Department for Applied Statistics
Johannes Kepler University Linz

IFAS Research Paper Series
2011-58

Filling and D-optimal designs for the correlated Generalized Exponential models

Juan M. Rodríguez-Díaz[a], Teresa Santos-Martín[a],
Helmut Waldl and Milan Stehlík

August 2011

[a] Department of Statistics, University of Salamanca, Spain

*Manuscript
Click here to view linked References

Filling and D-optimal designs for the correlated Generalized Exponential models

Rodríguez-Díaz, J. M.[a,*], Santos-Martín M.T.[a], Waldl, H.[b], Stehlík, M.[b]

[a] *Department of Statistics, University of Salamanca, Spain*
[b] *Department of Applied Statistics,Johannes Kepler University in Linz*

Abstract

The aim of this paper is to provide guidelines for the statistically efficient estimation of parameters of modified Arrhenius model for chemical kinetics. A modified Arrhenius model is used for instance by modeling a flux of methane in troposphere or by chemical kinetics for reactions at membranes. D-optimal and filling designs for the Generalized Exponential Model with correlated observations are studied, considering the exponential covariance with or without nugget effect. Both equidistant and exact designs for small samples are examined, studying the behavior of different types of filling designs when a greater number of observations is preferred. Probably the main lesson we can learn is that the D-optimal design is analytically peculiar and these designs can be practically obtained only by numerical computation; however, specially two point locally D-optimal designs are very interesting, since they may help us to find a reasonable range for filling designs. The latter ones are probably the only applicable when seeking for a higher number of design points. It is an interesting issue that very often the best designs do not use the whole design interval, but only a part of it; this should be taken into account by practitioners when they design their experiments. The second important observation is the large bias of the ML estimator of r. From the theoretical point of view this is not surprising since σ^2 and r are not simultaneously identifiable. We develop a bias reduction method and illustrate its effectiveness. We also provide practical implications for chemometrics.

Keywords: Arrhenius model, bias reduction, correlated observations, exponential model, filling designs, optimal design of experiments.

1. Introduction

The efficient estimation of the parameters of the modified Arrhenius model, quite popular in chemical kinetics, is the objective of the present work. A modified Arrhenius model is used for instance by modeling a flux of methane in troposphere Vaghjiani and Ravishankara (1991) or by chemical kinetics for reactions at membranes. Since troposphere, extensively studied because of greenhouse gas emission, provides a very exotic environment for kinetics of chemical reactions normally studied at the Earth, it can be argued that correlation parameter and modified Arrhenius

*Corresponding author. Faculty of Sciences. Plaza de los Caídos. 37008 Salamanca, Spain. Tel.:+34923294458; Fax: +34923294514
Email addresses: juanmrod@usal.es (Rodríguez-Díaz, J. M.), maysam@usal.es (Santos-Martín M.T.), helmut.waldl@jku.at (Waldl, H.), Milan.Stehlik@jku.at (Stehlík, M.)

equation can play a crucial role by its modeling. The Generalized Exponential (GE) model can be expressed as

$$Y = ax^m e^{-\beta x} + \epsilon = \eta(x, m, \beta) + \epsilon, \tag{1}$$

where a, $\beta \geq 0$, m are constants. Such a model has been studied for the case of uncorrelated errors by Dette and Sperlich (1994) from a Bayesian point of view, and by Rodríguez-Díaz and Santos-Martín (2009) for different efficiency functions, optimality criteria and restrictions on the design space. This last paper is devoted to the Modified-Arrhenius (MA) model, which is equivalent to the GE model through the change of variable X=1/T. The model has applications in experimental sciences, specially in chemical kinetic. There appears as a good alternative (in fact a generalization) to the widely used Arrhenius model to describe the influence of temperature on the rates of chemical processes, specially when the temperature range is not so small Laidler (1984). The influence of temperature on the rates of processes is usually given by terms of the Arrhenius equation for which optimal designs are studied in Rodríguez-Aragón and López-Fidalgo (2005). However, in some cases the Arrhenius model does not seem to be enough to describe the experiment, and the Modified Arrhenius (or GE model) appears as the right choice (see for instance Gierczak et al, 1997). The GE model belongs to the so called class of partially nonlinear regression models (see Hill, 1980). This particularly means that the D-optimal designs does not depend on the value of a, which is the linear parameter of the trend. An exact design will be a collection of points of the independent variable, $\{x_1, \ldots, x_n\}$, where n is the size of the design.

Other applications of model (1) in chemistry are related to the transition state theory (TST) of chemical reactions, IUPAC (2008). The Arrhenius like expressions appearing in TST take various forms, e.g. $k = K_B T/h \exp(-\delta G/(RT))$, where δG is the Gibbs free energy of activation, K_B is the Boltzman's constant and h is Planck's constant.

In practical chemical kinetic assessment people work in two steps: first the rates k_1 are estimated (typically with symmetric estimated error) and then modified Arrhenius model is fitted to the rates, $k_1 = a(1/T)^m e^{-\beta/T} + \epsilon$. Statistically correct would be to assess both steps by one optimal experimental planning. Throughout this paper we concentrate on the second phase: how to optimally distribute temperatures in order to obtain statistically efficient estimators of trend parameters a, β, m and correlation parameter r. Even the second step could be produced sequentially, at first step fitting the range of temperatures, for which 2 (or more)-point designs can be employed, designs that are studied in the following section. After fitting the minimum and maximum of our temperature range one can be interested in particular temperatures which shall be used for measurements. Here filling design would be useful, and reader can concentrate on Section 3. Optimal designing is typically studied by locally optimal designs, and thus is highly dependent on initial values of parameters m, β and r; therefore care should be taken on their guessing. In chemical kinetics trend parameters are the most important, but practitioners will rarely use 2- or 3-point designs. But these small point situations can be taken at least for statistical learning issues or as a benchmarks for locating the design interval.

In the present paper the case of correlated observations will be specially considered. Correlation is the natural dependence measure fitting for elliptically symmetric distributions (e.g. Gaussian). Two covariance structures will be studied:

Cov1) exponential covariance function $C(T_1, T_2, r) = \sigma^2 \exp(-dr), r > 0$ where r is the covariance parameter and $d = |T_1 - T_2|$ is the distance between temperatures.

Cov2) exponential with inverted arguments $C(T_1, T_2, r) = \sigma^2 \exp(-|\frac{1}{T_1} - \frac{1}{T_2}|r)$, $r > 0$.

The reason of using Cov2) is to avoid the ill conditioning of the covariance matrix for some situations. In general, for temperature ranges $[A, B]$ such that $0 <<< A < B$ we consider the

classical covariance structure Cov1).

A deeper discussion on the identifiability of the covariance parameters can be found in Müller and Stehlík (2009), where the role of the nugget effect is also discussed. It is well known that parameters r and σ^2 are not individually identifiable, which is observed also by simulation studies made in this work. The maximal likelihood estimators suffer from severe bias.

We center on D-optimality, which corresponds to the maximization of criterion function $\Phi(M) = \det M$, the determinant of the Fisher information matrix. In our setup θ stands for the trend parameters and r is the covariance parameter. $M_\theta(n)$ is the Fisher information matrix

$$\left(\frac{\partial \eta(x,\theta)}{\partial \theta_i} \frac{\partial \eta(x,\theta)}{\partial \theta_j} \right)_{i,j}$$

while

$$M_r(n) = \frac{1}{2} tr \left\{ C^{-1}(r) \frac{\partial C(r)}{\partial r} C^{-1}(r) \frac{\partial C(r)}{\partial r^T} \right\}$$

(see Pázman, 2007). For both parameters of interest we have

$$M(n)(\theta, r) = \begin{pmatrix} M_\theta(n) & 0 \\ 0 & M_r(n) \end{pmatrix}.$$

This method is "plugged" from the widely developed uncorrelated setup and further development is needed before it can be applied routinely in practice. Theoretical justifications for using the Fisher information for D-optimal design in the correlated setup can be found in Abt and Welch (1998) and Pázman (2007). For more references on the use of the Fisher information as design criterion in the correlated setup see e.g. Stehlík (2007), where the structure of the Fisher information matrices for a stationary process is studied.

The aim of this paper is in first place to provide a thorough study of small sample and asymptotical comparisons of the efficiencies and the impact of the correlation in the design (Section 2). We consider all possible combinations of parameters of interest, i.e. a plus β, $\{m, \beta\}$, $\{r, \beta\}$, and $\{m, \beta, r\}$; and center on exact designs for small samples, both equidistant and non-equidistant. There the analytical peculiarities of the exact design are studied, beginning with two-point designs, which play a crucial role for finding the range for filling designs. The behavior of $M_r(n)$ is studied as well. The explicit solutions are typically not available and numerical techniques should be used in order to obtain the D-optimal exact designs. However, some properties arise from the analysis of implicit solutions, and we can learn that correlation parameters play more influential role than trend parameters. This was also observed in the case of OU process (see Kiseľák and Stehlík, 2008) in an even more dramatic way. The other interesting fact is that the properties are mainly dependent through the quantity rd, e.g. the distance multiplied by the covariance parameter, what has been observed also in Kiseľák and Stehlík (2008) and Stehlík (2007). Practitioners usually prefer to take observations in a greater number of points (more than three) and thus, from Section 3 on, *filling designs* covering a chosen interval with several points are studied and their behavior compared through convenient examples. It can be seen that the upper bounds of temperature intervals are present in the support of optimal designs, which encompasses the natural fact that reaction kinetics for higher temperatures are speeder. In Section 4 the atmospheric lifetime of methane is taken as an example for studying the designs developed through the previous sections. Finally Section 5 introduces the bias reduction method for parameter r. To maintain the continuity of the explanation, proofs and technicalities are deferred to the Appendix.

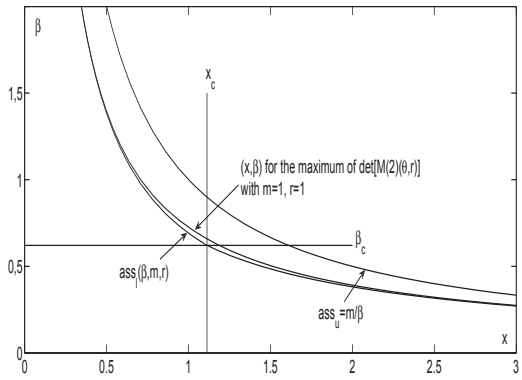

Figure 1: x and β for the maximum of $\det[M(2)(\theta, r)]$ for $m = 1$ and $r = 1$

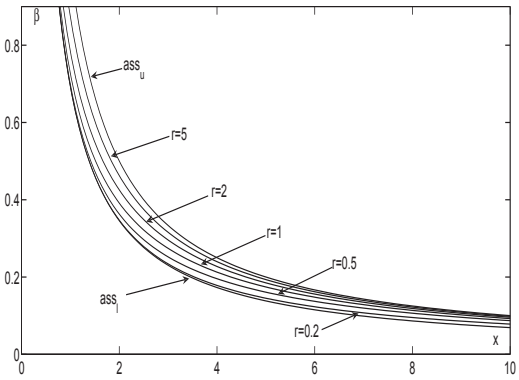

Figure 2: x and β for the maximum of $\det[M(2)(\theta, r)]$ for $m = 1$ and different values of r

2. Exact Designs

In this section exact designs with 2 or 3 support points, equidistant and non-equidistant, will be studied considering the cases of m being known and unknown.

2.1. Assuming m known

Some authors (see the discussion in Héberger et al (1987) advise to consider m known; in this case the only trend parameters are a and β. The non-correlated case has been already studied in Rodríguez-Díaz and Santos-Martín (2009), where it has been proved that for approximated designs a 2-point design is optimal. The covariance structure Cov1 corresponds to the Ornstein-Uhlenbeck process, for which $M_r(n)$ is known, (see Müller and Stehlík (2010) and Zagoraiou and Baldi-Antognini, 2009): $M_r(n) = \sum_{i=1}^{n-1} d_i^2(e^{2rd_i}+1)/(e^{2rd_i}-1)^2$, where $d_i = |T_{i+1} - T_i|$ are distances between neighboring temperatures. For $r \to 0+$ we have $M_r(n) \to \infty$ which encompasses the fact that neighboring points are important for efficient estimation of the correlation parameter (see e.g. Kiseľák and Stehlík, 2008).

Now let us consider two-point designs with the covariance Cov2, i.e. let $n = 2$, and $\{x, x+d\}$ be the design. The following Theorem, obtained by direct algebra, provides guidelines about how to compute two-point D-optimal designs.

