CONSTRUCTION OF ORTHOGONAL ARRAY-BASED LATIN HYPERCUBE DESIGNS FOR DETERMINISTIC COMPUTER EXPERIMENTS

Kazeem A. Osuolale, Waheed B. Yahya, Babatunde L. Adeleke

Department of Statistics, University of Ilorin, P.M.B 1515, Ilorin, Nigeria

Corresponding author: Kazeem A. Osuolale, whereisqosimadewale@gmail.com

ABSTRACT: In this work, orthogonal array-based Latin hypercube designs (OALHDs) for eight input variables deterministic computer experiments was proposed. The proposed class of designs has some merits over the random Latin hypercube design with respect to space-filling properties. In this study, a computer program was written to construct OALHDs. Orthogonal arrays (OAs) were generated following a mathematical theorem called Bush Construction Type I and the OAs constructed were subsequently used to construct the desired orthogonal array-based Latin hypercube designs which was used in the development of deterministic computer experiments. The OALHDs constructed have some space-filling properties and they can be used to develop a computer experiment. A MATLAB program was written to construct the design.
KEYWORDS: Computer experiments, Bush construction type I, Latin hypercube designs, Orthogonal array, Space-filling properties

1. INTRODUCTION

The rapidly increasing power of computers has made experimentation via computer modelling common in many areas of science, engineering and technology. Some physical phenomena that are either extremely complex or impossible to investigate using classical statistical experiments have been described by mathematical models implemented in complex computer codes (known as simulators). The first computer experiment was reported to have been conducted by Enrico Fermi and colleagues, Strogatz [Str03] at the Los Alamos Scientific Laboratory in 1953. Since then, scientists in diverse disciplines have embraced computer experiments as a powerful tool to understand their respective processes. For instance, Osuolale et al. [OYA14b] proposed a simple pendulum experiment as a demonstrative example for computer experiments where the time it took a pendulum bob to return to rest was determined as the output. The inputs to the computer code can be varied in order to determine the effect of different inputs on the response(s). The output(s) of such computer models or simulators serve as a proxy for the physical experimentation. A **computer experiment** is conducted using data obtained from a computer model in place of the physical process. Space-filling design like Latin hypercube design (LHD) is commonly used in designing computer experiments. In this study, Space-filling orthogonal array-based LHDs were constructed. Space-filling designs are designs that spread points evenly throughout the experimental region.

Bush construction type I method found in Hedayat et al. [HSS99] was adopted in the construction of the design. The Bush construction type I is based on Galois fields. In abstract algebra, a field is composed of a set F, and two binary operations that map F X F into F. A simple example is the set of non-negative integers along with the operations of ordinary addition and multiplication. A Galois field is one for which the set F is finite. If F is a finite set of the integers, it is clear that ordinary addition and multiplication cannot be the operations of the field since, for example, adding the largest element of F to itself doesn't result in an element of F. This method makes use of the elements of the Galois field over the irreducible polynomial of the field.

2. ORTHOGONAL ARRAYS

Orthogonal arrays (OAs) introduced by Rao [Rao46] and Bose and Bush [BB52] provide better statistical information and are majorly used in designing experiments. An orthogonal array of n runs, m factors, s levels, strength $t \geq 2$ and index λ is an n-by- m matrix with entries from a set of s levels, usually taken as 0. . . s-1 such that for every n-by-m matrix of s symbols, every subset of t columns from among the m columns when considered alone must contain each of the possible s^t ordered rows the same number of times. The variables n, m, s, t and λ are referred to as the parameters of the OA and such an array is denoted by OA (n, m, s, t). The variable $\lambda = n/s^t$ is called the index of the orthogonal array and is determined by the other parameters. Regular fractional factorial designs, as discussed in Wu and Hamada [WH00] are the most familiar examples of orthogonal arrays. The OA with

$s_1 = s_2 = \cdots s_n = s$ is symmetric, otherwise, the array is said to be asymmetric.
The construction of OALHDs is highly dependent upon the existence of orthogonal arrays. Another important problem in the study of OAs is to determine either the minimal number of rows n in any OA (n, m, s, t) for given values m, s and t or the maximal number of columns m for given values n, s and t . The solution to this problem is adapted from the celebrated inequalities found by Rao [Rao47] for the construction of orthogonal arrays.

