What Latin Hypercube Is Not

Oleg Mazonka and Charalambos Konstantinou

***Abstract*—Simulation methods continue to play a key role in designing models of theoretical or actual physical systems. Such models can reproduce the behaviour and provide insights into the operation of those systems. Well-determined input parameters are of prime importance for reliable simulation due to their impact to the performance sensitivity of the simulation design. Among various strategies for producing simulation samples is Latin Hypercube Design (LHD). LHDs are generated by Latin Hypercube Sampling (LHS), a type of stratified sampling that can be applied to multiple variables. LHS is proved to be an efficient and a popular method, however, it misses some important elements. While LHS focuses on parameter space aspect, this paper highlights five more aspects which may greatly impact the efficiency of sampling. In this paper, we do not provide solutions but rather bring up unanswered questions which could be missed during strategy planning on model simulation.**

***Index Terms*—Computer experiments, Latin Hypercube Sampling (LHS), space-filling designs, sequential sampling.**

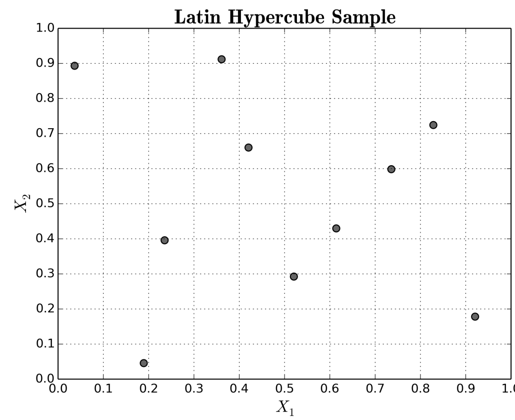


Fig. 1. A Latin Hypercube sample with $N = 10$, for two independent random variables X_1 and X_2 distributed uniformly on the unit square.

I. Introduction

Computer simulations aim to emulate a physical system through a computer model. The design and modelling of computer experiments allows to reproduce and capture the behaviour of the replicated system and thus make inferences about its characteristics. In science fields, many of these emulated processes need spatially distributed values to be used in computer simulation models as input. Latin Hypercube Sampling (LHS) has been used worldwide in such computer modelling applications.

LHS, introduced by McKay, Beckman, and Conover in 1979 [1], simultaneously stratifies on all input dimensions. In Latin Hypercube Designs (LHDs), a sample is maximally stratified in the scenario which the number of strata is equal to the sample size N. Fig. 1 shows a LHD sample drawn from two independent variables X_1 and X_2 distributed uniformly in $U[0, 1]$.

LHDs have always been particularly popular in computer simulations. For instance, LHDs have been used in a wide range of applications such as hurricane loss projection modelling, assessment for nuclear power plants and geologic isolation of radioactive waste, reliability analyses for manufacturing equipment, petroleum industry, transmission of HIV, etc [2]. There exist several reasons related with the popularity of LHS. For example, Latin Hypercube samples can cover design spaces regardless of their size and the requirements about the location and density of the data. The orthogonality of such samples allows also to construct non-collapsing, space-filling designs [3]. Hence, even if a subset of dimensions are forced to dropped out, the design remains a LHD allowing the samples to be reused. In addition, LHDs provide flexibility as they can adjust to the statistical assumptions of the experimental model.

LHDs often face problems to control the sampling of combinations of distributions and fill the design space. In many instances, optimization techniques to improve space-filling (e.g. minimizing some form of distance measure) or orthogonality (e.g. considering column-wise correlations) are challenging and expensive [4]. Other drawbacks of LHDs include the loss of statistical independence of sample values and the requirements in memory to hold the Latin Hypercube samples for each distribution. Also, there are practical scenarios which LHS does not appear to be significantly superior to random sampling for computational sensitivity analysis [5], [6]. LHD lacks ability to accommodate the prior knowledge of the model. To overcome this problem sequential sampling techniques, where a set of sample values are chosen so that certain criteria of the next experiment are satisfied, could be used. The yielded data is evaluated as it is collected, and further sampling is performed based on pre-defined rules of the resulted model performance.

In this paper we show that some aspects of analysis during strategy planning on model simulation can be missed. While LHS focuses on parameter space aspect, we present five more aspects worth considering before committing to a particular methodology of sampling design. We believe that there exist rigorous and correct solutions, of which the authors are not aware at the time of writing. Whether the solutions be known or ad hoc methods applied, they all remain outside of the scope of the paper.

