UNCECOMP 2015
1st ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Crete Island, Greece, 25–27 May 2015

AUDZE-EGLĀJS CRITERION FOR ORTHOGONAL AND REGULAR TRIANGULAR GRIDS

Magdalena Šmídová[1], Václav Sadílek[1], Jan Eliáš[1], Miroslav Vořechovský[1]

[1]Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology
Veveří 331/95, 602 00 Brno, Czech republic
e-mail: {smidova.m, sadilek.v, elias.j, vorechovsky.m}@fce.vutbr.cz

Keywords: Audze-Eglājs Criterion, Design of Experiments, Optimization, Full Factorial Design, Regular Triangular Design.

Abstract. *Computer experiments have become a powerful tool for the investigation of problems encompassing the randomness of observed phenomena. To use this tool, it is necessary to prepare a plan of the simulations that should be performed, i.e. what is known as Design of Experiments (DoE) should be carried out. Such a design has to fulfill certain requirements in order to provide applicable results. As designs are often stochastic, the fulfillment of the requirements is not assured. Therefore, the originally random design is optimized with respect to a selected criterion which should guarantee that the demanded properties of the design will be achieved.*

One of the criteria that can be used during optimization is the Audze-Eglājs (AE) criterion. It is a criterion that accentuates the space-filling property of the final design.

Depending on the chosen method of optimization, the knowledge of the lower bound of the optimization criterion might be necessary for efficient control over the optimization algorithm. As the lower bound is not known for the AE criterion, this article describes two types of deterministic designs that are virtually ideal with regard to the uniform filling of space, and therefore should be close to optimal from the point of view of the AE criterion. Consequently, they are supposed to provide a lower bound of the criterion, or its estimate where its value cannot be evaluated due to the limits of these deterministic designs regarding the number of design points.

The AE criterion values of these deterministic designs are compared to the values of various stochastic random and optimized designs.

Furthermore, the description of a closed-form exact formula that enables faster evaluation of the AE criterion exploiting the regularity of deterministic design is presented for one of the two deterministic designs described in the article.

1 INTRODUCTION

Recently, a growing number of research centers and also design offices in various engineering fields use computer experiments on daily basis for analysis and design processes concerning problems that include uncertain or random inputs.

One of the possible methods of dealing with such problems – achieving the optimal design or analyzing the dependence of a response on inputs – may be the response surface method first introduced by Box [1] in 1954 and since then improved and adapted in various ways (e.g. [2, 3]). While the real response of a complex system to a random input may be hard to evaluate, this method provides an estimation of the response using a response function. This function is in fact a model which has to be fitted in advance and whose complexity is determined by the researcher in accordance with their knowledge of the problem to be solved. Nevertheless, the evaluation of the response using such a model is usually significantly easier than that of the original problem which decreases the costs of the analysis.

The response function is usually determined by evaluating a certain number of simulations (N_s) of the real problem. The choice of the inputs for these simulations is generally denoted as a response surface design and various properties are demanded. Among them the most common is the orthogonality of such a design and the equal probability of the simulations. The first of these is usually fulfilled implicitly by optimizing the design for the other. Therefore this article will concentrate on the requirement for equal probability.

It would be difficult to attain such equality in the original domain, and therefore the input space is transformed to ensure the independence of the marginal random inputs. Afterwards, the design space of such a transformed domain is used so the equal probability of simulations is attained via the uniform filling of this design space, which is, in fact, a unit hypercube of a dimension identical to that of the original problem ($[0, 1]^{N_\mathrm{v}}$, where N_v is the number of random inputs of the original problem). Individual simulations are then represented by design points placed in the hypercube.

Methods of filling the unit hypercube with points are basically divided into deterministic and stochastic. Deterministic methods place the design points in regular patterns and may ensure perfect space-filling. One such example is full factorial design (FFD) [4]. The drawback of FFD is collapsibility: if the response function is insensitive to one or more input variables, a certain number of design points (proportional to the number of irrelevant input variables) collapse into identical evaluation.

The solution to this problem is the use of stochastic methods, namely the Monte Carlo method and its derivative – Latin Hypercube Sampling (LHS, [5]).

While using a stochastic sampling method, the space-filling property is not ensured automatically. Therefore, the initial random design should be optimized with respect to a selected criterion that prioritizes the uniform filling of the unit hypercube.

Various criteria have been defined to ensure the uniform filling of a given space. They are often based on the evaluation of discrepancy, e.g. Centered L_2-discrepancy [6], Wrap-Around L_2-discrepancy [7]. Maximin or miniMax criteria [8, 9] may also be used. The criterion that this article deals with is the Audze-Eglājs (AE) criterion [10, 11].

The value of any of these criteria for designs created by stochastic sampling methods is a random variable with a probability distribution dependent on the combination of $(N_\mathrm{s}, N_\mathrm{v})$ of the design. For each such combination a lower bound of the criterion also exists. The effective control of an optimization process is often conditional upon the knowledge of this lower bound (the minimum that can be reached during optimization) and the mean value.

Due to the regularity and perfect space-filling of certain deterministic designs it is reasonable to expect these designs to provide a value of the criterion that is minimal or close to minimum for a given configuration $(N_\mathrm{s}, N_\mathrm{v})$. Analysis of these designs can then give us the information necessary for the process of optimizing stochastic response surface designs.

The deterministic designs selected for analysis are full factorial design (FFD) and a new deterministic design proposed in this paper: regular triangular design (RTD). Both of them are described later together with an analysis of the AE criterion for various $(N_\mathrm{s}, N_\mathrm{v})$ configurations.

Since these deterministic designs are restricted to only specific numbers of design points ($N_\mathrm{s} = N^{N_\mathrm{v}}$, N being a natural number expressing the number of different input values for each random input variable), the article also anticipates the possibility of interpolation of the lower bound for an arbitrary N_s of a stochastic design.

2 THE AUDZE-EGLĀJS (AE) CRITERION

The AE criterion is one of various criteria, that can be used to assess the improvement of a design during optimization. This criterion is derived from the analogy of two systems. The first is the design of a computer experiment or a system of the design points in a *design space* $[0, 1]^{N_\mathrm{v}}$. The other is then a system of points with unit mass in the same design space. Such mass points interact via nonzero repulsive forces, and therefore the whole system cumulates potential energy. For better understanding these repulsive forces are often depicted as (nonlinear) springs between points (as illustrated for $N_\mathrm{v} = 2$, $N_\mathrm{s} = 4$ at Fig. 1). The amount of energy is assumed [10] to be inversely proportional to the square distance between each pair of points. The total cumulated potential energy can be calculated as a sum of energies accumulated in all pairs of points:

$$E^{\mathrm{AE}} = \sum_{i=1}^{N_\mathrm{s}} \sum_{j=i+1}^{N_\mathrm{s}} \frac{1}{L_{ij}^2} \tag{1}$$

where L_{ij} is the distance between points i and j.