Theorem 1. (Rao's [Rao47] **Inequalities)**

(i) $n \geq \sum_{i=0}^{u} \binom{m}{i} (s-1)^i$ *if t = 2u*

and

(i)(ii) $n \geq \sum_{i=0}^{u} \binom{m}{i} (s-1)^i + \binom{m-1}{u} (s-1)^{u+1}$,
if t = 2u+1 for u ≥ 0

These inequalities provide a scheme for determining either a lower bound on the number of experiments *n* in any OA (*n, m, s, t*) design for given values m, s and t or an upper bound on the number of factors m for given values n, s and t. The proof of the Theorems can be found in Hedayat et al. [HSS99]. Furthermore, the use of these Theorems depends on whether t is even or odd. The approach adopted in this work for the construction of orthogonal array-based Latin hypercube designs is somewhat different from the existing methods as this approach makes use of the mathematical theorem to construct orthogonal arrays which gave rise to the desired OALHDs.

3. ORTHOGONAL ARRAY-BASED LATIN HYPERCUBE DESIGNS (OALHDs)

Latin hypercube designs were the earlier design proposed mainly for computer experiments, McKay et al. [MBC79]. Latin hypercube designs ensure maximum of the minimum distance between design points and require uniform spacing of the levels of each input variable. LHD is a good choice with respect to some useful criteria such as maximin distance and orthogonality. Latin hypercube designs have become popular for a number of reasons. The reasons are due to the fact that it allows the creation of experimental designs with as many points as possible because of flexibility in terms of data density and location, and in addition, non-collapsing and space-filling properties. Morris and Mitchell [MM95] considered maximin Latin hypercube designs to further enhance the space-filling property of maximin distance designs.

Many applications involve a large number of input variables and as such finding the space-filling designs with a limited number of design points that provide a good coverage of the entire high dimensional input space is a futile task. To break this curse of dimensionality, the approach of constructing designs that are space-filling in the low dimensional projections has been discussed by several researchers. Randomized orthogonal arrays, Owen [Owe92b] and orthogonal array-based Latin hypercube, Tang [Tan93] are such designs. An excellent review of design and analysis of computer experiments can be found in Koehler and Owen [KO96] and Santner et al. [SWN14].
A Latin hypercube design of size *n* has

$$X_{ij} = \frac{(d_{ij} - U_{ij})}{n}, i = 1, \dots, n, j = 1, \dots, m \qquad (1)$$

where $d_{1j}, \dots, d_{nj}$ are random permutations of the integers 1,... , n. $U_{ij} \sim U[0,1]$ and the m permutations and nm uniform variates are mutually independent, Owen [Owe92a]. Many authors use the simpler Lattice sample following Patterson [Pat54], where

$$X_{ij} = \frac{(d_{ij} - 0.5)}{n}, i = 1, \dots, n, j = 1, \dots, m \qquad (2)$$

These are two natural ways of generating design points in the unit cube $[0,1]^m$ based on a given Latin hypercube. When projected onto each of the m variables, both methods have the property that one and only one of the n design points fall within each of the n small intervals defined by $[0, 1/n), [1/n, 2/n), \dots, [(n-1)/n, 1]$. The first method gives the points that are uniformly distributed in their corresponding intervals while the second method gives the mid-points of these intervals. The variables (n, m, s, t) do not have the same name as we do for an orthogonal array because they do not all have a clear statistical interpretation, Tang [Tan91]. For instance, s does not refer to the number of levels of the design as referred to in orthogonal array since the design has n levels. The other parameters n, m and t are interpreted in the same way.
Orthogonal array-based Latin hypercube designs were proposed by Tang [Tan93]. Osuolale et al. [OYA14a] also presented a paper on the algorithm for constructing space-filling designs for Hadamard matrices of Orders 4λ and 8λ. Osuolale et al. [OYA14c] also constructed space-filling designs for three input variables computer experiments and these designs achieve better space-filling properties. The two methods employed in Osuolale et al.

