# Optimal designs for regression models with autoregressive errors structure

Holger Dette
Fakultät für Mathematik
Ruhr-Universität Bochum
44780 Bochum
Germany

Andrey Pepelyshev, Anatoly Zhigljavsky
School of Mathematics
Cardiff University
Cardiff, CF24 4AG
UK

**Abstract**

In the one-parameter regression model with AR(1) and AR(2) errors we find explicit expressions and a continuous approximation of the optimal discrete design for the signed least square estimator. The results are used to derive the optimal variance of the best linear estimator in the continuous time model and to construct efficient estimators and corresponding optimal designs for finite samples. The resulting procedure (estimator and design) provides nearly the same efficiency as the weighted least squares and its variance is close to the optimal variance in the continuous time model. The results are illustrated by several examples demonstrating the feasibility of our approach.

Keywords and Phrases: linear regression; correlated observations; signed measures; optimal design; BLUE; AR processes; continuous autoregressive model
AMS Subject classification: Primary 62K05; Secondary 31A10

## 1 Introduction

Consider a linear regression model

$$y_j = \theta^T f(t_j) + \epsilon_j \quad (j = 1, \ldots, N), \tag{1.1}$$

where $\theta \in \mathbb{R}^m$ is a vector of unknown parameters, $f(t) = (f_1(t), \ldots, f_m(t))^T$ is a vector of linearly independent functions defined on some interval, say $[A, B]$, and $\epsilon_1, \ldots, \epsilon_N$ are random errors with $\mathbb{E}[\epsilon_j] = 0$ for all $j = 1, \ldots, N$ and covariances $\mathbb{E}[\epsilon_j \epsilon_k] = \rho(t_j - t_k)$. It is well known that the use of optimal or efficient designs yields to a reduction of costs by a statistical inference with a minimal number of experiments without loosing any accuracy. Optimal design theory has been studied intensively for the case when errors are uncorrelated using tools from convex optimization theory [see (Pukelsheim, 2006)], but the design problem in the case of dependent data is substantially harder because the corresponding optimization problems are usually non-convex. Most authors use asymptotic arguments to construct optimal designs, which do not solve the problem of non-convexity [see for example Sacks and Ylvisaker (1966, 1968), Bickel and Herzberg (1979), Näther (1985a), Zhigljavsky et al. (2010), Dette et al. (2015)]. Some optimal designs for the location model (in this case the optimization problems are in fact convex) and for a few one-parameter linear models have been discussed in Boltze and Näther (1982), Näther (1985a), Ch. 4, Näther (1985b), Pázman and Müller (2001) and Müller and Pázman (2003) among others]. Recently, for multi-parameter models, Dette et al. (2013) determined a necessary condition for the optimality of (asymptotic) designs for least squares estimation. Dette et al. (2014) studied nearly universally optimal designs, while Dette et al. (2016) constructed new matrix-weighted estimators with corresponding optimal designs, which are very close to the best linear unbiased estimator with corresponding optimal designs. Although these results are promising, they rely on certain structural assumptions on the covariance kernel. For example, Dette et al. (2013) assume that the regression functions in model (1.1) are eigenfunctions of an integral operator associated with the covariance kernel of the error process and Dette et al. (2016) assume that the covariance kernel is triangular [see Mehr and McFadden (1965) for an exact definition]. While these results cover the frequently used AR(1)-process as error structure, they are not applicable in models with autoregressive error processes of larger order.

The goal of the present paper is to give first insights in the optimal design problem for linear regression models with autoregressive error processes. We concentrate on a one-parameter linear regression model with an AR(1) and AR(2)-error process. In Section 2 we will introduce a signed least squares estimator and consider approximate designs on the design space $\mathcal{T} = \{t_1, \dots, t_N\}$, where the weights are not necessarily non-negative. We determine the optimal (signed) approximate design for signed least squares estimation, such that the signed least squares estimator has the same variance as the weighted least squares estimator based on observations at the experimental conditions $t_1, \dots, t_N$. In Section 3 we consider the one-parameter linear regression model with autoregressive errors of order 1 and study the asymptotic behavior of the signed least squares estimator with corresponding optimal design as the sample size tends to infinity. Section 4 is devoted to the case of an AR(2)-error process, where the situation is substantially more complicated. These results are then used in Section 5, where we consider the problem of constructing designs for signed least squares estimation in finite sample situations. We provide a procedure such that the signed least squares estimator with corresponding "optimal" design has nearly the same efficiency as the weighted least squares estimator with corresponding optimal design. Finally, the results are illustrated in several numerical examples.

## 2 Various least squares estimators

For estimating $\theta$, we use the following two estimators: the best linear unbiased estimator (BLUE)

$$\hat{\theta}_{\text{BLUE},N} = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}Y$$

and the signed least squares estimator (SLSE)

$$\widehat{\theta}_{\text{SLSE},N} = (X^T\mathbf{S}X)^{-1}X^T\mathbf{S}Y, \tag{2.1}$$

where $X = (f_i(x_j))_{j,i=1}^{N,m}$ is the design matrix of size $N \times m$, $\mathbf{S}$ is an $N \times N$ diagonal matrix with entries $+1$ and $-1$ on the diagonal and $\Sigma = (\rho(t_i - t_j))_{i,j=1}^{N}$ is the covariance matrix of observations. If $\mathbf{S}$ is the $N \times N$ identity matrix, then SLSE coincides with the ordinary least squares estimator (LSE). The covariance matrix of the BLUE and the SLSE are given by

$$\begin{aligned}\text{Var}(\hat{\theta}_{\text{BLUE},N}) &= (X^T\Sigma^{-1}X)^{-1} ,\\ \text{Var}(\hat{\theta}_{\text{SLSE},N}) &= (X^T\mathbf{S}X)^{-1}(X^T\mathbf{S}\Sigma\mathbf{S}X)(X^T\mathbf{S}X)^{-1},\end{aligned}$$

respectively. Throughout this paper we concentrate on the one-parameter regression model

$$y_j = \theta f(t_j) + \epsilon_j, \tag{2.2}$$

and remark that an extension to the multi-parameter model (1.1) could be performed following the discussion in Dette et al. (2016). A design on the (fixed) design space $\mathcal{T} = \{t_1, \ldots, t_N\}$ is an arbitrary discrete signed measure of the form $\xi = \{t_1, \ldots, t_N; w_1, \ldots, w_N\}$, where $w_i = s_i p_i$, $s_i \in \{-1, 1\}$, $p_i \geq 0$, $i = 1, \ldots, N$, and $\sum_{i=1}^N p_i = 1$. The variance of the SLSE for the design $\xi$ is given by

$$D(\xi) = \text{Var}(\hat{\theta}_{\text{SLSE},N}) = \sum_{i=1}^{N}\sum_{j=1}^{N} \rho(t_i - t_j) w_i w_j f_i f_j \Big/ \Big(\sum_{i=1}^{N} w_i f_i^2\Big)^2 , \tag{2.3}$$

where we use the notation $f_i = f(t_i)$ throughout this paper. The optimal design problem consists in the minimization of this expression with respect to the weights $w_1, \ldots, w_N$ assuming that the observation points $t_1, \ldots, t_N$ are fixed. Despite the fact that the functional $D$ in (2.3) is not convex as a function of $w_1, \ldots, w_N$, the problem of determining the optimal weights can be easily solved by a simple application of the Cauchy-Schwarz inequality. The proof of the following lemma is given in Dette et al. (2016); see also Theorem 5.3 in Näther (1985a), where this result was proved in a slightly different form.

**Lemma 2.1** *Assume that the matrix $\Sigma = (\rho(t_i - t_j))_{i,j=1,\ldots,N}$ is positive definite and $f_i \neq 0$ for all $i = 1, \ldots, N$. Then the optimal weights $w_1^*, \ldots, w_N^*$ minimizing the expression* (2.3) *are given by*

$$w_i^* = \mathbf{e}_i^T \Sigma^{-1} \mathbf{f} / f_i; \qquad i = 1, \ldots, N, \tag{2.4}$$

*where* $\mathbf{f} = (f_1, \ldots, f_N)^T$*,* $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T \in \mathbb{R}^N$ *is the i-th unit vector. Moreover, for the design* $\xi^* = \{t_1, \ldots, t_N; w_1^*, \ldots, w_N^*\}$ *with weights* (2.4) *we have* $D(\xi^*) = D^*$*, where* $D^* = 1/(\mathbf{f}^T \Sigma^{-1} \mathbf{f})$ *is the variance of the BLUE defined in* (2.2)*.*

Note that the optimal weights in Lemma 2.1 are not uniquely defined. In fact, they can always be multiplied by a constant without changing their optimality. In the following discussion we will consider the case where the points $t_i$ are given by the equidistant points on the interval $[A, B]$ and the sample size $N$ tends to infinity. Heuristically the BLUE converges in this case to the BLUE in the continuous time model, where the full trajectory of the stochastic process can be observed. Note that for any finite $N$ the SLSE with the optimal weights defined in Lemma 2.1 has the same variance as the BLUE.

Further we study the asymptotic properties of the SLSE and the optimal weights $w_i^*$ defined in (2.4) as the sample size increases. In many cases we will be able to approximate an $N$-point design $\xi = \{t_1, \ldots, t_N; w_1^*, \ldots, w_N^*\}$ with optimal weights defined in (2.4) by a signed measure (an approximate design) of the form

$$\xi(dt) = P_A \delta_A(dt) + P_B \delta_B(dt) + p(t)dt\,, \tag{2.5}$$

where $\delta_A(dt)$ and $\delta_B(dt)$ are Dirac-measures concentrated at the point $A$ and $B$, respectively, and $p(\cdot)$ is a density function (not necessarily non-negative) on the interval $[A, B]$. Approximate designs of the from (2.5) are easier to understand and analyze than discrete designs of the form $\xi = \{t_1, \ldots, t_N; w_1^*, \ldots, w_N^*\}$, and we will illustrate in Section 3 and 4 the derivation of the limits in the case of autoregressive error processes of order one and two, respectively.

