

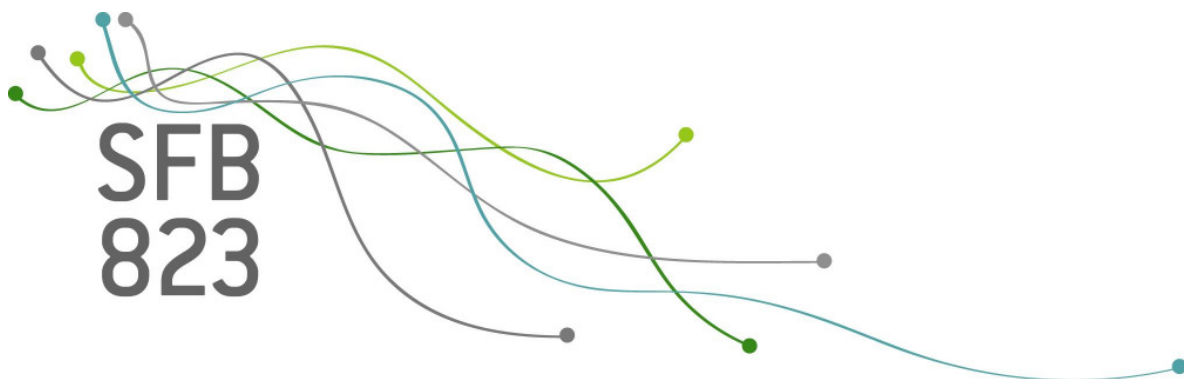
SFB  
823

# Generalized latin hypercube design for computer experiments

Holger Dette, Andrey Pepelyshev

Nr. 24/2009

Discussion Paper





# Generalized Latin hypercube design for computer experiments

Holger Dette

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

e-mail: holger.dette@rub.de

Andrey Pepelyshev

Sheffield University

Department of Probability & Statistics

Sheffield, UK

email: andrey@ap7236.spb.edu

September 16, 2009

## Abstract

Space filling designs, which satisfy a uniformity property, are widely used in computer experiments. In the present paper the performance of non-uniform experimental designs which locate more points in a neighborhood of the boundary of the design space is investigated. These designs are obtained by a quantile transformation of the one-dimensional projections of commonly used space filling designs. This transformation is motivated by logarithmic potential theory, which yields the arc-sine measure as equilibrium distribution. Alternative distance measures yield to Beta distributions, which put more weight in the interior of the design space. The methodology is illustrated for maximin Latin hypercube designs in several examples. In particular it is demonstrated that in many cases the new designs yield a smaller integrated mean square error for prediction. Moreover, the new designs yield to substantially better performance with respect to the entropy criterion.

Keywords and phrases: design for computer experiments, space filling designs, Latin hypercube designs, logarithmic potential, arc-sine distribution

# 1 Introduction

In modern scientific studies, mathematical computer models are widely used to describe complex processes. These models are a replacement for real (physical, chemical, biological, etc.) experiments which are too time consuming or too costly. Typical examples of computer models include the flow of air over an airplane wing, behavior of a metal structure under stress, combustion of gasses in a flame, VLSI-circuit and engine design, plan ecology, thermal-energy storage, health medical models and many other processes. Moreover, mathematical models may describe phenomena which can not be reproduced, for example, weather or climate change. Computer experiments usually involve a large number of input variables and are typically extremely time consuming. One run of a computer experiment may require several hours or longer. Therefore, the choice of an appropriate set of input conditions is of particular importance for making inference on the basis of a computer model. In the literature this rapidly growing field is called design for computer experiments [see Sacks, Welch, Mitchell and Wynn (1989), Santner, Williams and Notz (2003), Fang, Li, Sudjianto (2006) or Kleijnen (2008) among many others].

For computer experiments space-filling designs are very popular. The class of space-filling designs includes a wide variety of different types of designs such as Latin hypercube designs (LHD), designs based on random or pseudo-random sequences, sphere packing designs and others. Most of these designs are uniformly space-filling in the sense that the design points are uniformly scattered on the domain.

The purpose of the present note is to give some motivation for choosing non-uniform space-filling designs in computer experiments. Our approach applies an average measure of inter-point distances to each marginal projection of a space filling design. In particular, we consider Latin hypercube designs with marginals minimizing a logarithmic potential. This optimization problem yields to the arc-sine distribution and we propose to transform the marginals of maximin Latin hypercube designs by the corresponding quantile function. The transformed designs are called generalized maximin Latin hypercube designs and we show their efficiency for prediction with respect to the mean squared and entropy criterion in several numerical examples. Other functionals yield to Beta distributions, which are an interesting alternative if the prediction in the interior of the design space is of more interest.

The remaining part of the paper is organized as follows. In Section 2 we review some notations from the design of computer experiments and state two optimization problems, which have Beta distributions as solution. These results provide the motivation

for introducing generalized maximin Latin hypercube designs, which are obtained by transforming the one-dimensional projections of a maximin Latin hypercube design by the quantile function of a Beta-distribution. In particular, we consider the arc-sine distribution and the generalized Latin hypercube designs derived from this transformation which are compared with commonly used designs in several numerical examples in Section 3. Finally, some conclusions are given in Section 4.