[OYA14a] and Osuolale et al. [OYA14c] worked only for OA at two levels and three input variables respectively. Orthogonal arrays are used to construct LHDs in this study with maximin distance criterion to achieve better space-filling property and to extend the number of input variables in the experiment. Therefore, an OA based Latin hypercube design in the design space $[0, 1)^m$ can be generated.

Tang [Tan94] provided a way to obtain OA-based Latin hypercubes based on single replicated full factorial designs and proved that if the underlying orthogonal array is optimal with respect to the maximin distance criterion, so is the corresponding OA-based Latin hypercube. Leary et al. [LBK03] considered searching for optimal OA-based Latin hypercubes that minimize

$$\sum_{i=1}^{n}\sum_{j\neq i}\frac{1}{d_{ij}^{2}}$$

where d_{ij} is the Euclidean distance, defined as

$$d_{ij}=\frac{l_{ij}+{}^{(n-1)}\!/_{2}+u_{ij}}{n},\ i=1,\dots n,\ j=1,\dots m \tag{3}$$

with t = 2, between the i_{th} and j_{th} design points.

4. MATERIALS AND METHODS

A computer program was written in MATLAB to construct OALHDs.

Theorem 2 (Bush Construction Type I)

If $s \geq 2$ is a prime power then an $OA(s^t, s+1, s, t)$ of index unity exists whenever $s \geq t-1 \geq 0$. The level of the OA, s =7 while the strength, t=2 in this case.

Orthogonal arrays (OAs) were generated following Theorem 2 and they are thereafter used to construct the desired OALHDs using $X_{ij} = \frac{(d_{ij}-U_{ij})}{n}$. OA and OALHD as shown in the results section stand for the orthogonal array and the corresponding orthogonal array-based Latin hypercube designs.

5. RESULTS

Construction of OA (49, 8, 7, 2) LHD from Bush Construction Type I

[OA,OALHD] = oalhd_test(8,7)

OA=

0	0	0	0	0	0	0	0
0	1	1	2	3	5	1	6
0	2	2	4	6	3	2	5
0	3	3	6	2	1	3	4
0	4	4	1	5	6	4	3
0	5	5	3	1	4	5	2
0	6	6	5	4	2	6	1
1	0	1	1	2	3	5	1
1	1	2	3	5	1	6	0
1	2	3	5	1	6	0	6
1	3	4	0	4	4	1	5
1	4	5	2	0	2	2	4
1	5	6	4	3	0	3	3
1	6	0	6	6	5	4	2
2	0	2	2	4	6	3	2
2	1	3	4	0	4	4	1
2	2	4	6	3	2	5	0
2	3	5	1	6	0	6	6
2	4	6	3	2	5	0	5
2	5	0	5	5	3	1	4
2	6	1	0	1	1	2	3
3	0	3	3	6	2	1	3
3	1	4	5	2	0	2	2
3	2	5	0	5	5	3	1
3	3	6	2	1	3	4	0
3	4	0	4	4	1	5	6
3	5	1	6	0	6	6	5
3	6	2	1	3	4	0	4
4	0	4	4	1	5	6	4
4	1	5	6	4	3	0	3
4	2	6	1	0	1	1	2
4	3	0	3	3	6	2	1
4	4	1	5	6	4	3	0
4	5	2	0	2	2	4	6
4	6	3	2	5	0	5	5
5	0	5	5	3	1	4	5
5	1	6	0	6	6	5	4
5	2	0	2	2	4	6	3
5	3	1	4	5	2	0	2
5	4	2	6	1	0	1	1
5	5	3	1	4	5	2	0
5	6	4	3	0	3	3	6
6	0	6	6	5	4	2	6
6	1	0	1	1	2	3	5
6	2	1	3	4	0	4	4
6	3	2	5	0	5	5	3
6	4	3	0	3	3	6	2
6	5	4	2	6	1	0	1
6	6	5	4	2	6	1	0

