TENSOR-DECOMPOSITION-BASED SEQUENTIAL DESIGN OF EXPERIMENTS FOR COMPUTER SIMULATIONS

C. Alfaro-Isac*[†], A. Juan-Alejandre[†] and S. Izquierdo[†]

[†] Technological Institute of Aragon (ITAINNOVA)
C/ María de Luna 7-8, 50018 Zaragoza, Spain
Email: calfaro@itainnova.es - Web page: http://www. https://www.itainnova.es/

ABSTRACT

High-fidelity simulations aim at numerically replicating complex physical systems, usually involving several phenomena. Despite recent advances on digital computing, these simulations might take hours, even days to complete. To avoid performing a computationally expensive simulation for each case, surrogate models are used instead. A surrogate model links the inputs and outcomes of a set of computer experiments to approximate the response surface of a high fidelity model [1].

Sample selection has a direct influence on the accuracy of the resulting surrogate model. Classic design of experiments (DoE) include statistical methods (like factorial designs), or random distributions (Montecarlo, Latin Hypercube Sampling) [2]. In this work, a novel sequential sampling algorithm is described and tested using tensor decomposition as surrogate model aiming at obtaining the most precise model with the minimum design points.

While most classical DoE rely only on space-filling criteria, this particular DoE aims at balancing exploration and exploitation [3]. A gradient-based function is designed that locates the complex regions, accounting for the exploitation component. Since the gradient is not known in advance, the samples from which the surrogate model is constructed are iteratively removed to estimate the area where adding a new sample would have a greater impact. Canonical polyadic decomposition is performed to built the surrogate models, using TWINKLE library [4]. To avoid local oversampling, gradient-based function is compensated with a distance function, whose score is maximum in between sample points. The product of both functions constitutes the objective function to be maximized; the corresponding coordinates represent a sample candidate. The process is repeated for each existing sample point and the sample candidate whose objective function value is largest is selected as new sample point (see Fig. 1).

Integrating system knowledge into DoE requires that information to be updated as surrogate model is constructed; as opposite to static DoE's, points are added sequentially. An initial set of samples must be generated using existing sampling methods, in order to create a database on which performs the search for the optimal next sampling point. Additionally, stop criteria, such a maximum size of sample set or minimum model error, must be defined.