## 2 Efficient designs for computer experiments

In computer experiments, one of the primary objectives is to construct a meta-model for the estimation of the response at untried inputs. Correspondingly, appropriate optimality criteria for choosing a design in computer experiments should be based on the error of prediction. A very popular optimality criterion is the minimization of the mean square error over the design space. This criterion is adopted to lead to a design which should fill the entire design space uniformly at the initial stage of computer experiments. The examples of space-filling designs include Latin hypercube designs, sphere packing designs, distance based designs, uniform designs, minimal discrepancy designs, designs based on random or pseudo-random sequences, see Santner et al. (2003), Fang et al. (2006). A “good” design should yield a dense set in the projection on each coordinate and at the same time form a dense set in entire design space. By construction, Latin hypercube designs have projections that are spread out evenly over the values of each input variable. Unfortunately, these designs may have a poor filling of the entire hypercube (in the extreme case - if the design space is a square - all points are evenly spread out of the diagonal of this square); see e.g. Liefvendahl, Stocki (2006), Pepelyshev (2009). Several criteria of optimality are introduced in order to choose a “good” design in a class of Latin hypercube designs. Most criteria are based on a measure which quantifies how spread out a set of points is. A typical example is the maximin criterion, which determines the Latin hypercube design maximizing the minimal distance

$$\Psi_p(L) = \min_{\substack{i,j=1,\dots,n \\ i \neq j}} \|x_i - x_j\|_p = \min_{\substack{i,j=1,\dots,n \\ i \neq j}} \left( \sum_{s=1}^d |x_{s,i} - x_{s,j}|^p \right)^{1/p}, \quad (1)$$

where  $d$  denotes the dimension of the  $n$  input variables  $x_1, \dots, x_n$  and  $x_i = (x_{1,i}, \dots, x_{d,i})^T$  is  $i$ th point of the design  $L = \{x_1, \dots, x_n\}$ . Usually this criterion is used with  $p = 2$  for computational reasons or the isotropic argument. The maximin objective for the choice of points have been shown to be useful first in Niederreiter (1983).

A Latin hypercube design, which maximizes  $\Psi_p(L)$  in the class of all LHD, is called a maximin Latin hypercube design (MLHD). An alternative criterion was introduced by Audze, Eglaiss (1977) and is based on the analogy of minimizing forces between charged particles. These authors proposed the minimization of the function

$$\Psi_{AE}(L) = \sum_{i=1}^n \sum_{j=i+1}^n \frac{1}{\|x_i - x_j\|_2^2}. \quad (2)$$

Other criteria of uniformity are star  $L_2$ -discrepancy, centered  $L_2$ -discrepancy, wrap-around  $L_2$ -discrepancy which are motivated by quasi-Monte-Carlo methods and the Koksma-Hlawka inequality [see Hickernell (1998), Fang et al. (2000, 2002)].

Uniform designs are particularly important for inference about the mean  $\int G(x)dx$  of some known arbitrary integrable function  $G$  over the design space. In this case the sample mean of  $G$  computed from a uniform design minimizes a certain bound for the absolute error from the true mean [see for example Hickernell (1998)]. However in most computer experiments, one is typically not interested in estimating the mean of some function  $G$ . In many cases the output of a computer experiment is used for the construction of a meta-model for prediction. In this model the parameters have to be estimated and this objective is not taken into account by the choice of uniform designs. In order to address the problem of prediction and fitting, we propose to choose designs points which are more densely located near the boundary of the design space. In the following discussion we define a transformation for the projections of a Latin hypercube design. Our proposal is motivated by the following observations.

First, if for a (one dimensional) uniform design a sphere with center in the middle of the design space contains  $k$  points then, a sphere with the same radius and a center near the boundary of the design space contains fewer points. This effect is increasing with the dimension of the design space, which means that at the boundary of the design space there is substantially less information for prediction.

Second, note that in a Gaussian process analysis the uncertainty of the prediction at the point  $x$  is essentially of the form

$$V(x) = 1 - t^T(x)R^{-1}t(x),$$

where  $t(x) = (r(x, x_1), \dots, r(x, x_n))^T$  is the vector of correlations between the point  $x$  and the design points  $x_1, \dots, x_n$  and  $r(\cdot)$  is a given correlation function. For a one-dimensional input variable the typical form of the function  $V(x)$  with Gaussian correlation function  $r(\cdot)$  for the design points  $0, 0.1, \dots, 1$  is given in Figure 1. We observe that the uncertainty is larger near the boundary of the design space. Therefore,

to reduce the uncertainty (for example in a minimax approach) the design points should be located more densely near the boundary.

