Asymptotic optimal designs under long-range dependence error structure

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@rub.de

Nikolai Leonenko
Cardiff University
School of Mathematics
Cardiff CF24 4AG, UK
email: LeonenkoN@cf.ac.uk

Andrey Pepelyshev
St. Petersburg State University
Department of Mathematics
St. Petersburg, Russia
email: andrey@ap7236.spb.edu

Anatoly Zhigljavsky
Cardiff University
School of Mathematics
Cardiff CF24 4AG, UK
email: ZhigljavskyAA@cf.ac.uk

1 July 2009

Abstract

We discuss the optimal design problem in regression models with long range dependence error structure. Asymptotic optimal designs are derived and it is demonstrated that these designs depend only indirectly on the correlation function. Several examples are investigated to illustrate the theory. Finally the optimal designs are compared with asymptotic optimal designs, which were derived by Bickel and Herzberg (1979) for regression models with short range dependent error.

Keywords and phrases: long range dependence, asymptotic optimal designs, linear regression

AMS Subject Classification: 62K05

1 Introduction

Consider the common linear regression model

$$y(t) = \theta_1 f_1(t) + \ldots + \theta_p f_p(t) + \varepsilon(t) \tag{1.1}$$

where $f_1(t), \ldots, f_p(t)$ are known functions, $\varepsilon(t)$ is a random error, $\theta_1, \ldots, \theta_p$ denote the unknown parameters and t is the explanatory variable. We assume that N observations, say $y_1, \ldots, y_N$ can be taken at experimental conditions $-T \le t_1 \le \ldots \le t_N \le T$ to estimate the parameters in the linear regression model (1.1). If an appropriate estimate, say $\hat\theta = (\hat\theta_1, \ldots, \hat\theta_p)^T$ has been chosen, an optimal design minimizes a function of the variance-covariance matrix of this estimate, which is called optimality criterion [see e.g. Silvey (1980) or Pukelsheim (1993)].

Under the assumption of uncorrelated observations optimal designs have been studied by numerous authors [see the two books cited above and the textbooks of Fedorov (1972), Pázman (1986) and Atkinson and Donev (1992)]. However, less results are available for dependent observations, although this problem is of particular interest, because in many applications the variable t in the regression model (1.1) represents the time and all observations correspond to one subject. The reason for this is that optimal experimental designs for regression models with correlated observations have a very complicated structure and are difficult to find even in simple cases. Because explicit solutions are rarely available, an asymptotic theory was developed by Sacks and Ylvisaker (1966, 1968). In the Sacks–Ylvisaker approach, the design set is fixed and the number of design points in this set tends to infinity. As a result of this assumption, the design points become too close to each other and the corresponding asymptotic optimal designs depend only on the behavior of the correlation function in a neighborhood of the point 0.

Bickel and Herzberg (1979) and Bickel, Herzberg and Schilling (1981) considered a different model, where the design interval expands proportionally to the number of observation points and the correlation structure of errors is not used for construction of least squares estimate. The variance-covariance matrix of the estimate $\hat\theta$ is of order $O(1)$ in the model considered by Sacks and Ylvisaker (1966) and of order $1/N$ in the model discussed by Bickel and Herzberg (1979). Therefore the approach of Bickel and Herzberg makes the optimal designs derived for the dependent and independent cases more comparable. These authors assumed that the observations in model (1.1) have a correlation structure corresponding to a nondegenerate stationary process with short range dependence where a correlation function ρ satisfies $\rho(t) = o(1/t)$ if $t \to \infty$. As examples, in Bickel and Herzberg (1979) and Bickel, Herzberg and Schilling (1981) asymptotic optimal designs are derived for the linear regression model with and without intercept and for location model.

The purpose of the present paper is to extend the Bickel–Herzberg approach to the case of a stronger dependence of the errors in the linear regression model (1.1), which corresponds to an error process with long range dependence. Long range dependence is observed in many applications including hydrology, geophysics, turbulence, diffusion,

economics and finance. The phenomenon was already observed by Pearson (1902) in astronomy and by Smith (1938) in agriculture. Further examples where long range dependence was discussed can be found in Granger (1980), Mandelbrot (1973), Porter-Hudak (1990), Beran, Sherman, Taqqu and Willinger (1992), Barndorff-Nielsen, et al. (1990), Beran (1992), Metzler, et al. (1999) among many others. The interested reader is referred to the books of Beran (1994) and Doukhan, et al. (2003), which contain a good description of the basic properties of long range dependence processes and an extensive bibliography on this subject.
Most of the literature considers the estimation problem but - to the knowledge of the authors - design problems for regression models with long range dependence error structure have not been considered so far. In Section 2 we introduce the basic terminology and describe the optimal design problem. Our main results are given in Section 3, where we derive an asymptotic expression for the variance-covariance matrix, which is the basis for the construction of optimal designs in the regression model (1.1) with a long range dependent error structure. Finally, in Section 4 several asymptotic optimal designs are derived for the linear regression model and compared with the results obtained by Bickel and Herzberg (1979) under the assumption of a short range error structure.

2 Optimal designs for dependent observations

Consider the linear regression model (1.1), where the error process $\varepsilon(t)$ is the second-order process with

$$\mathbf{E}\varepsilon(t) = 0, \quad \mathbf{E}\varepsilon(t)\varepsilon(s) = \sigma^2 r(t,s), \tag{2.1}$$

and assume that

(C1) *the regression functions $f_1(t), \dots, f_p(t)$ are linearly independent and bounded on the interval $[-T,T]$ and satisfy a first order Lipschitz condition, that is $|f_i(t) - f_i(s)| \leq M|t-s|$ and $|f_i(t)| \leq M$ for all $t,s \in [-T,T]$, $i = 1,\dots,p$.*

Following Bickel and Herzberg (1979) we assume that $\varepsilon(t) = \varepsilon^{(1)}(t) + \varepsilon^{(2)}(t)$, where $\varepsilon^{(1)}(t)$ denotes a stationary process with correlation function $\rho(t)$ and $\varepsilon^{(2)}(t)$ is white noise. Consequently, we obtain

$$r(t,s) = \gamma\rho(t-s) + (1-\gamma)\delta_{t,s} \tag{2.2}$$

where δ denotes Kronecker's symbol. If N observations, say $y = (y_1, \dots, y_N)^T$ are available and the form of the correlation function is known, the vector of parameters can be estimated by the weighted least squares, i.e. $\hat{\theta} = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}y$ with $X^T = (f_i(t_j))_{i=1,\dots,p}^{j=1,\dots,N}$, and the variance-covariance matrix of this estimate is given by

$$\mathbf{D}(\hat{\theta}) = \sigma^2(X^T\Sigma^{-1}X)^{-1} \tag{2.3}$$

with $\Sigma = (\rho(t_i - t_j))_{i,j}$, $i,j = 1,\dots,N$. However in most applications knowledge about the correlation structure is not available and the unweighted least squares estimate $\tilde{\theta} = (X^T X)^{-1} X^T y$ is used. For this estimate the variance-covariance matrix is given by

$$\mathbf{D}(\tilde{\theta}) = \sigma^2 (X^T X)^{-1} X^T \Sigma X (X^T X)^{-1}. \tag{2.4}$$