OALHD=

0.0102	0.0102	0.0102	0.0102	0.0102	0.0102	0.0102	0.0102
0.0306	0.1531	0.1531	0.2959	0.4388	0.7245	0.1531	0.8673
0.0510	0.2959	0.2959	0.5816	0.8673	0.4388	0.2959	0.7245
0.0714	0.4388	0.4388	0.8673	0.2959	0.1531	0.4388	0.5816
0.0918	0.5816	0.5816	0.1531	0.7245	0.8673	0.5816	0.4388
0.1122	0.7245	0.7245	0.4388	0.1531	0.5816	0.7245	0.2959
0.1327	0.8673	0.8673	0.7245	0.5816	0.2959	0.8673	0.1531
0.1531	0.0306	0.1735	0.1735	0.3163	0.4592	0.7449	0.1735
0.1735	0.1735	0.3163	0.4592	0.7449	0.1735	0.8878	0.0306
0.1939	0.3163	0.4592	0.7449	0.1735	0.8878	0.0306	0.8878
0.2143	0.4592	0.6020	0.0306	0.6020	0.6020	0.1735	0.7449
0.2347	0.6020	0.7449	0.3163	0.0306	0.3163	0.3163	0.6020
0.2551	0.7449	0.8878	0.6020	0.4592	0.0306	0.4592	0.4592
0.2755	0.8878	0.0306	0.8878	0.8878	0.7449	0.6020	0.3163
0.2959	0.0510	0.3367	0.3367	0.6224	0.9082	0.4796	0.3367
0.3163	0.1939	0.4796	0.6224	0.0510	0.6224	0.6224	0.1939
0.3367	0.3367	0.6224	0.9082	0.4796	0.3367	0.7653	0.0510
0.3571	0.4796	0.7653	0.1939	0.9082	0.0510	0.9082	0.9082
0.3776	0.6224	0.9082	0.4796	0.3367	0.7653	0.0510	0.7653
0.3980	0.7653	0.0510	0.7653	0.7653	0.4796	0.1939	0.6224
0.4184	0.9082	0.1939	0.0510	0.1939	0.1939	0.3367	0.4796
0.4388	0.0714	0.5000	0.5000	0.9286	0.3571	0.2143	0.5000
0.4592	0.2143	0.6429	0.7857	0.3571	0.0714	0.3571	0.3571
0.4796	0.3571	0.7857	0.0714	0.7857	0.7857	0.5000	0.2143
0.5000	0.5000	0.9286	0.3571	0.2143	0.5000	0.6429	0.0714
0.5204	0.6429	0.0714	0.6429	0.6429	0.2143	0.7857	0.9286
0.5408	0.7857	0.2143	0.9286	0.0714	0.9286	0.9286	0.7857
0.5612	0.9286	0.3571	0.2143	0.5000	0.6429	0.0714	0.6429
0.5816	0.0918	0.6633	0.6633	0.2347	0.8061	0.9490	0.6633
0.6020	0.2347	0.8061	0.9490	0.6633	0.5204	0.0918	0.5204
0.6224	0.3776	0.9490	0.2347	0.0918	0.2347	0.2347	0.3776
0.6429	0.5204	0.0918	0.5204	0.5204	0.9490	0.3776	0.2347
0.6633	0.6633	0.2347	0.8061	0.9490	0.6633	0.5204	0.0918
0.6837	0.8061	0.3776	0.0918	0.3776	0.3776	0.6633	0.9490
0.7041	0.9490	0.5204	0.3776	0.8061	0.0918	0.8061	0.8061
0.7245	0.1122	0.8265	0.8265	0.5408	0.2551	0.6837	0.8265
0.7449	0.2551	0.9694	0.1122	0.9694	0.9694	0.8265	0.6837
0.7653	0.3980	0.1122	0.3980	0.3980	0.6837	0.9694	0.5408
0.7857	0.5408	0.2551	0.6837	0.8265	0.3980	0.1122	0.3980
0.8061	0.6837	0.3980	0.9694	0.2551	0.1122	0.2551	0.2551
0.8265	0.8265	0.5408	0.2551	0.6837	0.8265	0.3980	0.1122
0.8469	0.9694	0.6837	0.5408	0.1122	0.5408	0.5408	0.9694
0.8673	0.1327	0.9898	0.9898	0.8469	0.7041	0.4184	0.9898
0.8878	0.2755	0.1327	0.2755	0.2755	0.4184	0.5612	0.8469
0.9082	0.4184	0.2755	0.5612	0.7041	0.1327	0.7041	0.7041
0.9286	0.5612	0.4184	0.8469	0.1327	0.8469	0.8469	0.5612
0.9490	0.7041	0.5612	0.1327	0.5612	0.5612	0.9898	0.4184
0.9694	0.8469	0.7041	0.4184	0.9898	0.2755	0.1327	0.2755
0.9898	0.9898	0.8469	0.7041	0.4184	0.9898	0.2755	0.1327