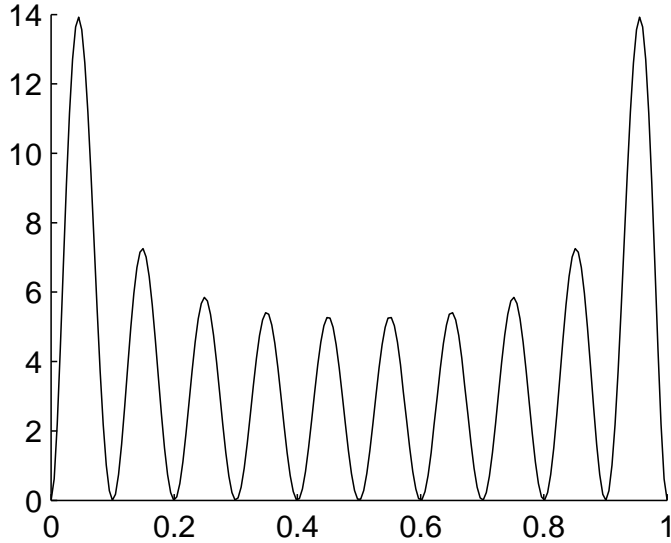


Figure 1: *The typical form of the uncertainty in Bayesian analysis for equidistant design points in the one-dimensional case.*

We now give a motivation for the specific choice of the transformation, which is slightly more formal than the heuristic arguments of the previous paragraphs. Recall the type of criterion in (2) and define a family of criteria by

$$\psi_a(L_{(s)}) = \begin{cases} \sum_{i=1}^n \sum_{j=i+1}^n \frac{1}{|z_i - z_j|^a} & \text{if } a \in (0, 1) \\ \sum_{i=1}^n \sum_{j=i+1}^n -\log |z_i - z_j| & \text{if } a = 0 \end{cases} \quad (3)$$

where  $L_{(s)} = \{z_1, \dots, z_n\}$  with  $z_i = x_{s,i}$  ( $s = 1, \dots, d$ ) denotes the projection of the initial design on the  $s$ th coordinate. Arguing as before, it is reasonable for the computer experiment to choose the design  $L = \{x_1, \dots, x_n\}$  which has one-dimensional marginals  $L_{(s)} = \{x_{j,1}, \dots, x_{j,n}\}$  minimizing the criterion (3). The solution to the case  $a = 0$  in (3) is well known [see Szegö (1975)] and given by the roots of the polynomial

$$z(z-1)P'_{n-1}(2z-1),$$

where  $P_{n-1}(z)$  denotes the  $(n-1)$ th Legendre polynomial orthogonal with respect to the Lebesgue measure on the interval  $[-1, 1]$ . Note that the uniform distribution on this

set is the  $D$ -optimal approximate design for a polynomial regression model of degree  $n$  [see Pukelsheim (2006)]. In particular it can be shown that for increasing sample size  $n$  the uniform distribution on this set converges weakly to the arc-sine distribution with density

$$p_{\frac{1}{2}}(x) = \frac{1}{\pi\sqrt{x(1-x)}}, \quad (4)$$

[see for example Dette and Studden (1995)].

The design  $\{z_1, \dots, z_n\}$  minimizing (3) for  $a \in (0, 1)$  can not be determined explicitly. However an approximate solution is possible by the following argument. If  $z_1, \dots, z_n$  are interpreted as realizations of independent identically distributed random variables with distribution  $\mu$ , then (with an appropriate normalization)  $\binom{n}{2}^{-1} \psi_a(L_{(s)})$  can be considered as an approximation of the integral

$$\phi_a(\mu) = \begin{cases} \int_0^1 \int_0^1 \frac{1}{|x-y|^a} \mu(dy)\mu(dx) & \text{if } a \in (0, 1] \\ - \int_0^1 \int_0^1 \log|x-y| \mu(dy)\mu(dx) & \text{if } a = 0 \end{cases} \quad (5)$$

It has recently been shown [see Dette, Pepelyshev, Zhigljavsky (2009)] that  $\phi_a$  is minimal for the Beta distribution

$$p_{\frac{1+a}{2}}(x) = \frac{1}{B(\frac{1+a}{2}, \frac{1+a}{2})} x^{\frac{a-1}{2}} (1-x)^{\frac{a-1}{2}} ; \quad a \in [0, 1]. \quad (6)$$

Note that for the choice  $a = 0$  we obtain the arc-sine distribution, which appears as the equilibrium distribution of the logarithmic potential [see Saff and Totik (1997)]. By the argument of the previous paragraph the arc-sine distribution is also the limit distribution of the  $D$ -optimal design for large degree polynomial regression [see Pukelsheim (2006), Ch. 9.5]. Moreover, Wiens and Zhou (1997) showed that  $V$ -optimal designs for linear models, which are robust against small departures from the assumed regression, converge weakly to the arc-sine distribution.

Our strategy of constructing a non-uniform design is now the following.

- (A) We generate Latin hypercube designs. For this purpose we recall an algorithm for the construction of Latin hypercube design, which was introduced in McKay et al. (1979). This algorithm generates  $n$  points in dimension  $d$  in the following manner.



