VERSITA

TATRA MOUNTAINS
Mathematical Publications

DOI: 10.2478/v10127-012-0011-2
Tatra Mt. Math. Publ. **51** (2012), 101–113

# PLANNING OF EXPERIMENTS FOR A NONAUTONOMOUS ORNSTEIN-UHLENBECK PROCESS

Vladimír Lacko

ABSTRACT. We study exact optimal designs for processes governed by mean--reversion stochastic differential equations with a time dependent volatility and known mean-reversion speed. It turns out that any mean-reversion Itō process has a product covariance structure. We prove the existence of a nondegenerate optimal sampling design for the parameter estimation and derive the information matrix corresponding to the observation of the full path. The results are demonstrated on a process with exponential volatility.

## 1. Introduction

The model underlying this paper is a nonautonomous nonstationary Ornstein--Uhlenbeck process, that is, an Itō process $\{X_t \mid t \geq 0\}$ governed by a stochastic differential equation (SDE) of the form

$$\begin{aligned} \mathrm{d}X_t &= \kappa(\bar{X} - X_t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t, \\ X_0 &\quad \text{unknown}, \end{aligned} \tag{1}$$

which can be observed at $n$ distinct times in the experimental domain $[T_*, T^*]$, $0 < T_* < T^*$. Here, the initial point $X_0$ and the asymptotic expectation $\bar{X}$ are unknown parameters, $\kappa > 0$ is known mean-reversion speed, $\sigma(\cdot) : \langle 0, \infty) \mapsto (0, \infty)$ is (up to a constant multiple) known deterministic and semicontinuous volatility function, and $W_t$ denotes a Wiener process.

The model (1) is motivated by the autonomous nonstationary Ornstein-Uhlenbeck process (ANOUP), which corresponds to its constant volatility version. The ANOUP has found many applications in different research fields such as


2010 Mathematics Subject Classification: Primary 62K05.
Keywords: exact design, product covariance structure, Itō stochastic differential equation, estimation, asymptotic information matrix, efficiency.
The research was supported by the Slovak VEGA Grant No. 2/0038/12 and the Comenius University Grant No. UK/383/2011.

physics or biology. In physics [9], [19], the ANOUP is a noise relaxation process, which describes the velocity of a particle under the influence of a friction. If we consider a Hookean spring, then the mean-reversion speed $\kappa$ is given by the ratio of the spring constant $k$ and the friction coefficient $\gamma$, and the volatility $\sigma(t)$ is constantly equal to $\sqrt{2k_{\mathrm{B}}T/\gamma}$, where $k_{\mathrm{B}}$ is the Boltzmann constant and $T$ is the temperature. In biology, the ANOUP is often employed for modelling neuronal response [15]. The governed variable $X_t$ expresses the voltage difference between the membrane and resting potentials at the trigger zone of the neuron, and we are interested in the initial and stationary difference in potentials. The value of $\kappa$ is the reciprocal of the membrane constant. However, the situation is quite different if we expect, for example, that in the problem of the Hookean spring the temperature $T$ is a function of the time (e.g., cooling or heating of the physical system), and we get $\sigma(t) = \sqrt{2k_{\mathrm{B}}T(t)/\gamma}$. Since the volatility is not constant, as it can be seen in the ANOUP, we should use the model (1) instead.

From the point of view of optimal design of experiments, stationary Ornstein-Uhlenbeck process was studied by K i s e ľ á k and S t e h l í k [7] and Z a g o r a i o u and A n t o g n i n i [20], who showed the optimality of equidistant sampling. The same result was achieved by H a r m a n and Š t u l a j t e r [3], [5] in the case of its nonstationary counterpart.

The purpose of this paper is to find an optimal $n$-point design for the model (1), that is, an $n$-vector $\boldsymbol{\tau} = (t_1, \dots, t_n)'$ of strictly increasing values from the experimental domain $[T_*, T^*]$, which is in a desired way optimal for estimating unknown parameters. In Section 2 we will formulate the model (1) in terms of linear regression with correlated errors. The information matrix and existence of the optimal designs are discussed in Section 3. A demonstration of the presented results on an example is given in Section 4.

We will use the following *notation*: By $\mathcal{T}_n = \{(t_1, \dots, t_n)' \mid T_* \le t_1 < t_2 < \cdots < t_n \le T^*\}$ we denote the set of all $n$-point design, $\overline{\mathcal{T}}_n$ denotes the closure of $\mathcal{T}_n$, and by $\mathcal{T}_n(\tilde{t})$ we denote the set of all $n$-point designs with $t_1 = \tilde{t}$. For a given design $\boldsymbol{\tau} \in \mathcal{T}_n$ we have $\boldsymbol{X}_{\boldsymbol{\tau}} = (X_{t_1}, \dots, X_{t_n})'$, $\mathbf{e}^{-\kappa\boldsymbol{\tau}} = (\mathrm{e}^{-\kappa t_1}, \dots, \mathrm{e}^{-\kappa t_n})'$, and $\boldsymbol{\varepsilon}_{\boldsymbol{\tau}} = (\varepsilon_{t_1}, \dots, \varepsilon_{t_n})'$. The symbol $\mathbf{1}_n$ denotes the $n$-vector $(1, \dots, 1)'$, and $\mathbf{0}_n$ denotes the $n$-dimensional zero vector.

