17th International Conference
ENGINEERING MECHANICS 2011
Svratka, Czech Republic, 9 – 12 May 2011

COMPARISON OF SPACE-FILLING DESIGN STRATEGIES

E. Myšáková*, M. Lepš*

Abstract: *Space-Filling Design Strategies create an essential part of a surrogate modeling. Two main objectives are usually placed on the resulting designs - orthogonality and space-filling properties. The last decade has witnessed the development of several methods for the latter objective. These methods are based on very different ideas and are characterized by distant complexities. Since the computing time can be the limiting constraint, we inspect the computing demands against the space-filling performances. In detail, our contribution presents and compares several different techniques of quazi-random numbers generators utilizing Latine Hypercube Sampling (LHS) methods as well as Delaunay triangulations.*

Keywords: ***Design of experiments, space-filling, Latin Hypercube Sampling, Delaunay triangulation, maximin.***

1. Introduction

The design of experiments (DoE) is an essential part of the development of any meta-model (surrogate) (Simpson et al., 2001, Jin, 2005). The aim is to gain maximum knowledge from a given system with a minimum number of designs. Since we assume that the final meta-model is a priori unknown, the design should be spread over the domain as uniformly as possible. The effectiveness of such DoE can be measured by several metrics aiming mainly at orthogonality or space-filling properties. See references (Cioppa and Lucas, 2007, Hofwing and Sternberg, 2010) for orthogonal and sources (Toropov et al., 2007, Crombecq et al., 2009) for space-filling metrics, respectively. We have selected the *maximin* metric (Mm) for its simplicity and easiness in visualization. The Mm is the minimal distance out of all distances between any two design points and is to be maximized. Using this metric, we would like to compare several methods that have been proposed in recent years to create good DoEs.

2. Space-Filling Methods

2.1. Latin Hypercube Sampling (L)

The LHS is one of the most popular space-filling algorithms, although its resulting space-filling properties can be very low. In the LHS, each variable is divided into n levels. Each level is selected randomly once, independently for each variable. This leads to a regular DoE, see Fig. 1. Note that the design can be placed in the middle of the level as well as everywhere (randomly) within the level.

The worst case LHS design is such that all points lie on a diagonal. To solve this deficiency, the simplest solution is to create a brand new LHS design, i.e. this is actually an application of a brute force method. Such an approach is used for example in MATLAB environment within the `lhs_design` routine. Hereafter, we denote this method as regular.

2.2. Optimized Latin Hypercube Sampling by Random Moves (HC)

A bad LHS design can be also improved by changing the positions of individual levels, see e.g. references (Novák and Lehký, 2006, Kučerová, 2007) for more details. At each iteration, in comparison to the above mentioned references where the Simulated Annealing algorithm is used, we accept only an improvement step here, and, therefore, this method is denoted as the random Hill Climbing algorithm.

* Eva Myšáková, Ing. Matěj Lepš, Ph.D., Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29 Prague 6, CZ, e-mail: leps@cml.fsv.cvut.cz

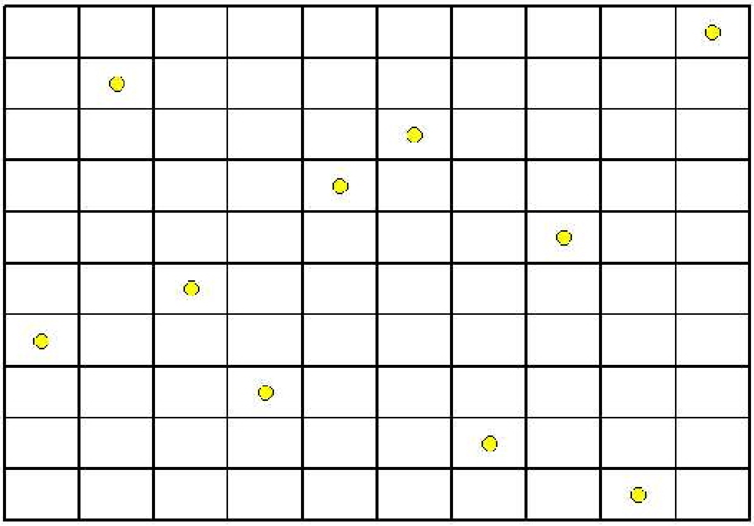

Fig. 1: Example of LHS design for two variables and ten levels with points selected in middle of levels.