As already mentioned in the introduction the AR(1) process corresponds to a triangular kernel and could also be treated with methodology developed in Dette et al. (2016). We discuss it here because for this case the arguments are simpler than for the AR(2). In fact, for the AR(2) error process the derivation of asymptotically optimal weights $w_1^*, \ldots, w_N^*$ of the form (2.4) as the sample size tends to infinity is substantially harder and we have to slightly modify the continuous approximation of the form (2.5) [see Section 4 for more details].

# 3 Autoregressive errors of order one

Consider the regression model (1.1) with $N$ equidistant points

$$t_j = A + (j-1)\Delta\ ,\ (j = 1, \ldots, N) \tag{3.1}$$

on the interval $[A, B]$, where $\Delta = (B - A)/(N - 1)$. Assume that the errors $\epsilon_1, \ldots, \epsilon_N$ in (2.2) satisfy the discrete AR(1) equation

$$\epsilon_j - a\epsilon_{j-1} = z_j \tag{3.2}$$

for some $0 < a < 1$, where $\epsilon_1 \sim N(0, \sigma^2)$ and $z_2, \ldots, z_N$ are Gaussian independent identically distributed random variables with mean 0 and variance $\sigma_z^2 = (1 - a^2)\sigma^2$. Without loss of generality, we assume $\sigma^2 = 1$.

**Remark 3.1** Note that discrete AR(1) processes (3.2) are usually considered for the parameter $-1 < a < 1$. For the subsequent discussion we need a continuous analogue, say $\{\varepsilon(t)\}_{t\in[A,B]}$, of the discrete AR(1) error process, which is in fact available in the case $0 < a < 1$; see Chan and Tong (1987). The corresponding process with drift is denoted by $y(t) = \theta f(t) + \varepsilon(t)$, $t \in [A, B]$. However, for $-1 < a < 0$ the discrete AR(1) process (3.2) does not have a continuous real-valued analogue and therefore in this case the limiting behavior of our estimators and designs is much harder to understand.
It is also worthwhile to mention that the autocovariance function of errors $\epsilon_1, \ldots, \epsilon_N$ is given by

$$\mathbb{E}[\epsilon_j \epsilon_k] = \rho(t_j - t_k) = e^{-\lambda|t_j - t_k|} = e^{\lambda t_j} e^{-\lambda t_k} \quad \text{if } t_j \le t_k,$$

where $\lambda = -\ln(a)/\Delta$. Thus, if $a \in (0, 1)$, the AR(1) error process has a triangular covariance kernel in the sense of Mehr and McFadden (1965), and the results of Dette et al. (2016) are applicable. In the following discussion we provide a different derivation of the asymptotically optimal weights, because the arguments will be useful for the discussion of an AR(2)-error process in Section 4.

For an AR(1)-error process, the inverse of the covariance matrix $\Sigma = (\rho(t_i - t_j))_{i,j=1}^N$ is given by the tridiagonal matrix

$$\Sigma^{-1} = \frac{1}{S} \begin{pmatrix} 1 & k_1 & 0 & 0 & \ldots & & \\ k_1 & k_0 & k_1 & 0 & \ldots & & \\ 0 & k_1 & k_0 & k_1 & 0 & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & 0 & k_1 & k_0 & k_1 & \\ & & 0 & 0 & k_1 & 1 & \end{pmatrix},$$

where $k_0 = 1 + a^2 = 1 + e^{-2\lambda\Delta}$, $k_1 = -a = -e^{-\lambda\Delta}$, $S = 1 - a^2 = 1 - e^{-2\lambda\Delta}$ and $\lambda = -\ln(a)/\Delta$. Recalling the definition of the optimal weights $w_i^*$, $i = 2, \ldots, N-1$, in (2.4) we have

$$\begin{aligned} S w_i^* f(t_i) &= k_1 f_{i-1} + k_0 f_i + k_1 f_{i+1} = (1 + a^2) f_i - a f_{i-1} - a f_{i+1} \\ &= a(2f_i - f_{i-1} - f_{i+1}) + (1 - 2a + a^2) f_i \\ &= a(2f_i - f_{i-1} - f_{i+1}) + (a - 1)^2 f_i\,. \end{aligned}$$

We now assume that $\lambda = -\ln(a)/\Delta$ is fixed and $\Delta = (B - A)/(N - 1) \to 0$. Since $S(\Delta) =$

$S'(0)\Delta + o(\Delta)$ with $S'(0) = 2\lambda$ and $a = 1 - \lambda\Delta + o(\Delta)$, we obtain

$$\begin{aligned} w_i^* f(t_i) &= \frac{\Delta}{S(\Delta)} \cdot \frac{a(2f_i - f_{i-1} - f_{i+1}) + (a-1)^2 f_i}{\Delta^2}\Delta \\ &= \frac{1}{S'(0)}[-f''(t_i) + \lambda^2 f(t_i)]\Delta + o(\Delta). \end{aligned}$$

Thus, we have

$$\frac{w_i^*}{\Delta} = \frac{1}{2\lambda f(t_i)}[-f''(t_i) + \lambda^2 f(t_i)] + O(\Delta).$$

Therefore, for small $\Delta$, the discrete signed measure $\{t_2, \ldots, t_{N-1}; w_2^*, \ldots, w_{N-1}^*\}$ is approximated by the continuous signed measure with density

$$p(t) = -\frac{1}{2\lambda f(t)}\Big(f''(t) - \lambda^2 f(t)\Big). \tag{3.3}$$

Now we consider the weights at the boundary points. For the left boundary weight, we obtain

$$\begin{aligned} w_1^* f(t_1) &= \frac{f_1 + k_1 f_2}{S(\Delta)} = \frac{\Delta}{S(\Delta)} \cdot \frac{f_1 - a f_2}{\Delta} \\ &= \frac{\Delta}{S(\Delta)}\Big[\frac{f_1 - f_2}{\Delta} + \frac{f_2 - a f_2}{\Delta}\Big] \\ &= \frac{1}{S'(0)}\,[-f'(t_1) - a'(0) f(t_1)] + O(\Delta). \end{aligned}$$

Since $t_1 = A$, for small $\Delta$, we have $w_1^* \approx P_A$, where

$$P_A = \frac{1}{f(A)S'(0)}\Big(-f'(A) - a'(0)f(A)\Big) = \frac{1}{2\lambda f(A)}\Big(-f'(A) + \lambda f(A)\Big). \tag{3.4}$$

Similarly, for the right boundary weight, we obtain

$$\begin{aligned} w_N^* f(t_N) &= \frac{f_N + k_1 f_{N-1}}{S(\Delta)} = \frac{\Delta}{S(\Delta)} \frac{f_N - a f_{N-1}}{\Delta} \\ &= \frac{\Delta}{S(\Delta)}\Big[\frac{f_N - f_{N-1}}{\Delta} + \frac{f_{N-1} - a f_{N-1}}{\Delta}\Big] \\ &= \frac{1}{S'(0)}\,[f'(t_N) - a'(0) f(t_{N-1})] + O(\Delta). \end{aligned}$$

Since $t_N = B$, for small $\Delta$, we have $w_N^* \approx P_B$, where

$$P_B = \frac{1}{f(B)S'(0)}\big(f'(B) - a'(0)f(B)\big) = \frac{1}{2\lambda f(B)}\big(f'(B) + \lambda f(B)\big). \tag{3.5}$$

Summarizing, we have proved the following result.

**Proposition 3.1** *Consider the one-parameter regression model* (2.2) *with* AR(1) *errors of the form* (3.2)*, where* $0 < a < 1$ *and* $f(\cdot)$ *is a twice continuously differentiable function such that* $f(t) \neq 0$ *for all* $t \in [A, B]$*. For large* $N$*, the optimal discrete SLSE (defined in Lemma 2.1) is approximated by the continuous SLSE*

$$\hat{\theta} = D^* \Big( P_A f(A) y(A) + P_B f(B) y(B) + \int_A^B p(t) f(t) y(t) dt \Big) \tag{3.6}$$

*where*

$$D^* = \Big( P_A f^2(A) + P_B f^2(B) + \int_A^B p(t) f^2(t) dt \Big)^{-1},$$

*and* $p(t)$*,* $P_A$ *and* $P_B$ *are defined in* (3.3), (3.4) *and* (3.5)*, respectively. For this approximation, we have*

$$D^* = \lim_{N\to\infty} \mathrm{Var}(\hat{\theta}_{\mathrm{SLSE},N}),$$

*i.e.* $D^*$ *is the limit of the variance* (2.3) *of the optimal discrete SLSE design as* $N \to \infty$*.*

Throughout the following discussion we call a triple $(p, P_A, P_B)$ containing a (signed) density $p$ and two weights $P_A$ and $P_B$, an approximate design for the *continuous* SLSE estimator defined in (3.6).