## 2. Corresponding linear regression model

After applying the I t ō lemma [6] to the transformation $\mathrm{e}^{\kappa t}(X_t - \bar{X})$ we obtain the solution to the SDE (1)

$$X_t = \mathrm{e}^{-\kappa t}X_0 + (1 - \mathrm{e}^{-\kappa t})\bar{X} + \int_0^t \mathrm{e}^{-\kappa(t-\nu)}\sigma(\nu)\,\mathrm{d}W_\nu. \tag{2}$$

Thus, the observations of the process driven by (1) at the design points $t_1, \dots, t_n$ satisfy the linear regression model

$$\boldsymbol{X}_{\boldsymbol{\tau}} = (\mathbf{e}^{-\kappa\boldsymbol{\tau}})X_0 + (\mathbf{1}_n - \mathbf{e}^{-\kappa\boldsymbol{\tau}})\bar{X} + \boldsymbol{\varepsilon}_{\boldsymbol{\tau}} = \mathbf{F}(\boldsymbol{\tau})\boldsymbol{\theta} + \boldsymbol{\varepsilon}_{\boldsymbol{\tau}}, \tag{3}$$

where $\mathbf{F}(\boldsymbol{\tau}) = (\mathbf{e}^{-\kappa\boldsymbol{\tau}}, \mathbf{1}_n - \mathbf{e}^{-\kappa\boldsymbol{\tau}})$ is the design matrix, $\boldsymbol{\theta} = (X_0, \bar{X})'$ is the vector of unknown parameters, and $\boldsymbol{\varepsilon}_{\boldsymbol{\tau}} = (\varepsilon_{t_1}, \dots, \varepsilon_{t_n})'$ is a vector of random errors such that

$$\mathbb{E}[\boldsymbol{\varepsilon}_{\boldsymbol{\tau}}] = \mathbf{0}_n \quad \text{and} \quad \mathbb{V}\mathrm{ar}[\boldsymbol{\varepsilon}_{\boldsymbol{\tau}}] = \boldsymbol{\Sigma}(\boldsymbol{\tau}). \tag{4}$$

We remark that the distribution of the vector $\boldsymbol{\varepsilon}_{\boldsymbol{\tau}}$ is Gaussian. In the sequel, we will derive the variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau})$, which is crucial for computing the information matrix $\mathbf{M}(\boldsymbol{\tau})$, and show that $\boldsymbol{\Sigma}(\boldsymbol{\tau})$ is positive definite for any $\boldsymbol{\tau} \in \mathcal{T}_n$.

We shall consider conditioning upon the value $x_0$ of $X_0$. A basic rule for covariance gives that $\mathbb{C}\mathrm{ov}[X_t, X_{t+s} \mid X_0 = x_0] = \mathbb{E}[X_t X_{t+s} \mid X_0 = x_0] - \mathbb{E}[X_t \mid X_0 = x_0]\mathbb{E}[X_{t+s} \mid X_0 = x_0]$. The expectations $\mathbb{E}[X_t \mid X_0 = x_0]$ and $\mathbb{E}[X_{t+s} \mid X_0 = x_0]$ are known (cf. (2)), henceforth we need to find $\mathbb{E}[X_t X_{t+s} \mid X_0 = x_0]$.

The key for computing $\mathbb{E}[X_t X_{t+s} \mid X_0 = x_0]$ is the transition kernel $f(x, t \mid x_\nu, \nu) = \frac{\mathrm{d}}{\mathrm{d}x}\Pr[X_t < x \mid X_\nu = x_\nu]$, $t > \nu$, of the process $X_t$ defined by the SDE (1), which solves the well-known Kolmogorov's forward equation [1].

Since $X_t$ is a Markov process, its transition density kernel satisfies

$$\forall_{\nu\in(0,t)} \ \Pr[X_t < x \mid X_0 = x_0] = \int_{-\infty}^{x} \left( \int_{\mathbb{R}} f(z, t-\nu \mid y, \nu) f(y, \nu \mid x_0, 0)\, \mathrm{d}y \right) \mathrm{d}z.$$

Consequently, using (2) we obtain

$$\begin{aligned}
\mathbb{E}[X_t X_{t+s} \mid X_0 = x_0] &= \int_{\mathbb{R}} x_1 f(x_1, t \mid x_0, 0) \left( \int_{\mathbb{R}} x_2 f(x_2, s \mid x_1, t)\, \mathrm{d}x_2 \right) \mathrm{d}x_1 \\
&= \int_{\mathbb{R}} x \mathbb{E}[X_{t+s} \mid X_t = x] f(x, t \mid x_0, 0)\, \mathrm{d}x \\
&= \mathrm{e}^{-\kappa s}\mathbb{E}[X_t^2 \mid X_0 = x_0] + \bar{X}(1 - \mathrm{e}^{-\kappa s})\mathbb{E}[X_t \mid X_0 = x_0] \\
&= \mathrm{e}^{-\kappa s}\mathbb{V}\mathrm{ar}[X_t \mid X_0 = x_0] + \mathrm{e}^{-\kappa s}\mathbb{E}^2[X_t \mid X_0 = x_0] \\
&\quad + \bar{X}(1 - \mathrm{e}^{-\kappa s})\mathbb{E}[X_t \mid X_0 = x_0],
\end{aligned}$$

which implies that

$$\mathbb{C}\mathrm{ov}[X_t, X_{t+s} \mid X_0 = x_0] = \mathrm{e}^{-\kappa s} D(t), \tag{5}$$

where $D(t) = \mathbb{V}\mathrm{ar}[X_t \mid X_0 = x_0]$. It can be shown [2] that the variance $D(t)$ of $X_t$ governed by SDE (1) follows the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}D(t) = -2\kappa D(t) + \sigma^2(t)$$

with the initial condition $D(0) = 0$. Using standard methods of solving ordinary differential equations we obtain the explicit solution

$$D(t) = \mathrm{e}^{-2\kappa t}\int_0^t \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu. \tag{6}$$

The same result follows from the Itō isometry [11]. Note that $D(t)$ is always positive for $t > 0$. We refer the reader to [11] for elements of the stochastic calculus.

The statement given in (5) holds in a more general situation if we assume the volatility function of the form $\sigma(X_t, t)$. Then the variance $D(X_0, t) = \mathbb{V}\mathrm{ar}[X_t \mid X_0 = x_0]$ might be influenced by the value $x_0$ of $X_0$, which results in a technically much more complicated model. However, this case is behind the scope of the presented paper.