- (1) Generate  $n$  uniform equidistant points  $x_{s,1}, \dots, x_{s,n}$  in the range of each input variable,  $s = 1, \dots, d$ . Without loss of generality, we assume that the range of each input is the interval  $[0, 1]$  and  $x_{s,j} \in \mathcal{R} = \{0, 1/(n-1), 2/(n-1), \dots, 1\}$ .
  - (2) Generate a matrix  $(p_{i,j})$  of size  $d \times n$  such that each row is a random permutation of the numbers  $1, \dots, n$  and these permutations are independent.
  - (3) Each column of the matrix  $(p_{i,j})$  corresponds to a design point, that is  $(x_{1,p_{1,j}}, \dots, x_{d,p_{d,j}})^T$  is the  $j$ th point of the Latin hypercube design.
- (B) We determine a maximin hypercube design maximizing the criterion (1). Several algorithms for this purpose can be found in Liefvendahl and Stocki (2006), Jin et al. (2005), Fang and Qin (2003) among many others.
- (C) For the definition of a generalized (non-uniform) design supported at the points  $z_1, \dots, z_n$ ,  $z_i = (z_{1,i}, \dots, z_{d,i})$ , we assume that the points  $x_{j,i}$  of the one-dimensional projection of a maximin Latin hypercube design are transformed by the quantile function of a continuous density  $p_{\frac{1+a}{2}}(t)$ , which minimizes the criterion (5), namely,  $z_{j,i}$  is a solution of the equation

$$x_{j,i} = \int_0^{z_{j,i}} p_{\frac{1+a}{2}}(t) dt$$

$i = 1, \dots, n$ ,  $j = 1, \dots, d$ . Note that the equally-spaced points correspond to the choice  $p_1(t) \equiv 1$ .

The resulting design will be called generalized maximin Latin hypercube design throughout this paper and by the previous discussion we expect that a good choice for the distribution in step (C) of the procedure is to use one of the densities  $\{p_{\frac{1+a}{2}}\}_{a \in [0,1]}$ . The parameter  $a$  can be considered as a tuning parameter, which specifies the importance of the boundary. The larger the parameter  $a \in [0, 1]$ , the more mass is put at the boundary of the interval  $[0, 1]$ . Note that the extreme case  $a = 1$  yields the uniform distribution, while the other extreme case is  $a = 0$  corresponding to the arc-sine distribution. In the following discussion we will concentrate on the arc-sine measure  $p_{\frac{1}{2}}$  which minimizes the logarithmic potential. For this density we obtain

$$z_{j,i} = \left( \cos \left( x_{j,i} \pi \right) + 1 \right) / 2, \quad i = 1, \dots, n.$$

In particular, we have

$$\int_0^{1/4} p_{\frac{1}{2}}(t) dt = \int_{1/2-1/4}^{1/2+1/4} p_{\frac{1}{2}}(t) dt = 1/3,$$

which indicates that – in contrast to the uniform distribution – the arc-sine distribution puts more mass in the neighbourhood of the boundary of the design space. In the following section we investigate the performance of these designs in several numerical examples.

### 3 Examples

In this section we compare two designs: the maximin Latin hypercube design (MLHD) and the generalized maximin Latin hypercube design (GMLHD) obtained from MLHD by applying the arc-sine transformation. We downloaded maximin Latin hypercube designs from the website <http://www.spacefillingdesigns.nl/>. The following illustrative examples of models are not hard time consuming computer models. However, they have a typical form of the response and allow to compute the accuracy of the meta-model for different designs.

We use the Bayesian framework, in which the meta-model is the mean of a posterior Gaussian process [see Sacks et al. (1989), Kennedy and O’Hagan (2001)]. More precisely, we consider a Gaussian process with linear mean and the Gaussian correlation function

$$r(x) = e^{-\sum_{s=1}^d x_s^2/\psi_s}$$

where  $\psi = (\psi_1, \dots, \psi_d)$  is a vector of correlation parameters. In the following examples, we estimate the hyper-parameter of the Gaussian process for each dataset by the algorithm given in Kennedy and O’Hagan (2001) and Bastos and O’Hagan (2008). Specifically, the parameters of the mean and variance function are determined by the integration of the likelihood with respect to the noninformative prior and the estimate of  $\psi$  is determined by the maximization of the posterior density.

To study the performance of the meta-model  $\hat{\eta}(x)$  for different designs, we use three criteria, two mean square error (MSE) criteria and an entropy criterion. The first mean square error criterion is defined by

$$\text{MSE}(L) = \int_{[0,1]^d} (\eta(x) - \hat{\eta}(x))^2 w(x) dx, \quad (7)$$

where  $L$  is the given design,  $w(x)$  is a weight function and  $\eta(x)$  is a computer model. For the sake of brevity we assume that  $w(x) \equiv 1$  and calculate the mean square error by Monte-Carlo sampling, that is

$$\text{MSE}_{[0,1]^d} = \frac{1}{N} \sum_{k=1}^N (\eta(x_k) - \hat{\eta}(x_k))^2$$

with  $N = 10000$ , where the points  $x_k$  are randomly chosen from  $[0, 1]^d$ . We also calculate the mean square error over  $2^d$  corner points, that is

$$\text{MSE}_{\{0,1\}^d} = \frac{1}{2^d} \sum_{x \in \{0,1\}^d} (\eta(x) - \hat{\eta}(x))^2,$$

and consider the entropy criterion, which is defined by

$$E = \log \det R, \tag{8}$$

where  $R = (r(x_i - x_j))_{i,j}$  is a correlation matrix for the design points. Note that several authors chose to maximize the entropy criterion in order to determine a space-filling design, see e.g. Sacks et al. (1989), Sebastiani and Wynn (2000).