**Remark 3.2** Observing the discussion in the second part of Remark 3.1 it is reasonable to compare Proposition 3.1 with Theorem 2.1 in Dette et al. (2016). Note that the expressions for the optimal signed density $p(\cdot)$ and optimal weights $P_A$ and $P_B$ at boundary points are particular cases of the general formulae

$$p(t) = -\frac{1}{f(t)v(t)} \Big[ \frac{h'(t)}{q'(t)} \Big]',$$

$$P_A = \frac{1}{f(A)v^2(A)q'(A)} \Big[ \frac{f(A)u'(A)}{u(A)} - f'(A) \Big], \quad P_B = \frac{h'(B)}{f(B)v(B)q'(B)}$$

with $u(t) = e^{\lambda t}$ and $v(s) = e^{-\lambda s}$, where $q(t) = u(t)/v(t)$ and $h(t) = f(t)/v(t)$. Indeed, we easily see that $h(t) = f(t)e^{\lambda t}$, $h'(t) = f'(t)e^{\lambda t} + f(t)\lambda e^{\lambda t}$, $q'(t) = 2\lambda e^{2\lambda t}$, $h'(t)/q'(t) = f'(t)e^{-\lambda t} + f(t)\lambda e^{-\lambda t}$ and, consequently,

$$\begin{aligned} p(t) &= -\frac{1}{f(t)e^{-\lambda t}} (f'(t)e^{-\lambda t} + f(t)\lambda e^{-\lambda t})' \\ &= -\frac{1}{f(t)e^{-\lambda t}} (f''(t)e^{-\lambda t} - \lambda f''(t)e^{-\lambda t} + f'(t)\lambda e^{-\lambda t} - f(t)\lambda^2 e^{-\lambda t}) \\ &= -\frac{1}{2\lambda f(t)} \Big( f''(t) - \lambda^2 f(t) \Big). \end{aligned}$$

as desired. Similarly, we have

$$\begin{aligned} P_A &= \frac{1}{f(A)e^{-2\lambda A}2\lambda e^{2\lambda A}}\Big[\frac{f(A)\lambda e^{\lambda A}}{e^{\lambda A}} - f'(A)\Big] = \frac{1}{2\lambda f(A)}\Big( - f'(A) + \lambda f(A)\Big) \\ P_B &= \frac{f'(B)e^{\lambda B} + f(B)\lambda e^{\lambda B}}{f(B)e^{-\lambda B}2\lambda e^{2\lambda B}} = \frac{1}{2\lambda f(B)}\Big(f'(B) + \lambda f(B)\Big) . \end{aligned}$$

# 4 Autoregressive errors of order two

In this section we assume that the observations in model (2.2) are taken at $N$ equidistant points of the form (3.1) and that the errors $\epsilon_1, \ldots, \epsilon_N$ satisfy the discrete AR(2) equation

$$\epsilon_j - a_1\epsilon_{j-1} - a_2\epsilon_{j-2} = z_j, \tag{4.1}$$

where $z_j$ are Gaussian independent identically distributed random variables with mean 0 and variance $\sigma_z^2 = \sigma^2(1+a_2)((1-a^2)-a_1^2)/(1-a_2)$. Here we make a usual assumption that (4.1) defines the AR(2) process for $j \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ but we only take the values such that $j \in \{1, 2, \ldots, N\}$. Let $r_k = \mathbb{E}[\epsilon_j \epsilon_{j+k}]$ be the autocovariance function of the AR(2) process $\{\epsilon_1, \ldots, \epsilon_N\}$ and assume without loss of generality that $\sigma^2 = 1$. It is well known that the inverse of the covariance matrix $\Sigma = (\mathbb{E}[\epsilon_j \epsilon_j])_{j,k}$ of the discrete AR(2) process is a five-diagonal matrix, i.e.

$$\Sigma^{-1} = \frac{1}{S}\begin{pmatrix} k_{11} & k_{12} & k_2 & 0 & 0 & 0 & \ldots & & \\ k_{21} & k_{22} & k_1 & k_2 & 0 & 0 & \ldots & & \\ k_2 & k_1 & k_0 & k_1 & k_2 & 0 & \ldots & & \\ 0 & k_2 & k_1 & k_0 & k_1 & k_2 & & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & 0 & k_2 & k_1 & k_0 & k_1 & k_2 & \\ & & 0 & 0 & k_2 & k_1 & k_{22} & k_{12} & \\ & & 0 & 0 & 0 & k_2 & k_{21} & k_{11} & \end{pmatrix}, \tag{4.2}$$

where the non-vanishing elements are given by $k_0 = 1 + a_1^2 + a_2^2$, $k_1 = -a_1 + a_1a_2$, $k_2 = -a_2$, $k_{11} = 1$, $k_{12} = k_{21} = -a_1$, $k_{22} = 1 + a_1^2$ and $S = (1+a_1-a_2)(1-a_1-a_2)(1+a_2)/(1-a_2)$. Using Lemma 2.1 and the explicit form (4.2) for $\Sigma^{-1}$ we immediately obtain the following result.

**Corollary 4.1** *Consider the linear regression model* (2.2) *with observations at $N$ equidistant points* (3.1) *and errors that follow the discrete AR*(2) *model* (4.1). *If $f_i = f(t_i) \neq 0$ for*

*$i = 1, \ldots, N$, then the optimal weights in (2.4) can be represented explicitly as follows:*

$$
\begin{aligned}
w_1^* &= \frac{1}{Sf_1}\left(k_{11}f_1 + k_{12}f_2 + k_2 f_3\right), \\
w_2^* &= \frac{1}{Sf_2}\left(k_{21}f_1 + k_{22}f_2 + k_1 f_3 + k_2 f_4\right), \\
w_N^* &= \frac{1}{Sf_N}\left(k_{11}f_N + k_{21}f_{N-1} + k_2 f_{N-2}\right), \\
w_{N-1}^* &= \frac{1}{Sf_{N-1}}\left(k_{12}f_N + k_{22}f_{N-1} + k_1 f_{N-2} + k_4 f_{N-3}\right), \\
w_i^* &= \frac{1}{Sf_i}\left(k_2 f_{i-2} + k_1 f_{i-1} + k_0 f_i + k_1 f_{i+1} + k_2 f_{i+2}\right)
\end{aligned}
$$

*for $i = 3, \ldots, N-2$.*

For the approximation of $w_i^*$, we have to study the behavior of the coefficients which depend on the autocovariance function $r_k$ of the AR(2) process (4.1). There are different types of autocovariance functions which will be introduced and discussed in the remaining part of this section.

Formally, a continuous AR(2) process is a solution of the linear stochastic differential equation of the form

$$d\varepsilon'(t) = \tilde{a}_1\varepsilon'(t) + \tilde{a}_2\varepsilon(t) + \sigma_0^2 dW(t),$$

where $W(t)$ is a standard Wiener process, [see Brockwell et al. (2007)]. Note that the process $\varepsilon(t)$ has the continuous derivative $\varepsilon'(t)$ and the continuous process with drift is again denoted by $y(t) = \theta f(t) + \varepsilon(t)$, $t \in [A, B]$. We also note that $y(t)$ is differentiable on the interval $[A, B]$. There are in fact three different forms of the autocovarince functions (note that we assume throughout $\sigma^2 = 1$) of continuous AR(2) processes [see e.g. formulas (14)–(16) in He and Wang (1989)], which are given by

$$\rho^{(1)}(t) = \frac{\lambda_2}{\lambda_2 - \lambda_1}e^{-\lambda_1|t|} - \frac{\lambda_1}{\lambda_2 - \lambda_1}e^{-\lambda_2|t|}, \tag{4.3}$$

where $\lambda_1 \neq \lambda_2$, $\lambda_1 > 0$, $\lambda_2 > 0$, by

$$\rho^{(2)}(t) = e^{-\lambda|t|}\Big\{\cos(q|t|) + \frac{\lambda}{q}\sin(q|t|)\Big\},$$

where $\lambda > 0$, $q > 0$, and by

$$\rho^{(3)}(t) = e^{-\lambda|t|}(1 + \lambda|t|),$$

where $\lambda > 0$. From formulas (11)–(13) in He and Wang (1989) we obtain that the corresponding three forms of the autocovariances of the discrete AR(2) process of the form (4.1) are given by

$$r_k^{(1)} = \mathbb{E}[\epsilon_j\epsilon_{j+k}] = Cp_1^k + (1-C)p_2^k, \quad C = \frac{(1-p_2^2)p_1}{(1-p_2^2)p_1 - (1-p_1^2)p_2}, \tag{4.4}$$

where $j \geq 0$, $p_1 \neq p_2$, $0 < |p_1|, |p_2| < 1$; by

$$r_k^{(2)} = p^k\big(\cos(bk) + C\sin(bk)\big), \quad C = \cot(b)\frac{1-p^2}{1+p^2}\,, \tag{4.5}$$

where $0 < p < 1$, $0 < b < 2\pi$ and $b \neq \pi$, and finally by

$$r_k^{(3)} = p^k\,(1 + kC), \quad C = \frac{1-p^2}{1+p^2}\,, \tag{4.6}$$

where $0 < |p| < 1$. In the following subsections we determine approximations for the optimal weights $w_i^*$ in Lemma 2.1 for the different types of autocovariance functions. All results will be summarized in Theorem 4.1 below.