The relations (5) and (6) yield:

**Lemma 1.** *The $ij$th element, $i \le j$, of the variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau})$ defined in* (4) *has the form*

$$\big\{\boldsymbol{\Sigma}(\boldsymbol{\tau})\big\}_{ij} = u(t_i)v(t_j), \tag{7}$$

*where*

$$u(t_i) = \mathrm{e}^{-\kappa t_i}\int_0^{t_i} \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu \quad \textit{and} \quad v(t_j) = \mathrm{e}^{-\kappa t_j}.$$

A standard approach to estimating parameters of a linear regression model is to employ the weighted least squares. The estimator is then given by

$$\hat{\boldsymbol{\theta}}(\boldsymbol{\tau}) = \mathbf{M}^{-1}(\boldsymbol{\tau})\mathbf{F}'(\boldsymbol{\tau})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\tau})\boldsymbol{X}_{\boldsymbol{\tau}}, \tag{8}$$

where

$$\mathbf{M}(\boldsymbol{\tau}) = \mathbf{F}'(\boldsymbol{\tau})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\tau})\mathbf{F}(\boldsymbol{\tau}),$$

and, according to the Gauss-Markov theorem [13], is the best linear unbiased estimator regardless of the distribution of the errors. Here $\mathbf{M}(\boldsymbol{\tau}) = \mathbb{V}\mathrm{ar}^{-1}\big[\hat{\boldsymbol{\theta}}(\boldsymbol{\tau})\big]$ is the information matrix, and so it is reasonable to expect an experimenter to choose a design $\tilde{\boldsymbol{\tau}} \in \mathcal{T}_n$, for which the matrix $\mathbf{M}(\tilde{\boldsymbol{\tau}})$ is "large". We will discuss the selection of an optimal design in the next section.

Notice that the least squares estimator and information matrix in (8) assumed the invertibility of the variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau})$. The next lemma states that $\boldsymbol{\Sigma}(\boldsymbol{\tau})$ is positive definite.

**Lemma 2.** *The variance-covariance matrix* $\mathbf{\Sigma}(\boldsymbol{\tau})$ *given by* (7) *is positive definite for any* $\boldsymbol{\tau} \in \mathcal{T}_n$, $n \geq 2$.

Proof. For a design $(t_1, t_2)' \in \mathcal{T}_2$, i.e., $t_1 < t_2$, we have $\big\{\mathbf{\Sigma}(t_1, t_2)\big\}_{11} = D(t_1) > 0$ and $\det\big[\mathbf{\Sigma}(t_1, t_2)\big] = D(t_1)D(t_2) - \mathrm{e}^{-2\kappa(t_2-t_1)}D^2(t_1) > 0$ because

$$\begin{aligned} D(t_2) &= \mathrm{e}^{-2\kappa t_2}\int_0^{t_2} \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu \\ &> \mathrm{e}^{-2\kappa(t_2-t_1)}\mathrm{e}^{-2\kappa t_1}\int_0^{t_1} \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu \qquad (9) \\ &= \mathrm{e}^{-2\kappa(t_2-t_1)}D(t_1). \end{aligned}$$

Now, assume that $\boldsymbol{\tau}_n = (t_1, \ldots, t_n)' \in \mathcal{T}_n$, $\mathbf{\Sigma}(\boldsymbol{\tau}_n)$ be positive definite, and, without loss of generality, $\boldsymbol{\tau}_{n+1} = (\boldsymbol{\tau}_n', t_{n+1})' \in \mathcal{T}_{n+1}$. Then

$$\mathbf{\Sigma}(\boldsymbol{\tau}_{n+1}) = \begin{pmatrix} \mathbf{\Sigma}(\boldsymbol{\tau}_n) & \boldsymbol{s} \\ \boldsymbol{s}' & D(t_{n+1}) \end{pmatrix},$$

where $\boldsymbol{s} = \big(\mathrm{e}^{-\kappa(t_{n+1}-t_1)}D(t_1), \ldots, \mathrm{e}^{-\kappa(t_{n+1}-t_n)}D(t_n)\big)'$. Since $\mathbf{\Sigma}(\boldsymbol{\tau}_n)$ is nonsingular, the matrix $\mathbf{\Sigma}(\boldsymbol{\tau}_{n+1})$ is row-equivalent to

$$\begin{pmatrix} \mathbf{\Sigma}(\boldsymbol{\tau}_n) & \boldsymbol{s} \\ \mathbf{0}_n' & D(t_{n+1}) - \boldsymbol{s}'\mathbf{\Sigma}^{-1}(\boldsymbol{\tau}_n)\boldsymbol{s} \end{pmatrix}.$$

The expression $\mathbf{\Sigma}^{-1}(\boldsymbol{\tau}_n)\boldsymbol{s}$ is equal to the $n$th unit vector $(0, \ldots, 0, 1)'$, cf. [3], hence $\boldsymbol{s}'\mathbf{\Sigma}^{-1}(\boldsymbol{\tau}_n)\boldsymbol{s} = \mathrm{e}^{-\kappa(t_{n+1}-t_n)}D(t_n)$. We can use the relation (9) to proof the positivity of $D(t_{n+1}) - \mathrm{e}^{-\kappa(t_{n+1}-t_n)}D(t_n)$. □

The covariance structure (7) has the so-called product form. Besides the model with the simplest product covariance structure—the Brownian motion with a time-dependent drift [4], [16]—we can find the product covariance structure also in other design problems, see, for instance, [3], [5], [10]. In particular, Harman and Štulajter [3], [5] analysed the autonomous nonstationary Ornstein-Uhlenbeck process discussed in Introduction, which in its original form coincides with the SDE (1) with a constant volatility $\sigma(t) \equiv \sigma$, and assumed the corresponding linear regression model to have a more general response function of the form $\mathbb{E}[X_t] = (a_1 + b_1\mathrm{e}^{-\kappa t})\theta_1 + (a_2 + b_2\mathrm{e}^{-\kappa t})\theta_2$ with $\boldsymbol{a} = (a_1, a_2)'$ and $\boldsymbol{b} = (b_1, b_2)'$ linearly independent.

## 3. Information matrix and optimal designs

In the previous section we stated that the design $\tilde{\boldsymbol{\tau}} \in \mathcal{T}_n$ is "good" if the information matrix $\mathbf{M}(\tilde{\boldsymbol{\tau}})$ is "large". While for one-parametric models the measure of information is straightforward, for multi-parametric models the situation is different.