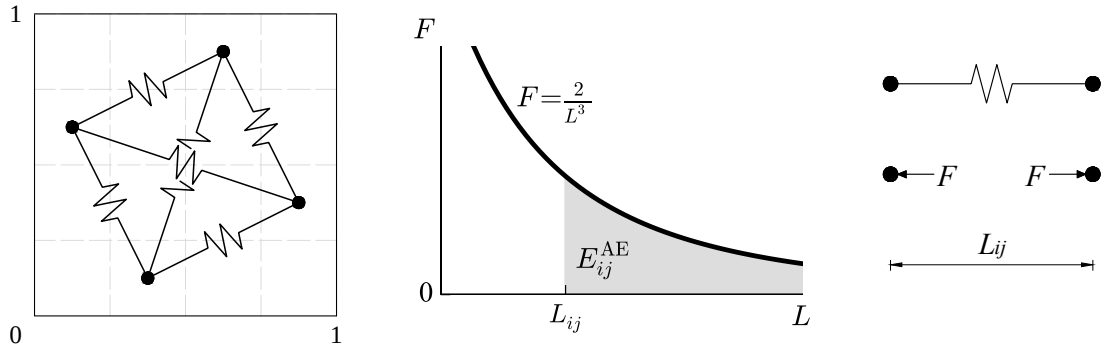

Figure 1: Physical analogy for the calculation of the AE criterion.

The number of pairs of N_s points is simply:

$$N_\mathrm{p} = \binom{N_\mathrm{s}}{2} = \frac{N_\mathrm{s}\,(N_\mathrm{s}-1)}{2} \tag{2}$$

Due to the physical analogy used, the design with the lowest potential energy is supposed [10, 11, 12, 13, 14, 15, 16, 17, 18] to represent the best possible placement of points in space with respect to uniformity of spread. Therefore, when optimizing Monte Carlo type designs to improve the projection properties of the final design, the optimization algorithm is controlled in such a way as to decrease the value of E^{AE}. The above is considered to be valid inversely as well. Thus, with points uniformly spread throughout a space, the value of the potential energy E^{AE} should be minimal.

In some methods for the selection of the coordinates of the sampling points (such as LHS), the AE optimality criterion cannot be used to vary the coordinates of the points along each variable. The only way to optimize such a type of design is to find the best mutual ordering of predefined coordinates for each separate variable. Heuristic techniques are often employed to perform this combinatorial optimization. Depending on the used optimization algorithm, knowledge of the minimum value of the used criterion might be essential for effective control over the optimization. This article therefore describes the improvement of the calculation of the AE criterion for full factorial designs – authors consider such designs to be ideally space-filling, hence the value of E^{AE} of such designs should be close to the minimum for the given N_{s} and N_{v} of the design.

3 FULL FACTORIAL DESIGN

Full factorial design (FFD), as used in this article, is a design that places the points in the unit hypercube regularly in a perfect orthogonal raster.

Each design point has N_{v} coordinates; therefore, instead of using a set of coordinates, the j-th point can be written as a vector $\mathbf{x}_j = (x_{1,j}, x_{2,j}, \ldots, x_{i,j}, \ldots, x_{N_{\mathrm{v}},j})$.

In order to create an FFD, the number of values N_i representing the i-th input variable ($i = 1, 2, \ldots, N_{\mathrm{v}}$) first must be determined. Then the coordinates of this variable are calculated by spreading N_i points uniformly along the edge of the unit hypercube (unit length $(0, 1)$). To obtain a vector of coordinates $\mathbf{x}_j$ for all the points in the FFD, it is necessary to create all permutations of these coordinates. This article will later deal only with designs considering the same number of points along all dimensions, thus $N_1 = N_2 = \ldots = N_{N_{\mathrm{v}}} = N$. Therefore, the total number of points in the unit hypercube is:

$$N_{\mathrm{s}} = N^{N_{\mathrm{v}}} \tag{3}$$

FFD examples with one, two or three input variables ($N_{\mathrm{v}} = 1,\ 2,\ 3$) are shown in Fig. 2 for $N = 5$.

4 EVALUATION OF THE AE CRITERION ON A FULL FACTORIAL DESIGN

Evaluation of the AE criterion as defined in section 2 requires the calculation of the distances between every pair of points occurring in the design. The total number of such pairs is N_{p}, see Eq. (2), and the direct application of Eq. (1) includes too many arithmetic operations. In the case of regular designs, the regularity may be exploited to simplify the calculation of the AE criterion because the distances between individual pairs of points are repeated.

Combinatorial analysis of the existing lengths and their corresponding frequency was performed for an FFD to accelerate the calculation.

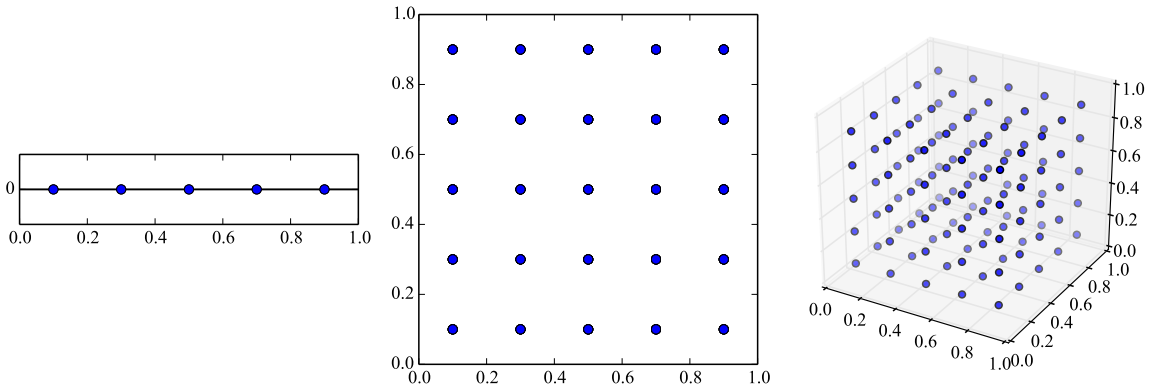

Figure 2: FFD: $N = 5$, $N_\mathrm{v} = 1,\ 2,\ 3$, $N_\mathrm{s} = N^{N_\mathrm{v}} = 5,\ 25,\ 125$.