Theorem 1. *The exact two-point D-optimal design for the GE model assuming Cov2 is* $\xi_2 = \{x, x+d\}$, *where* $x = (-\beta d + m + \sqrt{\beta^2 d^2 + m^2})/(2\beta)$ *and* d *is the zero of*

$$\beta^2\left(4^{2m+1}de^{2dr}r + 16^m dr + 2^{4m+1} + e^{4dr}\left(16^m dr - 2^{4m+1} + 16^m\sqrt{\beta^2d^2+m^2}\right) - 16^m\sqrt{\beta^2d^2+m^2}\right)d^2$$

$$+m\left(m+\sqrt{\beta^2d^2+m^2}\right)\left(4^{2m+1}de^{2dr}r + 16^m dr + 2^{4m+1} + e^{4dr}\left(16^m dr - 2^{4m+1}\right)\right) \quad (2)$$

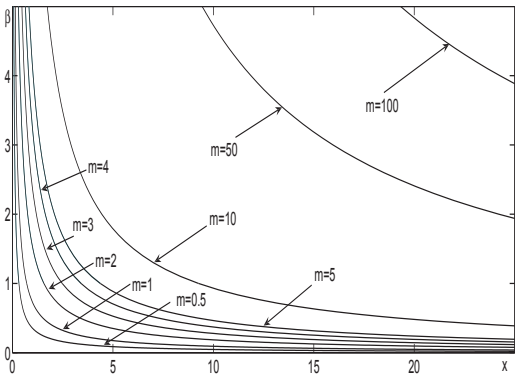

Figure 3: x and β for the maximum of $\det[M(2)(\theta, r)]$ for $r = 1$ and different values of m

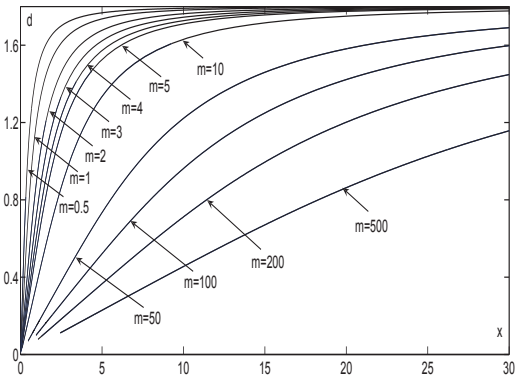

Figure 4: Solutions d for $r = 1$ and different values of m

For ξ_2 we have $d = 2x(m - x\beta)/(2x\beta - m)$; $m, \beta, r > 0$; $m/2\beta < x < m/\beta$. For all m and r the unique solution of x depending on β is bounded by the asymptotes

$$ass_l = \begin{cases} \frac{1+4m-\sqrt{1+4m}}{4\beta} & \text{for } \beta_c \leq \beta \\ \frac{m-D\beta+\sqrt{m^2+D^2\beta^2}}{2\beta} & \text{for } \beta \leq \beta_c \end{cases} \qquad ass_u = \frac{m}{\beta},$$

where $\beta_c = (1+4m-\sqrt{1+4m})/(2D(\sqrt{1+4m}-1))$ and $D = 1.8006/r$. The lower bound asymptote ass_l is already a good approximation for the exact solution, and for small x-values is very similar to the exact solution for the model with uncorrelated data analyzed in Rodríguez-Díaz and Santos-Martín (2009), where $x = (1 + 2m - \sqrt{1+2m})/(2\beta)$. Furthermore we have $\lim_{x\to\infty} d = D$. On the other hand, the design given by $(d, x) = (2ass_l(m - ass_l\beta)/(2ass_l\beta - m), ass_l)$ guarantees a minimum efficiency of 92% Rodríguez-Díaz et al. (2009).

The case of 2-point optimal designs with covariance Cov2 when only a and β are parameters of interest has been studied in Rodríguez-Díaz et al. (2009). The results for 3-point designs with correlation led to complicated expressions for both the general and even the equidistant cases.

Remark 1. It has already been proved that $x \leq m/\beta$. It is not difficult to check that the second point, $x + d$, verifies $x + d \geq m/\beta$, the same property noted in Rodríguez-Díaz and Santos-Martín (2009). Figure 1 shows the solution and the asymptotes for $m = 1$ and $r = 1$.

Changing the value of r only make the solutions (x, β) change slightly. The solution for x grows with r as we fix m and β, what is shown in Figure 2. Solutions for x with $r = 1$ fixed and different values of m and β can be seen in Figure 3.

Finally, in order to find the solutions for d we only need to plug in the solutions for x in $d = 2x(m - x\beta)/(2x\beta - m)$. Here we can observe that the solutions for d converge to the fixed value D as x increases, this limit only depending on r, $\lim_{x\to\infty} d = D = 1.8006/r$. The convergence rate grows with r and is also dependent on m as can be seen in Figure 4, and dr is bounded in the interval $(0, 1.8006)$.

The exact solution for the maximum of $\det[M(2)(\theta, r)]$ can be found by solving the equation $\partial \det[M(2)(\theta, r)]/\partial x = 0$ for β, which produces $\beta = m(d + 2x)/(2x(d + x))$. Then replacing β in

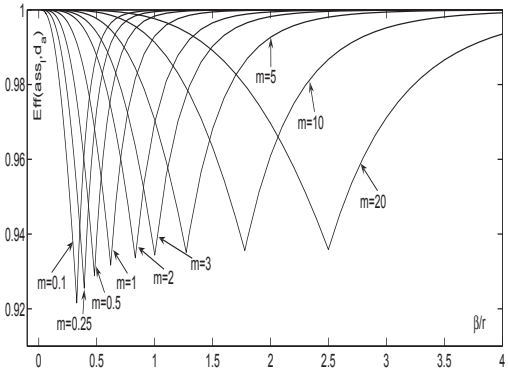

Figure 5: Efficiency of the approximation (ass_l, d_a) for(x, d)

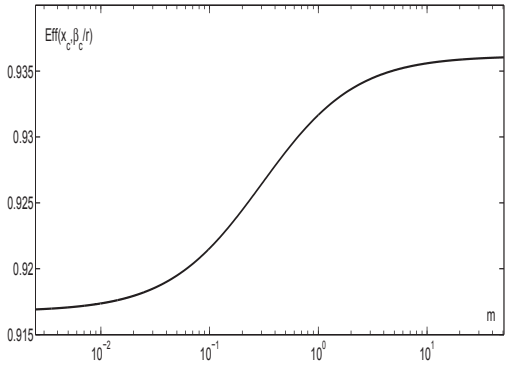

Figure 6: minimal efficiency for given m (log-scale)

$0 = \partial \det[M(2)(\theta, r)]/\partial d$ with this solution we get

$$x(d, m, r) = \frac{d}{2}\left(\sqrt{1 + \frac{4m}{dr(\tanh(dr) - 3\coth(dr)) + 4}} - 1\right)$$

$$\beta(d, m, r) = \frac{1}{2d}\left(dr(\tanh(dr) - 3\coth(dr)) + 4\right)\sqrt{1 + \frac{4m}{dr(\tanh(dr) - 3\coth(dr)) + 4}}$$

So we have found a parametrization of (x, β) for the maximum of $\det[M(2)(\theta, r)]$ with d, m and r as parameters. However we would like to have $(x(\beta, m, r), d(\beta, m, r))$, which could be achieved as follows: given β, m and r we can compute $ass_l(\beta, m, r)$ as an approximation of $x(\beta, m, r)$. Substituting ass_l for x in $d = 2x(m - x\beta)/(2x\beta - m)$ we get an approximation d_a of $d(\beta, m, r)$. We can then compute the exact values $(x_a(d_a, m, r), \beta_a(d_a, m, r))$ for the maximum of $\det[M(2)(\theta, r)]$. From this method we get the exact solution for β_a, that does not differ very much from the given β.

Very often the exact solution is not essential but a good approximation suffices in the majority of cases. In our problem the quality of the approximation can be measured with the efficiency

$$Eff(x, d) = \left(\frac{\det[M(2)(\theta, r|x, d)]}{\det[M(2)(\theta, r|x^*, d^*)]}\right)^{1/2}$$

where x^* and d^* are the exact solutions for the D-optimal design. The efficiency $Eff(ass_l, d_a)$ is dependent on m and the ratio β/r as can be seen in Figure 5. The minimal efficiency for fixed β/r is only dependent on m and is shown in Figure 6. It can be seen that the efficiency is always greater than 92 % which should be enough for practical purposes.

For $m = 0$ it is already known Rodríguez-Díaz and Santos-Martín (2009) that the first design point has to be the lower bound of the design space. This can be seen when computing

$$f(h) = \frac{\det[M(2)(\beta, r|x + h, d)]}{\det[M(2)(\beta, r|x, d)]} \quad ;$$

if $m = 0$ the function is $f(h) = e^{-4\beta h}$ and hence for $\beta \neq 0$ we have $f(h) < 1$ when $h > 0$, and $f(h)$ decreasing in h. Thus the maximal $\det[M(2)(\theta, r|x, d)]$ is obtained when taking the smallest x as design point.

For $m > 0$

$$f(h) = e^{-4\beta h} \left(\frac{(x+h)(x+h+d)}{x(x+d)} \right)^{2m},$$

and $\partial f(h)/\partial h > 0$ for $h = 0$.

Particularly, the interesting feature of this setup is that there exist a positive $d^\star$ that maximizes $\det[M(2)(\beta, r)]$. This behavior is different than the one observed by the OU process studied in Kiseľák and Stehlík (2008), Theorem 2. Therein is observed that when both parameters $\{\theta, r\}$ need to be estimated there exist D-optimal equidistant designs with finite d when the number of design points n is greater than 3, but for $n = 2, 3$ the collapsing should be compensated by a so-called nugget effect (see Stehlík et al., 2007), which decreases the smoothness of the process. The other differences are obtained by the evaluation of $\Phi_3(M)/\Phi_2(M)$ under the equidistant designs, where $\Phi_k(M)$ is the determinant of the Fisher information matrix for both parameters and the k-point equidistant design. When $r \to 0$, the correlation coefficient is $e^{-rd} = 1$ and $\Phi_3(M)/\Phi_2(M) \to \infty$. If $r \to \infty$, the correlation coefficient is $e^{-rd} = 0$ and $\lim_{r\to\infty} \Phi_3(M)/\Phi_2(M) > 2$.

In order to find the solutions for x and d it could be useful to study $M_\theta(2)$ alone, i.e. to maximize $\det[M_\theta(2)]$ for the correlated data. Assuming $a = \sigma = 1$ we have

$$\det[M_\theta(2)] = \frac{d^2 \exp(2dr - 2(d + 2x)\beta) x^{2m} (x + d)^{2m}}{\exp(2dr) - 1}$$

Also here when solving $0 = \partial \det[M_\theta(2)]/\partial x$ for d we get $d = 2x(m - x\beta)/(2x\beta - m)$; $m, \beta, r > 0$ $m/(2\beta) < x < m/\beta$. Replacing d in $0 = \partial \det[M_\theta(2)]/\partial d$ we finally need to solve

$$[(x\beta - m)(2(x\beta - m) - 1) - x\beta] \left[\exp\left(\frac{4rx(x\beta - m)}{m - 2x\beta} \right) - 1 \right] - 2rx(x\beta - m) = 0$$

for x, solution that can be found numerically. In Figure 7 this solution is compared with solutions for the maximum of $\det[M(2)(\theta, r)]$. It can be seen that the maximum of $\det[M_\theta(2)]$ is far away from the maximum of $\det[M(2)(\theta, r)]$, it is even smaller than the lower bound for the maximum of $\det[M(2)(\theta, r)]$.

The case when only a and β are parameters of interest and $m = 0$ was studied for exact designs and small number of points by Pepelyshev (2007). As is shown in Dette et al. (2008) $M_\beta(n)$ can be expressed in the form $X^T V^T V X$ where V is a 2-diagonal matrix given by $v_{i,i} = t_i, v_{i,i-1} = -s_i, t_1 = 1, s_1 = 0, X^T = (f(x_1), ..., f(x_n)), f(x) = (\partial\eta/\partial a, \partial\eta/\partial\beta)^T$ and

$$t_i = 1/\sqrt{1 - e^{-2r(x_i - x_{i-1})}}, \quad s_i = e^{-r(x_i - x_{i-1})} \cdot t_i, \quad i = 2, ..., n.$$

From the Cauchy-Binet formula we obtain for $m = 0$ the following expression for the determinant of the information matrix (see also Pepelyshev (2007), Dette et al., 2008)

$$\det M(a, \beta, r) = a^2 \sum_{1 \le i < j \le n} e^{-2\beta(x_i + x_j)} \psi^2(x_i, x_{i-1}, x_j, x_{j-1}),$$

where

$$\psi(x_i, x_{i-1}, x_j, x_{j-1}) = \frac{(1 - e^{-(r-\beta)d_i})(t_j - t_{j-1}e^{-(r-\beta)d_j}) - (1 - e^{-(r-\beta)d_j})(t_i - t_{i-1}e^{-(r-\beta)d_i})}{(1 - e^{-2rd_i})(1 - e^{-2rd_j})}$$

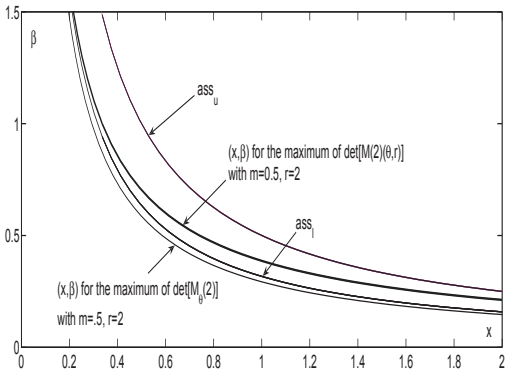

Figure 7: x and β for the maximum of $\det[M_\beta(2)]$ for $r = 2$ and $m = 0.5$

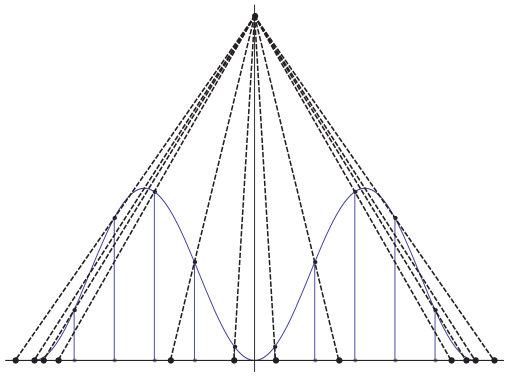

Figure 8: Construction of Trigonometric Designs when using the function $f(x) = \cos[s(x - \pi + v)] + 1$ for transformation of the initial points

with $d_i = x_i - x_{i-1}$.