### 3.1 Two-dimensional rational model

We start with a two-dimensional model which has the typical form of models in computer experiments and was studied recently in Bastos and O'Hagan (2008). The model is defined by

$$\eta(x) = (1 - e^{-0.5/x_2}) \frac{2300x_1^3 + 1900x_1^2 + 2092x_1 + 60}{100x_1^3 + 500x_1^2 + 4x_1 + 20}, \tag{9}$$

where the input variable  $x = (x_1, x_2)$  varies in the square  $[0, 1]^2$ . The results of the simulation study are depicted in Table 1, where we show the entropy and the square root of the simulated mean square error obtained by the different designs. From this table we see that the mean square error obtained by the MLHD is substantially larger than the mean square error obtained by the GMLHD. The entropy for the MLHD is smaller than the entropy obtained by the GMLHD. The improvement with respect to  $\text{MSE}_{[0,1]^2}$  varies between 53% ( $n = 30$ ) to 57% ( $n = 70$ ). The improvement with respect to the entropy criterion is about 15%. The improvement with respect to the accuracy of the prediction at the boundary, measured by  $\text{MSE}_{\{0,1\}^d}$ , is even larger (which is obvious by the construction of the GMLHD). Thus, in model (9) the GMLHD is substantially more efficient than the MLHD with respect to all criteria.

### 3.2 Borehole model

In this subsection we examine designs for a borehole model which was investigated in Worley (1987), Morris and Mitchell (1995), Ahn and Owen (2001), Fang and Lin

Table 1: Square root of mean square error and entropy for MLHD and GMLHD with different  $n$  for model (9).

	$n$	30	40	50	60	70
MLHD	$\text{MSE}_{[0,1]^2}$	0.2658	0.142	0.061	0.075	0.085
	$\text{MSE}_{\{0,1\}^2}$	1.28	1.10	0.57	0.51	0.57
	$E$	-53.415	-97.96	-113.4	-166.3	-98.26
GMLHD	$\text{MSE}_{[0,1]^2}$	0.1248	0.042	0.038	0.042	0.037
	$\text{MSE}_{\{0,1\}^2}$	0.13	0.12	0.06	0.05	0.06
	$E$	-44.337	-82.78	-91.22	-140.6	-85.47

(2003). The borehole model describes the flow rate through the borehole that is drilled from the ground surface through two aquifers. The flow rate is given by

$$\eta(x) = \frac{2\pi(H_u - H_l)}{\ln(r/r_w) \left( 1 + \frac{2LT_u}{\ln(r/r_w)r_w^2 K_w} + \frac{T_u}{T_l} \right)},$$

where the dimension of the vector of input variables is 8, that is

$$x = (r_w, r, T_u, T_l, H_u, H_l, L, K_w).$$

The input  $r_w \in [0.05, 0.15]$  is the radius of the borehole,  $r \in [100, 50000]$  is the radius of the influence,  $T_u \in [63070, 115600]$  is the transmissivity of the upper aquifer,  $T_l \in [63.1, 116]$  is the transmissivity of the lower aquifer,  $H_u \in [990, 1110]$  is the potentiometric head of the upper aquifer,  $H_l \in [700, 820]$  is the potentiometric head of the lower aquifer,  $L \in [1120, 1680]$  is the length of the borehole and  $K_w \in [9855, 12045]$  is the hydraulic conductivity of the borehole. The simulated square root of the mean square error is depicted in Table 2 and we see that the mean square error for the MLHD is a little smaller than the mean square error for the GMLHD if  $n = 60, 70, 80$ . This observation can be explained by the very sparse placement of (at most) 80 design points in a 8-dimensional design space.

The mean square error for the MLHD becomes larger than the mean square error for the GMLHD if the number of design points is increasing and in this case we observe a similar performance as in the 2-dimensional example. The mean square error at the corner points  $\text{MSE}_{\{0,1\}^d}$  is substantially larger for the MLHD than for the GMLHD. Similarly, the GMLHD performs substantially better than the MLHD with respect to entropy criterion.

Table 2: Square root of the mean square error and entropy for MLHD and GMLHD with different  $n$  for the borehole model.

	$n$	60	70	80	90	140	200
MLHD	$\text{MSE}_{[0,1]^s}$	1.05	0.90	0.61	0.53	0.51	0.37
	$\text{MSE}_{\{0,1\}^s}$	5.04	4.21	3.26	3.45	3.89	2.81
	$E$	-157.6	-193.8	-266.8	-273.2	-386	-561.2
GMLHD	$\text{MSE}_{[0,1]^s}$	1.21	1.04	0.65	0.54	0.48	0.32
	$\text{MSE}_{\{0,1\}^s}$	3.94	3.18	2.66	2.58	2.88	1.69
	$E$	-137.6	-170.7	-238.1	-241.2	-329.3	-473.4

### 3.3 Mechanical model

In this subsection, we examine designs for a non-linear mechanical model which was studied in Crisfield (1991, p. 2) and Liefvendahl and Stocki (2006). The model consists from the long bar and the spring as given in Figure 1. In the mechanical system, the angle  $\theta$  is assumed to be small.