4.1 Lengths occurring in the design

In this section the rank of a certain coordinate value is used instead of this value. An FFD with N coordinates for each input variable contains coordinates $(i - 0.5)/N$, $i = 1, 2, \ldots, N$, where i stands for the rank of the coordinate value. For example, if a design has $N = 5$

- the coordinate value 0.1 has a corresponding rank of 1,
- the coordinate value 0.3 has a corresponding rank of 2,
- ...

All the distances between the points in a hypercube can be rewritten using the rank differences. These lengths can be further divided into groups according to the mutual position of the points that define the distance:

- 1D distances – the pair of points lies on a line parallel to the edge of the hypercube. Subtracting the vector of ranks of coordinates of these two points provides a vector of differences with a single non-zero value, while all the other rank differences equal zero (e.g. 4D space – $(0, 0, 0, 1)$, $(0, 2, 0, 0)$, etc.). The number at the i-th position in the vector of differences expresses the difference in the rank of coordinates for the i-th input variable.
- 2D distances – both points lie in a plane parallel to one side of the hypercube. The vector of differences contains two non-zero numbers (e.g. $(0, 0, 1, 1)$, $(0, 2, 0, 1)$, etc.)
- ...

The total number of different *vectors of integer differences* (irrespective of the order of elements in the vector) for a given (N, N_v) design is equal to the number of combinations with repetition:

$$n_\mathrm{c} = \frac{(N + N_\mathrm{v} - 1)!}{N_\mathrm{v}!\,(N - 1)!} \tag{4}$$

For example, for $N = 3$, $N_\mathrm{v} = 4$, the vectors of differences are $(0, 0, 0, 0)$, $(0, 0, 0, 1)$, $(0, 0, 0, 2)$, $(0, 0, 1, 1)$, ..., $(2, 2, 2, 2)$ and their total number is $n_\mathrm{c} = 15$.

The vectors of differences will be marked as $\mathbf{c}_i$ ($i = 1, \ldots, n_\mathrm{c}$) and the entries in these vectors are denoted as $c_{i,j}$ ($j = 1, \ldots, N_\mathrm{v}$). The length l_i that corresponds to a vector $\mathbf{c}_i$ is calculated:

$$l_i = \frac{1}{N}\sqrt{\sum_{j=1}^{N_\mathrm{v}} c_{i,j}^2} \tag{5}$$

Although none of the vectors $\mathbf{c}_i$ are repeated, more of them may represent the same distance between points. The calculation of the AE criterion does not exploit this fact but it should be taken into consideration in case a list of the lengths occurring in the design is to be prepared.

The calculation of the frequency of individual lengths is described in the next section.

4.2 Frequency of lengths

The frequency of a certain length is calculated as by multiplying three coefficients (n_i^p, n_i^d and n_i^n) described in this section.

When calculating the AE criterion on an FFD, the *lengths* used in the calculation depend only on the *values* in the vectors of differences $c_{i,j}$ and not their ranks in the vector $\mathbf{c}_i$ (for example, the length belonging to the vector $(0, 0, 0, 1)$ is the same as the one belonging to $(0, 1, 0, 0)$). Therefore, it is necessary to calculate the first coefficient as a permutation with repetitions:

$$n_i^\mathrm{p} = \frac{N_\mathrm{v}!}{\prod_{k=0}^{N-1}\left[\left(N_\mathrm{v} - \sum_{j=1}^{N_\mathrm{v}} \left|\mathrm{sgn}\left(c_{i,j} - k\right)\right|\right)!\right]} \tag{6}$$

The visualization of the function of this coefficient is presented in Fig. 3, left, for a 2D design (the vector of differences $(3, 2)$ provides the same length as vector $(2, 3)$).

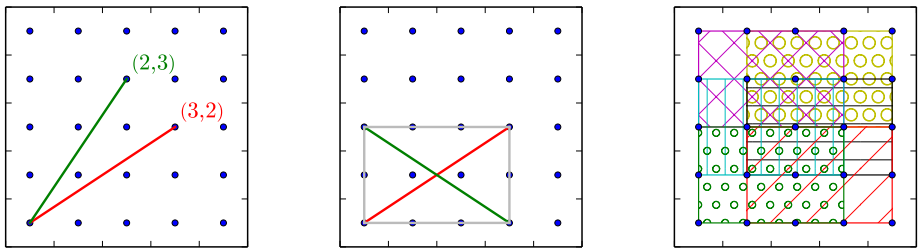

Figure 3: Visualization of the function of coefficients n_i^p, n_i^d and n_i^n in the calculation of the frequency of lengths.

The next coefficient takes into account the dimension of the length (as described in Sec. 4.1) for which the frequency is calculated. It is equal to the number of diagonals in the dimension and its function is displayed in Fig. 3, center.

The dimension can be quantified as:

$$d_i = \sum_{j=1}^{N_\mathrm{v}} \mathrm{sgn}\left(c_{i,j}\right) \tag{7}$$

and the coefficient itself is calculated:

$$n_i^\mathrm{d} = 2^{d_i - 1} \tag{8}$$

The last coefficient expresses the number of "sub-hyperbodies" that can fit into the unit hypercube. In Fig. 3, center, such a "hyperbody" is the grey rectangle; on the right it is represented by rectangles with various hatchings. Their number depends on N and the values $c_{i,j}$ in the vector of differences $\mathbf{c}_i$. The coefficient is:

$$n_i^{\mathrm{n}} = \prod_{j=1}^{N_{\mathrm{v}}} (N - c_{i,j}) \tag{9}$$

4.3 Calculation of the AE value

The calculation of E^{AE} using lengths and their frequencies as defined in equations 5–9 is conducted as follows:

$$E_{\mathrm{FFD}}^{\mathrm{AE}} = \sum_{i=2}^{n_{\mathrm{c}}} n_i^{\mathrm{p}} \cdot n_i^{\mathrm{d}} \cdot n_i^{\mathrm{n}} \cdot \frac{1}{l_i^2} = N^2 \sum_{i=2}^{n_{\mathrm{c}}} n_i^{\mathrm{p}} \cdot n_i^{\mathrm{d}} \cdot n_i^{\mathrm{n}} \cdot \frac{1}{\sum_{j=1}^{N_{\mathrm{v}}} c_{i,j}^2} \tag{10}$$

Since a vector of zero length (e.g. for 4D design it is the vector $(0,0,0,0)$) has been included in the number of various vectors of differences n_{c}, it is necessary to start the summation from $i = 2$ to prevent the evaluation of the distance of a point from itself.