Again using the Cauchy-Binet formula we obtain for $m \geq 0$ the expression for the determinant of the information matrix M_θ,:

$$\det M_\theta(n) = a^2 \sum_{1 \leq i < j \leq n} \frac{e^{-2\beta(x_i + x_j)}}{(1 - e^{-2rd_i})(1 - e^{-2rd_j})} \psi^2(x_i, x_{i-1}, x_j, x_{j-1}),$$

where $\psi(x_i, x_{i-1}, x_j, x_{j-1})$ has the form

$$e^{(\beta-r)d_j} x_i^m x_{j-1}^m (x_{j-1} - x_i) - x_i^m x_j^m (x_j - x_i) + e^{(\beta-r)d_i} x_{i-1}^m x_j^m (x_j - x_{i-1}) - e^{(\beta-r)(d_i+d_j)} x_{i-1}^m x_{j-1}^m (x_{j-1} - x_{i-1})$$

with $d_i = t_i - t_{i-1}, \quad x_0 = 0$ and $d_1 = \infty$

More details can be found in Rodríguez-Díaz et al. (2009).

2.2. Case m Unknown

Let us now study the case with m unknown, that needs to be estimated as well. Thus we have a three-parameter model.

2.2.1. Non-Correlated errors

The uncorrelated case will need at least three points for the design not being singular. Equidistant and non-equidistant designs will be analyzed and compared with previous results.

Let n=3 and consider the equidistant design $\{x, x + d, x + 2d\}$ when the parameters to be estimated are a, β, m for the uncorrelated setup.

PROOF. It is possible to obtain an expression for x as a function of d (or for d as a function of x) by solving the equality formed with the right parts of these previous equations:

$$x = \frac{3m - 3\beta d + 1}{3\beta} \qquad \left(\text{or } d = \frac{3m - 3\beta x + 1}{3\beta}\right) \quad .$$

For the case $x = (3m - 3\beta d + 1)/(3\beta)$ the value of the determinant is maximized when d is the solution of the equation

$$18\beta^2 d^2 + \left(9\beta^2 d^2(2m+1) - (3m+1)^2\right) \log\left(1 - \frac{9\beta^2 d^2}{(3m+1)^2}\right) = 0$$

with $0 < d < (1+3m)/(3\beta)$.

If we use $d = (3m - 3\beta x + 1)/(3\beta)$, the x giving the maximum determinant is the solution of the equation

$$2(3m-3\beta x+1)^2 + \left(2m(3m+1)^2 + 9\beta^2(2m+1)x^2 - 6\beta(m(6m+5)+1)x\right) \log\left(\frac{3\beta x(6m-3\beta x+2)}{(3m+1)^2}\right) = 0$$

with $0 < x < (2+6m)/(3\beta)$.

Example: If $m = \beta = a = 1$ the solutions are $x = 0.20689$ and $d = 1.12644$.

The case of non-equidistant designs is again troublesome. Details can be found in Rodríguez-Díaz et al. (2009).

2.2.2. Correlated Case

Now we will study the more complex case that assumes a non-trivial covariance structure. Let us take $n = 3$ and the equidistant design $\{x, x+d, x+2d\}$ when the parameters to be estimated are a, β, m. In this case the determinant of the information matrix will be

$$D_c = D_s \frac{2d^2 e^{4dr}(1+e^{2dr})}{(e^{2dr}-1)^4} \quad ,$$

where D_s is the determinant which does not depend on r.

The factor $f(d,r) = 2d^2 e^{4dr}(1+e^{2dr})/(e^{2dr}-1)^4$ verify $f(d,r) \to \infty$ when $r \to 0$ and $f(d,r) \to 0$ when $r \to \infty$.

Thus, the determinant can be expressed as

$$D_c = \frac{2a^4 d^4 e^{4dr-6\beta(d+x)} \left(1+e^{2dr}\right) x^{2m}(d+x)^{2m}(2d+x)^{2m} \log^2\left(\frac{x(2d+x)}{(d+x)^2}\right)}{\left(-1+e^{2dr}\right)^4}$$

After some algebra we get x as a function of d (and the rest):

$$x = \frac{-3\beta e^{4dr} d + 3\beta d - 5e^{2dr} rd - e^{4dr} rd - 2rd + 2e^{4dr} + 3e^{4dr} m - 3m - 2}{3\beta\left(-1+e^{4dr}\right)} \quad .$$

In order to obtain d we could proceed following one of the two ways:

- After knowing the initial values of β, m and r we get x as a function of d, and replacing it in the determinant we have this one as a function of d. We could compute d by solving $0 = \partial Det/\partial d$
- Replacing the value of x in one of the previous equations we get one equation on d (after replacing the inital values of β, m and r):

$$\log\left(\frac{\left(3m+5de^{2dr}r+2dr+e^{4dr}(-3m+dr-2)+2\right)^2 - 9\beta^2 d^2\left(-1+e^{4dr}\right)^2}{\left(3m+5de^{2dr}r+2dr+e^{4dr}(-3m+dr-2)+2\right)^2}\right) =$$

$$= 18\beta^2 d^2 \left(-1+e^{4dr}\right)^3 / \left[9\beta^2 d^2 \left(5de^{2dr}r + e^{4dr}(-2m+dr-2) + 2(m+dr+1)\right)\left(-1+e^{4dr}\right)^2 \right.$$
$$\left. + \left(-5de^{2dr}r - 2dr - e^{4dr}(dr-2) - 2\right)\left(3m + 5de^{2dr}r + 2dr + e^{4dr}(-3m+dr-2) + 2\right)^2\right]$$

And after knowing d and the initial values of β, m and r we could get x.

Example 1. For $a = \beta = m = 1$ and $r = 0.5$ the solution is $x = 0.205165$ and $d = 0.854672$.

A general non-equidistant 3-point design will be $\{x, x+d_1, x+d_1+d_2\}$. Following the same steps that in the previous sections, the determinant can be expressed as

$$Det = \frac{a^4 e^{2r(d_1+d_2)-2\beta(3x+2d_1+d_2)} x^{2m}(x+d_1)^{2m}(x+d_1+d_2)^{2m}}{\left(-1+e^{2rd_1}\right)^3\left(-1+e^{2rd_2}\right)^3}$$

$$\left[\left(1+e^{2rd_1}\right)\left(-1+e^{2rd_2}\right)^2 d_1^2 + \left(-1+e^{2rd_1}\right)^2\left(1+e^{2rd_2}\right)d_2^2\right]\log^2\left[\frac{x^{d_2}(x+d_1+d_2)^{d_1}}{(x+d_1)^{(d_1+d_2)}}\right]$$

Making the derivative respect to x equal to 0 and getting rid of the non-zero factors we get a very complicated expression which can be found, together further analysis, in Rodríguez-Díaz et al. (2009).

3. Filling Designs

The name *filling designs* refers to different types of designs that cover the design space with a specific number of points including both extremes. In previous works López-Fidalgo and Wong (2002) study several alternatives with the name *sequence designs* for the Michaelis-Menten model, with the distances between consecutive points increasing from the beginning of the interval. Later on, Rodríguez-Aragón and López-Fidalgo (2005) use them for the Arrhenius model, with the distances increasing from the center of the interval; and very recently Rodríguez-Díaz and Santos-Martín (2009) follow this last approach for study the Modified Arrhenius model, introducing the name *filling designs*. In the following, the main types of filling designs that appear in literature will be examined, as well as their behavior with respect to some optimality criteria. Definition of each type and formulae for the general case and for particular interval choices will be provided. The number of points in the design can be as well an important matter for the computations.

Let n denote the number of points and $\mathcal{T} = [A, B]$ the design interval, with length $l = B - A$. We will consider exact designs $\{t_1, ..., t_n\}$ with the points being all different. The extreme points of the interval will always be in the design, that is $t_1 = A$ and $t_n = B$, thus for all the types the two-point design will be $\{A, B\}$. The minimum distance between consecutive points will be denoted by d. When n is odd, the center of the interval, $c = (A+B)/2$, will be in the design, but maybe for the case of the Linear Inverse Design and the new types Parabolic and Trigonometric designs. The basic types only need the interval and the number of points for their construction.

3.1. Uniform Designs

Uniform distributed design: the Uniform design has as spacing parameter a constant value $d = l/(n-1)$ which is the distance between every pair of consecutive design points.

This is the simplest case. The i-th point will be given by

$$a_i = A + d(i-1) = A + \left(\frac{i-1}{n-1}\right) l \quad .$$

3.2. Arithmetic Designs

Arithmetical distributed design: in arithmetical designs, the spacing parameter grows in an arithmetical (id) progression as the distance from the center of the design interval $\mathcal{T}$ increases.

We will need to split the study in the cases of n being odd and even. Let $b = [n/2]$, where [...] denote the integer part. For n odd the design will be

$$\left\{ c,\ c \pm \sum_{i=1}^{b} \frac{i(i+1)}{2}\, d \right\}_{i=1,\dots,b} ,$$

with $d = l/(b + b^2)$, while for n even the points are given by

$$\left\{ c \pm \sum_{i=1}^{b} \frac{i^2 + i - 1}{2}\, d \right\}_{i=1,\dots,b} ,$$

being $d = l/(b^2 + b - 1))$.

3.3. Geometric Designs

Geometrical distributed design: in geometrical designs the spacing parameter grows in a geometrical (d^i) progression as the distance from the center of the design interval $\mathcal{T}$ increases.

Again we will study the cases of n odd and even separately. For n odd the design will be

$$\left\{ c,\ c \pm \sum_{i=1}^{b} \frac{d(d^{\mathrm{i}} - 1)}{d - 1} \right\}_{i=1,\dots,b} ,$$

with d the solution of the equation $2d(d^b - 1) = l(d - 1)$. For n even the points are given by

$$\left\{ c \pm \left(\sum_{i=1}^{b} \frac{d(d^{\mathrm{i}} - 1)}{d - 1} - \frac{d}{2} \right) \right\}_{i=1,\dots,b} ,$$

and in this case d is the solution of $2d(d^b - 1) = (l + d)(d - 1)$.

3.4. Linear Inverse Designs

Linear Inverse distributed design: in the linear inverse design, the interval $[f(A), f(B)]$ (where $f(t)$ is the model, $f(t) = at^{-m}e^{-\beta/t}$ in this case) is uniformly divided and the corresponding points in the design space $\mathcal{T}$ through the inverse regression function are taken as support points. The main problem in this type of designs is that we need to know initial values for the parameters in advance. Another additional issue is that in some cases the f function can be non-biunivoque in the interval $\mathcal{T}$, that is, the inverse could be not well defined.

For all these types of designs (uniform, arithmetic, geometric, lineal inverse) the samples are either uniformly distributed (directly, or inversely through the model), or more dense around the center of the interval. But maybe it would be preferred to concentrate samples near a point different from the center. Following this idea, a new type of filling design is presented, in this case depending on more parameters than the extremes of the interval and the number of points.

3.5. Parabolic Designs

In this case a parabolic transformation is used to spread a uniformly distributed set of points, adding more parameters in order to improve the existing types of filling designs. The procedure for constructing such a design is described in Figure 9.

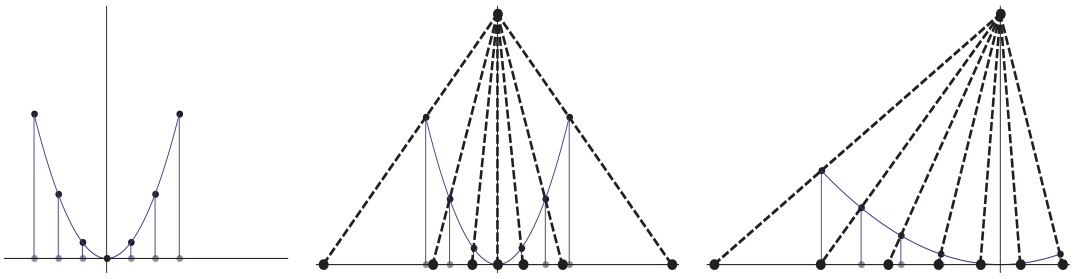

Figure 9: Construction of Parabolic Designs: first select a design in the x-axis and move it to the parabola (left), then project the points in the parabola again on the x-axis (center). The spreading can be modified by the choice of the projection point q, the parabola's curvature k and a suitable movement v of the initial design (right). The final step is the adaptation of these last points to the desired interval.