The following section introduces a few techniques employed in experimental designs. Our contribution of this paper – the

O. Mazonka is with the Defence Science and Technology (DST) Group, Department of Defence, Edinburgh SA 5111 Australia e-mail: Oleg.Mazonka@defence.gov.au.

C. Konstantinou is with the Department of Electrical and Computer Engineering, New York University School of Engineering, Brooklyn, NY, 11201 USA e-mail: ckonstantinou@nyu.edu.

Fig. 2. Randomized Complete Block Design (RCBD): each row represents a block in which the colours illustrate different treatments (4 blocks and 4 treatments within each block).

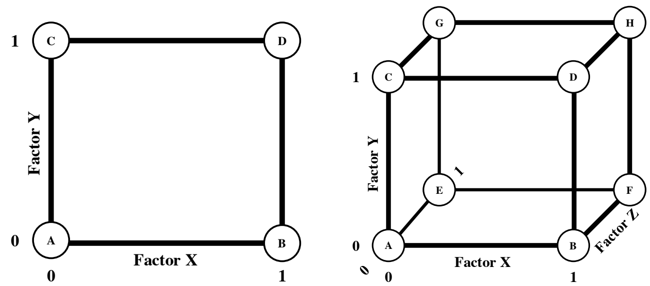
A	B	C
C	A	B
B	C	A

(a) 3×3 LS.

A	C	B	D
D	A	C	B
B	D	A	C
C	B	D	A

(b) 4×4 LS.

Fig. 3. Examples of a 3×3 and a 4×4 Latin Square (LS) designs.



(a) 2^2 full factorial.

(b) 2^3 full factorial.

Fig. 4. Examples of a 2^2 and a 2^3 full factorial designs; factors X, Y, Z.

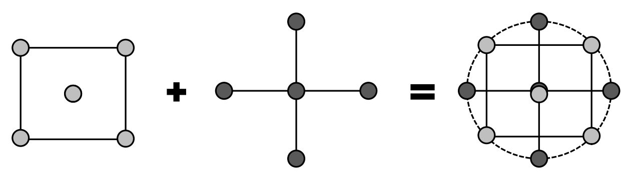


Fig. 5. Representation of the star points being added to factorial points.

different aspects of an experimental design are described in Section III.

II. Background and Overview

In experimental design we study a model (possibly stochastic) which depends on a number of parameters. These parameters are also called *factors*. Each factor can be set at a particular value. All such possible values are called *levels*. For example, a binary factor can take only two values: yes/no, true/false, etc; a continuous factor can take values within a range of real numbers. If all factors are fixed to some values, then this set is called a *design point*. And the whole set of design points represent the experimental design. A design point is a deterministic input to the model, to which the model outputs the result either deterministic or probabilistic.

In this section, we present the most notable techniques used for experimental designs.

1) Randomized Complete Block Design (RCBD): In RCBDs, used in agricultural experiments, the field is divided into blocks and each block is further divided into units. The number of units is equal to the number of treatments. Treatments are then assigned randomly to the subjects in the blocks so that a different treatment is applied to each unit. This procedure ensures that all treatments are observed within each block. A simple layout of a RCBD is presented in Fig. 2. The defining characteristic of RCBD is that each block sees each treatment exactly once [7].

2) Latin Square (LS): LS experimental designs follow the same concept of RCBDs with the main difference of conducting a single experiment in each block, so LS allows the existence of two factors [7]. In the previous agricultural example, the fertility of the land might variate in two directions (x-axis & y-axis) that can be used as the factors (e.g. due to soil type and water stagnation). The k levels of each factor can be represented in a square with k rows and k columns. In order to achieve a $k \times k$ LS design, then k treatments must be applied exactly once in each column and each row. Examples of a 3×3 and a 4×4 LS designs are shown in Fig. 3.

3) Factorial designs: Such experimental designs consider all possible combinations of levels across all the design factors, i.e. the sample size N of factorial designs is equal to $N = \prod_{i=1}^{n} L_i$, where n the number of factors and L_i the levels of each factor [7]. In some cases, each factor has two levels ($L = 2$), resulting in 2^n treatment combinations. Fig. 4 presents two such examples with two and three levels.

Although factorial designs include at least one trial for each possible combination of factors and levels, they come at a high cost in terms of experimental resources. In order to reduce the cost, a set of possible combinations can be excluded resulting in a fractional factorial design.

4) Central Composite Design (CCD): A CCD experimental design is a factorial or fractional factorial design with centre points, supplemented with a group of *star points* [8]. Those extra points allow estimation of the second order interactions between factors; otherwise only linear dependencies can be assumed.