The proposed Eq. 10 greatly simplifies the evaluation of the AE criterion for FFD design. The number of various arithmetical operations needed to exactly evaluate the AE criterion using Eq. (10) is much smaller than using Eq. (1). Table 1 presents a comparison of the number of operations using the two formulations for the case of $N_{\mathrm{v}} = 4$ and $N = 10$ (in total $N_{\mathrm{s}} = 10^4$ points). Each type of operation is associated with a certain weight – an average relative time needed to accomplish such an operation. We have used our Python computer code to obtain these relative times. Each weight can be multiplied by the number of corresponding operations and the total time is estimated by their sum. The ratio between the total time spent on the evaluation using Eq. (1) and Eq. (10) roughly estimates the "speed-up", and Fig. 4 presents these ratios. It is clear that using the proposed formula for FFD designs speeds-up the calculation considerably, especially for large numbers of grid points, N, and high dimensions, N_{v}. The actual times (or speed-ups) depend on the programming language, particular implementation and other circumstances. Using our implementation, the actual times correspond to the values presented in Tab. 1 and Fig. 4, i.e. to the times obtained by calculating the number of various types of operations.

5 REGULAR TRIANGULAR DESIGN

In this section we propose a new type of deterministic design that we call regular triangular design (RTD). The main idea of RTD is to reproduce the packing of hyperspheres. For 2-dimensional design, centers of optimally packed spheres create a mesh composed of equilateral triangles (one triangle being a basic building block of this mesh). In the case of three variables, the basic block is a regular tetrahedron, etc.; generally, for any dimension the building blocks always have $N_{\mathrm{v}} + 1$ vertices at a constant distance from one another.

Such a block is constructed first and then is copied several times to fill a certain space. Finally, due to the different size of the basic building block in each of the main directions, the blocks are adjusted to fill a hypercube in a manner described later.

5.1 Construction of the basic building block

Let us start with an initial point at coordinates $\mathbf{x}_0 = (x_{0,1}, x_{0,2}, x_{0,3}, \ldots, x_{0,N_{\mathrm{v}}}) = (0,0,0,\ldots,0)$. Then, another point is defined at a distance a from the initial point in the direc-

Table 1: Comparison of the number of necessary arithmetical operations for $N = 10$ and $N_\mathrm{v} = 4$.

operation	weight	number of operations	
		Eq. (1)	Eq. (10)
summation	1	199979999	25704
subtraction	1	199980000	39270
abs value	5.6	0	28560
signum	9	0	31416
multiplication	1	199980000	11428
division	2.7	49995000	1428
power	1	0	714
factorial	9	0	2861

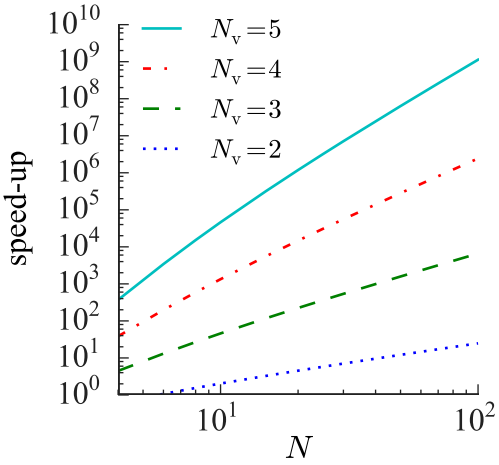

Figure 4: Speed-up of the evaluation of the AE criterion on FFD using Eq. (10).

tion of the first variable (a is an arbitrary distance unit). Its coordinates can be generally written as $\mathbf{x}_1 = (a_1, 0, 0, \ldots, 0)$. A simple equation describing that mutual distance equals the chosen unit a:

$$\sum_{v=1}^{N_\mathrm{v}} (x_{0,v} - x_{1,v})^2 = a^2 \tag{11}$$

gives the solution of a_1 being $\pm a$. Here, as well as in all the other calculations of the coordinates of the basic building block, there are always two options to choose from. We will systematically choose the positive one; nevertheless, an equivalent solution would be obtained by using the negative root for any of the coordinates.

The coordinates of the next vertex of the building block are $\mathbf{x}_2 = (b_1, a_2, 0, \ldots, 0)$. There are two unknowns, and therefore two equations have to be defined to determine b_1 and a_2. These equations again express that there is a constant distance between the third point and both previous vertices, which is equal to a:

$$\sum_{v=1}^{N_\mathrm{v}} (x_{0,v} - x_{2,v})^2 = a^2 \quad \Rightarrow \quad b_1^2 + a_2^2 = a^2 \tag{12}$$

$$\sum_{v=1}^{N_\mathrm{v}} (x_{1,v} - x_{2,v})^2 = a^2 \quad \Rightarrow \quad (a_1 - b_1)^2 + a_2^2 = a^2 \tag{13}$$

Subtracting Eq. (13) from Eq. (12), we obtain:

$$b_1^2 - \left(a_1^2 - 2a_1 b_1 + b_1^2\right) = 0 \tag{14}$$

This results in:

$$b_1 = \frac{a_1}{2} = \frac{a}{2} \tag{15}$$

and from Eq. (12) the value of a_2 can be calculated as:

$$a_2 = \sqrt{a^2 - b_1^2} = \frac{\sqrt{3}a}{2} \tag{16}$$

A general formula for an arbitrary coordinate $x_{i,j}$ can be developed as follows:

- For $j > i$:

$$x_{i,j} = 0 \tag{17}$$

- For $j < i$, the coordinate $x_{i,j}$ is calculated from the distance of $\mathbf{x}_i$ to $\mathbf{x}_0$ and $\mathbf{x}_j$:

$$\sum_{v=1}^{N_\mathrm{v}} \left(x_{0,v} - x_{i,v}\right)^2 = a^2 \quad \Rightarrow \quad \sum_{v=1}^{i} x_{i,v}^2 = a^2 \tag{18}$$

$$\sum_{v=1}^{N_\mathrm{v}} \left(x_{j,v} - x_{i,v}\right)^2 = a^2 \quad \Rightarrow \quad \sum_{v=1}^{j} \left(x_{j,v} - x_{i,v}\right)^2 + \sum_{v=j+1}^{i} x_{i,v}^2 = a^2 \tag{19}$$