Let us take for instance a uniformly distributed initial design in $[-1, 1]$, maybe moved by parameter v,

$$p_i = \frac{2(i-1)}{n-1} - 1 - v \qquad i = 1, ..., n;$$

then, using the parabola $y = kx^2$ and the projection point $Q = (0, q)$, the 'projected' design will be

$$P_i = \frac{p_i q}{q - k p_i^2} \qquad i = 1, ..., n \qquad .$$

Finally let us adapt this last design to fit in the interval $[A, B]$,

$$\left\{ A + \frac{P_i - P_1}{P_n - P_1} l \right\}_{i=1,...,n} \quad . \tag{3}$$

The parameters k and q regulate the dispersion of the points, and $-1 \leq v \leq 1$ controls the area where the concentration of points is greater (-1→ near A, 1→ near B, 0→ center of the design interval).

3.6. Trigonometric Designs

Parabolic designs offer the possibility of vary the place where the concentration of points is more dense, but for other hand only one concentration area can be chosen while sometimes it would be more convenient to have several ones. For this reason, a new type of filling design is introduced, allowing a much more rich concentration structure with the same number of parameters. It is based on the trigonometric functions and follows the same idea that the parabolic designs, that is, first transform an initial set of points through a convenient function (in this case the cosine function

has been chosen), then project from a specific point to get the desired dispersion, and finally make these last points fit in the chosen interval. Figure 8 shows the transformation in this case.

Now the function $f(x) = \cos[s(x - \pi + v)] + 1$ is chosen for the transformation. Beginning with a uniform design in $[-\pi, \pi]$,

$$p_i = \pi \frac{2i - n - 1}{n - 1} \qquad i = 1, ..., n;$$

the projected points from $Q = (0, q)$ are

$$P_i = \frac{p_i q}{q - f(p_i)} \qquad i = 1, ..., n \qquad .$$

The last step is the adaptation (3) to the interval $[A, B]$. In this case the parameter $q > 2$ regulates the dispersion of the points, $s > 1$ decides the number of dense areas and $0 \leq v \leq \pi$ the situation of these ones.

3.7. Comparison of filling designs: a real example

Let us focus on VG, the 62-point design in [195,420] obtained from data shown in Vaghjiani and Ravishankara (1991) and Gierczak et al (1997) and assume independent observations. From this last paper the initial values for the non-linear parameters of the model have been taken to be $m = -2.82$ and $\beta = 987$ (the vales shown in the first paper are very similar). The initial value of the linear parameter a has no influence in the D-optimal design and has been fixed to 1.85×10^{-7} to avoid extreme values of determinants during computations. Now at least three points are needed in the design, and it can be checked that the best three-point design is $\xi_3 = \{281, 374, 420\}$, with determinant 83.93 (approximate designs with equal weights will be considered in order to perform comparisons). Optimal four-point design are obtained when repeating one of the observations, getting a determinant value of 70.82 in every case. The determinant value does not increase when taking more points, thus it seems that the globally optimal design is ξ_3. When using the arithmetic design and compute the best interval where to set the design we get $t_1 = 258$, that is the interval [258,420]. The determinant using this setup is 14.13. The optimal uniform design takes $t_1 = 261$ and gets determinant 25.63. The optimal parabolic design has determinant 30.46 using $t_1 = 237$, $k = 1.2$, $q = 9$ and specially $v = 1$, meaning that the best design tries to concentrate the samples mainly at the upper extreme of the interval. The design is shown in Figure 10. Finally, the optimal

Figure 10: Parabolic optimal design for Gierczak example in [195,420]

trigonometric design needs to take $t_1 = 272$, $v = \pi/2$, $s = 1.7$ and $q = 4.7$, obtaining a determinant value of 55.59. Explicitly, the design is
{272., 272.05, 272.62, 272.64, 273.53, 274.12, 274.6, 275.71, 276.52, 276.72, 279.84, 283.99, 288.81, 294.06, 299.53, 305., 310.3, 315.31, 319.96, 324.22, 328.1, 331.61, 334.79, 337.68, 340.32, 342.77, 345.06, 347.24, 349.35, 351.43, 353.53, 355.67, 357.91, 360.3, 362.86, 365.65, 368.71, 372.07, 375.76, 379.79, 384.14, 388.75, 393.53, 398.33, 402.98, 407.3, 411.09, 414.23, 416.64, 418.33, 418.4, 418.4, 418.6, 418.64, 418.93, 419.16, 419.33, 419.36, 419.68, 419.83, 419.9, 420.}
and it is shown in Figure 11. The main characteristic is the fact that it has three concentration

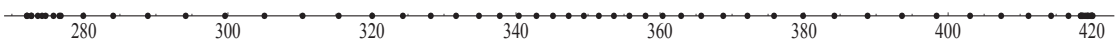

Figure 11: Trigonometric D-optimal design for VG example in [195,420]

areas, near the extremes and the center of the interval $[272, 420]$ and thus not far from ξ_3, and with the same shape.

Table 1 shows the behavior of the different types of filling designs for this example. practitioners. It can be seen that the using of Trigonometric designs can lead to a great improvement in efficiency with respect to other designs used in practice (20% of improvement respect to the Uniform or 30% respect to the Arithmetic), it is not very far from the optimal 3-point design and uses 62 different observations, which usually is preferred by practitioners. In any case, all of these designs improve the figures obtained with the VG design, that produces a determinant of 9.09 and thus a efficiency of 47.7%.

	Arith	Geom	Unif	L.Inv	Parab	Trigon	Opt 3-point
t_1^* $\{v, k, q\}$	258	270	261	238	237 $\{1, 1.2, 9\}$	272 $\{\pi/2, 1.7, 4.7\}$	$\{281, 374, 420\}$
Determinant	14.13	15.21	25.63	29.96	30.46	55.59	83.93
Efficiency	0.55	0.57	0.67	0.71	0.71	0.87	1

Table 1: Comparison of different designs for VG example in [195,420] assuming independent observations

Note that the optimal point t_1^* does not depend (or the variation is minimum) on the number of points that we want in the design. That is, when looking for instance for the best 30-point designs in $[195, 420]$, the optimal points for the first observation in each case are the same than those in Table 1 for the 62-point design. Nearly the same is true for the parameters of the Parabolic and Trigonometric designs. In this last one the maximum discrepancy was for the 10-point design, for which $t_1^* = 273$ and the best parameters set is $\{\pi/2, 1.8, 4.3\}$, both first point and parameters very close to the values shown in Table 1 for a much bigger design.

Another interesting fact is the following: within the same type of filling designs the determinant (and so the efficiency) is very much depending on the number of points that we want in the design. For instance, the best Trigonometric design in three points is $\{281, 374, 420\}$, exactly the optimal 3-point design, and thus it has maximum efficiency. However, the best 4-point Trigonometric design is $\{275, 287, 374, 420\}$ (using $v = \pi/2$, k=0.5 and q=2.3), losing 5.5% of efficiency. The decreasing of efficiency keeps when increasing the number of points, although it is not that sharp from that point on, neither is monotone. Figure 12 shows this behaviour.

Let us now study the case of non-independent observations, assuming the exponential covariance function given by $C(t_1, t_2, r) = \sigma^2 \exp(-r|t_1 - t_2|)$, where $r > 0$ is the covariance parameter. In this first approach r will be assumed known, and since is a non-linear parameter, nominal values of r will be necessary for computations. The best intervals for constructing each type of Filling Design in are computed for different values of r. The upper extreme is always the extreme of the design interval, 420, and the first point is given in the first row of each case in Table 2, where the rest of parameters for Parabolic and Trigonometric designs are shown as well.

Regarding the determinants, the values can vary greatly for the different values of r, but in

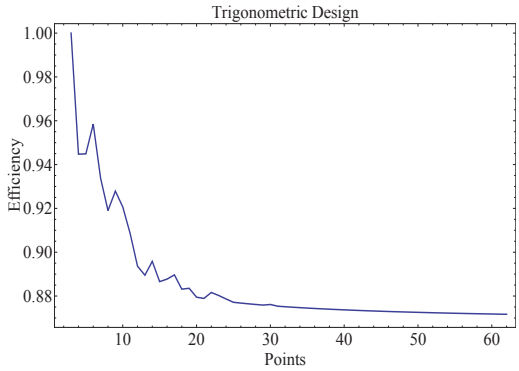

Figure 12: Decreasing of efficiency of Trigonometric Designs when increasing the number of points

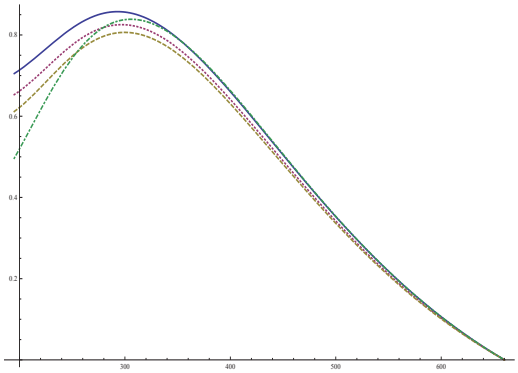

Figure 13: Efficiency of designs Uniform (full line), Arithmetic (dotted), Geometric (dashed) and Lineal Inverse (Dot-Dashed) for the uncorrelated case using the interval $[195, 660]$ and taking $t_2 = 660$, as functions of t_1

each case these values are similar in magnitude for all filling designs (see Table 3). In general, Trigonometric or Parabolic Designs are the best ones, and the differences between the design types are specially remarkable for greater values of r, that is, close to the independence between observations.

Optimal 3-point designs have been computed for the different values of r as well. For all the cases, the upper extreme 420 is one of the points, as expected. However, in this case the values of the determinants are extremely big (specially when r is small) when comparing with the ones of the filling designs and thus this comparison has no sense. Table 4 shows the two points that jointly with 420 form the design, as well as the value of the determinants obtained in the different cases.

The VG design has repeated measures. This means that the correlation matrix is singular and the criterion function cannot be computed. To avoid this fact, a procedure consisting in the addition of a random component with 0 mean to the repeated points could be employed. The procedure allows the obtaining of the criterion values for designs "very close" to the ones with repeated measures. The smaller the variance of the random component is, the closer will be the design to the original one. However, there is a problem concerning the correlation matrix for this case. To describe it let us focus for instance in the uncorrelated case. The case of independent observations is approximately equivalent to the correlated case when the r parameter is "big", let us say greater than 10, because in this case the covariance $C(t_1, t_2, r) = \sigma^2 \exp(-r|t_1 - t_2|)$ is in general very close to 0 for every pair of observations taken in the points $\{t_1, t_2\}$. But when employing the procedure described previously there will be pairs $\{t, t + \delta\}$, where δ is the variance of the random component, and when we choose δ small to get close to the repeated measures case the differences $|t_1 - t_2| = \delta$ are small as well. In this situation we need to take a very big r for the covariance to vanish. For instance, when δ follows a normal distribution with 0 mean and standard deviation 10^{-3}, then r needs to be taken greater than 10^6 for approaching the uncorrelated case (see Table 5). From this discussions follows that the design VG cannot be compared with the filling designs (which have non-repeated measures) for the correlated case.