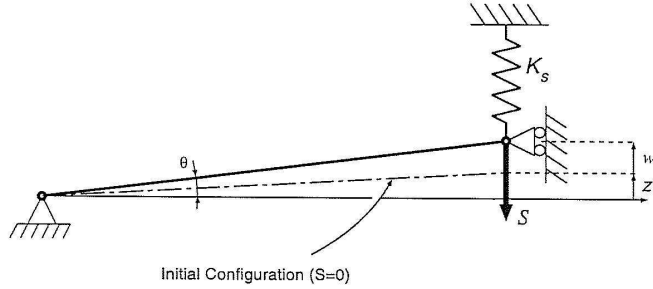


Figure 2: The mechanical system. The force  $S$  has negative sign.

We are interested in the vertical displacement  $w$  which is given by a solution of the equation

$$S = \frac{EA}{l^3}(z^2w + 1.5zw^2 + 0.5w^3) + K_s w$$

with the smallest absolute value, where the input  $E \in [4.8 \cdot 10^5, 5.2 \cdot 10^5]$  is the Young modulus of the bar material,  $A \in [95, 105]$  is the cross-sectional area of the bar,  $l \in [2650, 2750]$  is the length of the unloaded bar,  $z \in [24.5, 25.5]$  is the vertical coordinate of the right node of the bar when it is unloaded,  $K_s \in [0.88, 0.92]$  is the stiffness of the spring,  $S \in [-23, -22]$  is the vertical force applied to the right node of the bar. Thus

the model has six inputs where  $x = (E, A, l, z, K_s, S)$ . The simulated mean square error for the MLHD and GMLHD is displayed in Table 3. We see that the performance of the designs for the mechanical model is similar to the performance for the Borehole model. For a smaller number of design points the MLHD performs slightly better than the GMLHD with respect to the  $\text{MSE}_{[0,1]^6}$  criterion, while the opposite behaviour is observed if  $n$  is larger than 80. In both cases the differences between the two designs are rather small. On the other hand compared to the MLHD the GMLHD yields again a substantial improvement with respect to the entropy criterion and has a better performance in the corners of the design space.

Table 3: *Square root of the mean square error and entropy for MLHD and GMLHD with different  $n$  for the mechanical model.*

	$n$	50	60	70	80	90	140	200
MLHD	$\text{MSE}_{[0,1]^6}$	0.57	0.55	0.53	0.50	0.48	0.47	0.45
	$\text{MSE}_{\{0,1\}^6}$	1.54	1.51	1.49	1.26	1.22	1.23	1.17
	$E$	-31.5	-32.6	-38.5	-48.0	-56.7	-107.0	-145.5
GMLHD	$\text{MSE}_{[0,1]^6}$	0.63	0.60	0.57	0.52	0.47	0.43	0.41
	$\text{MSE}_{\{0,1\}^6}$	1.26	1.19	1.07	0.90	0.85	0.82	0.83
	$E$	-21.7	-21.5	-25.3	-31.7	-37.5	-71.1	-93.5

### 3.4 Neddermeyer's model

In this subsection, we examine the performance of the two designs for a model which was studied by Neddermeyer (1943) and Higdon et al. (2008). The model describes the internal radius of steel cylinder during a high explosive (HE) charge implosion. The internal radius as a function of time is given by a solution of the differential equation

$$y' = \left[ \frac{2}{R_1^2 f(y)^2} \left( \frac{v_0^2}{2} - \frac{s}{2\rho} g(y) \right) \right]^{-1/2},$$

where

$$f(y) = \frac{y^2}{(1 - \lambda^2)} \ln \frac{y^2 + 1 - \lambda^2}{y^2},$$

$$g(y) = \frac{1}{1 - \lambda^2} (y^2 + 1 - \lambda^2) \ln(y^2 + 1 - \lambda^2) - \lambda^2 \ln(\lambda^2),$$

$R_1 = 1.5$  is the initial outer radius of steel cylinder being imploded,  $\lambda = 2/3$  is the ratio of outer cylinder radius to the inner radius,  $\rho = 7.5$  is the specific density (relative to water),  $v_0 = 0.3 \cdot 10^5$  is the initial velocity imparted on the outer radius of the cylinder from the HE.