5) Box-Behnken Design (BBD): In contrast with the previous CCD design BBD experimental designs do not contain all the embedded factorial points [9]. Specifically, BBDs are a class of rotatable or nearly rotatable second-order designs based on three-level incomplete factorial designs. The points in such designs exist at the centre and at the midpoints of edges of the parameter space. Fig. 6 illustrates a BBD for three factors.

6) Plackett-Burman Design (PBD): PBD is an efficient screening design when only particular effects are of interest. This experimental method is often used in ruggedness tests with the purpose: *a)* to find the factors that strongly affect the outcome of an analytical procedure and *b)* to determine how closely one needs to control these factors [10]. PBD always involves $4k$ experiments ($k \geq 1$) and in each experiment the

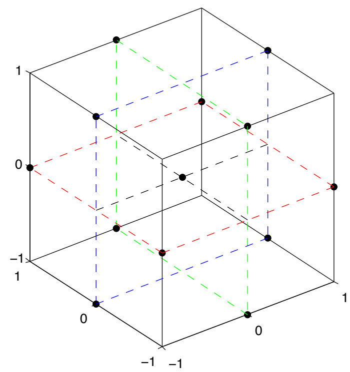


Fig. 6. The graphical representation of a Box-Behnken Design (BBD) for three factors.

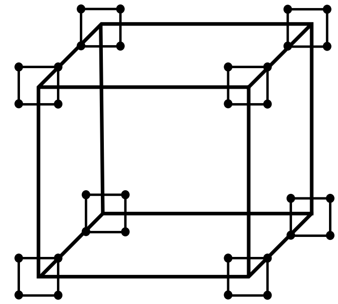


Fig. 7. Example of Taguchi design of experiment for 2^3 full factorial controllable factors and 2^2 full factorial noise factors.

TABLE I
A 12 RUN PB DESIGN MATRIX.

Run	A	B	C	D	E	F	G	H	I	J	K
1	+	+	−	+	+	+	−	−	−	+	−
2	−	+	+	−	+	+	+	−	−	−	+
3	+	−	+	+	−	+	+	+	−	−	−
4	−	+	−	+	+	−	+	+	+	−	−
5	−	−	+	−	+	+	−	+	+	+	−
6	−	−	−	+	−	+	+	−	+	+	+
7	+	−	−	−	+	−	+	+	−	+	+
8	+	+	−	−	−	+	−	+	+	−	+
9	+	+	+	−	−	−	+	−	+	+	−
10	−	+	+	+	−	−	−	+	−	+	+
11	+	−	+	+	+	−	−	−	+	−	+
12	−	−	−	−	−	−	−	−	−	−	−

maximum number of factors that can be studied is $4k-1$. For instance, in a 12-run design one can study up to 11 factors. A PBD matrix is shown in Table I in which the PB pattern is given in the first run. The other rows are generated by shifting one cell right the elements of the previous runs and at the end, a row of minus signs is added. PBDs utilize two levels for each factor, the higher level being denoted '$+$' and the lower '$-$'. Note, that for any pair of factors the number of combinations '$++$', '$+-$', '$-+$', and '$--$' is exactly the same.

7) Taguchi's idea: Taguchi procedures often refer to the experimental design as "off-line quality control" since they provide a good level of performance of products or processes occurred in the design space. Taguchi designs take into consideration that there exist factors (called noise factors) which cannot be controlled in normal operation of a product [11]. The other factors identifying as controllable parameters are able to control the variability and minimize the effect of the noise factors (robust parameter design problem).

In the world of simulations these parameters sometimes called *exploration parameters* (controllable factors) and *sensitivity analysis parameters* (noise/uncontrollable factors). The study then is aimed to compare the performance between different sets of exploration parameters as well as to compare the variation of performances due to noise factors. For example, the best performance may not be "the best" choice because of variability; where some suboptimal result is better because it is less affected by uncontrollable factors.

Fig. 7 shows a scenario of five parameters, two of which are uncontrollable and three are controllable. Two-levels full factorial experimental designs can be considered for both types of factors.

8) Monte Carlo: Such designs rely on random or pseudo-random samples of an input probability distribution [12]. The two main strengths on Monte Carlo methods are the following: first, the sampling process does not depend on number of factors – dimensionality, which means that any number of factors can be thrown into the design. Secondly, the data obtained from the running simulations can be analysed while new simulations are run. In this case to obtain better knowledge about the model one needs to run more simulations.