An experimental design $\xi = \{t_1, \dots, t_N\}$ is a vector of N points in the interval $[-T, T]$, which defines the time points or experimental conditions where observations are taken. Optimal designs minimize a functional of the variance-covariance matrix of the weighted or unweighted least squares estimate. Following Bickel and Herzberg (1979), we consider a correlation function which depends on the sample size N and is of the form $\rho_N(t) = \rho(Nt)$, where the function ρ satisfies $\rho(t) \to 0$ if $t \to \infty$; this corresponds to expanding the interval as the number of observations grows. The standard least squares estimate is considered in the following discussion because for computing this estimate, the form of the correlation function $\rho(t)$ is not used. Despite this, the least squares estimate often has good properties compared to the best linear unbiased estimate, see e.g. Beran (1994, p. 179).
For our asymptotic investigations we consider a sequence of designs $\xi_N = \{t_{1N}, \dots, t_{NN}\}$, which is generated by a continuous nondecreasing function

$$a : [0,1] \to [-T, T] \tag{2.5}$$

by

$$t_{iN} = a\left((i-1)/(N-1)\right), \; i = 1, \dots, N, \tag{2.6}$$

where the function $a(u)$ is the inverse of a distribution function. Note that the function a is obtained as the weak limit of ξ_N as $N \to \infty$ and that the equally-spaced design corresponds to the choice $a(u) = (2u-1)T$ $(u \in [0,1])$. We further assume several regularity conditions on the function a, which are required for the following asymptotic results. More precisely,

(C2) *let $a(u)$ be twice differentiable and assume that there exists a positive constant $M < \infty$ such that for all $u \in (0,1)$*

$$\frac{1}{M} \leq a'(u) \leq M, \quad |a''(u)| \leq M. \tag{2.7}$$

(C3) *the correlation function ρ is differentiable with bounded derivative, that is $|\rho'(t)| \leq M$, $t \in (0, \infty)$ and satisfies $\rho'(t) \leq 0$ for sufficiently large t.*

The last assumption implies that $\rho(t)$ is nonnegative for sufficiently large t. In contrast to Bickel and Herzberg (1979) we assume that

$$\int_0^\infty |\rho(t)| \, dt = \infty \tag{2.8}$$

and this assumption corresponds to the long-range dependence of the observations. Note that in this case it follows that

$$\int_0^\infty |\rho(t)|\, dt = \sum_{k=0}^\infty |\rho(k)| = \infty$$

where $\rho(k) = \mathrm{cov}(\varepsilon^{(1)}(t), \varepsilon^{(1)}(t+k))$. The correlation function of a stationary process with long range dependence can be written as

$$\rho_\alpha(t) = \frac{L(t)}{|t|^\alpha}\ ,\ \ |t| \to \infty \tag{2.9}$$

where $0 < \alpha \le 1$ and $L(t)$ is a slowly varying function (SVF) for large t (Doukhan et al., 2003) and satisfies

$$\rho_\alpha(t) = O(1/|t|^\alpha)\ ,\ \ |t| \to \infty.$$

In this case we will say that $\rho_\alpha(t)$ belongs to SVF family.

3 Main results

At first we introduce two parametric families of correlation functions which are important in applications.
The correlation function $\rho_\alpha(t)$ belongs to the Cauchy family if it is defined by

$$\rho_\alpha(t) = \frac{1}{(1+|t|^\beta)^{\alpha/\beta}}, \tag{3.1}$$

where $\beta > 0$, $0 < \alpha \le 1$ [see Gneiting (2000), Anh, et al. (2004), Barndorff-Nielsen, Leonenko (2005)]. This family includes

$$\rho_\alpha^{(1)}(t) = \frac{1}{(1+|t|^2)^{\alpha/2}},\ \ \rho_\alpha^{(2)}(t) = \frac{1}{1+|t|^\alpha},\ \ \rho_\alpha^{(3)}(t) = \frac{1}{(1+|t|)^\alpha},$$

which have a totally different shape at $t = 0$, but the same asymptotic behavior for large t [see Figure 1]. These three functions are known as characteristic functions of symmetric Bessel distribution, Linnik distribution and symmetric generalized Linnik distribution, correspondingly.
The correlation function $\rho_\alpha(t)$ belongs to the Mittag-Leffler family if it is defined by

$$\rho_\alpha(t) = E_{\nu,\beta}(-|t|^\alpha),\ \ \ E_{\nu,\beta}(-t) = \Gamma(\beta)\sum_{k=0}^\infty \frac{(-t)^k}{\Gamma(\nu k+\beta)}, \tag{3.2}$$

where $0 < \alpha \le 1$, $0 < \nu \le 1$, $\beta \ge \nu$ [see Schneider (1996), Barndorff-Nielsen, Leonenko (2005)]. This family is a smooth interpolation of long-range dependence ($0 < \alpha < 1$, $\beta = 1$) and short-range dependence ($\alpha = 1$). Note that the case $\alpha = 1$ corresponds to

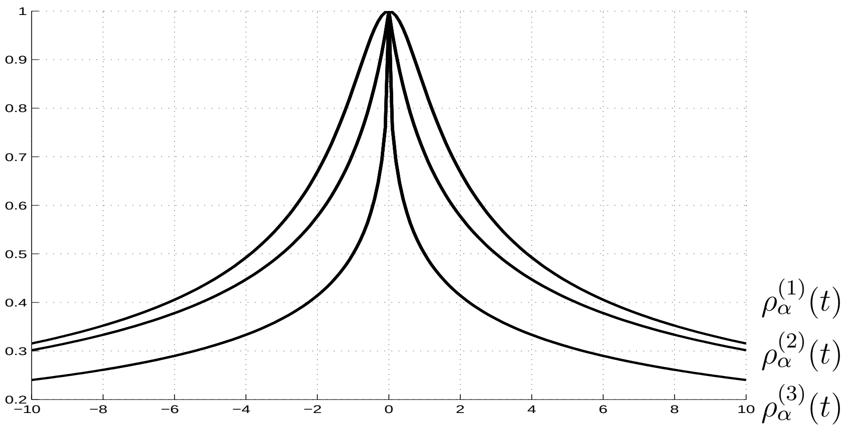

Figure 1: *The three correlation functions, where* $\alpha = 0.5$.