In this case we are interested in the internal radius after an implosion and the model is of the form  $\eta(x) = y(T, x)$  with three input variables  $x = (s, m, u_0)$ . Here  $T = 5 \cdot 10^5$  is a time moment after an implosion, the input variable  $s$  varies in the interval  $[1.85 \cdot 10^{10}, 2.25 \cdot 10^{10}]$  and represents the yield stress of steel, while  $m \in [0.29, 0.35]$  is the mass ratio between the HE and cylinder and  $u_0 \in [1.4 \cdot 10^{10}, 1.7 \cdot 10^{10}]$  is the detonation energy per gram of exploded gas from the HE. The simulated mean square error for the MLHD and GMLHD is shown in Table 4. In this case we observe a similar superiority of the GMLHD as in the first example. The GMLHD yields between 18% ( $n = 30$ ) and 39% ( $n = 150$ ) smaller values for the  $\text{MSE}_{[0,1]^3}$  than the MLHD. The improvement with respect to the entropy criterion is of similar size.

Table 4: *Square root of the mean square error and entropy for MLHD and GMLHD with different  $n$  for the Neddermeyer's model.*

	$n$	30	40	50	70	90	150
MLHD	$\text{MSE}_{[0,1]^3}$	1.34	1.13	1.08	1.07	0.86	0.90
	$\text{MSE}_{\{0,1\}^3}$	15.48	13.61	11.30	11.87	8.36	9.41
	$E$	-61.4	-86.9	-94.9	-175.4	-245.8	-271.5
GMLHD	$\text{MSE}_{[0,1]^3}$	1.10	0.99	0.86	0.82	0.68	0.55
	$\text{MSE}_{\{0,1\}^3}$	11.04	8.64	4.93	3.22	1.80	0.96
	$E$	-51.8	-72.6	-77.4	-145.5	-202.9	-214.1

## 4 Conclusions

In this paper we have proposed new designs for computer experiments which improve commonly used space-filling designs. The method is based on a quantile transformation of the projections of a commonly used space-filling design for computer experiments (such as maximin Latin hypercube, minimum distance, pseudo-random sequences, minimum discrepancy, etc.). For the quantile transformation we use the arc-sine distribution and as a consequence the new designs allocate more input variables in the neighborhood of the boundary of the design space. The usefulness of a transformation of the

uniform design has also been demonstrated in a different context by Al-Mharmah and Calvin (1996) who showed that a Beta(2/3, 2/3) density yields a better performance than the uniform density in random non-adaptive algorithms for the determination of the maximum of a continuous function (note that the arc-sine distribution is close to the Beta(2/3, 2/3) distribution).

Using (logarithmic) potential theory, some explanation is given why the new designs (obtained by the quantiles of the arc-sine distribution) should perform better than the original uniform space-filling designs. For maximin Latin hypercube designs the methodology is illustrated in four examples, which are commonly considered in the literature about designs for computer experiments. The transformed designs are called generalized maximin hypercube designs (GMLHD). In all examples the mean square error at the corner points obtained by the generalized maximin Latin hypercube designs is substantially smaller than the mean square error obtained by the maximin Latin hypercube designs. A similar superiority of generalized maximin Latin hypercube designs can be observed for the entropy criterion. If the mean square error over the full design space is considered the generalized maximin Latin hypercube designs show a better performance than the initial designs in most cases, in particular, if (compared to the dimension) the number of designs points is relatively large. If the sample sizes are small compared to the dimension (such as  $d = 8$  and  $n = 30$ ), the maximin Latin hypercube designs show a slightly better performance. However, in such cases an adequate comparison of two small sample designs is only possible within a narrow predetermined class of computer models. Consequently, results for small sample designs for a particular model can not be transferred to other models.

We finally note that the results are obtained for the uniform distribution  $w(x) \equiv 1$  in the mean square error criterion (7). Naturally, we can expect that if the weight  $w(x)$  is not uniform and we are interested in more precise prediction in the middle of the hypercube  $[0, 1]^d$ , other transformations of the initial design would yield to designs with a better performance. An appropriate transformation  $p$  could be obtained by incorporating the weight function  $w$  in the criterion (5). For example, if  $a = 0$  in (3), one could minimize the weighted logarithmic potential

$$\psi(\mu) = \sum_{i=1}^n \sum_{j=i+1}^n \log\left(\frac{1}{|z_i - z_j|}\right) - \sum_{i=1}^n \log w(z_i) \quad (10)$$

which has been studied extensively in the literature [see Szegö (1975)]. A typical weight function representing the importance of the interior of the design space is given by the



Beta density

$$w_{\alpha,\beta}(x) = \frac{1}{B(\alpha,\beta)} x^{\beta-1} (1-x)^{\alpha-1}, \alpha, \beta \geq 1 \quad (11)$$

where the case  $\alpha = \beta = 1$  corresponds to the uniform distribution considered in the present paper. It is well known [see Szegö (1975)] that for the choice (11) with  $\alpha > 1, \beta > 1$  (which corresponds to a nonuniform weight with more mass in the interior of the design space) the design  $\{z_1, \dots, z_n\}$  minimizing (10) is given by the roots of the polynomial  $P_n^{(2\alpha-3, 2\beta-3)}(2z-1)$ , where  $P_n^{(\gamma, \delta)}(z)$  denotes the  $n$ th Jacobi polynomial on the interval  $[-1, 1]$  orthogonal with respect to the measure  $(1-z)^\gamma(1+z)^\beta dz$ . The corresponding quantile transformations yield to generalized maximin Latin hypercube designs which put more design points in the interior of the design space  $[0, 1]^d$ .