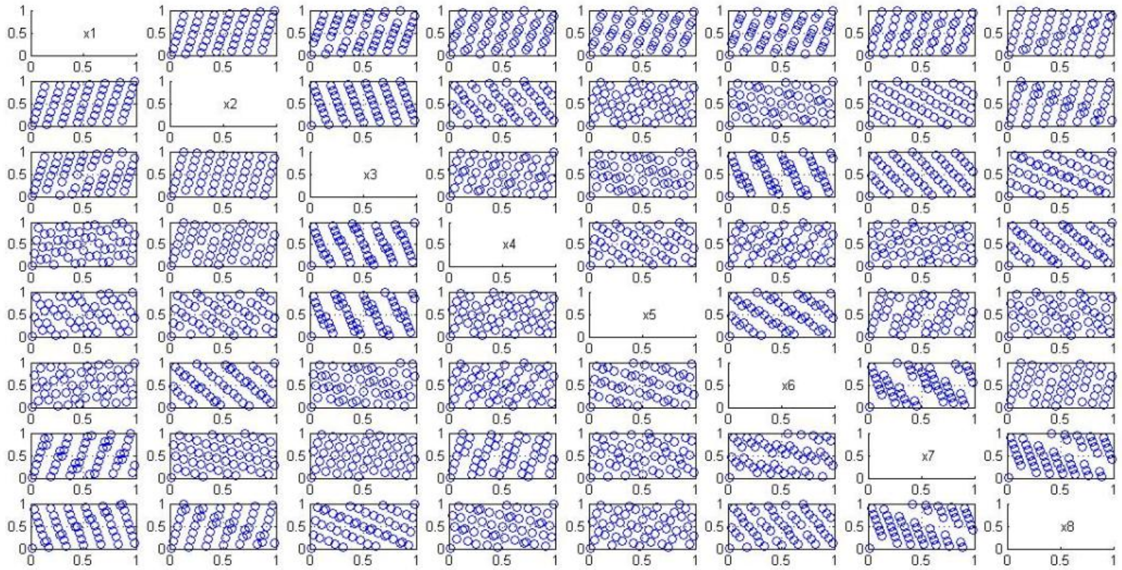
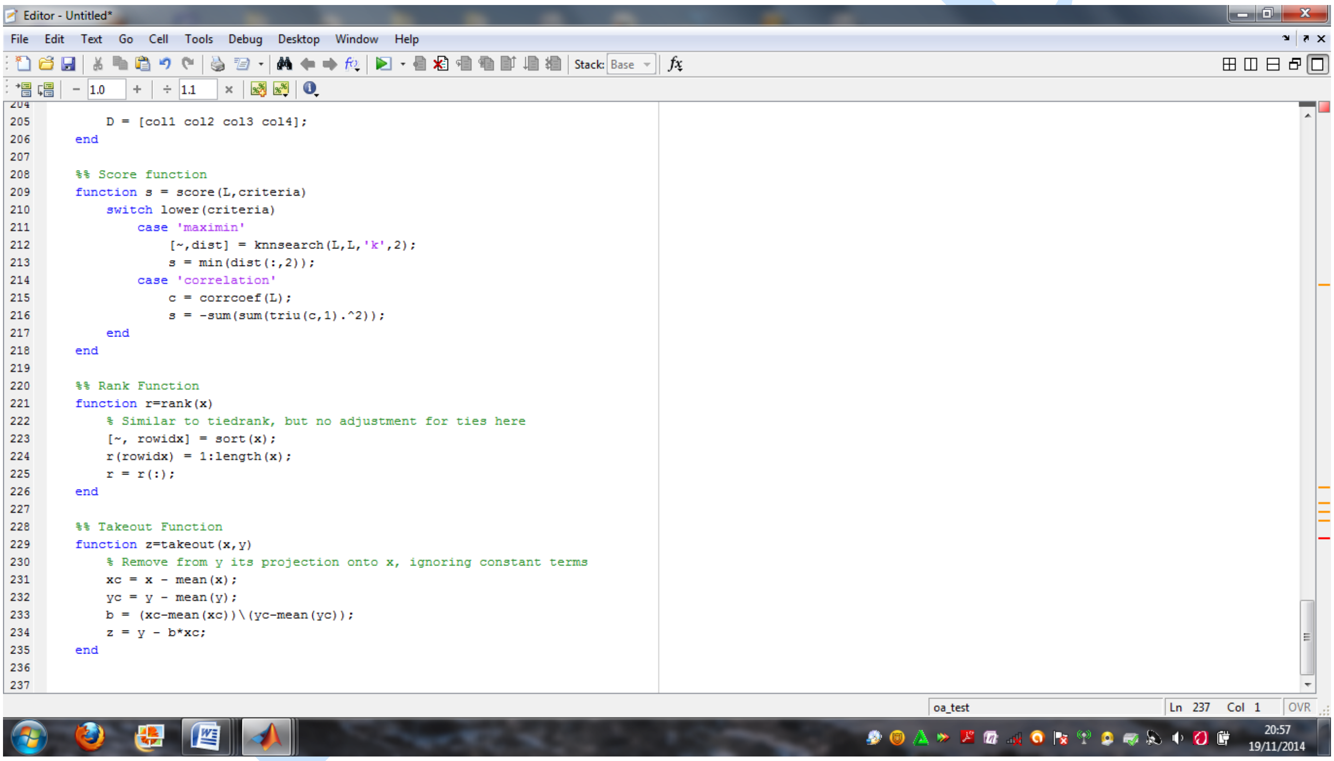