A usual way of comparing information matrices is using the Loewner ordering. Let $\mathcal{S}_+^m$ be the set of all symmetric positive semidefinite $(m \times m)$-matrices. Then, for $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{S}_+^m$ we say that $\mathbf{M}_1$ "Loewner dominates" $\mathbf{M}_2$, denoted by $\mathbf{M}_1 \succcurlyeq \mathbf{M}_2$, if $\mathbf{M}_1 - \mathbf{M}_2 \in \mathcal{S}_+^m$.

In practice the "size" of an information matrix is measured by optimality criteria, which mostly have some geometrical and statistical interpretation. By an optimality criterion we will call any function $\Phi(\cdot) : \mathcal{S}_+^m \mapsto [0, \infty)$ such that $\Phi$ is Loewner isotonic (i.e., if $\mathbf{M}_1 \succcurlyeq \mathbf{M}_2$, then $\Phi(\mathbf{M}_1) \geq \Phi(\mathbf{M}_2)$), nonconstant, positively homogeneous, concave and upper semicontinuous. Such optimality criteria are called "information functions". We refer the reader to [12] and [14] for more details on optimality criteria.

We will say that the design $\boldsymbol{\tau}_{n,\Phi}^*$ is a $\Phi$-optimal $n$-point design, if $\boldsymbol{\tau}_{n,\Phi}^*$ maximizes $\Phi\big[\mathbf{M}(\boldsymbol{\tau})\big]$ on $\mathcal{T}_n$, that is,

$$\Phi\big[\mathbf{M}(\boldsymbol{\tau}_{n,\Phi}^*)\big] = \sup_{\boldsymbol{\tau} \in \mathcal{T}_n} \Phi\big[\mathbf{M}(\boldsymbol{\tau})\big].$$

The fact that $\boldsymbol{\Sigma}(\boldsymbol{\tau})$ has a product structure and the results of H a r m a n and Š t u l a j t e r [4] imply that for $\boldsymbol{\tau} \in \mathcal{T}_n$

$$\mathbf{M}(\boldsymbol{\tau}) = \begin{pmatrix} \frac{\mathrm{e}^{-2\kappa t_1}}{D(t_1)} & \frac{\mathrm{e}^{-\kappa t_1}(1-\mathrm{e}^{-\kappa t_1})}{D(t_1)} \\ \frac{\mathrm{e}^{-\kappa t_1}(1-\mathrm{e}^{-\kappa t_1})}{D(t_1)} & \frac{(1-\mathrm{e}^{-\kappa t_1})^2}{D(t_1)} + \mathcal{C}(\boldsymbol{\tau}) \end{pmatrix}, \tag{10}$$

where

$$\mathcal{C}(\boldsymbol{\tau}) = \sum_{i=2}^{n} \frac{(\mathrm{e}^{\kappa t_i} - \mathrm{e}^{\kappa t_{i-1}})^2}{\mathrm{e}^{2\kappa t_i} D(t_i) - \mathrm{e}^{2\kappa t_{i-1}} D(t_{i-1})} = \sum_{i=2}^{n} \frac{(\mathrm{e}^{\kappa t_i} - \mathrm{e}^{\kappa t_{i-1}})^2}{\int_{t_{i-1}}^{t_i} \mathrm{e}^{2\kappa\nu} \sigma^2(\nu)\, \mathrm{d}\nu}. \tag{11}$$

We remark that for $\boldsymbol{\tau} \in \mathcal{T}_n$ the information matrix $\mathbf{M}(\boldsymbol{\tau})$ given in (10) is positive definite. Since the optimality criteria are usually continuous in $\mathbf{M}$ on the set of positive definite matrices and $\mathbf{M}(\boldsymbol{\tau})$ is continuous in $\boldsymbol{\tau}$ on $\mathcal{T}_n$, the function $\Phi\big[\mathbf{M}(\boldsymbol{\tau})\big]$ is continuous in $\boldsymbol{\tau}$ on $\mathcal{T}_n$.

Now, assume that $t_1^* \geq T_*$ is fixed. It follows from the Loewner isotonicity of optimality criteria that the design $\boldsymbol{\tau}^* \in \mathcal{T}_n(t_1^*)$ is optimal, if

$$\begin{aligned} \mathcal{C}(\boldsymbol{\tau}^*) \quad = \quad & \max_{\substack{t_2, \ldots, t_n \\ t_1^* < t_2, t_n \leq T^* \\ t_{i-1} < t_i, i = 3, \ldots, n}} \mathcal{C}(t_1^*, t_2, \ldots, t_n). \end{aligned} \tag{12}$$

Therefore, once we have chosen the value of $t_1^*$, the other design points solve the optimization problem (12). This approach is very suitable for a numerical optimization: firstly, we select the value of $t_1$ and then we find the maximum of $\mathcal{C}(\boldsymbol{\tau}^*)$ on $\mathcal{T}_n(t_1)$, which is used for evaluation of the optimality criterion. In this way we can find the maximum of the optimality criterion through $t_1$. Our numerical experience shows that this approach gives rather reliable results compared to a raw maximization of $\Phi\big[\mathbf{M}(\boldsymbol{\tau})\big]$ on the $n$-simplex $\mathcal{T}_n$, even if we employ heuristic methods like simulated annealing or genetic algorithms. For as simple volatility functions as exponential and quadratic the results suggested that it is quite challenging for heuristic methods to search for and find an optimal design through all $n$ design points (that is, $t_1$ is not fixed), although we performed a large number of simulation runs.

The following lemma ensures that an optimal design is not degenerated into one point.