Compound Designs

Finally, designs that use points of different types of filling designs are considered. The idea

r	Arith	Geom	Unif	L.Inv	Parab	Trigon
0.001	195	195	195	195	195 {1,0.6,5.4}	195 $\left\{\frac{3\pi}{4}, 1.3, 5.7\right\}$
	195	195	195	195	195 {1,1,9}	195 $\left\{\frac{\pi}{2}, 1.3, 8.3\right\}$
0.01	195	195	195	195	195 {1,0.9,8.4}	195 $\left\{\frac{\pi}{2}, 1.3, 8.8\right\}$
	195	195	195	195	195 {1,1.1,8.5}	195 $\left\{\frac{3\pi}{4}, 1.3, 5.4\right\}$
0.1	195	195	195	199	195 {1.,0.6,15.6}	195 $\left\{\frac{3\pi}{4}, 1.3, 15.6\right\}$
	202	203	202	207	199 {1,1.2,11.5}	200 $\left\{\frac{-3\pi}{4}, 1.3, 5.2\right\}$
0.5	223	210	215	205	208 {1,0.5,10.4}	209 $\left\{\frac{\pi}{2}, 1.2, 12.6\right\}$
	245	250	246	230	237 {1,0.9,10}	240 $\left\{\frac{\pi}{2}, 1.4, 5.5\right\}$
1	242	255	241	222	231 {1,0.5,8.9}	235 $\left\{\frac{\pi}{2}, 1.4, 11.4\right\}$
	267	277	286	270	237 {1,1.4,8.4}	255 $\left\{\frac{\pi}{2}, 1.9, 5.4\right\}$
10	257	270	261	238	237 {1,1.2,9.}	261 $\left\{\frac{\pi}{2}, 1.5, 5.4\right\}$
	305	392	396	395	267 {1,2.4,10.1}	279 $\left\{\frac{\pi}{2}, 1.9, 4.6\right\}$

Table 2: Best first point t_1^* and parameters (v, k, q) (when necessary) for filling designs and VG example (62 points) in [195,420], as functions of r. In each case the second line shows the figures when r is a parameter as well

r	Arith	Geom	Unif	L.Inv	Parab	Trigon
0.001	4.185	4.174	4.205	4.201	4.209	4.209
0.01	0.03527	0.03520	0.03543	0.03540	0.03544	0.03544
0.1	0.1990	0.1973	0.2042	0.1852	0.2045	0.2044
0.5	4.637	4.323	7.013	4.990	7.235	7.251
1	9.116	9.909	18.00	14.45	18.94	19.35
10	13.96	15.21	25.63	29.96	30.46	43.75

Table 3: Determinants for best filling designs depending on r for Gierczak example in [195,420]

is to mix the main characteristics of different types of designs. Two groups can be distinguished: Uniform, Arithmetic, Geometric and Lineal Inverse designs (Type I) select the points along the whole interval, while Parabolic and Trigonometric designs (Type II) tend to concentrate them in a specific number of areas. In general, the results show that the second approach is preferred, but taking points from both types may lead us to improve the efficiency of the joint design. In the construction of this last one the following characteristics have been taken into account:

- the designs will have the shape $\xi_\alpha = \alpha TypeI + (1-\alpha)TypeII$, $(0 \leq \alpha \leq 1)$, when the aim is to mix properties of filling designs of types $TypeI$ and $TypeII$
- the parameter α represents the weight of the $TypeI$ design in the global design ξ_α. When using exact designs this weight applies to the number n of points of ξ_α, that is, approximately αn points of ξ_α will be given by a design of type $TypeI$ and the rest will be obtained from a $TypeII$ design. For instance, when $\alpha = 0.5$ the 62-point design ξ_α will be obtained joining the points of a $TypeI$ and a $TypeII$ designs, with 31 points each.
- a common interval $[t_1, 420]$ has been used for the two sub-designs. Since for all types of filling designs the extremes of the interval are points of the design, at least these two points will be

$n \setminus r$	0.001	0.01	0.1	0.5	1	10
3	{307,389} (7036.38)	{293,386} (180.02)	{281,374} (83.94)	{281,374} (83.93)	{281,374} (83.93)	{281,374} (83.93)
3	{308, 390}	{300, 389}	{293, 406}	{296, 417}	{297, 418}	{297, 419}
4	{195, 297, 389}	{295, 371, 400}	{281, 374, 375}	{281, 374, 375}	{280, 281, 374}	{280, 281, 374}

Table 4: Optimal n-point designs for Gierczak example in [195,420] depending on r. The upper extreme 420 (not shown) appears in every design. The first row shows the figures for the case that r is not a parameter, with the determinants between brackets. The last two rows are the results when r is consider a parameter

r	0.001	0.1	1	100	10^4	10^6
Det(VG)	3.22131	0.0761318	0.181639	0.269954	5.22305	9.09322

Table 5: Determinants for VG design in [195,420] depending on r. A random component with 0 mean and standard deviation 10^{-3} has been added to the equal points in order avoid singular covariance matrix

repeated in ξ_α. For the correlated case, the procedure described previously adding a white noise with small variance to the repeated points has been used.

Initially, the more representative types of both tendencies of filling designs have been chosen: Uniform (homogeneous selection) and Trigonometric (concentration areas) designs. The compound design $\xi_{Uni-Trig,\alpha} = \alpha Uniform + (1-\alpha) Trigonometric$ has been studied for the uncorrelated case. The results show that there is a slight improvement for very small values of α: when α is between 0.025 and 0.1 the determinant of $\xi_{Uni-Trig,\alpha}$ is 55.59 greater than that of the Trigonometric design, taking the maximum value for the smallest of these α's. Talking about number of points, the design improves when between 2 and 6 points out of the 62 are selected with a uniform procedure, with the maximum determinant 56.82 obtained for just 2 'uniform' points. And in fact this result agrees with the impression given above that the concentration-areas designs are better, since the 2-point Uniform design takes simply the extremes of the interval, that is, increases the concentration that always exists at the extremes. However, the study for the correlated case does not show any improvement when using compound designs.

r **parameter**

The last step of the study is the consideration of r as a parameter that needs to be estimated as well. Thus now we deal with a four-parameter model and the information matrix is Pázman (2007)

$$M_{\theta,r}(n) = \begin{pmatrix} \frac{\partial \eta'(\theta)}{\partial \theta} C^{-1}(r) \frac{\partial \eta(\theta)}{\partial \theta'} & 0 \\ 0 & \frac{1}{2} tr \left\{ C^{-1}(r) \frac{\partial C(r)}{\partial r} C^{-1}(r) \frac{\partial C(r)}{\partial r^T} \right\} \end{pmatrix},$$

where η' denotes the transpose of the model. Table 2 (second row of each case) shows the optimal intervals for each type of design in this case. The biggest differences with the results in the first rows (when r is not estimated) are for great values of r, and especially important for the first four types (Type I designs). For $r = 10$ the best intervals for constructing the corresponding filling designs is can be very small.

Since now we have four parameters, the best designs in three and four points have been computed for different values of r. The procedure looks for observations with at least one-day gap between every pair of them. The upper extreme, 420, appears in every design. The rest of the points are

shown in Table 4 (last two rows), where it can be observed that for the highest values of r the designs tend to reduce the number of support points, that is, try to repeat one of the observations.

Compound designs in this setup produce little improvements for most of the values of r. Small values of α (between 0.04 and 0.12) are the best ones, with an only exception: the case $r = 0.5$, that gets the best improvement (but anyway less than 2%) when using $\alpha = 0.81$ and the set of parameters $t_1^* = 241$, $v = \pi/2$, $s = 1.5$ and $q = 5.1$, that is, 50 points using the uniform design and just 12 using the Trigonometric design.

As a conclusion about these results can be said that, respect to efficiency, filling designs appear as a good option, and when the observations are independent they provide alternative intervals for their construction that are included in the design interval and seems to be more efficient when taking the observations inside them. However, for the case of correlated observations the whole design interval is usually the best one for constructing the filling designs in, and the efficiency of these ones is very poor when comparing with the optimal 3-point designs. Anyway, these kind of designs (or close to them) are used by practitioners routinely. The designs that define concentration areas are usually preferred to those which place the points along the whole interval, and this idea is confirmed with the study of the compound design, that allows both approaches.

In the following, the basic types of Filling designs (Uniform, Arithmetic, Geometric and Linear Inverse) will be the ones used for comparison since they only need the interval for their construction.

4. Case study: Methane

Methane (CH_4) is an important greenhouse gas, whose concentration in the troposphere is steadily increasing. To estimate the flux of methane into the atmosphere and its atmospheric lifetime, its rate of removal needs to be accurately determined. The main loss process for atmospheric methane is the reaction with the hydroxyl radical, $OH + CH_4 \rightarrow CH_3 + H_2O$. This reaction has been extensively studied, and it can be expressed in the Arrhenius form as

$$k(T) = aT^{-m}exp(-\beta/T),$$

with $a_0 = 2.80 \times 10^{-14}$, $m_0 = -0.667$, and $\beta_0 = 1575$. This three-parameter fit may be preferred for lower stratosphere and upper troposphere calculations Jet Propulsion Laboratory Publication 06-2 (2006). The temperature dependence of this rate coefficient has been measured in different initial intervals by Vaghjiani and Ravishankara (1991), 223-420 K, Dunlop and Tully (1993), above 298 K, Finlayson-Pitts et al (1992), 278-378 K, Mellouki et al (1994), 233-343 K, Bonard et al (2002), 295-660 K, and Gierczak et al (1997), 195-300 K.

Considering independent and normally distributed errors with mean 0 and constant variance ($\sigma^2 = 1$) and m known, the parameters to be estimated are a and β. The locally D-optimal designs for the model are two-point designs given in Rodríguez-Díaz and Santos-Martín (2009).

4.1. Three-point designs

4.1.1. Uncorrelated case

Assuming the initial values m_0, and β_0 the designs depend on the upper bound of the design interval. For three-point designs there is no difference between the uniform, the arithmetic and the geometric designs, taking all of them the extreme points and the midpoint of the design interval. The efficiency of a design ξ with respect to a specific optimality criterion Φ expresses the proportion of observations of ξ that the Φ-optimal design needs to get estimators of the parameters with the

"same variance" that ξ, measuring the variance with the optimality criterion Φ. The D-efficiency of ξ is $Eff_D(\xi) = (detM(\xi)/detM(\xi^*))^{1/2}$, where ξ^* is the D-optimal design for the model. D-optimal (two-point) designs using the different intervals that appear in the bibliography are shown in Table 6, as well as the efficiencies of the 3-point filling designs when using the whole interval. The last row shows the point t_1 for which the equidistant 3-point design in $[t_1, t_{\max}]$ gives the maximum efficiency (over 89%). It can be seen that there are no differences between the filling designs respect to the efficiency when the whole interval is used. However, from the last row it seems sensible not to choose the whole interval but a part of it, avoiding the observations at the beginning of the interval, that seems to be less informative.

Initial interval	195-300	233-343	278-378	223-420	295-660
D-opt	{256, 300}	{287.2, 343}	{311.9, 378}	{340.6, 420}	{489.9, 660}
Eff(Uniform)	0.70	0.78	0.86	0.66	0.69
Eff(Lin.Inv.)	0.67	0.73	0.83	0.64	0.65
t_1	246.85	275.97	298.84	325.38	461.74

Table 6: D-optimal designs, efficiencies of 3-point filling designs and initial point of the best 3-point equidistant design for the reaction $OH + CH_4 \rightarrow CH_3 + H_2O$ when using different intervals (uncorrelated case)

4.1.2. Correlated case

Let us now assume Cov1 and study in first place the equidistant design. Fixing t_{max} as the interval upper bound, Table 7 shows the best lower bound t_1 for different r, first when r is not consider a parameter (upper row of each interval) and then taking r as a parameter (lower row). For r greater than 0.5, the covariance matrix gets closer the identity matrix and thus we have nearly the uncorrelated case. It can be seen that the differences between the two approaches increase with r and they may be very important.

$t_{min} - t_{max}$	$r = 0.001$	$r = 0.005$	$r = 0.01$	$r = 0.05$	$r = 0.1$	$r > 0.5$
195-300	199.04	258.11	258.02	239.37	241.87	246.85
	200	264.47	265.72	264.54	273.15	296.10
233-343	233	290.93	286.9	266.86	271.92	275.97
	233	298.63	298.66	300.60	313.46	339.05
278-378	278	315.78	308.56	289.22	295.73	298.84
	278	325.41	324.48	330.73	346.76	374.05
223-420	238.88	343.75	332.54	315.96	323.31	325.38
	241.55	356.43	354.39	367.78	387.16	416.03
295-660	295	471.26	445.49	458.61	461.70	461.74
	295	514	511.91	591.82	622.65	655.99

Table 7: Best t_1 fixing the upper extreme with correlation for different r (2-parameter model in first row of each interval, 3-parameter model in second row) for the equidistant design

However, when considering a general 3-point design, the behavior is different. Table 8 shows the design in three points that maximize the determinant of the information matrix $M_\theta(3)$ for the example of Methane and different values of r; the upper extreme of the interval is always one of the

points, and thus it is not shown. Now the case when r is a parameter produces similar figures, see Rodríguez-Díaz et al. (2009) for details. In both cases, when r is big (equivalent to the uncorrelated case) we get the two support points of the D-optimal designs from Table 6, repeating one of the points. As r decreases we get three different points, which become the two extremes and a middle point for the smallest values of the parameter. Finally, let us note that the designs for the extreme values $r < 0.001$ are the same in both tables, thus in these cases there is no difference between r being estimated or not.

r	195-300	233-343	223-420	295-660
660				
10	255.7, 256.5	286.7, 287.8	340.0, 341.1	489.3, 490.6
660				
1	252.8, 259.2	283.6, 290.5	336.9, 344.3	485.7, 494.3
497.4, 660				
0.1	242.1, 267.3	270.8, 299.4	323.3, 357.7	467.7, 513.5
525.8, 660				
0.01	257.7, 281.3	287.6, 317.6	332.1, 378.9	445.7, 546.8
565.3, 660				
0.001	195.0, 270.2	233.0, 306.7	223.0, 362.0	295.0, 523.3
660				

Table 8: Optimal design assuming correlation when r is not a parameter

4.2. Designs with more than three points

Four is the lowest number of points for which Uniform, Arithmetic and Geometric designs are different. For this reason four-point designs will be now studied in order to compare these three types of designs.

The D-efficiency of the filling designs supported in four points for the uncorrelated model is similar for the basic types (see Rodríguez-Díaz et al., 2009). However, there are more differences when looking for the best interval for constructing the designs in.