**Acknowledgments.** This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG) and the NIH grant award IR01GM072876:01A1. Andrey Pepelyshev also acknowledge the financial support provided by MUCM project (EPSRC grant EP/D048893/1). We would also like to thank Anthony O'Hagan, Jeremy Oakley and Anatoly Zhigljavsky for their very perceptive comments.

## References

- Al-Mharmah H., Calvin J. M. (1996). Optimal random non-adaptive algorithm for global optimization of brownian motion. *J. Global Optim.* 8, 81-90.
- Audze P., Eglais V. (1977). New approach for planning out of experiments, *Problems of Dynamics and Strengths*, 35, 104-1-07.
- Ahn J., Owen A. (2001). Quasi-regression. *Journal of Complexity*, 17, 588-607.
- Bastos L., O'Hagan A. (2008). Diagnostics for gaussian process emulators. Submitted to *Technometrics*.
- Crisfield M.A. (1991). Non-linear finite element analysis of solids and structures.
- Detle H., Pepelyshev A., Zhigljavsky A. (2009). A new approach to optimal designs for models with correlated observations, submitted. <http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik3/publications/pzcorr10.pdf>.
- Detle, H. and Studden, W. J. (1995). Some new asymptotic properties for the zeros of Jacobi, Laguerre and Hermite polynomials. *Constructive Approximation*, 11, 227-238.

- Fang K.-T., Lin D.K.J. (2003). Uniform experimental designs and their application in industry. In Handbook of Statistics, Ed. by R. Khattree and C.R. Rao, 131-170.
- Fang K.-T., Qin H. (2003). A note on construction of nearly uniform designs with large number of runs. *Statist. Probab. Lett.* 61, no. 2, 215–224.
- Fang K.-T., Lin D.K.J., Winker P., Zhang Y. (2000). Uniform design: theory and application. *Technometrics* 42, no. 3, 237–248.
- Fang K.-T. Li R., Sudjianto A. (2006). Design and modeling for computer experiments. Chapman & Hall/CRC.
- Fang K.-T., Ma C.-X., Winker P. (2002). Centered  $L_2$ -discrepancy of random sampling and Latin hypercube design, and construction of uniform designs. *Math. Comp.* 71, no. 237, 275–296.
- Hickernell F.J. (1998). A generalized discrepancy and quadrature error bound. *Math. Comp.* 67, no. 221, 299–322.
- Higdon D., Gattiker J., Williams B., Rightley M. (2008). Computer model calibration using high-dimensional output. *J. Amer. Stat. Ass.*, 103, 570–583.
- Jin R., Chen, W., Sudjianto A. (2005). An efficient algorithm for constructing optimal design of computer experiments. *J. Statist. Plann. Inf.* 134, 268–287.
- Kennedy M.C., O’Hagan A. (2001). Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B* 63, no. 3, 425–464.
- Kleijnen J.P.C. (2008). Design and Analysis of Simulation Experiments. Springer, New York.
- Liefvendahl M., Stocki R. (2006). A study on algorithms for optimization of Latin hypercubes. *J. Statist. Plann. Inference* 136, 3231–3247.
- McKay M. D., Beckman R. J., Conover W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, no. 2, 239–245.
- Morris M.D., Mitchell T.J. (1995). Exploratory designs for computer experiments *J. Stat. Plan. Inf.*, 43, 381–402.
- Neddermeyer S. (1943). Collapse of hollow steel cylinders. Los Alamos Report, LA-18, Los Alamos National Laboratory.
- Niederreiter H. (1983). A quasi-Monte Carlo method for the approximate computation of the extreme values of a function. *Studies in pure mathematics*, 523–529, Birkhauser, Basel.

- Pepelyshev A. (2009). Improvement of random LHD for high dimensions. Proceedings of the 6th St. Petersburg Workshop on Simulation, VVM, St. Petersburg, 1091-1096.
- Pukelsheim F. (2006) Optimal Design of Experiments. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.
- Saff E.B., Totik V. (1997). Logarithmic potentials with external fields. Springer-Verlag, Berlin.
- Sacks J., Welch W.J., Mitchell T.J., Wynn H.P. (1989). Design and analysis of computer experiments. With comments and a rejoinder by the authors. *Statist. Sci.* 4, no. 4, 409–435.
- Santner T.J., Williams B.J., Notz W. (2003). The Design and Analysis of Computer Experiments.
- Sebastiani P., Wynn, H. P. (2000). Maximum entropy sampling and optimal Bayesian experimental design. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 62(1):145-157.
- Szegö, G. (1975). Orthogonal Polynomials. Amer. Math. Soc. Colloq. Publ., Vol. 23, Amer. Math. Soc., NY.
- Wiens D.P., Zhou J. (1997). Robust Designs Based on the Infinitesimal Approach. *Journal of the American Statistical Association* 92, 1503–1511.
- Worley B.A. (1987). Deterministic uncertainty analysis. OSTI ID: 6104504.