2.3. Optimized Latin Hypercube Sampling by Heuristic Moves (H)

Since we know, which pair of points creates the worst value within the Mm metric, we have applied heuristic procedure, where we try to change only the position of levels of those two bad designs. The algorithm is changing the inappropriate level position with one randomly chosen position until an improvement occurs (and then follows with the next bad position).

2.4. Removing Points from LHS Design (R)

The last method, based again on the LHS design, is utilizing the speed of the LHS design. The creation of the LHS design with more points does not waste too much time. Therefore, we construct a design with more points than needed, and then, we are repeatedly removing the point that creates the worst Mm distance until the original number of points is attained.

2.5. Delaunay Triangulation Based methods (DC, DT)

The last sort of methods is based on the triangulation of an admissible domain by simplexes (Crombecq et al., 2009). Because it is relatively simple to compute a volume of a simplex, we have a rough estimation, where is the biggest unsampled region. We are starting from a basic LHS design with a few points. Then, our procedure iteratively adds a centroid of the biggest simplex into the designs. There are two variants that differ in whether (after each step) a whole triangulation is repeated or not. The first possibility computes better estimates of an unsampled space, but is more time consuming; the second possibility is unprecise but faster.

3. Results and Conclusions

It is evident from Fig. 2 that the regular LHS and the LHS with random moves cannot compete with others. It is also clear that Delaunay-based methods cannot compete with the heuristic method. Moreover, with the growing number of dimensions Delaunay-based methods are more time consuming than LHS-based methods. Our experience shows that at four dimensions the heuristic method attains, at the same time, four times bigger maximin values.

Finally, our contribution has shown results for rectangular domains. Our next research will inspect the abilities of presented methods for designs within irregular domains. From the first view, the LHS-based methods seem unsuitable in comparison to those based on Delaunay triangulation. One can describe almost any shape by a triangulation, and therefore, the application of triangulation-based methods seems reasonable.