For instance, Figure 13 shows the respective plots of the efficiency, as function of t_1, when the initial interval $[195, 660]$ is used and t_2 fixed to be the upper limit, for β_0, m_0 and the designs Uniform (full line), Arithmetic (dotted) and Geometric (dashed), in the uncorrelated case. It can be seen that the differences are most important for the lowest values of t_1, that is, when using almost the whole interval. These differences can be observed as well in Table 9, that shows the best t_1 for Uniform, Arithmetic and Geometric designs, when t_2 is fixed to be $t_{\max}$ and we assume independent observations.

Table 10 shows the best t_1 when fixing the upper extreme of the interval in the correlated case for different r, for the Uniform (U), Arithmetic (A) and Geometric (G) designs, considering that r is not a parameter. It can be observed that for small r the best efficiency is obtained when taking the filling designs in the whole initial interval. However, for independent observations it is more efficient to take the observations in part of the initial interval. The results when r is consider a parameter are similar to these ones, and can be checked in Rodríguez-Díaz et al. (2009).

$t_{min} - t_{max}$	Uniform	Arithmetic	Geometric
195-300	245.15	241.90	240.06
233-343	273.79	269.81	267.42
278-378	296.24	291.66	288.79
223-420	322.24	316.93	313.50
295-660	454.70	445.47	439.00

Table 9: Best t_1 fixing the upper extreme (n=4)

		$r = 0.001$	$r = 0.005$	$r = 0.01$	$r = 0.05$	$r = 0.1$	$r > 0.5$
195-300	U	195	195	251.38	228.88	232.70	245.15
	A	195	204.35	253.27	229.90	231.68	241.61
	G	197.67	250.31	256.08	235.87	237.59	238.81
233-343	U	233	233	279.05	254.24	262.04	273.80
	A	233	233	281.05	254.60	259.81	269.70
	G	233	287.49	284.75	262.68	266.80	266.09
278-378	U	278	278	299.28	274.98	285.63	296.24
	A	278	305.28	301.47	274.81	282.23	291.62
	G	278	312.8	306.16	284.56	289.77	287.44
223-420	U	223	329.73	321.21	300.07	313.38	322.23
	A	227.61	334.01	323.64	299.09	308.49	316.91
	G	237.20	340.91	329.81	310.71	316.1	312.16
295-660	U	295	450.56	418.99	440.33	453.60	454.70
	A	295	454.28	421.88	431.85	442.67	445.47
	G	295	467.55	440.61	448.89	443.99	438.22

Table 10: Best t_1 for filling designs when fixing the upper extreme, for different r (not a parameter)

The case $n > 4$ does not produce special results. For instance, the comparison of D-efficiency for $n = 6$ shows similar figures for the different types of filling designs, slightly worse than for $n = 4$ (see Rodríguez-Díaz et al. (2009) for more details).

4.3. Goodness of estimators

From recent computations in Matlab, the ML estimator of r assuming Cov1 for the data in Gierczak et al (1997) and Vaghjiani and Ravishankara (1991) is $\hat{r} = 0.03643$. The first attempts to estimate the r-parameter suffered from numerical difficulties which shaped up to be caused by the scaling of the k_5 data. Transforming the data according to $y = k_5 \cdot 10^{15}$ made the problems disappear.

The ML estimator of r and σ^2 simultaneously was computed as follows: taking the 18 different temperatures t and the 62 according measurements of y (k_5-values multiplied by 10^{15}), the expectation of y was modeled according to Gierczak et al (1997) and assuming an exponential covariance function:

$$E(y|t) = 1.85 \cdot 10^{-5} t^{2.82} \exp\left(-\frac{987}{t}\right)$$

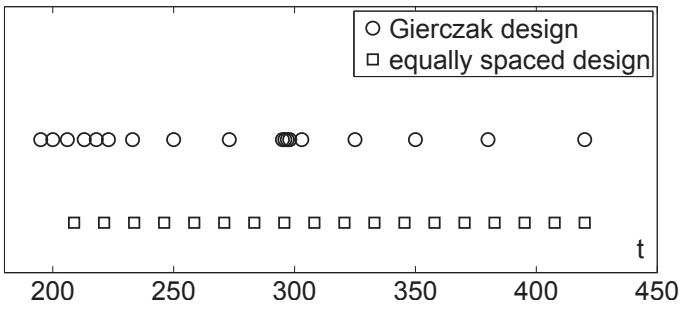

Figure 14: Different designs

$$Cov(y_i, y_j|d, r) = \sigma^2 \cdot \exp(-d \cdot r) \qquad d = d_{ij} = |t_i - t_j|$$

The observation y was assumed to be normally distributed and the log-likelihood was maximized with respect to r and σ^2. Maximization succeeded numerically using a grid of points at which the log-likelihood was evaluated and zooming in on the maximum of the log-likelihood. As final estimates we got $\hat{\sigma}_y^2 = 0.2365$ and $\hat{r} = 0.03643$. Actually, the numerical difficulties caused by the repeated measurements of k_1 at single temperatures were eliminated by employing a nugget effect i.e. adding a constant α to the trace of the covariance matrix of y. Setting $\alpha = \hat{\sigma}^2 \times 10^{-3}$ the log likelihood stabilized sufficiently. The bias caused by adding the nugget α was negligible because the ML estimate of r and σ^2 may be considerably biased without nugget (see later), and the main conclusion from these computations is that correlation plays an important role in this example.

Now, the design effect on the ML estimation of m, β, r and σ^2 will be analyzed. The 18 different design points for k_5 are not selected systematically. The question is whether different designs yield more precise ML estimates (see Figure 14).

In a first approach we compared the Vaghjiani-Gierczak (VG) design with an 18- point equidistant design on the interval $[208.8; 420]$. For both designs we simulated 1000 k_5-samples and for each of the simulated samples we computed the ML estimate of m, β, r and σ^2 simultaneously. For the simulations we set the parameter values to the ML estimates from the original 62 point sample.

We compared the empirical distributions of the ML-estimates of the two designs for each parameter. The results can be seen in Table 11 and Figures 15 to 18. As an estimate for the D-optimality criterion we also computed the determinant of the inverted empirical covariance matrix of the parameter estimates for each design: for the VG design we got $\det M_{VG} = 2.3984 \cdot 10^7$, for the equidistant design we got $\det M_{equidist} = 6.3271 \cdot 10^6$.

Based on these simulations it seems that the two designs are roughly equivalent. We notice a remarkable estimation bias - especially for $\hat{r}_{ML}$ - which should be analyzed in more details. We should also deliberate whether the simultaneous ML estimation of several parameters is advantageous or not.

The empirical correlation between the ML-estimates for m and β is extremely high in both designs ($R_{m,\beta} = 0.98$). In the VG design the correlation between r and σ^2 is also remarkably high ($R_{r,\sigma^2} = -0.48$) whereas in the equidistant design it is clearly smaller ($R_{r,\sigma^2} = -0.26$). So far this is the most significant difference between the two designs inspected.

Table 11: statistics of the ML estimates of the model parameters for gierczak design and equidistant design (18 points)

	true	Gierczak design			equidistant design		
	value	mean	bias	variance	mean	bias	variance
m	2.82	2.829	.009056	$8.592 \cdot 10^{-5}$	2.829	.008723	$8.094 \cdot 10^{-5}$
β	987	1009	21.77	530.9	1008	21.17	493.1
σ^2	.2365	.1904	-.04615	.01021	.1936	-.04285	.01610
r	.03643	.06560	.02917	.003477	.07058	.03415	008176

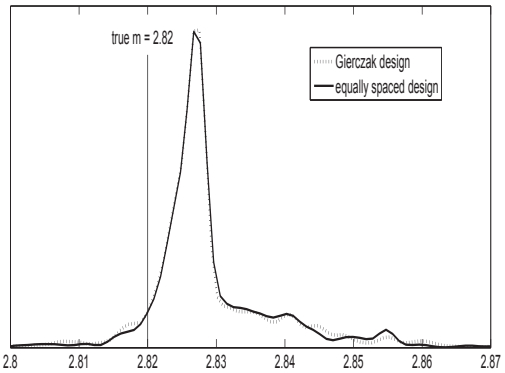

Figure 15: Empirical distribution of $\hat{m}$ for different designs

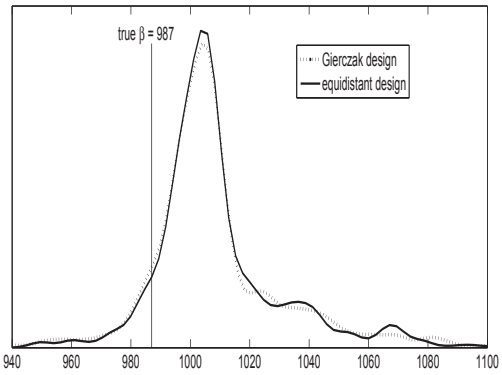

Figure 16: Empirical distribution of $\hat{\beta}$ for different designs

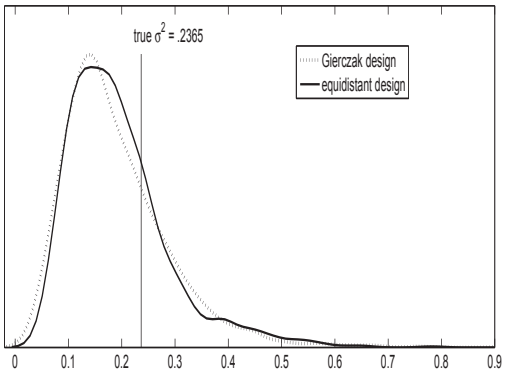

Figure 17: Empirical distribution of $\hat{\sigma}^2$ for different designs

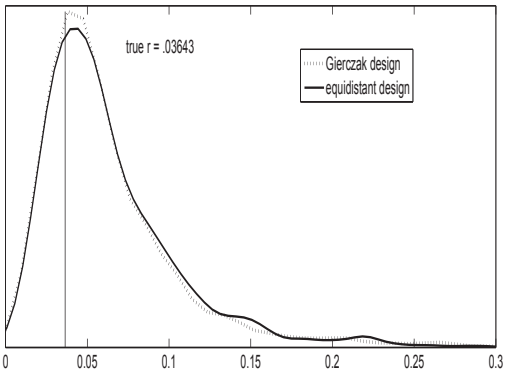

Figure 18: Empirical distribution of $\hat{r}$ for different designs

5. Bias reduction methods

There are many issues related to the reduction of bias. One of special importance is how to appropriately choose the nugget effect.

We may consider the following explanation of the ML estimator behavior. Let us consider the semivariogram $\gamma(t) = 0.5E(X(s+t) - X(s))^2$. We have $\gamma(t) \to c_\gamma = c_0 + c_{ms}; t \to 0$ as considered in Kim et al (2010). c_γ is known as the nugget effect, c_0 denotes the nugget effect of a micro scale process and c_{ms} denotes the measurement error variance. These are two sources of variability: one source derives from the spatial variability at a scale smaller than the minimum lag distance, and the other source is experimental error. In such a model we may well explain the observed facts:

1) With a small number of design points we observe $\hat{r} = \infty$ with relatively high frequency. This may be understood as effort of the very smooth process to decrease the dimension of the process or as an effort of observations to be independent. The reason is too rough measuring of the smooth process. If we add a positive nugget effect to such a process, it slightly improves this behavior (slightly lower frequency of observing $\hat{r} = \infty$), however, we obtain a relatively high frequency of $\hat{r} = 0$. This may be understand as a "deterministic" limit of our equations for ML estimates with positive probability. By adding a negative nugget, we have found interestingly that we got empirically almost never $\hat{r} = \infty$ or $\hat{r} = 0$. However, it is not easy to make a theoretical explanation of these results.

Bias is decreasing with n. In our model, after having $n > 20$, bias is practically negligible. The bias reduction method using the nuggets is based on the following idea: we are measuring with the finite set of nuggets $\tau_1 < .. < \tau_m$ (also negative ones), having $\tau_s = 0$ for some $1 < s < m$, and computing $r_{MLi}, i = 1, .., m$; then we make a regression $bias_i = r_{MLi} - \tau_i r_{ML} n$. We are plotting cases $n = 3, 5$. Then the $0.5(r_{ML1} + r_{MLs})$ is less biased than r_{ML}. A interesting issue is the behavior for $m \to \infty$.

5.1. Bayesian approach for bias reduction of the MLE

Here we compute the maximum likelihood estimate (MLE) of the parameter r of the Generalized Exponential model (GE):

$$y_i = a \cdot t_i^m \cdot \exp\left(-\frac{\beta}{t_i}\right) + \varepsilon_i = \eta_i(t_i, m, \beta) + \varepsilon_i \qquad i = 1, \ldots, n$$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \sim \mathbf{N}(\eta(\mathbf{t}, m, \beta); \Sigma(\mathbf{t}, r)) \qquad \eta(\mathbf{t}, m, \beta) = \begin{pmatrix} \eta_1(t_1, m, \beta) \\ \vdots \\ \eta_n(t_n, m, \beta) \end{pmatrix}$$

$$\Sigma(\mathbf{t}, r) = (Cov_{ij})_{i,j=1,\ldots,n} \qquad Cov_{ij} = Cov(t_i, t_j, r) = \sigma^2 \cdot \exp(-|t_i - t_j| \cdot r) \qquad r > 0$$

5.1.1. Standardized design interval

Let $\mathbf{t} = (t_1, \ldots, t_n)^T$ be the n design points in the design interval, $\mathbf{t} \in [T_l; T_u]$. Since $Cov_{ij} = \sigma^2 \cdot \exp(-|t_i - t_j| \cdot r)$ we may always standardize our design interval to $[0; 1]$: the standardized design points are $s_i = \frac{t_i - T_l}{L}$ where L is the length of $[T_l; T_u]$. The parameter r in $[T_l; T_u]$ corresponds to $L \cdot r$ in the standardized interval $[0; 1]$. From now on we will assume a standardized design interval.