ordinary diffusion while the case $0 < \alpha < 1$, $\beta = 1$ corresponds to subdiffusion or slow diffusion [see Metzler, Klafter (2000)]. In particular,

$$E_{1,1}(-t) = e^{-t},\ E_{1,2}(-t) = (1 - e^{-t})/t,\ E_{1,3}(-t) = 2(e^{-t} - 1 + t)/t^2,$$

$$E_{1/2,1}(-t) = e^{t^2}\left(1 - \frac{2}{\sqrt{\pi}}\int_0^t e^{-u^2}\,du\right).$$

In the following discussion we derive optimal designs for the three families of correlation functions, which are given by (2.9), (3.1) and (3.2). The function $Q(t) = \sum_{j=1}^{\infty} \rho(jt)$ plays an important role in the asymptotic analysis by Bickel and Herzberg (1979), but in the case of long range dependence this function is infinite. For an asymptotic analysis under long range dependence we introduce the function

$$Q_\alpha(t) = \lim_{N\to\infty} \frac{1}{d_\alpha(N)} \sum_{j=1}^{N} \rho_\alpha(jt), \tag{3.3}$$

where the normalizing sequence is given by

$$d_\alpha(N) = \begin{cases} N^{1-\alpha} & \text{if } \alpha < 1 \text{ and } \rho_\alpha \text{ has the form (3.1) or (3.2)} \\ \ln N & \text{if } \alpha = 1 \text{ and } \rho_\alpha \text{ has the form (3.1) or (3.2)} \\ L(N)N^{1-\alpha} & \text{if } \alpha < 1 \text{ and } \rho_\alpha \text{ has the form (2.9)} \\ L(N)\ln N & \text{if } \alpha = 1 \text{ and } \rho_\alpha \text{ has the form (2.9)} \end{cases}$$

and show in Lemma 1 below that the function $Q_\alpha(t)$ is well defined.

Remark. Consider the correlation function

$$\rho(t) = (1 - t^2)e^{-t^2/2} \tag{3.4}$$

which is used to describe scaling laws in the Burgers turbulence problem with weakly dependent random initial conditions [see Leonenko, Woyczynski (1999)]. This function is singular since it is not positive, its spectral density equals $t^2 e^{-t^2/2}$ and $\int \rho(t)\,dt = 0$. Note that function $Q(t)$ is finite for the choice (3.4), but the rate of decrease of the variance of the least squares estimate is much faster than for short range dependence. For example, let $p = 1$, $f(t) \equiv 1$, $T = 1$, $t_i = -1 + 2(i-1)/(N-1)$, $i = 1, \dots, N$, $\rho_N(t) = \rho((N-2)t/2)$. Then we have $X^T X = N$ and the variance of the least squares estimate (2.4) has the form $\mathbf{D}(\theta) = \sigma^2 \sum_{i,j=1}^N \rho_N(t_i - t_j)/N^2 \le 6\sigma^2/N^2$. For $f(t) \equiv t$ and the same points we have $\mathbf{D}(\theta) \le 20\sigma^2/N^2$.

Lemma 1. *If the correlation function $\rho_\alpha(t)$ belongs either to the Cauchy, Mittag-Leffler or SVF family, then the limit in (3.3) exists and is given by*

$$Q_\alpha(t) = \begin{cases} \dfrac{c}{(1-\alpha)|t|^\alpha}\,, & 0 < \alpha < 1, \\ \frac{c}{|t|}\,, & \alpha = 1, \end{cases}$$

where

$$c = \begin{cases} \frac{\Gamma(\beta)}{\Gamma(\beta-\nu)}\,, & \text{if } \rho_\alpha(t) \text{ belongs to the Mittag-Leffler family}, \\ 1\,, & \text{otherwise}. \end{cases}$$

Proof. Define the function

$$Q_{\alpha,N}(t) = \frac{1}{d_\alpha(N)} \sum_{j=1}^N \rho_\alpha(jt),$$

and assume that the correlation function $\rho_\alpha(t)$ is an element of the Cauchy family. Since the function $\rho_\alpha(t)$ defined in (3.1) is positive and decreasing for $0 < \alpha < 1$ we have

$$\begin{aligned} Q_{\alpha,N}(t) &= \frac{1}{N^{1-\alpha}} \int_0^N \frac{1}{(1+|st|^\beta)^{\alpha/\beta}}\,ds + O\left(\frac{1}{N^{1-\alpha}}\right) \\ &= \frac{1}{N^{1-\alpha}} N \int_0^N \frac{d(s/N)}{N^\alpha (1/N^\beta + |st/N|^\beta)^{\alpha/\beta}} + O\left(\frac{1}{N^{1-\alpha}}\right) \\ &= \int_0^1 \frac{dv}{(|vt|^\beta)^{\alpha/\beta}} + O\left(\frac{1}{N^{\alpha-\alpha^2}}\right) = \frac{1}{(1-\alpha)|t|^\alpha} + O\left(\frac{1}{N^{\alpha-\alpha^2}}\right) \\ &= Q_\alpha(t) + O\left(\frac{1}{N^{\alpha-\alpha^2}}\right). \end{aligned}$$

For $\alpha = 1$ we obtain

$$\begin{aligned} Q_1(t) &= \lim_N \frac{1}{\ln N} \sum_{j=1}^N \frac{1}{(1+|jt|^\beta)^{1/\beta}} = \lim_N \frac{1}{|t| \ln N} \int_0^N \frac{1}{(1+|st|^\beta)^{1/\beta}}\,d(st) \\ &= \lim_N \frac{1}{|t| \ln N} \int_0^{Nt} \frac{1}{1+|v|}\,dv = \frac{1}{|t|}, \end{aligned}$$

which completes the proof for the case where $\rho_\alpha(t)$ belongs to the Cauchy family. Now assume that the correlation function $\rho_\alpha(t)$ is an element of the Mittag-Leffler family. Since

$$E_{\nu,1}(-|t|^\alpha) \sim \frac{1}{|t|^\alpha \Gamma(1-\nu)}$$

as $t \to \infty$ and $0 < \alpha < 1$ [see Djrbashian (1993)] we have

$$Q_\alpha(t) = \lim_N \frac{1}{N^{1-\alpha}} \sum_{j=1}^N E_{\nu,1}(-|t|^\alpha) = \frac{1}{(1-\alpha)\Gamma(1-\nu)|t|^\alpha}.$$

Observing

$$E_{\nu,\beta}(-|t|) \sim \frac{\Gamma(\beta)}{|t|\Gamma(\beta-\nu)}$$

for $t \to \infty$ and $\alpha = 1$ [see Djrbashian (1993)] we obtain

$$Q_1(t) = \lim_N \frac{1}{\ln N} \sum_{j=1}^N \rho_1(jt) = \frac{\Gamma(\beta)}{|t|\Gamma(\beta-\nu)}.$$