## 4.1 Autocovariances of the form (4.4)

From Corollary 4.1 we obtain that

$$\begin{aligned} Sw_i^* f_i &= -a_2 f_{i-2} + (a_1a_2 - a_1)f_{i-1} + (1 + a_1^2 + a_2^2)f_i + (a_1a_2 - a_1)f_{i+1} - a_2 f_{i+2} \\ &= a_2(2f_i - f_{i-2} - f_{i+2}) - (a_1a_2 - a_1)(2f_i - f_{i-1} - f_{i+1}) \\ &\quad +(1 + a_1^2 + a_2^2 - 2a_2 + 2a_1a_2 - 2a_1)f_i \\ &= a_2(2f_i - f_{i-2} - f_{i+2}) - (a_1a_2 - a_1)(2f_i - f_{i-1} - f_{i+1}) \\ &\quad +(a_1 + a_2 - 1)^2 f_i \end{aligned}$$

for $i = 3, 4, \ldots, N-2$. Now consider the case when the autocovariance structure of the errors has the form (4.4) for fixed $N$. Suppose that the parameters of the autocovariance function (4.4) satisfy $p_1 \neq p_2$, $0 < p_1, p_2 < 1$. We do not discuss the case with negative $p_1$ or negative $p_2$ because discrete AR(2) processes with such parameters do not have continuous real-valued analogues. From the Yule-Walker equations we obtain that the coefficients $a_1$ and $a_2$ in (4.1) are given by

$$a_1 = r_1\frac{1-r_2}{1-r_1^2}, \quad a_2 = \frac{r_2 - r_1^2}{1-r_1^2}\,, \tag{4.7}$$

where $r_1 = r_1^{(1)}$ and $r_2 = r_2^{(1)}$ are defined by (4.4). With the notation $\lambda_1 = -\log(p_1)/\Delta$ and $\lambda_2 = -\log(p_2)/\Delta$ with $\Delta = (B-A)/N$ we obtain

$$p_1 = e^{-\lambda_1\Delta}, \quad p_2 = e^{-\lambda_2\Delta}. \tag{4.8}$$

We will assume that $\lambda_1$ and $\lambda_2$ are fixed but $\Delta$ is small and consider the properties of different quantities as $\Delta \to 0$. By a straightforward Taylor expansion we obtain the approximations

$$\begin{aligned} a_1 = a_1(\Delta) &= 2 - (\lambda_1 + \lambda_2)\Delta + (\lambda_1^2 + \lambda_2^2)\Delta^2/2 + O(\Delta^3), \\ a_2 = a_2(\Delta) &= -1 + (\lambda_1 + \lambda_2)\Delta - (\lambda_1 + \lambda_2)^2\Delta^2/2 + O(\Delta^3), \\ S = S(\Delta) &= 2\lambda_1\lambda_2(\lambda_1 + \lambda_2)\Delta^3 + O(\Delta^4), \\ C = C(\Delta) &= \frac{\lambda_2}{\lambda_2 - \lambda_1} + \frac{1}{6}\lambda_1\lambda_2\frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}\Delta^2 + O(\Delta^4). \end{aligned} \tag{4.9}$$

Consequently (observing (4.8) and (4.9)), for large $N$ the continuous AR(2) process with autocovariances (4.3) can be considered as an approximation to the discrete AR(2) process with autocovariances (4.4).
Since $S = O(\Delta^3)$, $a_1 = 2 + O(\Delta)$ and $a_2 = -1 + O(\Delta)$, it follows

$$
\begin{aligned}
S\frac{w_i^*}{\Delta^4}f_i &= -4a_2\frac{1}{\Delta^2}f''(t_i) + (a_1a_2 - a_1)\frac{1}{\Delta^2}f''(t_i) + \frac{1}{\Delta^4}(a_1 + a_2 - 1)^2 f_i + O(\Delta) \\
&= \frac{1}{\Delta^2}(a_1a_2 - a_1 - 4a_2)f''(t_i) + \frac{1}{\Delta^4}(a_1 + a_2 - 1)^2 f_i + O(\Delta) \\
&= -(\lambda_1^2 + \lambda_2^2)f''(t_i) + \lambda_1^2\lambda_2^2 f_i + O(\Delta).
\end{aligned}
$$

Thus, the optimal weights $w_i^*$, $i = 3, \ldots, N-2$, are approximated by the signed density

$$
p(t) = -\frac{1}{s_3 f(t)}\big((\lambda_1^2 + \lambda_2^2)f''(t) - \lambda_1^2\lambda_2^2 f(t)\big), \tag{4.10}
$$

where $s_3 = 2\lambda_1\lambda_2(\lambda_1 + \lambda_2)$. For the boundary points we obtain

$$
\begin{aligned}
Sw_1^*f_1 &= f_1 - a_1f_2 - a_2f_3 \\
&= (-2f_2 + f_3 + f_1) + (\lambda_1 + \lambda_2)(f_2 - f_3)\Delta \\
&\quad +((-1/2f_2 + 1/2f_3)\lambda_1^2 + f_3\lambda_1\lambda_2 + (-1/2f_2 + 1/2f_3)\lambda_2^2)\Delta^2 \\
&\quad +((1/6f_2 - 1/6f_3)\lambda_1^3 - 1/2f_3\lambda_1^2\lambda_2 - 1/2f_3\lambda_1\lambda_2^2 + (1/6f_2 - 1/6f_3)\lambda_2^3)\Delta^3 + O(\Delta^4) \\
&= \big(f''(t_2) - (\lambda_1 + \lambda_2)f'(t_2) + f_3\lambda_1\lambda_2\big)\Delta^2 + O(\Delta^3)
\end{aligned}
$$

and

$$
\begin{aligned}
Sw_2^*f_2 &= -a_1f_1 + (1 + a_1^2)f_2 + (a_1a_2 - a_1)f_3 - a_2f_4 \\
&= (-2f_1 + f_4 + 5f_2 - 4f_3) + (\lambda_1 + \lambda_2)(f_1 - 4f_2 + 4f_3 - f_4)\Delta \\
&\quad +((-1/2f_1 + 1/2f_4 - 3f_3 + 3f_2)\lambda_1^2 + (2f_2 - 4f_3 + f_4)\lambda_2\lambda_1 \\
&\qquad + (-1/2f_1 + 1/2f_4 - 3f_3 + 3f_2)\lambda_2^2)\Delta^2 \\
&\quad +((1/6f_1 - 5/3f_2 + 5/3f_3 - 1/6f_4)\lambda_1^3 + (-f_2 + 3f_3 - 1/2f_4)\lambda_2\lambda_1^2 \\
&\qquad + (-f_2 + 3f_3 - 1/2f_4)\lambda_2^2\lambda_1 + (1/6f_1 - 5/3f_2 + 5/3f_3 - 1/6f_4)\lambda_2^3)\Delta^3 + O(\Delta^4) \\
&= \big(f''(t_3) - 2f''(t_2) + (\lambda_1 + \lambda_2)(3f'(t_2) - f'(t_1) - f'(t_3)) - f_3\lambda_1\lambda_2\big)\Delta^2 + O(\Delta^3) \\
&= \big(-f''(t_2) + (\lambda_1 + \lambda_2)f'(t_2) - f_3\lambda_1\lambda_2\big)\Delta^2 + O(\Delta^3)
\end{aligned}
$$

Thus, we can see that

$$
w_1^* = -w_2^* + O(1) = Q_A\frac{1}{\Delta} + O(1),
$$

where

$$
Q_A = \frac{1}{s_3 f(A)}\big(f''(A) - (\lambda_1 + \lambda_2)f'(A) + \lambda_1\lambda_2 f(A)\big). \tag{4.11}
$$

This means that the coefficients $w_1^*$ and $w_2^*$ at $t_1$ and $t_2$ are large in absolute value and have different signs. Similarly, we have

$$w_N^* = -w_{N-1}^* + O(1) = Q_B \frac{1}{\Delta} + O(1)$$

where

$$Q_B = \frac{1}{s_3 f(B)} \big( f''(B) + (\lambda_1 + \lambda_2) f'(B) + \lambda_1 \lambda_2 f(B) \big). \tag{4.12}$$

To do a finer approximation, we have to investigate the quantity

$$g := S w_1^* f_1 + S w_2^* f_2,$$

which is of order $O(1)$. Indeed, we have

$$\begin{aligned} g &= (3f_2 - 3f_3 - f_1 + f_4) + (\lambda_2 + \lambda_1)(f_1 - 3f_2 + 3f_3 - f_4)\Delta \\ &\quad + ((-f_1 + f_4 - 5f_3 + 5f_2)/2(\lambda_1^2 + \lambda_2^2) + (2f_2 - 3f_3 + f_4)\lambda_2\lambda_1)\Delta^2 \\ &\quad + ((f_1 - 9f_2 + 9f_3 - f_4)/6(\lambda_1^3 + \lambda_2^3) \\ &\quad + (-2f_2 + 5f_3 - f_4)/2(\lambda_1^2\lambda_2 + \lambda_1\lambda_2^2))\Delta^3 + O(\Delta^4) \\ &= f'''(t_1)\Delta^3 + O(\Delta^4) + (-f'(t_1)(\lambda_1^2 + \lambda_2^2) - f'(t_1)\lambda_2\lambda_1)\Delta^3 \\ &\quad + f(t_1)(\lambda_1^2\lambda_2 + \lambda_1\lambda_2^2)\Delta^3 + O(\Delta^4) \\ &= \big( f'''(t_1) - (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) f'(t_1) + \lambda_1\lambda_2(\lambda_1 + \lambda_2) f(t_1) \big)\Delta^3 + O(\Delta^4) \end{aligned}$$

and, consequently,

$$w_1^* f_1 + w_2^* f_2 = \frac{1}{s_3}\big( f'''(t_1) - (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) f'(t_1) + \lambda_1\lambda_2(\lambda_1 + \lambda_2) f(t_1) \big) + O(\Delta),$$

where $s_3 = 2\lambda_1\lambda_2(\lambda_1 + \lambda_2)$. Therefore, if $\Delta \to 0$, it follows that $w_1^* + w_2^* \approx P_A$, where

$$P_A = \frac{1}{s_3 f(A)} \big( f'''(A) - (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) f'(A) + \lambda_1\lambda_2(\lambda_1 + \lambda_2) f(A) \big). \tag{4.13}$$

Similarly, we obtain $w_N^* + w_{N-1}^* \approx P_B$ if $\Delta \to 0$, where

$$P_B = \frac{1}{s_3 f(B)} \big( -f'''(B) + (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) f'(B) + \lambda_1\lambda_2(\lambda_1 + \lambda_2) f(B) \big). \tag{4.14}$$

Summarizing, we have proved the following result.