  Subtracting these two equations from each other provides, after some modifications:

$$x_{i,j} = \frac{\left(\sum_{v=1}^{j-1} x_{j,v}^2 - 2x_{j,v}x_{i,v}\right) + x_{j,j}^2}{2x_{j,j}} \tag{20}$$

  Therefore, the first coordinate ($j = 1$) is constant for all the points $i > 1$ and is equal to:

$$b_1 = x_{i,1} = \frac{x_{1,1}}{2} = \frac{a_1}{2} = \frac{a}{2} \tag{21}$$

  The second coordinate ($j = 2$) is constant as well (for $i > 2$) and equal to:

$$b_2 = \frac{(-b_1^2 + a_2^2)}{2a_2} \tag{22}$$

  where $a_2 = x_{2,2}$. A general formula for the j-th coordinate is therefore:

$$b_j = \frac{\left(\sum_{v=1}^{j-1} -b_v^2\right) + a_j^2}{2a_j} \tag{23}$$

- Finally, the coordinate $x_{j,j}$, defined as a_j, will be determined. Considering the distance of point $\mathbf{x}_j$ from $\mathbf{x}_0$, we obtain:

$$a_j = \sqrt{a^2 - \sum_{v=1}^{j-1} b_v^2} \tag{24}$$

The basic building block has in total $N_\mathrm{v} + 1$ vertices at the coordinates:

$$\begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \\ \vdots \\ \mathbf{x}_{N_\mathrm{v}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ a_1 & 0 & 0 & 0 & \cdots & 0 \\ b_1 & a_2 & 0 & 0 & \cdots & 0 \\ b_1 & b_2 & a_3 & 0 & \cdots & 0 \\ b_1 & b_2 & b_3 & a_4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & b_3 & b_4 & \cdots & a_{N_\mathrm{v}} \end{bmatrix} \tag{25}$$

where coefficients a_i and b_i are evaluated using Eqs. 24 and 23.

5.2 Periodic repetition, scaling and periodicity

Having the basic building block, the filling of the space can be carried out by repeating the block along the coordinates/variables. The number of repetitions along a coordinate v is marked r_v and can be chosen independently for every variable.

Starting with the basic point with coordinates $\mathbf{x}_0 = (0, 0, 0, \ldots, 0)$, any other i-th vertex of the basic building block can be selected and added repetitively along its own direction $r_i - 1$ times. Then, a j-th vertex is added $r_j - 1$ times to every previously placed point. This is repeated until all vertices from the basic block are inserted. The described process is sketched in Fig. 5 for $N_\mathrm{v} = 3$ and $r_v = 4$ for $v = 1,\ 2,\ 3$.

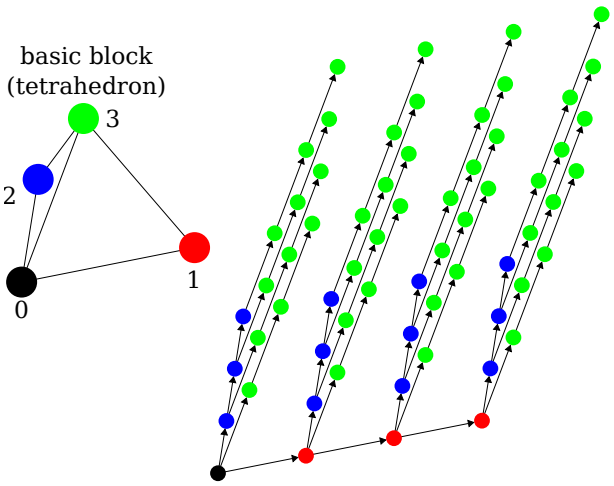

Figure 5: Repetition of the basic building block.

The generated points have to be rescaled and modified to fit into the unit hypercube. The size of the whole block of points created along a coordinate i is $a_i r_i$. To make the size of this block equal 1, the coordinates have to be scaled by factor $(a_i r_i)^{-1}$. Due to the fact that the hypercube is based on the orthogonal directions while the RTD is not, some of the points created lie outside the unit hypercube. The position of these points is updated by changing every $x_{i,j}$ coordinate in the design to $x_{i,j} - \mathrm{floor}\,(x_{i,j})$, where the floor function returns the largest previous integer.

During further analysis only the designs with constant r_v were considered for all input variables, $r_1 = r_2 = \ldots = r_{N_\mathrm{v}} = N$. The total number of simulations in such a design is then $N_\mathrm{s} = N^{N_\mathrm{v}}$. This might not be the "optimal" solution as the size of the basic block differs in every direction (the parameters a_i that express the size of the basic block along the coordinate axes are a, $0.8660a$, $0.8165a$, $0.7906a$, etc.); therefore, a better filling might be obtained using a different number of repetitions along individual axes.

RTD examples for $N = 6$ and $N_\mathrm{v} = 2$, 3 and 5 are shown in Figures 6, 7 and 8.

6 COMPARISON OF FULL FACTORIAL, REGULAR TRIANGULAR AND RANDOM DESIGNS

In this section, the AE criterion values for various designs are studied and compared. In particular, the compared types of designs include:

- two deterministic grids (FFD and RTD designs)

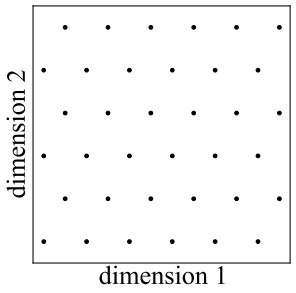

Figure 6: Design points for $N_\mathrm{v} = 2$ and $N = 6$.

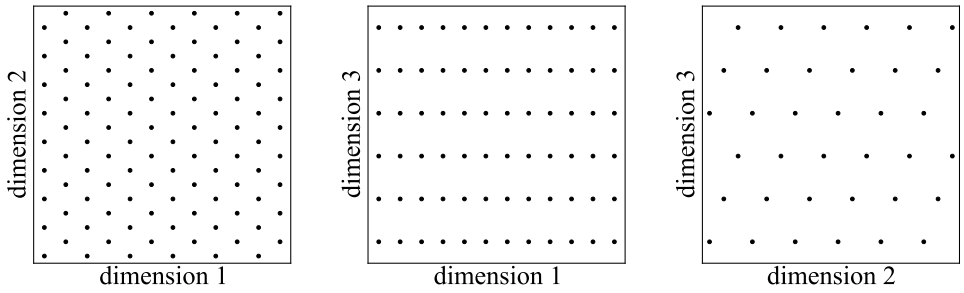

Figure 7: Design points for $N_\mathrm{v} = 3$ and $N = 6$ ($N_\mathrm{s} = N^{N_\mathrm{v}} = 6^3 = 216$) – 2D projections of the points.