Finally, assume that the correlation function $\rho_\alpha(t)$ is an element of the SVF family. Then we obtain

$$\begin{aligned}
Q_\alpha(t) &= \lim_N \frac{1}{L(N)N^{1-\alpha}} \int_0^N \frac{L(st)}{|st|^\alpha}\, ds \\
&= \lim_N \frac{1}{L(N)N^{1-\alpha}} N \int_0^N \frac{L(Nts/N)\, d(s/N)}{N^\alpha |st/N|^\alpha} \\
&= \lim_N \int_0^1 \frac{L(Ntv)dv}{L(N)|vt|^\alpha} = \int_0^1 \frac{dv}{|vt|^\alpha} = \frac{1}{(1-\alpha)|t|^\alpha},
\end{aligned}$$

where we have used Theorem 2.6 from Seneta (1976) in last line. For $\alpha = 1$ we have

$$\begin{aligned}
Q_1(t) &= \lim_N \frac{1}{L(N)\ln N} \int_1^N \frac{L(st)}{|st|}\, ds = \lim_N \frac{1}{\ln N} \int_1^N \frac{L(st)/L(N)}{st}\, ds \\
&= \lim_N \frac{1}{\ln N} \int_1^N \frac{1}{|st|}\, ds + \lim_N \frac{1}{\ln N} \int_1^N \frac{L(st)/L(N) - 1}{|st|}\, ds \\
&= \frac{1}{|t|} + \lim_N \frac{1}{\ln N} \int_{1/N}^1 \frac{L(Nvt)/L(N) - 1}{|vt|}\, dv = \frac{1}{|t|}
\end{aligned}$$

which completes proof of Lemma 1. □

Next we find a comfortable asymptotic representation for the main term in the variance-covariance matrix of the least squares estimates.

Lemma 2. *Assume that the correlation function $\rho_\alpha(t)$ belongs either to the Cauchy, Mittag-Leffler or SVF family, such that*

$$\int_0^1 Q_\alpha(a'(t))\, dt < \infty, \tag{3.5}$$

and that the regularity conditions (C1)-(C3) in Section 2 and 3 are satisfied. We have

$$\frac{1}{d_\alpha(N)N}\sum_{i\neq j} f_s(t_{iN})f_r(t_{jN})\rho_\alpha(N(t_{jN}-t_{iN})) = 2\int_0^1 f_s(a(u))f_r(a(u))Q_\alpha(a'(u))\,du + o(1)$$

as $N\to\infty$ *for all* $s,r=1,\ldots,p$, $0<\alpha\le 1$.

Proof. We only give a proof for the correlation function from the Cauchy family and $0<\alpha<1$, the proof for the other cases is similar. We use the notation $f=f_s$, $g=f_r$, $\rho=\rho_\alpha$ and the decomposition

$$N^{\alpha-2}\sum_{i\neq j} f(t_{iN})g(t_{jN})\rho(N(t_{jN}-t_{iN})) = S_1+S_2\ ,$$

where

$$S_1 \;=\; 2N^{\alpha-2}\sum_{i=1}^N f(t_{iN})g(t_{iN})\sum_{j=i+1}^N \rho(N(t_{jN}-t_{iN})), \tag{3.6}$$

$$S_2 \;=\; 2N^{\alpha-2}\sum_{i=1}^N f(t_{iN})\sum_{j=i+1}^N (g(t_{jN})-g(t_{iN}))\rho(N(t_{jN}-t_{iN})). \tag{3.7}$$

With the notation $i_N=(i-1)/(N-1)$ we obtain from the differentiability of the functions a and ρ

$$\rho(N(t_{jN}-t_{iN})) = \rho(N(a(j_N)-a(i_N))) = \rho(a'(i_N)(j-i)) + \nu\frac{(j-i)^2}{N-1}$$

where $|\nu|\le M^2/2$. Let r_N denote a sequence such that $r_N\to\infty$ slowly as $o(N^{(1-\alpha)/3})$ and consider the cases $i\le r_N$ and $i>r_N$ in (3.6) and (3.7) separately. Note that

$$\begin{aligned}\left|\sum_{j=i+r_N}^N \rho(N(t_{jN}-t_{iN}))\right| &= \sum_{j=i+r_N}^N \rho(N(a(j_N)-a(i_N)))\\ \le \tilde M \sum_{j=i+r_N}^N \rho((j-i)/M) &\le \tilde M\sum_{k=r_N}^\infty \rho(k/M) = o(N^{1-\alpha})\end{aligned}$$

as $N\to\infty$ uniformly with respect to j, where $\tilde M$ is a constant and we have used the fact that the function $a'(u)$ is bounded from below and Lemma 1. Similarly, we obtain

$$\left|\sum_{j=i+r_N}^N (g(t_{jN})-g(t_{iN}))\rho(N(t_{jN}-t_{iN}))\right| \;\le\; 2MT\sum_{j=i+r_N}^\infty |\rho(N(t_{jN}-t_{iN}))| = o(N^{1-\alpha})$$

as $N\to\infty$ uniformly with respect to j, because the function g is bounded. This implies

$$S_1 \;=\; 2N^{\alpha-2}\sum_{i=1}^N f(t_{iN})g(t_{iN})\sum_{j=i+1}^{i+r_N} \rho(N(t_{jN}-t_{iN})) + o(1), \tag{3.8}$$

$$S_2 \;=\; 2N^{\alpha-2}\sum_{i=1}^N f(t_{iN})\sum_{j=i+1}^{i+r_N} (g(t_{jN})-g(t_{iN}))\rho(N(t_{jN}-t_{iN})) + o(1) \tag{3.9}$$

as $N \to \infty$. For the the first term on the right hand side of (3.9) we obtain the estimate

$$\begin{aligned}
\tilde{S}_2 &= N^{\alpha-2} \left| \sum_{i=1}^{N} f(t_{iN}) \sum_{j=i+1}^{i+r_N} (g(t_{jN}) - g(t_{iN}))\rho(N(t_{jN} - t_{iN})) \right| \\
&\leq 2N^{\alpha-1} M^2 T \sum_{j=i+1}^{i+r_N} |\rho(N(t_{jN} - t_{iN}))| \\
&\leq 2N^{\alpha-1} M^2 T \sum_{j=i+1}^{i+r_N} \left(|\rho(a'(i_N)(j-i))| + M^2 \frac{(j-i)^2}{N-1} \right) \\
&\leq 2N^{\alpha-1} M^2 T (M r_N + M^2 r_N^3 / N) = o(1)
\end{aligned}$$

as $N \to \infty$ while the dominating term on the right hand side of (3.8) is given by