**Proposition 4.1** *Consider the one-parameter model* (2.2) *such that the errors follow the* AR(2) *model with autocovariance function* (4.4). *Assume that $f(\cdot)$ is a three times continuously differentiable and $f(t) \neq 0$ for all $t \in [A, B]$. Then for large $N$, the optimal discrete SLSE (defined in Lemma 2.1) can be approximated by the continuous SLSE*

$$\hat\theta = D^* \Big( Q_B f(B) y'(B) - Q_A f(A) y'(A) + P_A f(A) y(A) + P_B f(B) y(B) + \int_A^B p(t) f(t) y(t) dt \Big) \tag{4.15}$$

*where*

$$D^* = \Big(Q_B f(B) f'(B) - Q_A f(A) f'(A) + P_A f^2(A) + P_B f^2(B) + \int_A^B p(t) f^2(t) dt\Big)^{-1}$$

*and* $p(t)$, $Q_A$, $Q_B$, $P_A$ *and* $P_B$ *defined in* (4.10), (4.11), (4.12), (4.13) *and* (4.14) *respectively. For this approximation, we have* $D^* = \lim_{N\to\infty} \mathrm{Var}(\hat{\theta}_{\mathrm{SLSE},N})$*, i.e.* $D^*$ *is the limit of the variance* (2.3) *of the optimal discrete SLSE design as* $N \to \infty$*.*

In the following discussion we call a tuple $(p, Q_A, Q_B, P_A, P_B)$ which contains a (signed) density $p(\cdot)$ and four weights $Q_A, Q_B, P_A, P_B$, an approximate design for the continuous SLSE estimator defined in (4.15).

## 4.2 Autocovariances of the form (4.5)

Consider the autocovariance function of the form (4.5), then the coefficients $a_1$ and $a_2$ are given by (4.7) where $r_1 = r_1^{(2)}$ and $r_2 = r_2^{(2)}$ are defined by (4.5). With the notations $\lambda = -\log p/\Delta$ and $q = b/\Delta$ (or equivalently $p = e^{-\lambda\Delta}$ and $b = q\Delta$) we obtain by a Taylor expansion

$$\begin{aligned} a_1 &= 2 - 2\lambda\Delta + (\lambda^2 - q^2)\Delta^2 + O(\Delta^3), \\ a_2 &= -1 + 2\lambda\Delta - 2\lambda^2\Delta^2 + O(\Delta^3), \\ S &= 4\lambda(\lambda^2 + q^2)\Delta^3 + O(\Delta^4) \end{aligned}$$

and

$$C = \frac{\lambda}{q} - \frac{\lambda(\lambda^2 + q^2)}{3q}\Delta^2 + O(\Delta^4)$$

as $\Delta \to 0$. Similarly, we have

$$\begin{aligned} S\frac{w_i^*}{\Delta^4} f_i &= \frac{1}{\Delta^2}(a_1 a_2 - a_1 - 4a_2) f''(t_i) + \frac{1}{\Delta^4}(a_1 + a_2 - 1)^2 f_i + O(\Delta) \\ &= -2(\lambda^2 - q^2) f''(t_i) + (\lambda^2 + q^2)^2 f_i + O(\Delta). \end{aligned}$$

Thus, the optimal weights $w_i^*$, $i = 3, \ldots, N-2$, are approximated by the signed density

$$p(t) = -\frac{1}{s_3 f(t)}\big(2(\lambda^2 - q^2) f''(t) - (\lambda^2 + q^2)^2 f(t)\big), \tag{4.16}$$

where $s_3 = 4\lambda(\lambda^2 + q^2)$. Similarly, we obtain that

$$\begin{aligned} w_1^* &= -w_2^* + O(1) = Q_A \frac{1}{\Delta} + O(1), \\ w_N^* &= -w_{N-1}^* + O(1) = Q_B \frac{1}{\Delta} + O(1), \end{aligned}$$

where

$$Q_A = \frac{1}{s_3 f(A)}\big(f''(A) - 2\lambda f'(A) + (\lambda^2+q^2)f(A)\big), \quad (4.17)$$

$$Q_B = \frac{1}{s_3 f(B)}\big(f''(B) + 2\lambda f'(B) + (\lambda^2+q^2)f(B)\big). \quad (4.18)$$

Calculating $g := Sw_1^* f_1 + Sw_2^* f_2$ we have

$$\begin{aligned} g &= (3f_2 - f_1 - 3f_3 + f_4) + 2\lambda(f_1 - 3f_2 + 3f_3 - f_4)\Delta \\ &\quad +((-f_1 + 7f_2 - 8f_3 + 2f_4)\lambda^2 + q^2(f_1 - 3f_2 + 2f_3))\Delta^2 \\ &\quad +((-f_1 + 7f_2 - 4f_3)\lambda q^2 + (f_1 - 15f_2 + 24f_3 - 4f_4)/3\lambda^3)\Delta^3 + O(\Delta^4) \\ &= f'''(t_1)\Delta^3 - (3\lambda^2 - q^2)f'(t_1)\Delta^3 + 2\lambda(\lambda^2+q^2)f(t_1)\Delta^3 + O(\Delta^4). \end{aligned}$$

Therefore, it follows that $w_1^* + w_2^* \approx P_A$ if $\Delta \to 0$, where

$$P_A = \frac{1}{s_3 f(A)}\big(f'''(A) - (3\lambda^2 - q^2)f'(A) + 2\lambda(\lambda^2+q^2)f(A)\big), \quad (4.19)$$

and $s_3 = 4\lambda(\lambda^2+q^2)$. Similarly, we obtain the approximation $w_N^* + w_{N-1}^* \approx P_B$ if $\Delta \to 0$, where

$$P_B = \frac{1}{s_3 f(B)}\big(-f'''(B) + (3\lambda^2 - q^2)f'(B) + 2\lambda(\lambda^2+q^2)f(B)\big). \quad (4.20)$$

Summarizing, we have proved the following result.

**Proposition 4.2** *Consider the one-parameter model* (2.2) *such that the errors follow the* AR(2) *model with autocovariance function* (4.5). *Assume that $f(\cdot)$ is a three times continuously differentiable and $f(t) \neq 0$ for all $t \in [A, B]$. Then for large $N$, the optimal discrete SLSE (defined in Lemma 2.1) can be approximated by the continuous SLSE* (4.15), *where the tuple $(p, Q_A, Q_B, P_A, P_B)$ is defined by* (4.16), (4.17), (4.18), (4.19) *and* (4.20) *respectively.*

## 4.3 Autocovariances of the form (4.6)

For the autocovariance function (4.6) the coefficients $a_1$ and $a_2$ in the AR(2) process are given by (4.7) where $r_1 = r_1^{(3)}$ and $r_2 = r_2^{(3)}$ are defined by (4.6). With the notation $\lambda = -\log p/\Delta$ (or equivalently $p = e^{-\lambda\Delta}$) we obtain the Taylor expansions

$$\begin{aligned} a_1 &= 2 - 2\lambda\Delta + \lambda^2\Delta^2 + O(\Delta^3), \\ a_2 &= -1 + 2\lambda\Delta - 2\lambda^2\Delta^2 + O(\Delta^3), \\ S &= 4\lambda^3\Delta^3 + O(\Delta^4), \\ C &= \lambda\Delta - \frac{\lambda^3}{3}\Delta^3 + O(\Delta^5) \end{aligned}$$

as $\Delta \to 0$. Similar calculations as given in the previous paragraphs give

$$\begin{aligned} S\frac{w_i^*}{\Delta^4} f_i &= \frac{1}{\Delta^2}(a_1 a_2 - a_1 - 4a_2) f''(t_i) + \frac{1}{\Delta^4}(a_1 + a_2 - 1)^2 f_i + O(\Delta) \\ &= -2\lambda^2 f''(t_i) + \lambda^4 f_i + O(\Delta). \end{aligned}$$

Thus, the optimal weights $w_i^*$, $i = 3, \ldots, N-2$, are approximated by the signed density

$$p(t) = -\frac{1}{s_3 f(t)}\big(2\lambda^2 f''(t) - \lambda^4 f(t)\big), \tag{4.21}$$

where $s_3 = 4\lambda^3$. For the remaining weighs $w_1^*$ and $w_2^*$ we obtain

$$\begin{aligned} w_1^* &= -w_2^* + O(1) = Q_A \frac{1}{\Delta} + O(1), \\ w_N^* &= -w_{N-1}^* + O(1) = Q_B \frac{1}{\Delta} + O(1), \end{aligned}$$

with

$$\begin{aligned} Q_A &= \frac{1}{s_3 f(A)}\big(f''(A) - 2\lambda f'(A) + \lambda^2 f(A)\big), && (4.22) \\ Q_B &= \frac{1}{s_3 f(B)}\big(f''(B) + 2\lambda f'(B) + \lambda^2 f(B)\big). && (4.23) \end{aligned}$$

Calculating $g := Sw_1^* f_1 + Sw_2^* f_2$ we have

$$\begin{aligned} g &= (3f_2 - 3f_3 - f_1 + f_4) + 2\lambda(f_1 - 3f_2 + 3f_3 - f_4)\Delta \\ &\quad -\lambda^2(f_1 - 7f_2 + 8f_3 - 2f_4)\Delta^2 \\ &\quad +1/3\lambda^3(f_1 - 15f_2 + 24f_3 - 4f_4)\Delta^3 + O(\Delta^4) \\ &= f'''(t_1)\Delta^3 - 3\lambda^2 f'(t_1)\Delta^3 + 2\lambda^3 f(t_1)\Delta^3 + O(\Delta^4). \end{aligned}$$

Therefore, if $\Delta \to 0$, it follows that $w_1^* + w_2^* \approx P_A$, where

$$P_A = \frac{1}{s_3 f(A)}\big(f'''(A) - 3\lambda^2 f'(A) + 2\lambda^3 f(A)\big), \tag{4.24}$$

and $s_3 = 4\lambda^3$. Similarly, we obtain the approximation $w_N^* + w_{N-1}^* \approx P_B$ if $\Delta \to 0$, where

$$P_B = \frac{1}{s_3 f(B)}\big(-f'''(B) + 3\lambda^2 f'(B) + 2\lambda^3 f(B)\big). \tag{4.25}$$

Summarizing, we have proved the following result.