9) Latin Hypercube Designs (LHDs): A type of stratified Monte Carlo sampling is LHS which simultaneously stratifies on all input dimensions [8]. LHDs, already presented in Section I, require fewer iterations compared to Monte Carlo methods in order to recreate the input distribution. To recapitulate, the features of LHDs rely on both random and stratified sampling.

III. EXPERIMENTAL DESIGN CONSIDERATIONS

The purpose of experimental designs is to study the performance of systems and processes. The choice of the experiments as a series of tests is performed in a way to maximize efficiency. The problem in such design models is how to select a number of certain experiments from a possible infinite number of options in order to *a)* gain as much valuable information as possible and *b)* satisfy the constrains of limited resources because each experiment bears cost and requires time.

Actual applications utilized in computer models are typically characterized by the input parameters called factors. All factors can be seen as dimensions of a parameter space S. A single experiment requires all factors to be set at particular values. Such set of values correspond to a single point $\xi \in S$ – a design point. LHD generates a set of design points uniformly distributed within the multidimensional box in S, which is defined by the ranges of parameter values. Schematically this can be expressed via the following formula:

$$E[S] \to D$$

where E is a mathematical model which takes the definition of parameter space S as input and produces the experimental design D. LHD is an example of such engine.

The choice of point ξ in the experimental space determines the outcomes $\{\phi_i\}$ of the experiment from a variety of observable results ϕ in space O. Each replication, the repetition of the experiment aiming to acquire a more precise outcome distribution, is represented by the index i. For each selected ξ several replications can be executed and each replication results in possibly different ϕ_i. In stochastic simulations this variability is achieved by supplying different random seeds. In physical experiments this variability may come from measurement errors or the inability to control all factors that determine the outcome. For each point ξ the desired, and practically not achievable, outcome is to draw the real probability distribution of ϕ. The probability distribution can be approached with the growth of the number of replications executed at point ξ.

There is a dilemma however, in how to obtain the distribution of the observable outcomes: whether to increase number of replications in each point ξ or to increase number of points ξ. In the first case, more replications at fewer points ξ can result in higher granularity information. On the other hand, a greater number of points with fewer replications can provide a better illustration of the overall information. Finding an optimal balance between these two elements is outside of the scope of the paper. Had this question been resolved, a new engine,

$$E[S, O] \to D$$

taking into account the space of observables O, would provide a better design.

Besides the trade-off between the number of points ξ and the number of replications, the simulation cost C is also important. If different points ξ and different replications require the same amount of resources, then the cost is proportional to the number of simulations. In the general case however, not all the areas of the experimental space S would have the same cost effect: certain zones can be less expensively explored than others in terms of cost $C(\xi)$. Additionally, in complex experimental designs the cost $C(\xi)$ may not be known in advance but obtained as one of the results O. Obviously, the knowledge of the cost function would influence the optimal design:

$$E[S, O, C] \to D$$

Similarly with the simulation cost C, the value $V(\xi)$ of information obtained as a result of simulation at point ξ defines another aspect influencing D. Some areas of the experimental space S for instance, might be of greater importance for the system compared to other points ξ in the design. In addition, as with the simulation cost case, the value $V(\xi)$ of particular areas may depend on the obtained experimental results. For example, in combat simulations a balance between forces strongly governs winning or losing outcomes [13]. In the parameter space there may be vast areas of no particular interest because the outcome is predictable. The valuable information is the transition area where the probability to lose or win is not definite. An optimal design would have all design points in this transition area; then the engine E including function $V(\xi)$ becomes:

$$E[S, O, C, V] \to D$$

At first glance the value V constrain can be interpreted as quite artificial. One can assume to tolerate some extra experiments in not important areas as trade-off for simplicity of having a multidimensional box with flat boundaries. However, multidimensional spaces can be very counter-intuitive. For example, an n-dimensional unit ball submerged into a n-dimensional cube with edge equal to 2 (the ball touches each face of the cube) has its volume about 80% of the volume of the cube in 2-dimensions – a circle and a square. In 3-dimensional space this value is around 50%, in 10-dimensions $\approx 1.6\%$, and in 20-dimensional space the ratio of the ball's volume to the cube's volume is $2.5 \cdot 10^{-8}$. For the latter case of the 20-parameter space, a randomly chosen point within the smallest enclosing box will almost always be outside of the ball. In this example, having 20 factors requires more than 40 million points in order to get at least one point inside the ball representing a constrain on factors.