$$\begin{aligned}
\tilde{S}_1 &= N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) \sum_{j=i+1}^{i+r_N} \rho(N(t_{jN} - t_{iN})) \\
&= N^{\alpha-2} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) \sum_{j=i+1}^{i+r_N} \rho(a'(i_N)(j-i))) + o(1) \\
&= N^{-1} \sum_{i=1}^{N} f(t_{iN}) g(t_{iN}) Q_\alpha(a'(i_N)) + o(1) \\
&= \int_0^1 f(a(u)) g(a(u)) Q_\alpha(a'(u)) \, du + o(1).
\end{aligned}$$

as $N \to \infty$ which proves the assertion of Lemma 2. $\square$

Theorem 1. *Let the correlation function $\rho_\alpha(t)$ be either an element of the Cauchy, Mittag-Leffler or SVF family. If (3.5) and the regularity assumptions* (C1)–(C3) *stated in Sections 2 and 3 are satisfied, then we obtain for the variance-covariance matrix of the least squares estimate defined in (2.4)*

$$\sigma^2 \frac{N}{d_\alpha(N)} \mathbf{D}(\tilde{\theta}) = 2\gamma W^{-1}(a) R_\alpha(a) W^{-1}(a) + O(1/d_\alpha(N)),$$

where the matrices W and R_α are given by

$$W(a) = \left(\int_0^1 f_i(a(u)) f_j(a(u)) \, du \right)_{i,j=1}^{p},$$

$$R_\alpha(a) = \left(\int_0^1 f_i(a(u)) f_j(a(u)) Q_\alpha(a'(u)) \, du \right)_{i,j=1}^{p}.$$

Proof. In view of (2.2) we obtain that

$$X^T\Sigma X = \left(\gamma \sum_{i\neq j} f_k(t_{iN})f_l(t_{jN})\rho_\alpha(N(t_{jN}-t_{iN})) + \sum_{i=1}^{N} f_k(t_{iN})f_l(t_{jN})\right)_{k,l=1}^{p}$$

where $X^T = (f_i(t_{jN}))_{i=1,\ldots,p}^{j=1,\ldots,N}$ and $t_{iN} = a\left((i-1)/(N-1)\right),\ i=1,\ldots,N.$ An application of Lemma 2 yields

$$\frac{X^TX}{N} = W(a) + O\left(\frac{1}{N}\right), \quad \frac{X^T\Sigma X}{d_\alpha(N)N} = 2\gamma R_\alpha(a) + O\left(\frac{1}{d_\alpha(N)}\right).$$

The assertion of the theorem now follows by inserting these limits into (2.4). $\square$

Note that the constant γ only appears as a factor in the asymptotic variance-covariance matrices of the least squares estimate. Because most optimality criteria are positively homogeneous [see e.g. Pukelsheim (1993)] it is reasonable to consider the matrix

$$W^{-1}(a)R_\alpha(a)W^{-1}(a),$$

which is proportional to the asymptotic variance-covariance matrix of the least squares estimate. Moreover, if the function a corresponds to a continuous distribution with a density, say ϕ, then $a'(t) = 1/\phi(t)$ and the asymptotic variance-covariance matrix of the least squares estimate is proportional to the matrix

$$\Psi_\alpha(\phi) = W^{-1}(\phi)R_\alpha(\phi)W^{-1}(\phi),$$

where the matrices $W(\phi)$ and $R_\alpha(\phi)$ are given by

$$\begin{aligned} W(\phi) &= \left(\int_{-T}^{T} f_i(t)f_j(t)\phi(t)\,dt\right)_{i,j=1,\ldots,p}, \\ R_\alpha(\phi) &= \left(\int_{-T}^{T} f_i(t)f_j(t)Q_\alpha(1/\phi(t))\phi(t)\,dt\right)_{i,j=1,\ldots,p} \\ &= \frac{c}{1-\alpha}\left(\int_{-T}^{T} f_i(t)f_j(t)\phi^{1+\alpha}(t)\,dt\right)_{i,j=1,\ldots,p}, \end{aligned}$$

and we have used the representation $Q_\alpha(t) = c/((1-\alpha)|t|^\alpha)$ for the last identity. An (asymptotic) optimal design minimizes an appropriate function of the matrix $\Psi_\alpha(\phi)$ (for classical least squares estimation). Note that under long range dependence the variance-covariance matrix of the least squares estimate converges slower to zero than in the case of independent or short-range dependent errors. In the case of short-range dependence, no other normalization is necessary apart from normalizing the variance-covariance matrix. Under long-range dependence an additional factor $d_\alpha(N)/N$ is needed. Moreover, it is worthwhile to note that under long range dependence the asymptotic variance-covariance matrix is fully determined by the function $Q_\alpha(t)$ and does not otherwise depend on the particular correlation function $\rho_\alpha(t)$. In the following section we discuss several examples in order to illustrate the concept.

4 Examples

In most cases, the asymptotic optimal designs for the regression model (1.1) have to be found numerically; explicit solutions are only possible for very simple models. In this section we consider models with one or two parameters.

4.1 Optimal designs for linear models with one parameter

Consider the linear regression model with $p = 1$, that is $y(t) = \theta f(t) + \varepsilon(t)$ $(\theta \in \mathbb{R})$. In this case the problem of minimizing the asymptotic variance-covariance of the least squares estimate reduces to the minimization of the function

$$\Psi_\alpha(p) = \frac{\int f^2(t) Q_\alpha(1/p(t)) p(t)\, dt}{\left(\int f^2(t) p(t)\, dt\right)^2} \int p(t)\, dt$$

in the class of all nonnegative functions $p(t)$ on the interval $[-T, T]$. Because $Q_\alpha(t)$ is strictly convex on $(0, \infty)$ it follows from Theorem 3.1 in Bickel and Herzberg (1979) that a minimizer, say $p^*(t)$, exists and that $\phi^*(t) = p^*(t)/\int p^*(t)\, dt$ is the asymptotic optimal density. For the minimizing function p^* we obtain

$$\frac{\partial}{\partial \epsilon} \Psi_\alpha(p^* + \epsilon(p - p^*)) \Big|_{\epsilon=0} \geq 0$$

for all nonnegative functions p on the interval $[-T, T]$. Consequently the asymptotic optimal density should satisfy $\int p^*(t)\, dt = 1$ and

$$\int \Big(f^2(t) \big(H_\alpha(1/p^*(t)) - \mu \big) + \tilde{\tau} \Big) (p(t) - p^*(t))\, dt \geq 0 \tag{4.1}$$

for all nonnegative functions p on the interval $[-T, T]$, where the function $H_\alpha : (0, \infty) \to \mathbb{R}^+$ is given by

$$\begin{aligned} H_\alpha(t) &= Q_\alpha(t) - t Q'_\alpha(t) = \begin{cases} \frac{1+\alpha}{1-\alpha}/t^\alpha, & 0 < \alpha < 1 \\ 2/t & \alpha = 1 \end{cases} \\ \mu &= 2 \frac{\int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt}{\int f^2(t) p^*(t)\, dt}, \\ \tilde{\tau} &= \int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt. \end{aligned} \tag{4.2}$$