**Proposition 4.3** *Consider the one-parameter model* (2.2) *such that the errors follow the* AR(2) *model with autocovariance function* (4.6). *Then for large $N$, the optimal discrete SLSE (defined in Lemma 2.1) can be approximated by the continuous SLSE* (4.15)*, where the tuple $(p, Q_A, Q_B, P_A, P_B)$ is defined in* (4.21), (4.22), (4.23), (4.24) *and* (4.25) *respectively.*

### 4.4 General statement

Propositions 4.1 - 4.3 can be combined in the following statement.

**Theorem 4.1** *Consider the one-parameter model* (2.2) *such that the errors follow the* AR(2) *model. Assume that* $f(\cdot)$ *is a three times continuously differentiable and* $f(t) \neq 0$ *for all* $t \in [A, B]$. *Define the following constants depending on the form of the autocovariance function* $r_k$. *If* $r_k$ *is of the form* (4.4), *set*

$$\begin{aligned}\lambda_1 &= -\frac{\ln(p_1)}{\Delta},\ \lambda_2 = -\frac{\ln(p_2)}{\Delta},\\ \tau_0 &= \lambda_1^2\lambda_2^2,\ \tau_2 = \lambda_1^2+\lambda_2^2,\ \beta_1 = \lambda_1+\lambda_2,\ \beta_0 = \lambda_1\lambda_2,\\ \gamma_1 &= \lambda_1^2+\lambda_1\lambda_2+\lambda_2^2\ ,\gamma_0 = \lambda_1\lambda_2(\lambda_1+\lambda_2),\ s_3 = 2\lambda_1\lambda_2(\lambda_1+\lambda_2).\end{aligned}$$

*If* $r_k$ *is of the form* (4.5), *set*

$$\begin{aligned}\lambda &= -\frac{\ln(p)}{\Delta},\ q = -\frac{b}{\Delta},\\ \tau_0 &= (\lambda^2+q^2)^2,\ \tau_2 = 2(\lambda^2-q^2),\ \beta_1 = 2\lambda,\ \beta_0 = \lambda^2+q^2,\\ \gamma_1 &= (3\lambda^2-q^2)\ ,\gamma_0 = 2\lambda(\lambda^2+q^2),\ s_3 = 4\lambda(\lambda^2+q^2).\end{aligned}$$

*If* $r_k$ *is of the form* (4.6), *set*

$$\begin{aligned}\lambda &= -\frac{\ln(p)}{\Delta},\ \tau_0 = \lambda^4,\ \tau_2 = 2\lambda^2,\ \beta_1 = 2\lambda,\ \beta_0 = \lambda^2,\\ \gamma_1 &= 3\lambda^2\ ,\gamma_0 = 2\lambda^3,\ s_3 = 4\lambda^3.\end{aligned}$$

*For large* $N$, *the optimal discrete SLSE (defined in Lemma 2.1) can be approximated by the continuous SLSE*

$$\hat{\theta} = D^*\big(Q_B f(B)y'(B) - Q_A f(A)y'(A) + P_A f(A)y(A) + P_B f(B)y(B) + \int_A^B p(t)f(t)y(t)dt\big)$$

*where*

$$D^* = \Big(Q_B f(B)f'(B) - Q_A f(A)f'(A) + P_A f^2(A) + P_B f^2(B) + \int_A^B p(t)f^2(t)dt\Big)^{-1}.$$

*For this approximation, we have* $D^* = \lim_{N\to\infty}\mathrm{Var}(\hat{\theta}_{SLSE,N})$, *i.e.* $D^*$ *is the limit of the variance* (2.3) *of the optimal discrete SLSE design as* $N \to \infty$. *Here the quantities* $p(t)$, $Q_A$, $Q_B$, $P_A$

*and $P_B$ in the continuous SLSE are defined by*

$$
\begin{aligned}
p(t) &= -\frac{1}{s_3 f(t)}\big(\tau_2 f''(t) - \tau_0 f(t)\big), && (4.26)\\
P_A &= \frac{1}{s_3 f(A)}\big(f'''(A) - \gamma_1 f'(A) + \gamma_0 f(A)\big),\\
P_B &= \frac{1}{s_3 f(B)}\big(-f'''(B) + \gamma_1 f'(B) + \gamma_0 f(B)\big),\\
Q_A &= \frac{1}{s_3 f(A)}\big(f''(A) - \beta_1 f'(A) + \beta_0 f(A)\big),\\
Q_B &= \frac{1}{s_3 f(B)}\big(f''(B) + \beta_1 f'(B) + \beta_0 f(B)\big). && (4.27)
\end{aligned}
$$

# 5 Examples

## 5.1 Approximations of the discrete SLSE

Consider the one-parameter model with $f(t) = t^\alpha$ and AR(1) errors ($0 < a < 1$). The design space is given by an interval $[A, B]$ such that $f(t) \neq 0$ for all $t \in [A, B]$. Then the optimal discrete design for the SLSE is approximated by a design of the form (2.5), where the density $p(t)$, and the weights $P_A$ and $P_B$ are defined by

$$
\begin{aligned}
p(t) &= -\frac{1}{2\lambda}\big(\alpha(\alpha-1)t^{-2} - \lambda^2\big),\\
P_A &= \frac{1}{2\lambda}\big(-\alpha A^{-1} + \lambda\big),\\
P_B &= \frac{1}{2\lambda}\big(\alpha B^{-1} + \lambda\big).
\end{aligned}
$$

In Table 1 we display values of $p(t)$, $P(A)$ and $P_B$ for several exponents $\alpha$ and also for the regression function $f(t) = e^t$ For example, if $f(t) = e^t$ we observe that $P_A$ is positive for $\lambda > 1$ and negative for $0 < \lambda < 1$, $P_B$ is positive for $\lambda > 0$, $p(t)$ is positive for $\lambda > 1$ and negative for $\lambda \in (0, 1)$. For large $\lambda$, the contribution of observations at the interval $(A, B)$ to the continuous SLSE is significant. For the location model $f(t) = 1$, we can see that $P_B = P_B = 1/2$ and $p(t) = \lambda/2$. This implies that for small $\lambda$ the contribution of observations at boundary points to the continuous SLSE is large and the contribution of observations at the interval $(A, B)$ to the continuous SLSE is small. For large $\lambda$, the contribution of observations at the interval $(A, B)$ to the continuous SLSE is essential.

Next we consider the same models with an AR(2) error process. For example, if $f(t) = t^\alpha$ the

Table 1: *The function $p(t)$ and the weights $P_A$ and $P_B$ of the continuous SLSE for several functions $f(t)$ and an AR(1) error process.*

| $f(t)$ | $P_A$ | $P_B$ | $p(t)$ |
|---|---|---|---|
| $1$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{\lambda}{2}$ |
| $t$ | $\frac{1}{2}-\frac{1}{2A\lambda}$ | $\frac{1}{2}+\frac{1}{2B\lambda}$ | $\frac{\lambda}{2}$ |
| $t^2$ | $\frac{1}{2}-\frac{1}{A\lambda}$ | $\frac{1}{2}+\frac{1}{B\lambda}$ | $\frac{\lambda}{2}-\frac{1}{\lambda t^2}$ |
| $t^3$ | $\frac{1}{2}-\frac{3}{2A\lambda}$ | $\frac{1}{2}+\frac{3}{2B\lambda}$ | $\frac{\lambda}{2}-\frac{3}{\lambda t^2}$ |
| $t^4$ | $\frac{1}{2}-\frac{2}{A\lambda}$ | $\frac{1}{2}+\frac{2}{B\lambda}$ | $\frac{\lambda}{2}-\frac{6}{\lambda t^2}$ |
| $e^t$ | $\frac{1}{2}-\frac{1}{2\lambda}$ | $\frac{1}{2}+\frac{1}{2\lambda}$ | $\frac{\lambda}{2}-\frac{1}{2\lambda}$ |

SLSE is approximated by the continuous SLSE of the form (2.5), where

$$
\begin{aligned}
p(t) &= -\frac{1}{s_3}\big(\tau_2\alpha(\alpha-1)t^{-2}-\tau_0\big),\\
P_A &= \frac{1}{s_3}\big(\alpha(\alpha-1)(\alpha-2)A^{-3}-\gamma_1\alpha A^{-1}+\gamma_0\big),\\
P_B &= \frac{1}{s_3}\big(-\alpha(\alpha-1)(\alpha-2)B^{-3}+\gamma_1\alpha B^{-1}+\gamma_0\big),\\
Q_A &= \frac{1}{s_3}\big(\alpha(\alpha-1)A^{-2}-\beta_1\alpha A^{-1}+\beta_0\big),\\
Q_B &= \frac{1}{s_3}\big(\alpha(\alpha-1)B^{-2}+\beta_1\alpha B^{-1}+\beta_0\big).
\end{aligned}
$$

Note that signs of $p(t)$, $Q_A$, $Q_B$, $P_A$ and $P_B$ depend on the form of the autocovariance function and its parameters. For the form (4.6), we provide values of $p(t)$, $Q_A$, $Q_B$, $P_A$ and $P_B$ for several functions $f(t)$ in Table 2. The other cases can be obtained similarly and are not displayed for the sake of brevity.