**Lemma 3.** *Let $n \geq 3$ and $\boldsymbol{\tau}_0 = t_n \mathbf{1}_n$. Then there exists $\boldsymbol{\tau}_1 = (t_1, \ldots, t_n)' \in \overline{\mathcal{T}}_n$ such that $t_1 < t_n$ and $\mathbf{M}(\boldsymbol{\tau}_1) \succcurlyeq \mathbf{M}(\boldsymbol{\tau}_0)$.*

Proof. Let $\boldsymbol{\tau}_1 = (t_1, \ldots, t_n)'$ with $t_1 < t_2 = \ldots = t_n$. To prove the statement of the lemma it is sufficient to show that the matrix $\mathbf{M}(\boldsymbol{\tau}_1) - \mathbf{M}(\boldsymbol{\tau}_0)$ equal to

$$\begin{pmatrix} \frac{\mathrm{e}^{-2\kappa t_1}}{D(t_1)} - \frac{\mathrm{e}^{-2\kappa t_n}}{D(t_n)} & \frac{\mathrm{e}^{-\kappa t_1}(1-\mathrm{e}^{-\kappa t_1})}{D(t_1)} - \frac{\mathrm{e}^{-\kappa t_n}(1-\mathrm{e}^{-\kappa t_n})}{D(t_n)} \\ \frac{\mathrm{e}^{-\kappa t_1}(1-\mathrm{e}^{-\kappa t_1})}{D(t_1)} - \frac{\mathrm{e}^{-\kappa t_n}(1-\mathrm{e}^{-\kappa t_1})}{D(t_n)} & \frac{(1-\mathrm{e}^{-\kappa t_1})^2}{D(t_1)} - \frac{(1-\mathrm{e}^{-\kappa t_n})^2}{D(t_n)} + \mathcal{C}(\boldsymbol{\tau}_1) \end{pmatrix}$$

is positive semidefinite. Inequality (9) gives that $D(t_n) > \mathrm{e}^{-2\kappa(t_n - t_1)} D(t_1)$ for $t_1 < t_n$, which implies the positivity of $\{\mathbf{M}(\boldsymbol{\tau}_1) - \mathbf{M}(\boldsymbol{\tau}_0)\}_{11}$. Next,

$$\det\big[\mathbf{M}(\boldsymbol{\tau}_1) - \mathbf{M}(\boldsymbol{\tau}_0)\big] = 0,$$

since

$$\mathcal{C}(\boldsymbol{\tau}_1) = \frac{(\mathrm{e}^{\kappa t_n} - \mathrm{e}^{\kappa t_1})^2}{\mathrm{e}^{2\kappa t_n} D(t_n) - \mathrm{e}^{2\kappa t_1} D(t_1)}.$$

□

**Theorem 4.** *Under the assumption of the model* (3) *with the covariance structure* (7)*, there always exists a (feasible) $\Phi$-optimal $n$-point design $\boldsymbol{\tau}_{n,\Phi}^* \in \mathcal{T}_n$.*

Proof. From Lemma 3 we get that $t_1 \neq t_n$, and the position of the design points $t_2, \ldots, t_n$ results from the optimization problem (12). Hence, we need to show that for any $t_i \in (t_{i-1}, t_{i+1})$, $i = 2, \ldots, n-1$,

$$\frac{(\mathrm{e}^{\kappa t_i} - \mathrm{e}^{\kappa t_{i-1}})^2}{\int_{t_{i-1}}^{t_i} \mathrm{e}^{2\kappa\nu} \sigma^2(\nu)\,\mathrm{d}\nu} + \frac{(\mathrm{e}^{\kappa t_{i+1}} - \mathrm{e}^{\kappa t_i})^2}{\int_{t_i}^{t_{i+1}} \mathrm{e}^{2\kappa\nu} \sigma^2(\nu)\,\mathrm{d}\nu} \geq \frac{(\mathrm{e}^{\kappa t_{i+1}} - \mathrm{e}^{\kappa t_{i-1}})^2}{\int_{t_{i-1}}^{t_{i+1}} \mathrm{e}^{2\kappa\nu} \sigma^2(\nu)\,\mathrm{d}\nu}, \tag{13}$$

that is, we can place a design point $t_i$ between any two design points $t_{i-1}$ and $t_{i+1}$ in a way that $t_i$ is optimal and distinct from $t_{i-1}$ and $t_{i+1}$.

If we set

$$A = \mathrm{e}^{\kappa t_i} - \mathrm{e}^{\kappa t_{i-1}}, \quad B = \mathrm{e}^{\kappa t_{i+1}} - \mathrm{e}^{\kappa t_i},$$

$$a = \int_{t_{i-1}}^{t_i} \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu \quad \text{and} \quad b = \int_{t_i}^{t_{i+1}} \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu,$$

then the inequality (13) is equivalent to the statement

$$\forall_{A,B,a,b>0}\ \frac{A^2}{a} + \frac{B^2}{b} \geq \frac{(A+B)^2}{a+b}.$$

After some algebraic manipulation we obtain that $b^2A^2 - 2bAaB + a^2B^2 \geq 0$, which is true. Moreover, the inequality in (13) becomes an equality if and only if $t_i \in \{t_{i-1}, t_{i+1}\}$. □

In some applications it is natural to estimate the unknown parameters using the most recent observations. The derivative of $\mathcal{C}(\boldsymbol{\tau})$ with respect to $t_n$

$$\frac{\partial\mathcal{C}(\boldsymbol{\tau})}{\partial t_n} = \big[2\kappa - \sigma^2(t_n)\mathcal{U}(t_n, t_{n-1})\big]\mathcal{U}(t_n, t_{n-1}),$$

where

$$\mathcal{U}(x,y) = \frac{\mathrm{e}^{2\kappa x} - \mathrm{e}^{\kappa(x+y)}}{\int_y^x \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu},$$

yields

**Proposition 1.** *If $\sigma(t)$ is a nonincreasing function, then $t_n^* = T^*$ is optimal in the model* (3) *with the covariance structure* (7).