Assume first that $0 < \alpha < 1$. Note that the function H_α is strictly decreasing with $H_\alpha(+0) = \infty$, $H_\alpha(\infty) = 0$ and that its inverse is given by

$$H_\alpha^-(t) = \left(\frac{1 + \alpha}{t(1 - \alpha)} \right)^{1/\alpha}.$$

Hence the solution of (4.1) has the form

$$p^*(t) = \begin{cases} \dfrac{1}{H_\alpha^{-1}(\mu - \tau/f^2(t))} = \Big(\dfrac{1-\alpha}{1+\alpha}(\mu - \tau/f^2(t))\Big)^{\frac{1}{\alpha}}, & \mu - \tau/f^2(t) \geq 0, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3}$$

where μ is defined by (4.2) and

$$\tau = \int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt + \int f^2(t) Q'_\alpha(1/p^*(t))\, dt. \tag{4.4}$$

Note that τ is a solution of equation $\int p(t)\, dt = 1$. Indeed, multiplying $f^2(t) H_\alpha(1/p^*(t)) \equiv \mu f^2(t) - \tau$ with $p^*(t)$ and integrating with respect to t yields

$$\int f^2(t) H_\alpha(1/p^*(t)) p^*(t)\, dt = \int (\mu f^2(t) - \tau) p^*(t)\, dt = \mu \int f^2(t) p^*(t)\, dt - \tau$$

Now the definition of $H_\alpha(t)$ and μ gives

$$\int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt - \int f^2(t) Q'_\alpha(1/p^*(t))\, dt = 2 \int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt - \tau$$

which yields (4.4). Consequently we proved the following result.

Theorem 2. *Assume that the correlation function $\rho_\alpha(t)$ is either an element of the Cauchy, Mittag-Leffler or SVF family. Then, for the one-parameter linear regression model, the asymptotic optimal design exists, it is absolute continuous with respect to the Lebesgue measure and has the density $p^*(t)$ defined in* (4.3)*, where μ and τ are given by* (4.2) *and* (4.4)*, respectively.*

We now consider two special cases, which are of particular importance. If $p = 1$ and $f(t) \equiv 1$ we obtain the location model and the asymptotic optimal density is the uniform density, that is

$$p^*(t) = \begin{cases} \frac{1}{2T}, & |t| \leq T, \\ 0, & \text{otherwise.} \end{cases} \tag{4.5}$$

Similarly, in the linear regression through the origin we have $p = 1$, $f(t) \equiv t$, and the asymptotic optimal density is given by

$$p(t) = \begin{cases} 0, & |t| \leq \sqrt{\tau/\mu}, \\ \big(\frac{1-\alpha}{1+\alpha}(\mu - \tau/t^2)\big)^{1/\alpha}, & \sqrt{\tau/\mu} \leq |t| \leq T, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\mu = 2\frac{\int t^2 p^{1+\alpha}(t)\, dt}{(1-\alpha)\int t^2 p(t)\, dt}, \quad \tau = \int t^2 p^{1+\alpha}(t)\, dt,$$

and α is the parameter of the correlation function. The above formulas are given for $0 < \alpha < 1$. For $\alpha = 1$ and $f(t) = t$, the asymptotic optimal density is the uniform density (4.5). The optimal densities for the parameters $\alpha = 1/4, 1/2, 3/4, 0.95$ and $T = 1$ are displayed in Figure 2. The parameters μ and τ and the efficiency of uniform design are shown in Table 1. We observe that the uniform design is rather inefficient for small values of the parameter α. The uniform design has a reasonable efficiency only if α is close to 1.

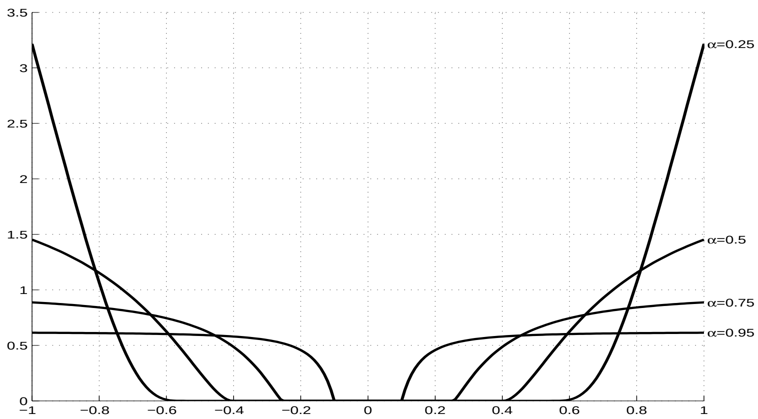

Figure 2: *Asymptotic optimal design densities for the linear regression through the origin,* $T = 1$.

Table 1: *Parameters of the asymptotic optimal design density for the linear regression through the origin and the efficiency of uniform design* (4.5)*,* $T = 1$*. The optimal density for* $\alpha = 1$ *is* $p(t) = 1/2$*,* $-1 \le t \le 1$.

α	μ	τ	$\sqrt{\tau/\mu}$	eff_{uni}
0.05	2.34	1.06	0.67	0.40
0.25	3.19	0.96	0.55	0.59
0.50	4.32	0.70	0.40	0.78
0.75	6.84	0.44	0.25	0.93
0.95	24.78	0.25	0.10	0.99

It is worthwhile to mention that the asymptotic optimal designs derived so far depend sensitively on the parameter α, which is usually not available before the experiment. Because misspecification of this parameter can yield to a substantial loss of efficiency of the optimal design we propose to construct robust designs, which are less sensitive with respect to such misspecifications. More precisely, we denote by $p^*_\alpha(t)$ the optimal density design for parameter α. Following Dette (1995) or Müller and Pázman (1998) a robust version of the optimality criterion is of the form

$$\Psi_{\mathcal{A}}(p) = \min_{\alpha \in \mathcal{A}} \mathrm{eff}(p, \alpha) = \min_{\alpha \in \mathcal{A}} \frac{\Psi_\alpha(p^*_\alpha)}{\Psi_\alpha(p)}$$

where p_α^* is the optimal design for the correlation function ρ_α and $\mathcal{A}$ is set of possible α values specified by the experimenter. A design maximizing $\Psi_\mathcal{A}$ is called standardized maximin optimal. Numerical optimization of this function for the set $\mathcal{A} = \{0.1, 0.2, \ldots, 0.9\}$ shows that standardized maximin optimal design has a density which can be approximated by the function

$$p_\mathcal{A}^*(t) = (5.7275t^2 - 1.16963 - 3.0264t^4)_+ \tag{4.6}$$

which is close to the optimal density p_α^* for $\alpha = 0.44$. In Table 2 we show the efficiency of this design for various values of α. We observe that the design $p_\mathcal{A}^*$ is very efficient for all elements in the set $\mathcal{A}$.