- Latin Hypercube designs optimized by the AE criterion (LHS-AE) and LHS designs with random ordering (LHS-RAND)
- Crude Monte Carlo sampling optimized by the AE criterion (MC-AE) and MC designs with random ordering (MC-RAND)

These designs are compared using various graphs. The deterministic designs were generated just once. The other designs are random (they depend on a pseudorandom generator) and thus they were generated $N_\mathrm{run} = 1000$ times to obtain information about the variability of the results. First of all, the AE criterion is computed for all designs and standardized by dividing the criterion by the number of considered pairs of points:

$$r(N_\mathrm{s}, N_\mathrm{v}) = \frac{E^\mathrm{AE}}{\binom{N_\mathrm{s}}{2}} \tag{26}$$

The reason for standardizing the AE norm is that the values for different sample sizes N_s become of comparable magnitude. The value of the AE criterion is, in fact, a sum of inverse squared lengths. The lengths between all pairs of points can be arranged in a triangular matrix, $\mathbf{T}$: $T_{i,j} = 1/L_{i,j}^2$, $i, j = 1, \ldots, N_\mathrm{s}$, $i \neq j$. The number of such elements is $\binom{N_\mathrm{s}}{2}$. The norm E^AE is just a sum of these elements. Therefore, the value of the norm r can be seen as an average of the inverse squared length between all pairs of points in the hypercube.

Fig. 9 left compares the studied norm r for $N_\mathrm{v} = 2$ random variables. It is evident that the FFD design provides the lowest possible value of the norm for the selected sample sizes. RTD provides approximately the same values of r. The highest values of the norm and also

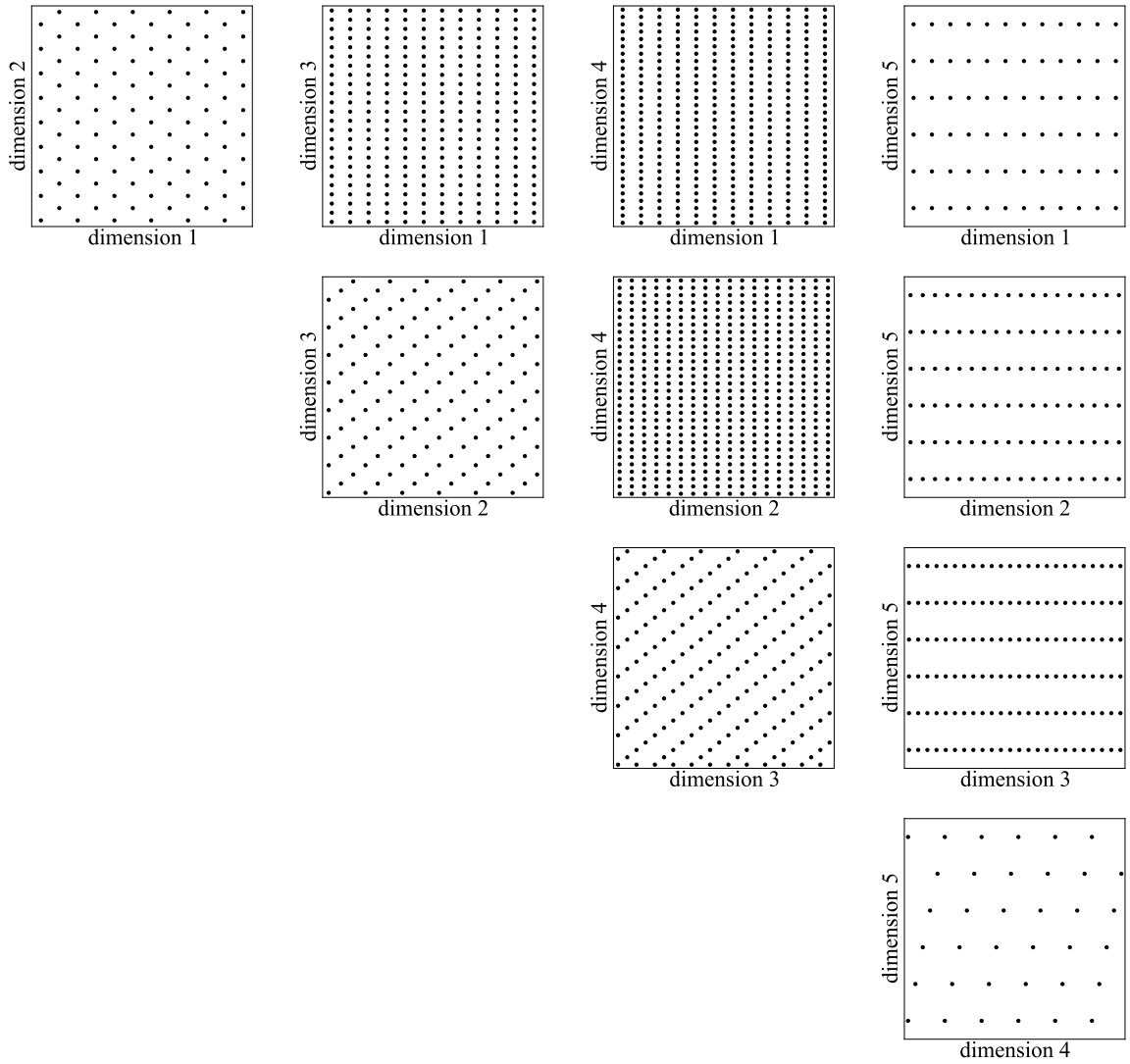

Figure 8: Design points for $N_\mathrm{v} = 5$ and $N = 6$ ($N_\mathrm{s} = N^{N_\mathrm{v}} = 6^5 = 7776$) – 2D projections of the points.

the highest variance of the value are obtained for random Monte Carlo samples (MC-RAND). Employing the combinatorial optimization of MC samples (MC-AE) improves the situation considerably and for large samples the norm has a very low variance and is approximately equal to the values of FFD and RTD. LHS-RAND is clearly better then MC-RAND for small sample sizes. When the mutual ordering of samples in LHS is optimized via the AE criterion, the results are approximately as good as for the deterministic designs FFD and RTD.

