation of the mean of stationary and nonstationary Ornstein–Uhlenbeck processes and sheets. Comput. Math. Appl. 45, 563–579.

Baran, S., Sikolya, K., 2012. Parameter estimation in linear regression driven by a Gaussian sheet. Acta Sci. Math. (Szeged) 78, 689–713.

Baran, S., Stehlík, M., 2013. Optimal designs for parameters of shifted Ornstein–Uhlenbeck sheets measured on monotonic sets. Appl. Stoch. Model. Bus. (submitted for publication).

Crary, S.B., 2002. Design of computer experiments for metamodel generation. Analog Integr. Circuits Signal Process. 32, 7–16.

Kiselák, J., Stehlík, M., 2008. Equidistant D-optimal designs for parameters of Ornstein–Uhlenbeck process. Statist. Probab. Lett. 78, 1388–1396.

Koizumi, H., Jin, Y.H., 2012. Performance enhancement of a latent heat thermal energy storage system using curved-slab containers. Appl. Therm. Eng. 37, 145–153.

Minty, G.J., 1962. Monotone (non-linear) operators in Hilbert space. Duke Math. J. 29, 341–346.

Minty, G.J., 1963. On a monotonicity method for the solution of nonlinear equations in Banach spaces. Proc. Natl. Acad. Sci. US 50, 1038–1041.

Pronzato, L., Müller, W.G., 2012. Design of computer experiments: space filling and beyond. Stat. Comput. 22, 681–701.

Sacks, J., Schiller, S.B., Welch, W.J., 1989. Design for computer experiments. Technometrics 31, 41–47.

Santner, T.J., Williams, B.J., Notz, W.I., 2003. The Design and Analysis of Computer Experiments. Springer-Verlag, New York.

Shewry, M.C., Wynn, H.P., 1987. Maximum entropy sampling. J. Appl. Stat. 14, 165–170.

Stehlík, M., 2004. Some properties of D-optimal designs for random fields with different variograms. Research Report Series Report No. 4, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Austria.