Table 2: *Efficiency of the standardized maximin optimal design* $p_\mathcal{A}^*$ *defined by* (4.6) *in the linear regression through the origin. The correlation structure is given by the SVF family with parameter* α*.*

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\text{eff}(p^*, \alpha)$	0.84	0.92	0.97	0.99	0.99	0.97	0.94	0.89	0.84

4.2 Linear regression

Consider the case $p = 2$, $f_1(t) = 1$, $f_2(t) = t$, which corresponds to the linear regression model. In this case the asymptotic variance-covariance matrix is proportional to

$$\Psi_\alpha(p) = \begin{pmatrix} 1 & \int tp(t)\,dt \\ \int tp(t)\,dt & \int t^2p(t)\,dt \end{pmatrix}^{-1} R(p) \begin{pmatrix} 1 & \int tp(t)\,dt \\ \int tp(t)\,dt & \int t^2p(t)\,dt \end{pmatrix}^{-1}$$

where

$$R(p) = \begin{pmatrix} \int Q_\alpha(1/p(t))p(t)\,dt & \int tQ_\alpha(1/p(t))p(t)\,dt \\ \int tQ_\alpha(1/p(t))p(t)\,dt & \int t^2Q_\alpha(1/p(t))p(t)\,dt \end{pmatrix}$$

For a symmetric density this matrix is diagonal and

$$\Psi_\alpha(p) = \text{diag}\left(\int Q_\alpha(1/p(t))p(t)\,dt, \frac{\int t^2Q_\alpha(1/p(t))p(t)\,dt}{\left(\int t^2p(t)\,dt\right)^2}\right).$$

Consequently the optimal design for estimating the slope in the linear regression has the density (4.3) where μ and τ are defined in (4.2) and (4.4) (this follows from the fact that the element in the position $(2, 2)$ of the matrix $\Psi_\alpha(p)$ corresponds to the optimality criterion for the linear regression through the origin).
The D-optimal designs for the linear regression model have to be determined numerically in all cases. Some D-optimal design densities corresponding to the parameters $\alpha = 1/4, 1/2, 3/4, 0.95$ and $T = 1$ are displayed in Figure 3.

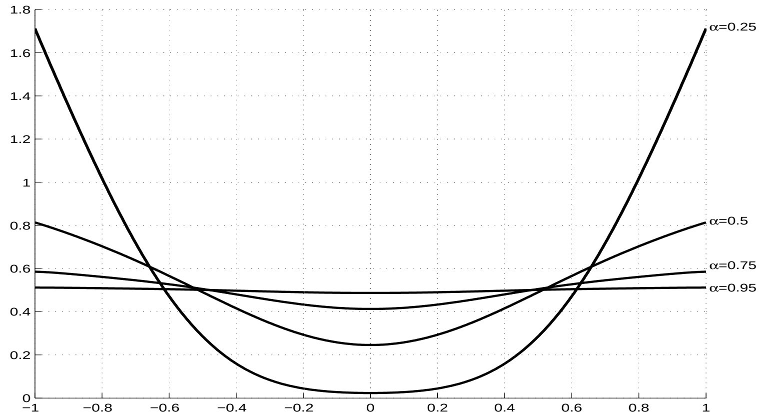

Figure 3: *Asymptotic D-optimal design densities for the linear regression,* $T = 1$.

4.3 Comparison of optimal designs under long and short range dependence

It is of some interest to compare the asymptotic optimal designs under short and long range dependence. For this purpose we consider again the linear regression model with no intercept. Bickel and Herzberg (1979) discussed the correlation function $\rho_\lambda(t) = e^{-\lambda|t|}$. The asymptotic optimal designs are given by

$$p(t) = \begin{cases} 0, & |t| \leq \sqrt{\tau/\mu}, \\ \frac{1}{H^{-}(\mu-\tau/t^2)}, & \sqrt{\tau/\mu} \leq |t| \leq T, \\ 0, & \text{otherwise}, \end{cases}$$

where the quantities μ, τ are defined by

$$\mu = \frac{1}{2\gamma} + 2\frac{\int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt}{\int f^2(t) p^*(t)\, dt},$$

$$\tau = \frac{1}{2\gamma} \int f^2(t) p^*(t)\, dt + \int f^2(t) Q_\alpha(1/p^*(t)) p^*(t)\, dt + \int f^2(t) Q'_\alpha(1/p^*(t))\, dt,$$

respectively, [see Bickel, Herzberg and Schilling (1981)] and depend on the parameters λ and γ defined in (2.2). Some of these designs are shown in Figure 4, while the relevant parameters are given in Table 3, which also contains the efficiency of the uniform design. We observe that - in contrast to the case of long range dependence - the uniform design is rather efficient provided that either the parameter λ is not too large or γ is not too small.
We now compare asymptotic optimal designs derived under the assumption of a long-range dependence with asymptotic optimal designs under short range dependence. In Table 4 we show the efficiency of a design derived under the assumption of short range dependence, in the situation, where the "true" correlation structure is a member of the

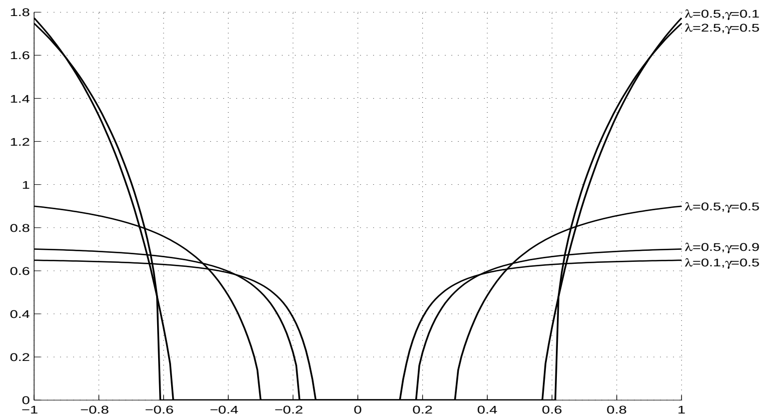

Figure 4: *Asymptotic optimal design densities for the linear regression through the origin, where the correlation function is given by* $\rho_\lambda(t) = e^{-\lambda|t|}$.

Table 3: *Parameters of the asymptotic optimal design density for the linear regression through the origin, where the correlation function is given by* $\rho_\lambda(t) = e^{-\lambda|t|}$. *The last column of the table shows the efficiency of the uniform design* (4.5).