For example, if $f(t)=e^t$ we can see that both $P_A$ and $Q_A$ are positive for all $\lambda \neq 1$, $P_B$ is positive for $\lambda > 0.5$ and negative for $\lambda \in (0, 0.5)$, $p(t)$ is positive for $\lambda > \sqrt{2}$ and negative for $\lambda \in (0, \sqrt{2})$. For large $\lambda$, the contribution of observations at the interval $(A, B)$ to the continuous SLSE is notable. For the location model $f(t)=1$, we can see that $P_A = P_B = 1/2$, $Q_A = Q_B = 1/(4\lambda)$ and $p(t) = \lambda/4$. This implies that for small $\lambda$ the contribution of observations at boundary points to the continuous SLSE is very large and the contribution of observations at the interval

$(A, B)$ to the continuous SLSE is small. For large $\lambda$, the contribution of observations at the interval $(A, B)$ to the continuous SLSE is essential.

Table 2: *The function $p(t)$ and the weights $P_A$, $P_B$, $Q_A$ and $Q_B$ in the continuous SLSE for several functions $f(t)$ and an AR(2) error process with the autocovariance function* (4.6).

| $f(t)$ | $P_A$ | $P_B$ | $p(t)$ | $Q_A$ | $Q_B$ |
|---|---|---|---|---|---|
| $1$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{\lambda}{4}$ | $\frac{1}{4\lambda}$ | $\frac{1}{4\lambda}$ |
| $t$ | $\frac{1}{2} - \frac{3}{4A\lambda}$ | $\frac{1}{2} + \frac{3}{4B\lambda}$ | $\frac{\lambda}{4}$ | $\frac{1}{4\lambda} - \frac{1}{2A\lambda^2}$ | $\frac{1}{4\lambda} + \frac{1}{2B\lambda^2}$ |
| $t^2$ | $\frac{1}{2} - \frac{3}{2A\lambda}$ | $\frac{1}{2} + \frac{3}{2B\lambda}$ | $\frac{\lambda}{4} - \frac{1}{\lambda t^2}$ | $\frac{1}{4\lambda} - \frac{1}{A\lambda^2} + \frac{1}{2A^2\lambda^3}$ | $\frac{1}{4\lambda} + \frac{1}{B\lambda^2} + \frac{1}{2B^2\lambda^3}$ |
| $t^3$ | $\frac{1}{2} - \frac{9}{4A\lambda} + \frac{3}{2A^3\lambda^3}$ | $\frac{1}{2} + \frac{9}{4B\lambda} - \frac{3}{2B^3\lambda^3}$ | $\frac{\lambda}{4} - \frac{3}{\lambda t^2}$ | $\frac{1}{4\lambda} - \frac{3}{2A\lambda^2} + \frac{3}{2A^2\lambda^3}$ | $\frac{1}{4\lambda} + \frac{3}{2B\lambda^2} + \frac{3}{2B^2\lambda^3}$ |
| $t^4$ | $\frac{1}{2} - \frac{3}{A\lambda} + \frac{6}{A^3\lambda^3}$ | $\frac{1}{2} + \frac{3}{B\lambda} - \frac{6}{B^3\lambda^3}$ | $\frac{\lambda}{4} - \frac{6}{\lambda t^2}$ | $\frac{1}{4\lambda} - \frac{2}{A\lambda^2} + \frac{3}{A^2\lambda^3}$ | $\frac{1}{4\lambda} + \frac{2}{B\lambda^2} + \frac{3}{B^2\lambda^3}$ |
| $e^t$ | $\frac{1}{2} - \frac{3}{4\lambda} + \frac{1}{4\lambda^3}$ | $\frac{1}{2} + \frac{3}{4\lambda} - \frac{1}{4\lambda^3}$ | $\frac{\lambda}{4} - \frac{1}{2\lambda}$ | $\frac{1}{4\lambda} - \frac{1}{2\lambda^2} + \frac{1}{4\lambda^3}$ | $\frac{1}{4\lambda} + \frac{1}{2\lambda^2} + \frac{1}{4\lambda^3}$ |

## 5.2 Practical implementation

Suppose that the $N$ equidistant points defined in (3.1) are the potential observation points. Let $K+2$ be the number of observations actually taken in the experiment and that we want to construct a discrete design, which can be implemented in practice. Suppose that $K$ is small and $N$ is large, then efficient designs and corresponding estimators for the model (2.2) can be derived from the continuous approximations, which have been developed in the previous sections.

In Dette et al. (2016) a procedure with a good finite sample performance is proposed. It consists of a slight modification of the SLSE given in (2.1) and a discretization of the density $p(t)$ defined in (3.3) for AR(1) errors and (4.27) for AR(2) errors. To be precise consider a continuous SLSE with weights at the points $A$ and $B$ (the end-points of the interval $[A, B]$), which correspond to the masses $P_A$ and $P_B$ and, for the AR(2) errors, $Q_A$ and $Q_B$ as well. We thus only need to approximate the continuous part of the design, which has a density on $(A, B)$, by a $K$-point design with equal masses.

We assume that the density $p(\cdot)$ is not identically zero on the interval $(A, B)$. Define $\varphi(t) = \kappa |p(t)|$ for $t \in (A, B)$ and choose the constant $\kappa$ such that $\int_A^B \varphi(t)dt = 1$, that is,

$$\kappa = \frac{1}{\int_A^B |p(t)|dt}.$$

Denote by $F(t) = \int_A^t \varphi(s)ds$ the corresponding cumulative distribution function. As $K$-point design we use a $K$-point approximation to the measure with density $\varphi(t)$, that is $\hat{\xi}_K =$

$\{t_{1,K},\dots,t_{K,K};1/K,\dots,1/K\}$, where $t_{i,K} = R(F^{-1}(i/(K+1)))$ $i = 1,2,\dots,K$. Here $R(t)$ is the operator of rounding a number $t$ towards the set of points defined by (3.1), that is $R(F(i/(K+1)) = t_{i,K} := A + (i^* - 1)\Delta$, where

$$|F(i/(K+1)) - A + (i^* - 1)\Delta| = \min\{|F(i/(K+1)) - A + (j-1)\Delta|;\quad j = 1,\dots,N\}.$$

If $p(t) = 0$ on a sub-interval of $[A,B]$ and $F^{-1}(i/(K+1))$ is not uniquely defined then we choose the smallest element from the set $R(F^{-1}(i/(K+1))$ as $t_{i,K}$. Also we define $s_{i,K} = \mathrm{sign}(p(t_{i,K}))$ and obtain from the representation of the continuous SLSE for AR(1) errors in Proposition 3.1 a reasonable estimator with corresponding design. To be precise, $y_1,\dots,y_{K+2}$ should be observed at experimental conditions $A, t_{1,K}, t_{2,K},\dots, t_{K,K}, B$, respectively, and the parameter $\theta$ has to be estimated by the *modified* SLSE

$$\hat\theta_{K+2} = D_{K+2}\Big(P_A f(A) y_A + P_B f(B) y_B + \frac{B-A}{\kappa K}\sum_{i=1}^{K} s_{i,K} f(t_{i,K}) y_i\Big),$$

where

$$D_{K+2} = \Big(P_A f^2(A) + P_B f^2(B) + \frac{B-A}{\kappa K}\sum_{i=1}^{K} s_{i,K} f^2(t_{i,K})\Big)^{-1}.$$

It follows from the discussion of the previous paragraph that $\mathrm{Var}(\hat\theta_{K+2}) \approx D^*$, where $D^*$ is defined in (3.6). Similarly, the modified SLSE for AR(2) errors is defined by

$$\begin{aligned}\hat\theta_{K+2} = D_{K+2}\Big(&\, Q_B f(B) y'(B) - Q_A f(A) y'(A)\\ &+ P_A f(A) y_A + P_B f(B) y_B + \frac{B-A}{\kappa K}\sum_{i=1}^{K} s_{i,K} f(t_{i,K}) y(t_{i,K})\Big)\end{aligned}\tag{4.28}$$

where

$$D_{K+2} = \Big(Q_B f(B) f'(B) - Q_A f(A) f'(A) + P_A f^2(A) + P_B f^2(B) + \frac{B-A}{\kappa K}\sum_{i=1}^{K} s_{i,K} f^2(t_{i,K})\Big)^{-1}.$$

In (4.28) the expressions are the derivatives $y'(A)$ and $y'(B)$ of the continuous approximation $\{y(t)\}_{t\in[A,B]}$, which are usually not available in practice. Therefore, we recommend to make two additional observations at the points $A+\Delta$ and $B-\Delta$ and to replace the derivatives by their approximations $(y_{A+\Delta} - y_A)/\Delta$ and $(y_B - y_{B-\Delta})/\Delta$. Thus, we replace the estimator (4.28) by the weighted least squares estimator (WLSE)

$$\tilde\theta_{K+4} = (X^T W X)^{-1} X^T W Y,\tag{4.29}$$

where $Y = (y_A, y_{A+\Delta}, y_{t_{1,K}},\dots,y_{t_{K,K}}, y_{B-\Delta}, y_B)^T$ and the matrix $W$ is defined by

$$W = \mathrm{diag}\Big\{\frac{P_A}{2} + \frac{Q_A}{\Delta}, \frac{P_A}{2} - \frac{Q_A}{\Delta}, s_{1,K}\frac{B-A}{\kappa K},\dots, s_{K,K}\frac{B-A}{\kappa K}, \frac{P_B}{2} - \frac{Q_B}{\Delta}, \frac{P_B}{2} + \frac{Q_B}{\Delta}\Big\}.\tag{4.30}$$

Note that the variance of $\tilde\theta_{K+4}$ is given by

$$\mathrm{Var}(\tilde\theta_{K+4}) = (X^T W X)^{-1}(X^T W \Sigma W X)(X^T W X)^{-1}.$$