The algorithm is tested on two-dimensional Rosenbrock function (fig. 2.a) and ternary-mixture density computation (fig. 2.b), with four input dimensions (temperature, pressure, two component fractions). In the first case it is compared to latin hypercube sampling over identical initial datasets; in the second, the initial sampling is two-level full factorial and LHS up to 25 samples. The next samples are calculated through sequential algorithm or continuing with latin hypercube sampling. In both cases the mean error decreases faster with increasing sample size using the sequential DoE, outperforming classical DoE's, independently of the dimensionality.
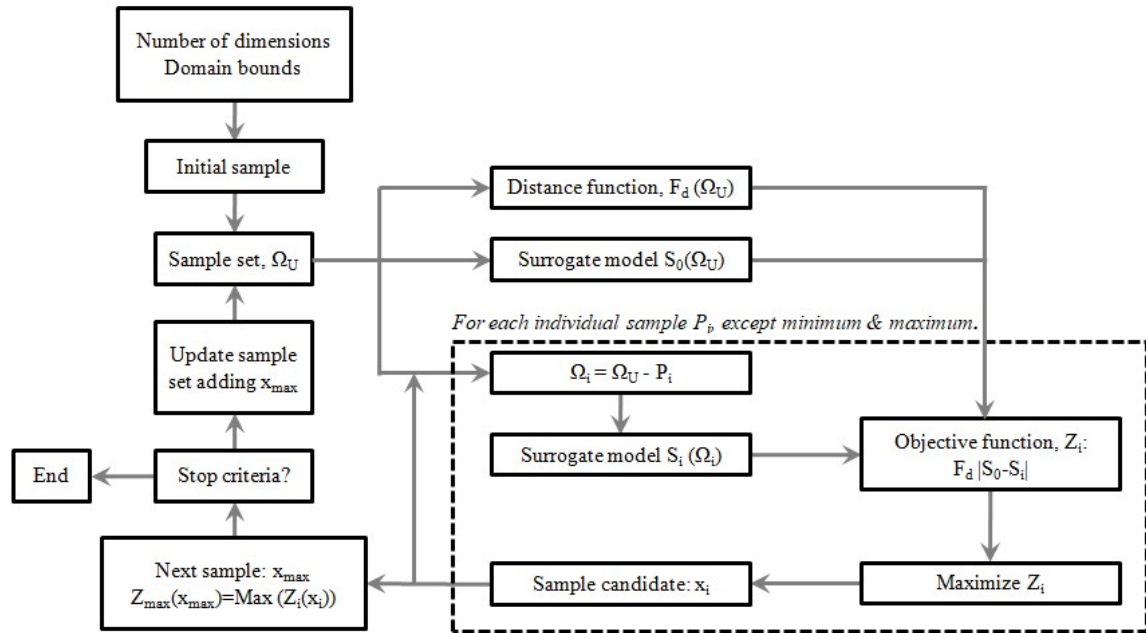
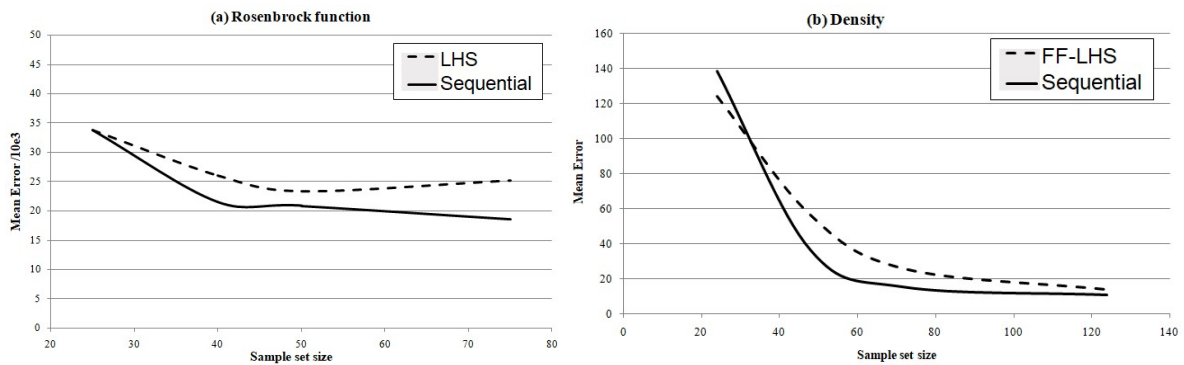*Fig. 1: Flow chart of sequential algorithm.*

Fig. 2: Illustrative examples: (a) Rosenbrock function; (b) ternary-mixture density.

REFERENCES

[1]   B. A. Williams and S. Cremaschi, "Surrogate Model Selection for Design Space Approximation And Surrogatebased Optimization," *Comput. Aided Chem. Eng.*, vol. 47, pp. 353–358, 2019, doi: 10.1016/B978-0-12-818597-1.50056-4.

[2]   S. S. Garud, I. A. Karimi, and M. Kraft, "Design of computer experiments: A review," *Comput. Chem. Eng.*, vol. 106, pp. 71–95, 2017, doi: 10.1016/j.compchemeng.2017.05.010.

[3]   P. Singh, D. Deschrijver, and T. Dhaene, "A balanced sequential design strategy for global surrogate modeling," *Proc. 2013 Winter Simul. Conf. - Simul. Mak. Decis. a Complex World, WSC 2013*, no. December, pp. 2172–2179, 2013, doi: 10.1109/WSC.2013.6721594.

[4]   V. Zambrano and R. Rodr, "TWINKLE : A Digital-Twin-Building Kernel for Real-Time Computer-Aided Engineering."