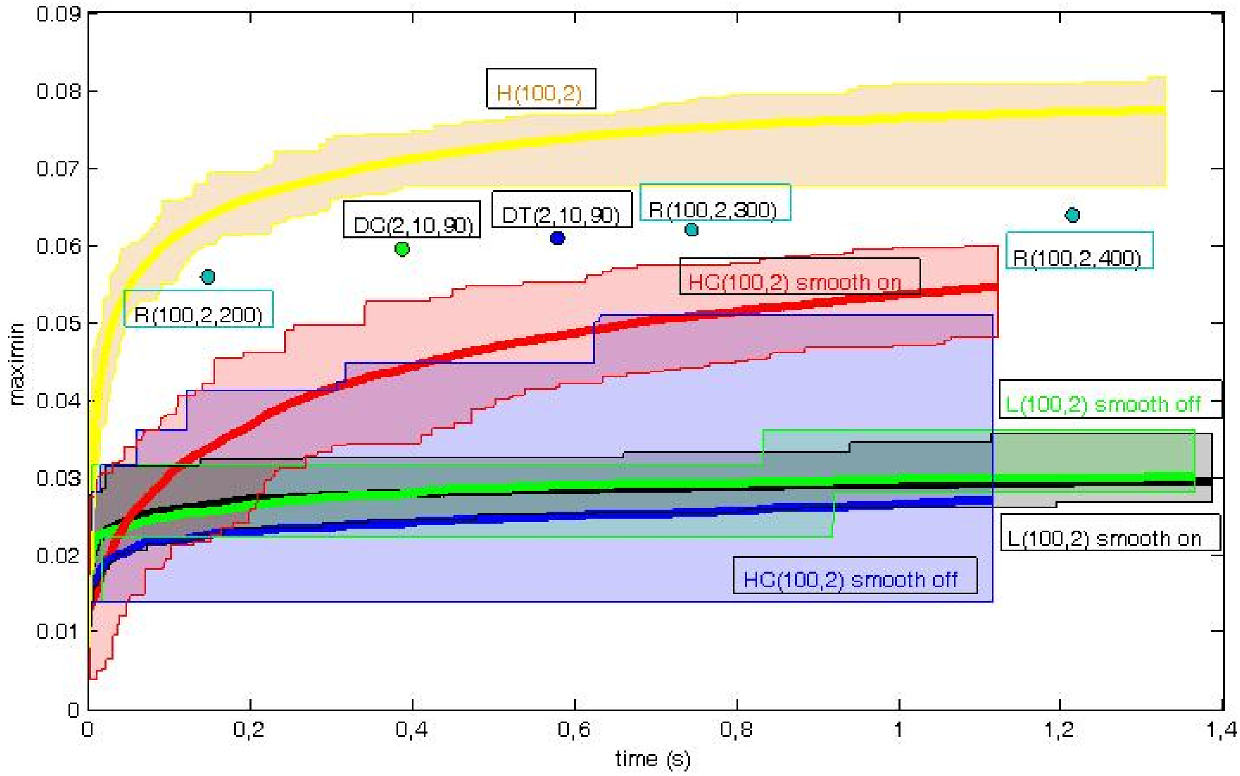

Fig. 2: Comparison of speed (time) vs. quality (maximin - higher is better) of spacefilling algorithms. Key: Yellow lines/H = Heuristic method, Cyan dots/R = Removing points, Green dot/DC = Delaunay triangulation without re-meshing, Blue dot/DT = Multiple Delaunay triangulations, Red and Blue lines/HC = random Hill Climbing algorithm, Green and Black lines/L = repeatedly created regular LHS design; smooth on generates random points within LHS level, smooth off in the middle; -(100, 2) stems for 100 points in two dimensions, (2,10, 90) means 2 dimensions, 10 random starting points and 90 points added by triangulation and (100, 2, x) equals to 100 points in two dimensions remaining from x original points; color boundary lines stem for obtained maximum and minimum values, bold line is for averages.

Acknowledgement

The authors gratefully acknowledge the financial support from the Ministry of Education, Youth and Sports MSM 6840770003 (Algorithms for computer simulation and application in engineering).

References

Cioppa, T. M. & Lucas, T. (2007) Efficient nearly orthogonal and space-filling latin hypercubes. Technometrics, 49(1):45-55.

Crombecq, K., Couckuyt, I., Gorissen, D., & Dhaene, T. (2009) Space-filling sequential design strategies for adaptive surrogate modelling. in: Proceedings of the First International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering (Topping, B. H. V. & Tsompanakis, Y., eds), Civil-Comp Press, Stirlingshire, UK.

Hofwing, M. & Strmberg, N. (2010) D-optimality of non-regular design spaces by using a Bayesian modification and a hybrid method. Structural and Multidisciplinary Optimization, 42:73-88.

Jin, Y. (2005) A comprehensive survey of fitness approximation in evolutionary computation. Soft Computing, 9:3-12.

Kučerová, A. (2007) *Identification of nonlinear mechanical model parameters based on softcomputing methods.* PhD thesis, Ecole Normale Superieure de Cachan, Laboratoire de Mecanique et Technologie.

Novák, D. & Lehký, D. (2006) ANN inverse analysis based on stochastic small-sample training set simulation. Engineering Applications of Artificial Intelligence, 19(7):731-740.

Simpson, T. W., Peplinski, J. D., Koch, P. N., & Allen, J. K. (2001) Metamodels for computer-based engineering design: Survey and recommendations. Engineering with Computers, 17:129-150.

Toropov, V. V., Bates, S. J., & Querin, O. M. (2007) Generation of extended uniform latin hypercube designs of experiments. In Topping, B. H. V., editor, Proceedings of the Ninth International Conference on the Application of Artificial Intelligence to Civil, Structural and Environmental Engineering. Civil-Comp Press, Stirlingshire, UK.