### 5.3 Practical performance

Consider the regression model (2.2) with $f(t) = 1$, $[A, B] = [0, 1]$ and AR(2) errors. Suppose that $N = 101$ so that $t_i = i/100$, $i = 0, 1, \ldots, N$, are potential observation points. We also assume that the autocorrelation function $r_k$ is of the form (4.6) with $\lambda = 1$. We investigate the design $\xi_{K+2}$ with $(K+2)$ points $0, t_{1,K}, t_{2,K}, \ldots, t_{K,K}, 1$ and the design $\xi_{K+4}$ with $(K+4)$ points $0, 0.01, t_{1,K}, t_{2,K}, \ldots, t_{K,K}, 0.99, 1$. The points $t_{1,K}, t_{2,K}, \ldots, t_{K,K}$ are shown in the second column of Table 3. In this table we also display the variances of the WLSE $\tilde{\theta}_{K+4}$, defined by (4.30), the LSE $\hat{\theta}_{\text{LSE},K+2}$ based on the design $\xi_{K+2}$ and the BLUE $\hat{\theta}_{\text{BLUE},K+2}$ and $\hat{\theta}_{\text{BLUE},K+4}$ for the designs $\xi_{K+2}$ and $\xi_{K+4}$, respectively. Let $\hat{\theta}_{\text{BLUE}}$ denote the BLUE based on 101 observations at the points $\{\frac{i}{100} |\ i = 0, \ldots, 100\}$, then we observe that $0.80158449 = \text{Var}(\hat{\theta}_{\text{BLUE}}) \approx D^* = 0.8$ that is in agreement with Theorem 4.1. We also observe that $\text{Var}(\hat{\theta}_{\text{BLUE},K+4}) \cong \text{Var}(\hat{\theta}_{\text{BLUE}})$ and $\text{Var}(\hat{\theta}_{\text{BLUE},K+2}) \ncong \text{Var}(\hat{\theta}_{\text{BLUE}})$ showing the importance of taking one additional observation at each boundary point $A$ and $B$. Note that the proposed estimator $\tilde{\theta}_{K+4}$ defined in (4.30) is nearly as accurate as the BLUE $\hat{\theta}_{\text{BLUE},K+4}$ at the same points and that the LSE $\hat{\theta}_{\text{LSE},K+2}$ is about $10 - 15\%$ worse than the BLUE.

Table 3: *The variances of the LSE, the WLSE defined by* (4.30) *and the BLUE for designs with* $K+2$ *and* $K+4$ *points.* $f(t) = 1$, $[A, B] = [0, 1]$, $N = 101$, *the autocovariance structure is given by* (4.6) *with* $\lambda = 1$, *which yields* $D^* = 0.80000$ *and* $\text{Var}(\hat{\theta}_{\text{BLUE}}) = 0.80158449$.

| $K$ | $t_{1,K}, \ldots, t_{K,K}$ | $\text{Var}(\hat{\theta}_{\text{LSE},K+2})$ | $\text{Var}(\tilde{\theta}_{K+4})$ | $\text{Var}(\hat{\theta}_{\text{BLUE},K+2})$ | $\text{Var}(\hat{\theta}_{\text{BLUE},K+4})$ |
|---|---|---|---|---|---|
| 2 | 0.33, 0.67 | 0.914 | 0.80170 | 0.82663 | 0.80158714 |
| 3 | 0.25, 0.5, 0.75 | 0.921 | 0.80165 | 0.82022 | 0.80158533 |
| 4 | 0.2, 0.4, 0.6, 0.8 | 0.925 | 0.80162 | 0.81681 | 0.80158484 |
| 5 | 0.17, 0.33, 0.5, 0.67, 0.83 | 0.928 | 0.80161 | 0.81443 | 0.80158466 |

As a second example, consider the regression model (2.2) with $f(t) = t^2$, $[A, B] = [0.1, 1.1]$ and AR(2) errors. Suppose that $N = 101$ so that $t_i = 0.1 + i/100$, $i = 0, 1, \ldots, N$, are potential observation points. We also assume that the autocorrelation function $r_j$ is of the form (4.6) with $\lambda = 2$. We investigate the design $\xi_{K+2}$ with $(K+2)$ points $0.1, t_{1,K}, t_{2,K}, \ldots, t_{K,K}, 1.1$ and the design $\xi_{K+4}$ with $(K+4)$ points $0.1, 0.11, t_{1,K}, t_{2,K}, \ldots, t_{K,K}, 1.09, 1.1$. The non-trivial points are shown in the second column of Table 4. In the other columns we display the variances of the different estimators introduced in the previous paragraph. We observe again that $0.37055791 = \text{Var}(\hat{\theta}_{\text{BLUE}}) \approx D^* = 0.36543$ that is in line with Theorem 4.1. Note also that $\text{Var}(\hat{\theta}_{\text{BLUE},K+4}) \cong \text{Var}(\hat{\theta}_{\text{BLUE}})$ and the estimator $\hat{\theta}_{\text{BLUE},K+2}$ without the two additional observations at the boundary is not efficient. Again the proposed estimator $\tilde{\theta}_{K+4}$ is nearly as accurate as the BLUE at the same points but the LSE $\hat{\theta}_{\text{LSE},K+2}$ is dramatically worse than the

BLUE.

Table 4: *The variances of the LSE, the WLSE and the BLUE for designs with* $K+2$ *and* $K+4$ *points.* $f(t) = t^2$, $[A, B] = [0.1, 1.1]$, $N = 101$ *and the autocovariance is given by* (4.6) *with* $\lambda = 2$, *which yields* $D^* = 60000/164189 \cong 0.36543$ *and* $\mathrm{Var}(\hat{\theta}_{\mathrm{BLUE}}) = 0.37055791$.

| $K$ | $t_{1,K}, \ldots, t_{K,K}$ | $\mathrm{Var}(\hat{\theta}_{\mathrm{LSE},K+2})$ | $\mathrm{Var}(\tilde{\theta}_{K+4})$ | $\mathrm{Var}(\hat{\theta}_{\mathrm{BLUE},K+2})$ | $\mathrm{Var}(\hat{\theta}_{\mathrm{BLUE},K+4})$ |
|---|---|---|---|---|---|
| 2 | 0.14, 0.22 | 0.723 | 0.40218 | 0.53175 | 0.37079053 |
| 3 | 0.12, 0.17, 0.27 | 0.751 | 0.40204 | 0.52509 | 0.37072082 |
| 4 | 0.12, 0.15, 0.20, 0.30 | 0.783 | 0.40176 | 0.52089 | 0.37068565 |
| 5 | 0.12, 0.14, 0.17, 0.22, 0.33 | 0.818 | 0.40139 | 0.51689 | 0.37065785 |

**Acknowledgements.** This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Project C2) of the German Research Foundation (DFG) and the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The work of Andrey Pepelyshev was partly supported by the project "Actual problems of design and analysis for regression models" (6.38.435.2015) of St. Petersburg State University. The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise.

# References

Bickel, P. J. and Herzberg, A. M. (1979). Robustness of design against autocorrelation in time I: Asymptotic theory, optimality for location and linear regression. *Annals of Statistics*, 7(1):77–95.

Boltze, L. and Näther, W. (1982). On effective observation methods in regression models with correlated errors. *Math. Operationsforsch. Statist. Ser. Statist.*, 13:507–519.

Brockwell, P., Davis, R., and Yang, Y. (2007). Continuous-time gaussian autoregression. *Statistica Sinica*, 17(1):63.

Chan, K. S. and Tong, H. (1987). A note on embedding a discrete parameter ARMA model in a continuous parameter ARMA model. *J. Time Ser. Anal.*, 8(3):277–281.

Dette, H., Pepelyshev, A., and Zhigljavsky, A. (2013). Optimal design for linear models with correlated observations. *The Annals of Statistics*, 41(1):143–176.

Dette, H., Pepelyshev, A., and Zhigljavsky, A. (2014). Nearly universally optimal designs for models with correlated observations. *Computational Statistics & Data Analysis*, 71:1103–1112.

Dette, H., Pepelyshev, A., and Zhigljavsky, A. (2015). Design for linear regression models with correlated errors. *Handbook of Design and Analysis of Experiments*, 7:237–273.

Dette, H., Pepelyshev, A., and Zhigljavsky, A. (2016). Optimal designs in regression with correlated errors. *The Annals of Statistics*, 44(1):113–152.

He, S. and Wang, J. (1989). On embedding a discrete-parameter arma model in a continuous-parameter arma model. *Journal of Time Series Analysis*, 10(4):315–323.

Mehr, C. and McFadden, J. (1965). Certain properties of Gaussian processes and their first-passage times. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 505–522.

Müller, W. G. and Pázman, A. (2003). Measures for designs in experiments with correlated errors. *Biometrika*, 90:423–434.

Näther, W. (1985a). *Effective Observation of Random Fields*. Teubner Verlagsgesellschaft, Leipzig.

Näther, W. (1985b). Exact design for regression models with correlated errors. *Statistics*, 16:479–484.

Pázman, A. and Müller, W. G. (2001). Optimal design of experiments subject to correlated errors. *Statistics and Probability Letters*, 52:29–34.

Pukelsheim, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia.

Sacks, J. and Ylvisaker, N. D. (1966). Designs for regression problems with correlated errors. *Annals of Mathematical Statistics*, 37:66–89.

Sacks, J. and Ylvisaker, N. D. (1968). Designs for regression problems with correlated errors; many parameters. *Annals of Mathematical Statistics*, 39:49–69.

Zhigljavsky, A., Dette, H., and Pepelyshev, A. (2010). A new approach to optimal design for linear models with correlated observations. *Journal of the American Statistical Association*, 105:1093–1103.