Fig. 9 right shows the same kind of graphs for $N_\mathrm{v} = 5$ (a five-dimensional unit hypercube). The reader might be confused by the fact that the deterministic grids FFD and RTD do not provide the lower bound of the norm. This is caused by the fact that the definition of the AE criterion leads to non-uniform design. Even though the AE criterion was developed to achieve a *uniform* spread of points in the hypercube, the opposite is true. As shown in [19], the unexpected behavior of the AE criterion is due to the presence of the boundaries of the hypercube. A remedy has been proposed which involves considering the periodic repetition of the design domain along all directions; based on this idea a simple modification of the criterion

has been proposed. The problems with the standard definition of the AE criterion become more pronounced in higher dimensions. Therefore, the FFD designs do not seem to provide the lowest possible value of the norm.

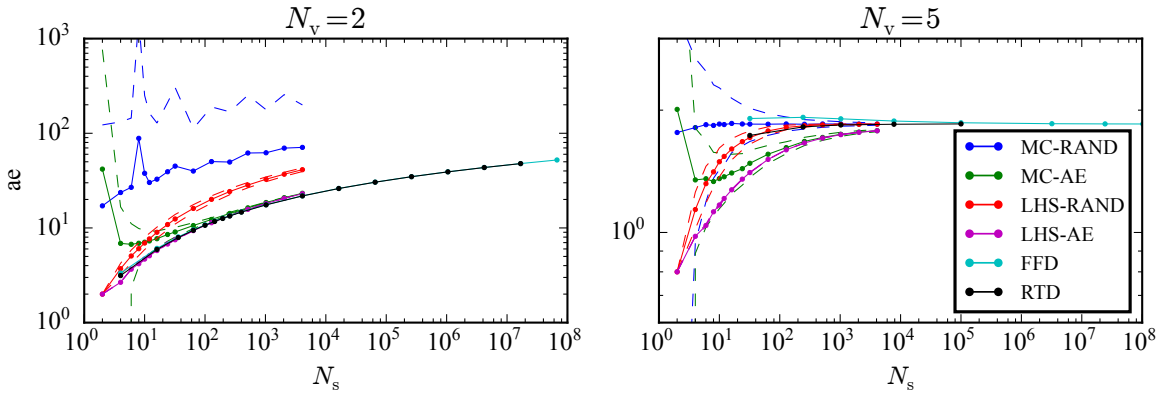

Figure 9: Comparison of the AE criterion value – dependence on the method used for design creation.

Another feature that can be studied and compared for various kinds of designs is the distribution of the shortest lengths $\Delta_{\min}$ – the values calculated for each point $j = 1, \ldots, N_s$ as its distance to its nearest neighbor. These N_s values of $\Delta_{\min}$ can be viewed as the highest values in the columns (rows) in the **T** matrix. In the cases of the deterministic designs FFD and RTD, the value of $\Delta_{\min}$ is identical for all points as the distance to the nearest neighbor is always the same. In the case of randomized designs the shortest distances show a certain scatter and probabilistic distribution. Clearly, the AE-optimized variants LHS-AE and MC-AE have more uniform $\Delta_{\min}$ distribution than the unoptimized designs; see Fig. 10.

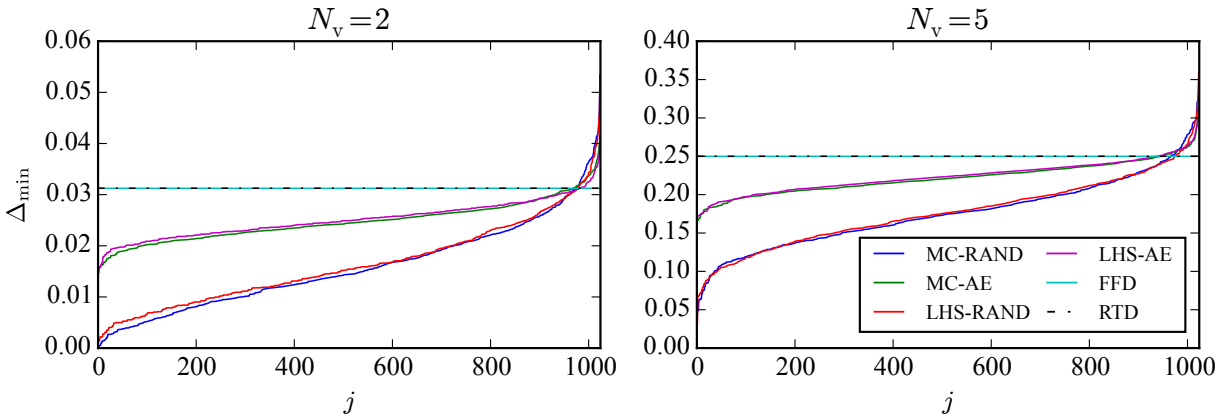

Figure 10: Sorted minimal lengths in designs optimized via various methods, $N_s = 1024$.

The last type of comparison selected for presentation is a histogram of all lengths $\Delta_{i,j}$, i.e. the distances between all N_p pairs of points. Fig. 11 shows a comparison of the histograms. It can be seen that the histograms do not differ considerably while the methods result in quite different AE criterion values. The reason for this discrepancy is that the AE criterion is based on inverted

squared lengths and so is very sensitive to the left tail (small values $\Delta_{i,j}$). It is interesting to mention that the distribution of the squared length $\Delta_{i,j}^2$ tends to a Gaussian distribution as N_v (and N) grow large because $\Delta_{i,j}^2$ is the sum of a high number of asymptotically independent random contributions along separated dimensions.

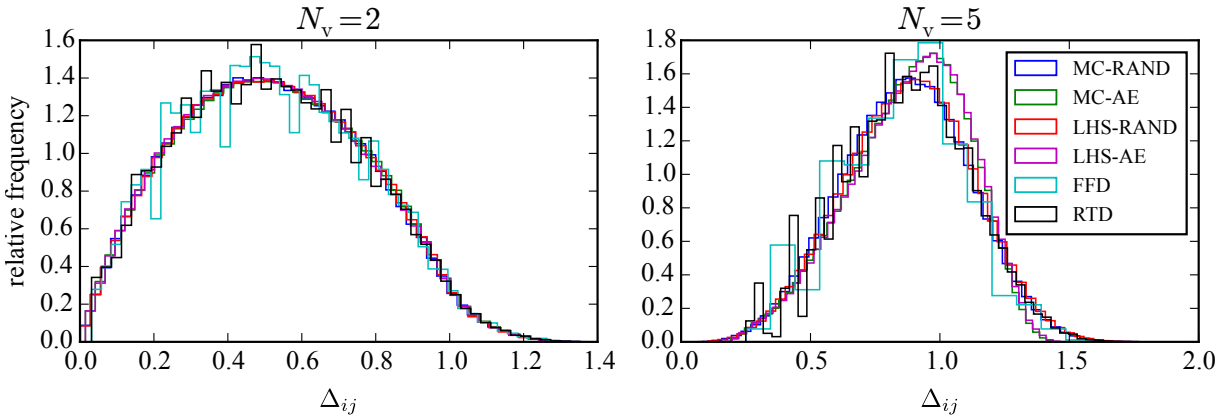

Figure 11: Histogram of lengths featured in designs optimized via various methods, $N_\mathrm{s} = 1024$.