Another aspect that needs to be taken into account is the information I already obtained from the experiment. This information depends on the set of points which have been evaluated $\{\xi_i \to \{\phi\}\}$ plus the set of points which are being evaluated but do not have yet the result $\{\xi_j\}$. In general the information I is a function of these two sets: $I = I(\{\xi_i \to \{\phi\}\}, \{\xi_j\})$. Design engine E has to wisely produce more points where not enough information has been or expecting to be accumulated and less where the space is already explored:

$$E[S, O, C, V, I] \to D$$

This idea is known as sequential sampling.

The last aspect we consider in this paper is the resources R required to conduct the new experiment. In the simple scenario which the cost C follows a uniform distribution, the resources R are equivalent to the total number of samples, i.e. the sum of the number of replications N_ξ over the number of points ξ: $R = \sum_\xi N_\xi$.

The resource R in many cases influence the implementation of the design E. When the design points are evaluated sequentially, the order of evaluation is important, because depending on the intermediate outcomes the evaluation can be simply interrupted, re-designed, and/or re-evaluated. The design engine E therefore should include this aspect as well:

$$E[S, O, C, V, I, R] \to D$$

In its simplicity, LHD, besides the non-collapsing and filling properties in the experimental space S, ignores all the other aspects:

$$E[S, \cancel{O}, \cancel{C}, \cancel{V}, \cancel{I}, \cancel{R}] \to D$$

observational space O, cost C and value V functions, prior information I, and the experimentation resources R. From this point of view, applicability of LHD appears sound only after considering and consciously ignoring these five aspects.

IV. Conclusions

LHD is a useful and practical tool which can provide exceptional benefits in experimental design. In this paper we have highlighted aspects which are important in experimental design but, except parameter space S, are not present in the Latin Hypercube method. Awareness of those aspects brings justification of proper usage of LHD and sets up the boundaries where LHD can be safely utilized.

Considering these aspects brings up a question: Is there a method taking all known information and giving the optimal design? The authors believe that there exist the solutions – the perfect engine E – derived from the basic principles of probability theory. At the same time, it is possible that each aspect raised in this paper can be addressed by a simple ad hoc solution which may not be very far from optimal.

References

[1] M. McKay, R. Beckman, and W. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 42, no. 1, pp. 55–61, 2000.

[2] E.L. Melnick and B.S. Everitt, *Encyclopedia of Quantitative Risk Analysis and Assessment*, Number v. 1 in Encyclopedia of Quantitative Risk Analysis and Assessment. Wiley, 2008.

[3] J. Kleijnen, S. Sanchez, et al., "A userss guide to the brave new world of designing simulation experiments. tilburg university," *Center for Economic Research Discussion Paper*, , no. 1, 2003.

[4] F. AC Viana, "Things you wanted to know about the latin hypercube design and were afraid to ask," in *10th World Congress on Structural and Multidisciplinary Optimization, Orlando, Florida, USA (cf. p. 69)*, 2013.

[5] D.-A. Becker, "Overview of advantages and drawbacks of different methods for sensitivity analysis in the context of performance assessment," in *5th IGD-TP Technical Exchange Forum, Kalmar, Sweden*, 2014.

[6] A. Saltelli, K. Chan, et al., *Sensitivity analysis*, vol. 1, Wiley NY, 2000.

[7] B. J. Winer, D. R. Brown, and K. M. Michels, *Statistical principles in experimental design*, vol. 2, McGraw-Hill, 1971.

[8] A. Giunta, S. Wojtkiewicz, et al., "Overview of modern design of experiments methods for computational simulations," in *Proceedings of the 41st AIAA aerospace sciences meeting and exhibit, AIAA-2003-0649*, 2003.

[9] SL C. Ferreira, RE Bruns, et al., "Box-behnken design: an alternative for the optimization of analytical methods," *Analytica chimica acta*, vol. 597, no. 2, pp. 179–186, 2007.

[10] K Vanaja and RH Shobha Rani, "Design of experiments: concept and applications of plackett burman design," *Clinical research and regulatory affairs*, vol. 24, no. 1, pp. 1–23, 2007.

[11] Ranjit K Roy, *Design of experiments using the Taguchi approach: 16 steps to product and process improvement*, John Wiley & Sons, 2001.

[12] P. Paxton, P. Curran, et al., "Monte carlo experiments: Design and implementation," *Structural Equation Modeling*, vol. 8, no. 2, pp. 287–312, 2001.

[13] PB Williams and FDJ Bowden, "Dynamic morphological exploration," in *Proceedings of the 22nd National Conference of the Australian Society for Operations Research, Australia*, 2013.