Proof. Clearly, the function $\mathcal{U}(x,y)$ is continuous and positive. For any

$$t_{n-1} < t_n \leq T^*$$

we have

$$\sigma^2(t_n)\mathcal{U}(t_n, t_{n-1}) = \frac{\mathrm{e}^{2\kappa t_n} - \mathrm{e}^{\kappa(t_n+t_{n-1})}}{\int_{t_{n-1}}^{t_n} \mathrm{e}^{2\kappa\nu}\frac{\sigma^2(\nu)}{\sigma^2(t_n)}\,\mathrm{d}\nu} \leq \frac{\mathrm{e}^{2\kappa t_n} - \mathrm{e}^{\kappa(t_n+t_{n-1})}}{\int_{t_{n-1}}^{t_n} \mathrm{e}^{2\kappa\nu}\,\mathrm{d}\nu}$$

$$= 2\kappa\frac{\mathrm{e}^{2\kappa t_n} - \mathrm{e}^{\kappa(t_n+t_{n-1})}}{\mathrm{e}^{2\kappa t_n} - \mathrm{e}^{2\kappa t_{n-1}}} < 2\kappa,$$

That is $\frac{\partial\mathcal{C}(\boldsymbol{\tau})}{\partial t_n} > 0$ for any $\boldsymbol{\tau}$. □

Let us take into consideration that we can perform measurements at every point in the experimental domain

$$[T_*, T^*], \quad \text{and} \quad \boldsymbol{\tau}_n = (t_1, \ldots, t_n)' \quad \text{with} \quad t_1 = T_* \quad \text{and} \quad t_n = T^*.$$

From the Taylor series expansion of (11) and by setting $t_i - t_{i-1} = \Delta$, $i = 2, \dots, n$ we get that

$$\mathcal{C}(\boldsymbol{\tau}_n) = \kappa^2 \sum_{i=2}^{n} \frac{\left(\int_{t_{i-1}}^{t_{i-1}+\Delta} \mathrm{e}^{\kappa\nu}\,\mathrm{d}\nu\right)^2}{\int_{t_{i-1}}^{t_{i-1}+\Delta} \mathrm{e}^{2\kappa\nu}\sigma^2(\nu)\,\mathrm{d}\nu} = \kappa^2 \sum_{i=2}^{n} \frac{\left(\mathrm{e}^{\kappa t_{i-1}}\Delta + \mathrm{o}(\Delta)\right)^2}{\mathrm{e}^{2\kappa t_{i-1}}\sigma^2(t_{i-1})\Delta + \mathrm{o}(\Delta)}$$
$$= \kappa^2 \sum_{i=2}^{n} \frac{\mathrm{e}^{2\kappa t_{i-1}}\Delta^2 + \mathrm{o}(\Delta^2)}{\mathrm{e}^{2\kappa t_{i-1}}\sigma^2(t_{i-1})\Delta + \mathrm{o}(\Delta)} = \kappa^2 \sum_{i=2}^{n} \frac{\Delta + \mathrm{o}(\Delta)}{\sigma^2(t_{i-1}) + \mathrm{o}(\Delta)/\Delta}.$$

Consequently,

$$\mathcal{C}(\boldsymbol{\tau}_n) \to \mathcal{C}_\infty(T_*, T^*) = \kappa^2 \int_{T_*}^{T^*} \frac{\mathrm{d}\nu}{\sigma^2(\nu)}, \qquad \text{for} \quad n \to \infty \quad \text{and} \quad \Delta \to 0,$$

which, using the relation (13) and the technique in the proof of Lemma 3, leads to

**Proposition 2.** *The information matrix given by the observation of the full path in the model* (1) *is*

$$\mathbf{M}_\infty(T_*, T^*) = \begin{pmatrix} \frac{\mathrm{e}^{-2\kappa T_*}}{D(T_*)} & \frac{\mathrm{e}^{-\kappa T_*}(1-\mathrm{e}^{-\kappa T_*})}{D(T_*)} \\ \frac{\mathrm{e}^{-\kappa T_*}(1-\mathrm{e}^{-\kappa T_*})}{D(T_*)} & \frac{(1-\mathrm{e}^{-\kappa T_*})^2}{D(T_*)} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathcal{C}_\infty(T_*, T^*). \quad (14)$$

*Moreover, for any design* $\boldsymbol{\tau} = (t_1, \dots, t_n)' \in \mathcal{T}_n$ *with* $T_* \le t_1$ *and* $t_n \le T^*$ *holds:*

i) $\mathcal{C}(\boldsymbol{\tau}) \le \mathcal{C}_\infty(T_*, T^*)$,

ii) $\mathbf{M}(\boldsymbol{\tau}) \preccurlyeq \mathbf{M}_\infty(t_1, t_2) \preccurlyeq \mathbf{M}_\infty(T_*, T^*)$.

The formula stated in (14) has an intuitive physical interpretation. In the Hookean spring problem with a time-dependent temperature formulated in Introduction, up to a constant multiple the function $\sigma(t)$ reflects the square root of the temperature. Thus the lower the temperature the higher the information about the unknown parameters. In the physical view, low temperature causes small fluctuations, so the measurements are more precise.

Another theoretical contribution of the formula (14) is the information contained in a subdomain interval. We shall explain this in detail. Let us consider that we perform one measurement at $T_*$ and then we can observe the full trajectory of the process on a subdomain interval $(a, a + \Delta]$ of the fixed length $\Delta$, where $a \in [T_*, T^* - \Delta]$. Then it is optimal to perform the measurements on such interval (determined by $a$), which maximizes $\int_a^{a+\Delta} \frac{\mathrm{d}\nu}{\sigma^2(\nu)}$. That is, the areas with low $\sigma(t)$ are more informative, and the measurements should be more concentrated in such areas. This effect is demonstrated on an example in the next section.

In the limit information matrix $\mathbf{M}_\infty(T_*, T^*)$ we can find even more information useful for statistical planning. Being given a $\Phi$-optimal $n$-point design $\boldsymbol{\tau}^*_{n,\Phi}$

we can measure the $\Phi$-efficiency (see, e.g., [14]) of the design $\boldsymbol{\tau}^*_{n,\Phi}$ with respect to the maximum possible information defined by

$$\mathrm{eff}_\Phi(\boldsymbol{\tau}^*_{n,\Phi}) = \frac{\Phi[\mathbf{M}(\boldsymbol{\tau}^*_{n,\Phi})]}{\Phi[\mathbf{M}_\infty(T_*, T^*)]}, \tag{15}$$

which enables us to investigate how many observations $n$ are necessary to reach a sufficient efficiency, and wether an additional observation brings a significant increase in information. For instance, in the next section we will discuss an example, for which we show that the mentioned efficiency is relatively high even for designs with small number of observations. A similar idea has been introduced by H a r m a n and Š t u l a j t e r [4], and highlights the importance of exact optimal designs in contrast to asymptotic designs, which were studied in the pioneering papers of S a c k s and Y l v i s a k e r [16]–[18].