7 CONCLUSIONS

The paper focuses on the Audze-Eglājs (AE) criterion of experimental design optimality. The main focus is on two types of deterministic designs that are virtually ideal as regards space-fillingness. The paper presents the calculation of the AE criterion for these deterministic designs. As the designs are close to optimal, they are supposed to provide a lower bound of the criterion, or its estimate, where its value cannot be evaluated due to the limits of deterministic designs regarding the number of design points. This lower bound (minimum) of the criterion is often required in order to obtain efficient control over an optimization algorithm.

The main results are:

- An algorithm for a new kind of deterministic design was developed to achieve the lowest possible AE criterion value: regular triangular design.
- A closed-form exact formula for computation of the AE criterion of full factorial design (orthogonal grid). The proposed formula exploits the fact that many distances between points inside the unit hypercube are repeated and the numbers of repetitions for each distance are calculated exactly. The evaluation of the AE criterion is much faster compared to the general formulation. The AE criterion value for an FFD design is believed to yield the lower bound on the AE criterion for the sample size $N_\mathrm{s} = N^{N_\mathrm{v}}$.
- Comparison of AE criteria, minimal lengths and histograms of lengths for various designs (including the two deterministic designs, LHS and Monte Carlo designs – optimized and random).

Future work will focus on the interpolation of the derived formula for the (supposedly) lower bound on the AE criterion and also on the newly developed [19] PAE criterion, which is based on the AE criterion but considers periodic boundary conditions in order to remove the non-uniformity of designs optimized by the standard AE criterion.

ACKNOWLEDGEMENT

The authors acknowledge financial support provided by the Czech Ministry of Education, Youth and Sports under project No. LO1408 "AdMaS UP - Advanced Materials, Structures and Technologies" under "National Sustainability Programme I", and also support provided by the Czech Science Foundation under project No. 15-07730S (FIRBO) and Specific University Research project MŠMT No. FAST-J-15-2868. Additionally, the work of the second author has been supported by Brno University of Technology project CZ.1.07/2.3.00/30.0005.

REFERENCES

[1] G.E.P. Box, The exploration and exploitation of response surfaces: Some general considerations and examples. *Biometrix*, **10(1)**, 16–60, 1954.

[2] C.G. Bucher and U. Bourgund, A fast and efficient response surface approach for structural reliability problems. *Structural Safety*, **7**, 57–66, 1990.

[3] S. Gupta and C.S. Manohar, An improved response surface method for the determination of failure probability and importance measures. *Structural Safety*, **26**, 123–139, 2004.

[4] D.C. Montgomery, *Design and Analysis of Experiments, 8th edition*. John Wiley & Sons, Inc., 2012.

[5] W.J. Conover, On a better method for selecting input variables. Unpublished Los Alamos National Laboratories manuscript, reproduced as Appendix A of "Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems" by J.C. Helton and F.J. Davis, Sandia National Laboratories report SAND2001-0417, 1975.

[6] K.T. Fang, C.X. Ma, P. Winker, Centered L_2-discrepancy of random sampling and Latin hypercube design, and construction of uniform designs. *Mathematics of Computation*, **71(237)**, 275–296, 2000.

[7] K.T. Fang, C.X. Ma, Wrap-around L_2-discrepancy of random sampling, Latin hypercube and uniform designs. *Journal of Complexity*, **17(4)**, 608–624, 2001.

[8] M.E. Johnson, L.M. Moore, D. Ylvisaker, Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, **2(26)**, 131–148, 1990.

[9] B.G.M. Husslage, *Maximin Designs for Computer Experiments*. PhD thesis, Tilburg University, 2006.

[10] P. Audze, V. Eglājs, New approach for planning out of experiments. *Problems of Dynamics and Strengths*, **35**, 104–107, 1977. (in Russian).

[11] S.J. Bates, J. Sienz, D.S. Langley, Formulation of the Audze–Eglais uniform Latin Hypercube design of experiments. *Advances in Engineering Software*, **34(8)**, 493–506, 2003.

[12] M. Liefvendahl and R. Stocki, A study on algorithms for optimization of Latin hypercubes. *Journal of Statistical Planning and Inference*, **136(9)**, 3231–3247, 2006.

[13] S.J. Bates V.V. Toropov, O.M. Querin, Generation of Extended Uniform Latin Hypercube Designs of Experiments. B.H.V. Topping, editor. *The Ninth International Conference on the Application of Artificial Intelligence to Civil, Structural and Environmental Engineering*. Civil-Comp Press, Stirlingshire, Scotland, 2007.

[14] B.G.M. Husslage, G. Rennen, E.R. van Dam, D. den Hertog, Space-filling Latin hypercube designs for computer experiments. CentER Discussion Paper 2008-104, Tilburg University, 2008.

[15] B.G.M. Husslage, G. Rennen, E.R. van Dam, D. den Hertog, Space-filling Latin hypercube designs for computer experiments. *Optimization and Engineering*, **12(4)**, 611–630, 2011.

[16] A. Kovalovs, S. Rucevskis, Identification of elastic properties of composite plate. *Functional Materials and Nanotechnologies (FM&NT)*, volume 23 of *IOP Conf. Series: Materials Science and Engineering*. IOP Publishing, Ltd, 2011.

[17] E. Janouchová, A. Kučerová, Competitive comparison of optimal designs of experiments for sampling-based sensitivity analysis. *Computers & Structures*, **124(0)**, 47–60, 2013. Special Issue.

[18] H.M. Vu, J.P. Forth, D.V. Dao, V.V. Toropov, The use of optimisation for enhancing the development of a novel sustainable masonry unit. *Applied Mathematical Modelling*, **38(3)**, 853–863, 2014.

[19] J. Eliáš, M. Vořechovský, Modification of the Audze–Eglājs criterion to achieve a uniform distribution of sampling points. *Computer Methods in Applied Mechanics and Engineering*, in review, 2015.