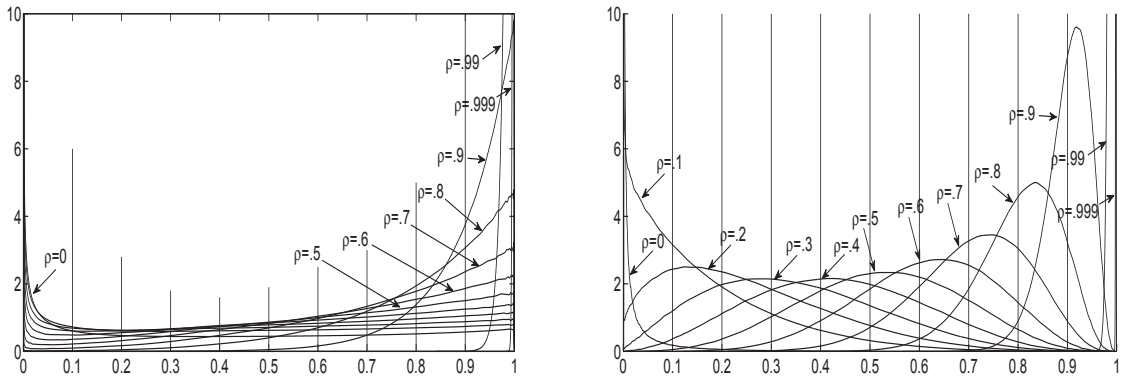

Figure 19: Empirical distributions of $\hat{\rho}$ for $n = 3$ (left panel) and $n = 10$ (right panel) for $\rho = 0, .1, \ldots, .9, .99, .999$

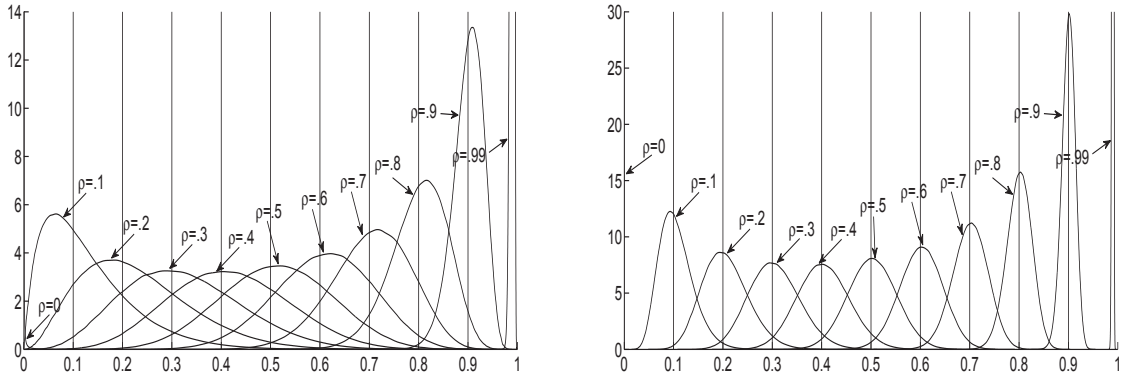

Figure 20: Empirical distributions of $\hat{\rho}$ for $n = 20$ (left panel) and $n = 100$ (right panel) for $\rho = 0, .1, \ldots, .9, .99$

5.1.2. Equidistant design

For the sake of simplicity let us assume all parameters known except r and use an equidistant design with design points $t_i = \frac{i-1}{n-1} \quad i = 1, \ldots, n$. Here we show only the results for $n = 3$, i.e. our design is $t_1 = 0 \quad t_2 = \frac{1}{2} \quad t_3 = 1$.

5.1.3. Empirical distribution of the MLE of ρ

Since we frequently get $\hat{r} = \infty$ it is better to describe the empirical distribution of $\hat{\rho} = \widehat{\exp(-r)}$. In our simulations we set $n = 3$ and $\rho = 0, 0.01, 0.03, 0.05, 0.1, 0.2, 0.3, \ldots, .8, .9, .99, .999$, for each simulated sample (sample size 1.000.000) we computed the MLE of ρ (see Rodríguez-Díaz et al. (2011) for the derivation of the MLE for r). In Figure 19 (left panel) we see the empirical distributions of $\hat{\rho}$. In Table 12 we see the main statistics for each simulation.

From the simulations we see that the MLE $\hat{\rho}$ performs poor for $n = 3$ especially with small ρ-values. Though it may be noticed that with increasing n also the performance of $\hat{\rho}$ improves very fast, this can be seen in Figure 19 (right panel for $n = 10$) and in Figure 20 ($n = 20$ and $n = 100$). From the simulations it is clear that the MLE of ρ is a consistent estimator.

Obviously the MLE of ρ is a biased estimator, the bias depending on the design size n and the

ρ	$\sharp$ of $\hat{\rho}=0$	mean($\hat{\rho}$)	bias	sd($\hat{\rho}$)	MSE($\hat{\rho}$)
.0	420835	0.2563	0.25633	0.328909	0.173887
.01	376872	0.2770	0.26698	0.332783	0.182020
.03	343881	0.2953	0.26526	0.335783	0.183113
.05	319808	0.3101	0.26014	0.337787	0.181772
.1	277920	0.3399	0.23988	0.340417	0.173425
.2	217788	0.3937	0.19372	0.341890	0.154416
.3	170271	0.4476	0.14758	0.339010	0.136707
.4	129046	0.5052	0.10520	0.329928	0.119920
.5	93081	0.5678	0.06777	0.314233	0.103335
.6	59407	0.6384	0.03845	0.287537	0.084156
.7	30305	0.7181	0.01810	0.245092	0.060398
.8	8632	0.8068	0.00684	0.181106	0.032846
.9	227	0.9020	0.00199	0.094720	0.008976
.99	0	0.9900	$2.727 \cdot 10^{-5}$	0.009918	$9.836 \cdot 10^{-5}$
.999	0	0.9990	$7.135 \cdot 10^{-7}$	$9.992 \cdot 10^{-4}$	$9.984 \cdot 10^{-7}$

Table 12: Expectation, bias and MSE of the MLE for ρ and $n = 3$ (equidistant designs, 1.000.000 simulated samples each)

true parameter ρ. We now confine ourselves to the case $n = 3$ (the case $n = 2$, where the MLE $\hat{\rho}$ is the most biased, is studied in Waldl, 2011). For $n > 3$ the following may be done analogously. In Figure 21 we see the expectations $E(\hat{\rho}|\rho)$ (right panel) and the expected bias for different values of ρ (left panel). From the left panel we also see, that the MLE performs quite well if $\rho > 0.65$.

For the bias reduction we choose a Bayesian approach: Since we have no information about ρ we have to assume ρ to be distributed uniformly on $(0; 1)$. Under this assumption we try to find an unbiased estimate of ρ, i.e. we are looking for an estimate $\tilde{\rho}$ on the basis of the MLE $\hat{\rho}$ with $E(\tilde{\rho}) = \rho$ via the expectation $E(\rho|\hat{\rho})$.

From the simulations we know the empirical distributions of $\hat{\rho}$ for a given sequence of discrete ρ-values ρ_i, i.e. $\mathcal{P}(\hat{\rho}|\rho_i)$. To get an approximation of the distribution of ρ for a given $\hat{\rho} = \rho_0$ we have to compute approximations of the pdf $f(\frac{1}{2}(\rho_{i-1} + \rho_i)|\hat{\rho} = \rho_0)$ first. This may be done by:

$$\hat{P}(\rho \in [\rho_{i-1}, \rho_i]|\hat{\rho} = \rho_0) \propto \frac{1}{2} \left(P(\hat{\rho} \in [\rho_0 \pm \Delta]|\rho_{i-1}) + P(\hat{\rho} \in [\rho_0 \pm \Delta]|\rho_i)\right) \cdot (\rho_i - \rho_{i-1})$$

for $\rho_0 = 0$ we may set $\Delta = 0$.

For the computation of the expectation $E(\rho|\hat{\rho} = \rho_0)$ the discrete approximation of $\int_0^1 \rho \cdot f(\rho|\hat{\rho} = \rho_0)\, d\rho$ will be used. By doing this e.g. for $n = 3$ and $\rho_0 = 0$ we get: $E(\rho|\hat{\rho} = 0) = 0.2493$, i.e. if we get the MLE $\hat{\rho} = 0$ and have no additional information about ρ, we expect ρ to be 0.2493. Figure 21 shows the expectation of $\mathcal{P}(\rho|\hat{\rho})$ when ρ is uniformly distributed on $(0; 1)$. This conditional expectation is bounded on the interval $[0.24; 0.995]$ although $\rho \in [0; 1]$, the reason for this is the big dispersion of $\mathcal{P}(\hat{\rho}|\rho)$.

Now an unbiased estimate $\tilde{\rho}$ may easily be defined:

$$\tilde{\rho} := \hat{E}(\rho|\hat{\rho})$$

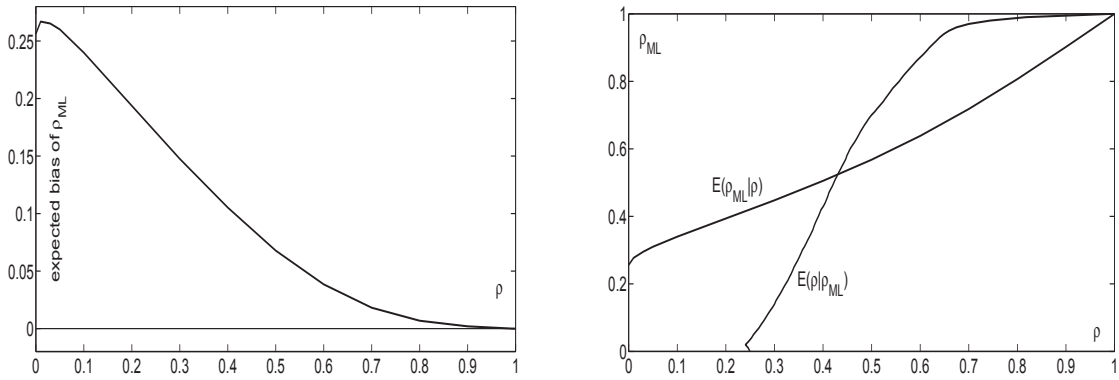

Figure 21: Expected bias of $\hat{\rho}$ for $n = 3$ and different ρ-values (left panel). Expectations $E(\hat{\rho}|\rho)$ and $E(\rho|\hat{\rho})$ with ρ uniformly distributed for $n = 3$ (right panel)

5.1.4. Properties of $\tilde{\rho}$

To compare the characteristics of this unbiased estimator and the MLE, 10.000.000 uniformly distributed ρ-values were simulated and for each of the ρ-values we simulated a sample of size 3 according to the GE model and computed the MLE for ρ and the unbiased $\tilde{\rho}$. The empirical distributions of the biases of these estimates can be seen in Figure 22. From the boundedness of $\tilde{\rho}$ it is clear that the bias of $\tilde{\rho}$ is also bounded on $[1 - 0.24; .995] = [-.76; .995]$. Besides we have discontinuities in the empirical pdf of $\tilde{\rho}$ in 0.2493 and 0. The empirical pdf of the distribution of $\hat{\rho}$ has a discontinuity in 0.

The mean bias of the MLE was $E(\hat{\rho} - \rho) = 0.0953$. The mean bias of the unbiased estimator was $E(\tilde{\rho} - \rho) = 0.0014$, the deviation from 0 is due to the inaccuracy of the approximations. The mean square errors of the two estimates are:

$$MSE(\hat{\rho}) = 0.0967 \qquad MSE(\tilde{\rho}) = 0.0535$$

So also from this point of view $\tilde{\rho}$ clearly improves $\hat{\rho}$.

$\tilde{\rho}$ is also better than the trivial unbiased estimator $\breve{\rho} = 0.5$ (no matter what data are observed) with $MSE(\breve{\rho}) = 0.0833$. It is remarkable that for $n = 3$ the trivial estimate $\breve{\rho} = 0.5$ is better than the MLE $\hat{\rho}$.

6. Conclusions

Probably the main lesson we can learn is that the D-optimal design is analytically peculiar and these designs can be practically obtained only by numerical computation. However, especially two point locally D-optimal designs are of interest, since they may help us to find a reasonable range for a bigger design, particularly for filling designs. The latter ones are probably the only applicable when seeking for a higher number of design points.