## 4. Example

To give a simple demonstration of the previously presented results we will focus on $D$-optimal designs for

$$\mathrm{d}X_t = \kappa(\bar{X} - X_t)\,\mathrm{d}t + \mathrm{e}^{-\lambda t}\mathrm{d}W_t, \tag{16}$$

where $\kappa > 0$ is the mean-reversion speed and $\lambda \in \mathbb{R}$ is a known constant. If $\lambda$ is positive, then we have a system with exponentially decreasing temperature, and in the case of a negative value of $\lambda$ the system is being heated. There are two reasons for use of an exponential function for the volatility in this example: firstly, an exponential function fits the evolution of the temperature of an object put into an environment with different but constant temperature and, secondly, we obtain simple closed-form formulas for information and efficiency.

Using the relations (6) and (10) we obtain that

$$\mathbf{M}_\lambda(\boldsymbol{\tau}) = \begin{pmatrix} \frac{2(\kappa-\lambda)}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1} & \frac{2(\kappa-\lambda)(\mathrm{e}^{\kappa t_1}-1)}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1} \\ \frac{2(\kappa-\lambda)(\mathrm{e}^{\kappa t_1}-1)}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1} & \frac{2(\kappa-\lambda)(\mathrm{e}^{\kappa t_1}-1)^2}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1} + \mathcal{C}_\lambda(\boldsymbol{\tau}) \end{pmatrix},$$

where

$$\mathcal{C}_\lambda(\boldsymbol{\tau}) = 2(\kappa-\lambda)\sum_{i=2}^{n} \frac{(\mathrm{e}^{\kappa t_i} - \mathrm{e}^{\kappa t_{i-1}})^2}{\mathrm{e}^{2(\kappa-\lambda)t_i} - \mathrm{e}^{2(\kappa-\lambda)t_{i-1}}}.$$

Obviously, for $\lambda \to \kappa$ we have

$$\mathbf{M}_\kappa(\boldsymbol{\tau}) = \begin{pmatrix} \frac{1}{t_1} & \frac{\mathrm{e}^{\kappa t_1}-1}{t_1} \\ \frac{\mathrm{e}^{\kappa t_1}-1}{t_1} & \frac{(\mathrm{e}^{\kappa t_1}-1)^2}{t_1} + \mathcal{C}_\kappa(\boldsymbol{\tau}) \end{pmatrix}, \quad \text{with} \quad \mathcal{C}_\kappa(\boldsymbol{\tau}) = \sum_{i=2}^{n} \frac{(\mathrm{e}^{\kappa t_i} - \mathrm{e}^{\kappa t_{i-1}})^2}{t_i - t_{i-1}}.$$

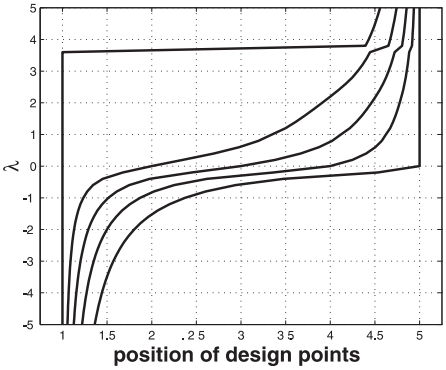


FIGURE 1. $D$-optimal 5-point designs for the model (16) with $\kappa$ =2, $T_* = 1$, $T^* = 5$ and different values of $\lambda$. For a given $\lambda$ the horizontal cut gives the optimal position of design points.

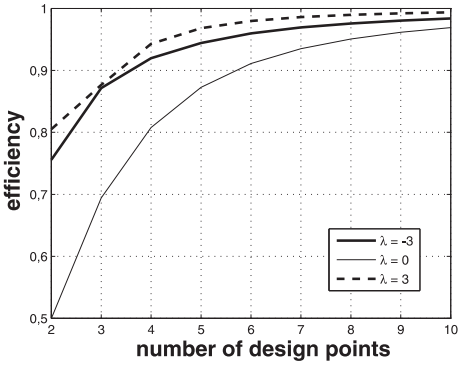


FIGURE 2. The relation between the size of a design sample $n$ and efficiency of $D$-optimal $n$-point design with respect to the maximum possible information for the model (16) with $\kappa = 2$, $T_*$=1, $T^* = 5$ and $\lambda = -3, 0, 3$.

Taking the $D$-optimality criterion $\Phi_D[\mathbf{M}] = \det^{\mathbf{1/2}}(\mathbf{M})$ we get

$$\det{}^{1/2}\big(\mathbf{M}(\boldsymbol{\tau})\big) = \begin{cases} \left[\frac{2(\kappa-\lambda)}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1}\mathcal{C}_\lambda(\boldsymbol{\tau})\right]^{1/2}, & \lambda \neq \kappa, \\ \left[\frac{1}{t_1}\mathcal{C}_\kappa(\boldsymbol{\tau})\right]^{1/2}, & \lambda = \kappa. \end{cases}$$

We remark that for $\kappa = 2$ and $\lambda = 1$ the function $\mathcal{C}_\lambda(\boldsymbol{\tau})$ depends only on $t_1$ and $t_n$.