λ	γ	μ	τ	$\sqrt{\tau/\mu}$	eff_{uni}
0.5	0.5	3.41	0.32	0.30	0.89
0.5	0.1	9.82	3.23	0.57	0.63
0.5	0.9	2.38	0.08	0.18	0.97
0.1	0.5	12.70	0.22	0.13	0.99
2.5	0.5	1.45	0.54	0.61	0.57

SVF family. We observe that the loss of efficiency is only substantial if the parameter α is small. The opposite situation is displayed in Table 5 which shows the efficiency of the asymptotic optimal design under long range dependence (from the the SVF family) but the "true" correlation structure is in fact of exponential type. Again, the asymptotic optimal designs derived under the long range dependence are rather efficient, except when the parameter α is very small.

Table 4: *Efficiency of the asymptotic optimal design density for the correlation function* $\rho_\lambda(t) = e^{-\lambda|t|}$ *in the linear regression through the origin, while the "true" correlation function belongs to the SVF family.*

	α	0.05	0.25	0.50	0.75	0.95
$\lambda=0.5$	$\gamma=0.5$	0.62	0.82	0.96	1.00	0.97
$\lambda=0.5$	$\gamma=0.1$	0.81	0.97	0.99	0.89	0.77
$\lambda=0.5$	$\gamma=0.9$	0.53	0.73	0.90	0.99	1.00
$\lambda=0.1$	$\gamma=0.5$	0.50	0.70	0.88	0.98	1.00
$\lambda=2.5$	$\gamma=0.5$	0.81	0.97	0.98	0.89	0.77

Table 5: *Efficiency of asymptotic optimal design density for a long range dependence error structure in the linear regression through the origin, while the "true" correlation function is given by* $\rho_\lambda(t) = e^{-\lambda|t|}$.

λ	0.5	0.5	0.5	0.1	2.5
γ	0.5	0.1	0.9	0.5	0.5
$\alpha=0.05$	0.19	0.40	0.15	0.15	0.35
$\alpha=0.25$	0.69	0.94	0.59	0.58	0.93
$\alpha=0.50$	0.94	0.98	0.87	0.86	0.98
$\alpha=0.75$	1.00	0.88	0.98	0.98	0.85
$\alpha=0.95$	0.95	0.73	0.99	1.00	0.68

Acknowledgements. The support of the Deutsche Forschungsgemeinschaft (SFB 823, "Statistik nichtlinearer dynamischer Prozesse ") is gratefully acknowledged. The work of H. Dette was also supported in part by an NIH grant award IR01GM072876:01A1 and by the BMBF project SKAVOE.

References

Anh V. V., Knopova V. P., Leonenko N. N. Continuous-time stochastic processes with cyclical long-range dependence. Aust. N. Z. J. Stat. 46, (2004), no. 2, 275–296.

Atkinson A.C., Donev A.N. *Optimum experimental designs.* Clarendon Press, 1992.

Barndorff-Nielsen O. E., Jensen J. L., Sorensen M. Parametric Modelling of Turbulence. Philosophical Transactions: Physical Sciences and Engineering, Vol. 332, No. 1627, (1990), 439–455.

Barndorff-Nielsen O. E., Leonenko N. N. Spectral properties of superpositions of Ornstein-Uhlenbeck type processes. Methodol. Comput. Appl. Probab. 7, (2005), no. 3, 335–352.

Beran J. Statistical methods for data with long-range dependence. Statist. Sci. 7(4) (1992), 404-427.

Beran J. *Statistics for Long-Memory Processes.* Chapman and Hall, 1994.

Beran J., Sherman R., Taqqu M.S., Willinger W. Variable-bit-rate video traffic and long-range dependence. Unpublished manuscript. (1992).

Bickel P.J., Herzberg A.M. Robustness of design against autocorrelation in time. I. Asymptotic theory, optimality for location and linear regression. Ann. Statist. 7 (1979), no. 1, 77–95.

Bickel P.J., Herzberg, A.M., Schilling M.F. Robustness of design against autocorrelation in time. II. Optimality, theoretical and numerical results for the first-order autoregressive process. J. Amer. Statist. Assoc. 76 (1981), no. 376, 870–877.

Dette H. Designing of experiments with respect to "standardized" optimality criteria. Journal of the Royal Statistical Society, Ser. B, 59, (1995). 97–110.

Djrbashian M. M. *Harmonic analysis and boundary value problems in the complex domain.* Birkhauser Verlag, 1993.

Doukhan P., Oppenheim G., Taqqu M. (eds), *Theory and Applications of Long-Range Dependence.* Birkhauser Press, 2003.

Fedorov V. V. *Theory of optimal experiments.* Academic Press, New York, 1972.

Gneiting T. Power-law correlations, related models for long-range dependence and their simulation. J. Appl. Probab. 37, (2000), no. 4, 1104–1109.

Granger C. W. J. Long memory relationships and the aggregation of dynamic models. J. Econometrics 14, (1980), no. 2, 227–238.

Leonenko N. N., Woyczynski W. A. Parameter identification for singular random fields arising in Burgers' turbulence. J. Statist. Plann. Inf. 80, (1999), no. 1-2, 1–13.

Mandelbrot B.B. Le problème de la rétalité des cycle lents et le syndrome de Joseph. Economie Appliquéee, (1973), 26, 349–365.

Metzler R., Barkai E., Klafter J. Anomalous diffusion and relaxation close to thermal equilibrium : a fractional fokker-planck equation approach. Phys. Rev. Lett., v. 82, n. 18, (1999), 3563–3567.

Metzler R., Klafter J. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Rep. 339 (2000), no. 1, 1–77.

Pázman A., Müller, C. Applications of necessary and sufficient conditions for maximum efficient design. Metrika 48, (1998), 1–19.

Pázman A. *Foundations of optimum experimental design.* Reidel, Dordrecht, 1986.

Pearson K. On the mathematical theory of errors of judgement, with special reference to the personal equation. Philos. Trans. Roy. Soc. Ser. A (1902), 198, 235–299.

Porter-Hudak S. An application of the seasonal fractionally differenced model to the monetary aggregates. J. Amer. Statist. Assoc. (1990), 85, 338–344.

Pukelsheim F. *Optimal design of experiments.* John Wiley & Sons Inc., New York, 1993.

Sacks J., Ylvisaker N.D. Designs for regression problems with correlated errors. Ann. Math. Statist. 37, (1966), 66–89.

Sacks J., Ylvisaker N.D. Designs for regression problems with correlated errors; many parameters. Ann. Math. Statist. 39, (1968), 49–69.

Seneta E. *Regularly varying functions.* Lecture Notes in Mathematics, Vol. 508. Springer-Verlag, 1976.

Schneider W. R. Completely monotone generalized Mittag-Leffler functions. Exposition. Math. 14, (1996), no. 1, 3–16.

Silvey S. D. *Optimal design.* Chapman & Hall, 1980.

Smith H.F. An empirical law describing heterogeneity in the yields of agricultural crops. Journal of Agricultural Science, (1938), 28, 1–23.