It is an interesting issue that very often the best designs do not use the whole design interval, but only a part of it. This idea should be taken into account by practitioners when designing experiments. The best interval for taking the observations in depends on the initial interval we are using and the value of the correlation parameter r. In general, it can be said that when the

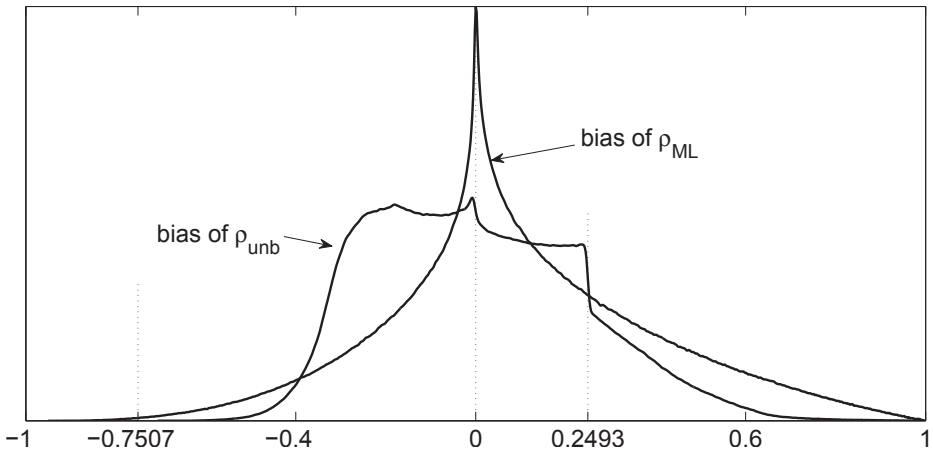

Figure 22: Empirical distributions of the biases of $\hat{\rho}$ and $\tilde{\rho}$ for $n = 3$ and uniformly distributed ρ

correlation is high (small r) a good experiment should take samples within the whole interval, while for independent observations the sensible choice would be concentrate the samples in a fraction of it. The question whether to consider or not r as a parameter that needs to be estimated can also be important for some cases (e.g. equidistant designs).

Another important observation is the large bias of the ML estimator of r. From the theoretical point of view this is not surprising since σ^2 and r are not simultaneously identifiable. Therefore we developed bias reduction method which can help practitioners to obtain better locally D-optimal designs.

Appendix

Proof of Proposition 1

The determinant of the information matrix will be in this case:

$$D_s = \det(x, d) = a^4 d^2 e^{-6\beta(d+x)} x^{2m} (d+x)^{2m} (2d+x)^{2m} [\log(x) - 2\log(d+x) + \log(2d+x)]^2$$

We obtain the critical points solving the following system of equations:

$$2a^4 d^2 e^{-6\beta(d+x)} x^{2m-1} (d+x)^{2m-1} (2d+x)^{2m-1} (\log(x) - 2\log(d+x) + \log(2d+x))$$

$$\left(2d^2 + \left(2(m-3\beta x)d^2 + 3x(2m-3\beta x)d + 3x^2(m-\beta x)\right)(\log(x) - 2\log(d+x) + \log(2d+x))\right) = 0$$

$$-2a^4 d e^{-6\beta(d+x)} x^{2m} (d+x)^{2m-1} (2d+x)^{2m-1} (\log(x) - 2\log(d+x) + \log(2d+x))$$

$$\left(2d^2 + \left(6\beta d^3 + (-4m+9\beta x-2)d^2 + 3x(-m+\beta x-1)d - x^2\right)(\log(x) - 2\log(d+x) + \log(2d+x))\right) = 0$$

The solutions $x = 0$, $d = 0$, $x = -d$ or $x = -2d$ do not maximize the determinant. The points giving the maximum verify the equations:

$$-2d^2 = \left(2(m-3\beta x)d^2 + 3x(2m-3\beta x)d + 3x^2(m-\beta x)\right)(\log(x) - 2\log(d+x) + \log(2d+x))$$

$$-2d^2 = \left(6\beta d^3 + (-4m + 9\beta x - 2)d^2 + 3x(-m + \beta x - 1)d - x^2\right)(\log(x) - 2\log(d+x) + \log(2d+x))$$

□

Proof of Theorem 1

We have $X = (\nabla\eta(x,\beta)^t, \nabla\eta(x+d,\beta)^t) = \begin{pmatrix} x^m e^{-\beta x} & -ax^{m+1}e^{-\beta x} \\ (x+d)^m e^{-\beta(x+d)} & -a(x+d)^{m+1}e^{-\beta(x+d)} \end{pmatrix}$

Then $M_\beta(2) = X^T C^{-1}(r) X$.

So for both parameters of interest we have $M(2)(\theta, r) = \begin{pmatrix} M_\theta(2) & 0 \\ 0 & M_r(2) \end{pmatrix}$ where $M_r(2) = d^2 \frac{e^{2rd}+1}{(e^{2rd}-1)^2}$ (see Kiseľák and Stehlík, 2008), with determinant

$$\det[M(2)(\theta, r)] = \frac{a^2 d^4 e^{2dr - 2\beta(d+2x)} \left(1 + e^{2dr}\right) x^{2m} (d+x)^{2m}}{\left(-1 + e^{2dr}\right)^3} \tag{4}$$

or equivalently

$$\det[M(2)(\theta, r)] = \frac{1}{4} a^2 d^4 e^{-2\beta(d+2x)} x^{2m} (d+x)^{2m} \coth(dr) csch^2(dr)$$

When we fix x the value of the determinant is maximized for $d = d_x^*$, the solution of the non-linear equation

$$d = \frac{-B + \sqrt{B^2 - 4AZ}}{2A}$$

where

$$A = \beta\left(-1 + e^{4dr}\right) + \left(1 + 4e^{2dr} + e^{4dr}\right) r \quad ,$$

$$B = m - \beta x + 4e^{2dr} rx + rx + e^{4dr}(-m + \beta x + rx - 2) + 2 \quad ,$$

$$Z = -2\left(-1 + e^{4dr}\right) x \quad .$$

However, a general solution can be found if, conversely, d is assumed initially fixed. In that case $\det[M(2)]$ is a continuous function in x with two stationary points given by

$$\frac{-\beta d + m \pm \sqrt{\beta^2 d^2 + m^2}}{2b} \quad .$$

Since the smallest one is not positive, the value of x that maximizes the determinant for a specific d is

$$x^* = \frac{-\beta d + m + \sqrt{\beta^2 d^2 + m^2}}{2b} \quad ,$$

with a similar structure that the corresponding points in Rodríguez-Torreblanca and Rodríguez-Díaz (2007) or Rodríguez-Díaz and Santos-Martín (2009), Theorem 1. Replacing this expression in (4) the optimal value of d, d^*, is the solution of the non-linear equation (2).

□

Acknowledgement

We acknowledge the helpful discussion with Licesio J. Rodríguez-Aragón from the University of Castilla La Mancha (Spain), and the useful comments from Prof. Šárka Klementová from Dept. of Mathematics, Physics and Chemistry, Univ. of South Bohemia. Research was supported by

'Acciones Integradas 2008-09 (Project Nr. ES 18/2008), Aktion 53p19, the Spanish Junta de Castilla y León (Project 'SA071A09') and the Spanish Ministry of Education and Science (Projects 'MTM 2007-672111-C03-01' and 'Ingenio Mathematica (i-MATH)' No. CSD2006-00032, Consolider - Ingenio 2010).

References

Abt, M. and Welch, W.J. 1998. Fisher information andmaximum-likelihood estimation of covariance parameters in Gaussian stochastic processes. The Canadian Journal of Statistics 26, 127-137.

Bonard A., Daële V., Delfau J.-L. and Vovelle C. 2002. Kinetics of OH radical reactions with methane in the temperature range 295-660 K and with dimethyl ether and methyl-tert-butyl ether in the temperature range 295-618 K. J. Phys. Chem. A 106, 4384-4389.

Dette H., Kunert J. and Pepelyshev A. 2008. Exact optimal designs for weighted least squares analysis with correlated errors. Statistica Sinica 18, 135-154.

Dette, H. and Sperlich, S. 1994. A note on Bayesian D-optimal designs for a generalization of the exponential growth model. South African Statistical Journal 28, 103-117.

Dunlop J.R. and Tully F.P. 1993. A kinetic study of OH radical reactions with methane and perdeuterated methane. J. Phys. Chem. 97, 11148-11150.

Finlayson-Pitts B.J., Ezell M.J., Jayaweera T.M., Berko H.N. and Lai C.C. 1992. Kinetics of the reactions of OH with methyl chloroform and methane – Implications for global tropospheric OH and the methane budget. Geophys. Res. Lett.19, 13, 1371-1374.

Gierczak T., Talukdar R.K., Herndon S.C., Vaghjiani G.L. and Ravishankara A.R. 1997. Rate coefficients for the reactions of hydroxyl radical with methane and deuterated methanes. J. Phys. Chem. A 101, 3125-3134.

Héberger K., Kemény S. and Vidóczy T. (1987). On the errors of Arrhenius Parameters and Estimated Rate Constant Values. International Journal of Chemical Kinetics 19, 171-178.

Hill P.D.H. 1980. D-optimal designs for partially nonlinear regression models. Technometrics 22, 275-276.

International Union of Pure and Applied Chemistry (IUPAC). Transition State Theory. http://goldbook.iupac.org/T06470.html (accessed November 23, 2008)

Jet Propulsion Laboratory 2006. Chemical kinetics and photochemical data for use in atmospheric studies. Evaluation Number 15. NASA Panel for Data Evaluation, California Institute of Technology, Pasadena, California.

Kim T.Y., Park J.S. and Song G.M. 2010. An asymptotic theory for the nugget estimator in spatial models. Journal of Nonparametric Statistics 22, 2, 181-195.

Kiseľák J. and Stehlík M. 2008. Equidistant D-optimal designs for parameters of Ornstein-Uhlenbeck process. Statistics and Probability Letters 78, 1388-1396.

Laidler K. J. 1984. The Development of the Arrhenius Equation. Journal of Chemical Education 61, 494-498.

López-Fidalgo J. and Wong W.K. 2002. Design Issues for the Michaelis-Menten Model. Journal of Theoretical Biology 215, 1-11.

Mellouki A., Téton S., Laverdet G., Quilgars A., and LeBras G. 1994. Kinetic studies of OH reactions with H_2O_2, C_3H_8 and CH_4 using the pulsed laser photolysis - laser induced fluorescence method. J. Chim. Phys. 91, 473.

Müller W.G. and Stehlík M. 2009. Issues in the Optimal Design of Computer Simulation Experiments. Appl. Stochastic Models Bus. Ind. 25, 163-177.

Müller W.G. and Stehlík M. 2010. Compound optimal spatial designs. Environmetrics 21, 354-364.

Olkin, I. and Pratt, J. W. 1958. Unbiased estimation of certain correlation coefficients. Annals of Mathematical Statistics 29, 201-211.

Pázman. 2007. Criteria for Optimal Design for Small-Sample Experiments with Correlated Observations. Kybernetika 43, 4, 453-462.

Pepelyshev A. 2007. Optimal designs for the exponential model with correlated observations, In: López-Fidalgo, J., Rodríguez-Díaz, J.M., Torsney, B. (Eds.), Proc. mODa8. 165-171.

Rodríguez-Aragón L.J. and López-Fidalgo J. 2005. Optimal designs for the Arrhenius equation. Chemometr. Intell. Lab. Syst. 77, 131-138.

Rodríguez-Díaz J.M. and Santos-Martín M.T. 2009. Study of the best designs for modifications of the Arrhenius equation. em Chemometr. Intell. Lab. Syst. 95, 199-208.

Rodríguez-Díaz J.M., Santos-Martín M.T., Stehlík M. and Waldl H. 2009. Filling and D-optimal designs for the correlated Generalized Exponential models. Ifas Research Report.

J.M. Rodríguez-Díaz, T. Santos-Martín, M. Stehlík and H. Waldl. 2011. Filling and D-optimal designs for the correlated Generalized Exponential models - extended version, Research Report, Department of Applied Statistics, Johannes Kepler University, Linz.

Rodríguez-Torreblanca C. and Rodríguez-Díaz J.M. 2007. Locally D- and c-optimal designs for Poisson and Negative Binomial regression models. Metrika 66, 161-172.

Stehlík M. 2007. D-optimal designs and equidistant designs for stationary processes. In: López-Fidalgo, J., Rodríguez-Díaz, J.M., Torsney, B. (Eds.), Proc. mODa8 205-212.

Stehlík, M., Rodríguez-Díaz, J.M., Müller, W.G. and López-Fidalgo, J. 2007. Optimal allocation of bioassays in the case of parametrized covariance functions: an application in lung's retention of radioactive particles. Test 17, 56-68.

Vaghjiani G.L. and Ravishankara A.R. 1991. New measurement of the rate coefficient for the reaction of OH with methane. Nature 350, 406-409.

H. Waldl. 2011. On the ML Estimation of the Correlation Parameter in Generalized Exponential Models. In: T.E. Simos (Eds.), Proc. ICNAAM 2011.

Zagoraiou M. and Baldi-Antognini A. 2009 Optimal designs for parameter estimation of the Ornstein-Uhlenbeck process. Appl. Stochastic Models Bus. Ind.