The $\Phi_D$-efficiency of a design with respect to the maximal possible information (cf. (15)) is given by

$$\mathrm{eff}_{\Phi_D}(\boldsymbol{\tau}) = \begin{cases} \left[\frac{2\lambda}{\kappa^2}\cdot\frac{\mathrm{e}^{2(\kappa-\lambda)T_*}-1}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1}\cdot\frac{\mathcal{C}_\lambda(\boldsymbol{\tau})}{\mathrm{e}^{2\lambda T^*}-\mathrm{e}^{2\lambda T_*}}\right]^{1/2}, & \kappa \neq \lambda,\ \lambda \neq 0, \\ \left[\frac{1}{\kappa^2(T^*-T_*)}\cdot\frac{\mathrm{e}^{2(\kappa-\lambda)T_*}-1}{\mathrm{e}^{2(\kappa-\lambda)t_1}-1}\mathcal{C}_\lambda(\boldsymbol{\tau})\right]^{1/2}, & \kappa \neq \lambda,\ \lambda = 0, \\ \left[\frac{2\lambda}{\kappa^2}\cdot\frac{T_*}{t_1}\cdot\frac{\mathcal{C}_\kappa(\boldsymbol{\tau})}{\mathrm{e}^{2\lambda T^*}-\mathrm{e}^{2\lambda T_*}}\right]^{1/2}, & \kappa = \lambda. \end{cases}$$

In the sequel, we give some numerical results for the model (16) with the mean-reversion speed $\kappa = 2$ and bounds for the experimental domain $T_* = 1$ and $T^* = 5$.

Figure 4 depicts $D$-optimal 5-point designs. More precisely, for a particular $\lambda$ a horizontal cut gives optimal positions of the design points. For $\lambda = 0$ the design is equidistant, which is a known result already shown by H a r m a n and Š t u l a j t e r [5]. For $\lambda > 0$, in accord with Proposition 1, the position of the

last design point is equal to $T^* = 5$. The explanation for the "jump" in optimal position of $t_1$ around $\lambda = 3.8$ can be found in flatness of the function $\varphi(t) = \max_{\boldsymbol{\tau} \in \mathcal{T}_n(t)} \det^{1/2}\big(\mathbf{M}(\boldsymbol{\tau})\big)$, so the numerical optimization through $t_1$ is less accurate and more sensitive to roundoff errors. In both cases, $\lambda > 0$ and $\lambda < 0$, we can notice that if $|\lambda|$ is large, then the design points are more concentrated around $T^*$ and $T_*$, respectively, where the fluctuations are smaller.

In the previous section we noted that small sample designs can be quite efficient with respect to the maximum possible information, and the contributions of additional measurements to the information are not significant. This is illustrated in Figure 4, which displays the dependence of the efficiency of the $D$-optimal $n$-point designs on the size $n$ of the design sample for the model (16).

**Acknowledgements.** The author would like to thank Assoc. Prof. R a d o s l a v H a r m a n and Prof. D a n i e l Š e v č o v i č as well as the anonymous referee for their helpful comments.

REFERENCES

[1] ALLEN, E.: *Modeling with Itō Stochastic Differential Equations*. Springer, Dordrecht, 2007.
[2] ARNOLD, L.: *Stochastic Differential Equations: Theory and Applications*. John Wiley & Sons, New York, 1974.
[3] HARMAN, R.—ŠTULAJTER, F.: *Optimality of equidistant sampling designs for a non-stationary Ornstein-Uhlenbeck process*, in: Proc. of the 6th St. Petersburg Workshop on Simulation, Vol. 2 (S. M. Ermakov et al. eds.), St. Petersburg State University, St. Petersburg, 2009, pp. 1097–1101.
[4] HARMAN, R.—ŠTULAJTER, F.: *Optimal sampling designs for the Brownian motion with a quadratic drift*, J. Statist. Plann. Inference **141** (2011), 2750-2758.
[5] HARMAN, R.—ŠTULAJTER, F.: *Optimal sampling designs for a two-parametric Ornstein-Uhlenbeck process* (submitted).
[6] ITŌ, K.: *On a formula concerning stochastic differentials*, Nagoya Math. J. **3** (1951), 55–65.
[7] KISEĽÁK, J.—STEHLÍK, M.: *Equidistant and D-optimal designs for parameters of Ornstein-Uhlenbeck process*, Statist. Probab. Lett. **78** (2008), 1388–1396.
[8] KUNZE, M.—LORENZI, L.—LUNARDI, A.: *Nonautonomous Kolmogorov parabolic equations with unbounded coefficients*, Trans. Amer. Math. Soc. **362** (2010), 169–198.
[9] LEMONS, D. S.: *An Introduction to Stochastic Processes in Physics*. Johns Hopkins University Press, Baltimore, 2002.
[10] MUKHERJEE, B.: *Exactly optimal sampling designs for processes with a product covariance structure*, Canad. J. Statist. **31** (2003), 69–87.
[11] ØKSENDAL, B.: *Stochastic Differential Equations: An Introduction with Applications* (6th ed.), Springer, Berlin, 2003.
[12] PÁZMAN, A.: *Foundations of Optimum Experimental Design*. Riedel, Dordrecht, 1986.
[13] PÁZMAN, A.: *Nonlinear Statistical Models*. Kluwer, Dordrecht, 1993.
[14] PUKELSHEIM, F.: *Optimal Design of Experiments*. John Wiley & Sons, New York, 1993.

[15] RICCARDI, L. M.—SACERDOTE, L.: *The Ornstein-Uhlenbeck process as a model for neuronal activity*. Biol. Cybern. **35** (1979), 1–9.
[16] SACKS, J.—YLVISAKER, D.: *Designs for regression problems with correlated errors*. Ann. Math. Stat. **37** (1966), 66–89.
[17] SACKS, J.—YLVISAKER, D.: *Designs for regression problems with correlated errors: many parameters*, Ann. Math. Stat. **39** (1968), 49–69.
[18] SACKS, J.—YLVISAKER, D.: *Designs for regression problems with correlated errors III*, Ann. Math. Stat. **41** (1970), 2057–2074.
[19] SCHUSS, Z.: *Theory and Applications of Stochastic Processes: An Analytical Approach*. Springer, New York, 2010.
[20] ZAGORAIOU, M.—ANTOGNINI, A. B.: *Optimal designs for parameter estimation of the Ornstein-Uhlenbeck process*, Appl. Stoch. Models Bus. Ind. **25** (2009), 583-600.

Received October 31, 2011

*Department of Applied Mathematics and Statistics*
*Faculty of Mathematics, Physics and Informatics*
*Comenius University in Bratislava*
*Mlynská dolina*
*SK–842-48 Bratislava*
*SLOVAKIA*
*E-mail*: lackovladimir@gmail.com