Figure 1. The Bivariate Projections among the 8 factors OA (49, 8, 7, 2) LHD depicting space-filling properties of the design



```
    D = [col1 col2 col3 col4];
end

%% Score function
function s = score(L,criteria)
    switch lower(criteria)
        case 'maximin'
            [~,dist] = knnsearch(L,L,'k',2);
            s = min(dist(:,2));
        case 'correlation'
            c = corrcoef(L);
            s = -sum(sum(triu(c,1).^2));
    end
end

%% Rank Function
function r=rank(x)
    % Similar to tiedrank, but no adjustment for ties here
    [~, rowidx] = sort(x);
    r(rowidx) = 1:length(x);
    r = r(:);
end

%% Takeout Function
function z=takeout(x,y)
    % Remove from y its projection onto x, ignoring constant terms
    xc = x - mean(x);
    yc = y - mean(y);
    b = (xc-mean(xc))\(yc-mean(yc));
    z = y - b*xc;
end
```

Figure 2. MATLAB Code for the Construction of Orthogonal Array-Based Latin Hypercube Designs

6. DISCUSSION

From our results, the following were observed as shown in the above results. Section 5 shows the construction of OA (49, 8, 7, 2)-LHD using Bush Construction Type I. OA (49, 8, 7, 2)-LHD contains 49 rows (runs) with 8 columns (factors) .We have been able to construct OALHDs for 8 factors with the strength of 2. The mathematical theorem can work for designs of various runs and different number of factors. It works for s = 3, 5, 7, 9, 11, 13, 15, … with factors, m = 4, 6, 8, 10, 12, 14, 16, … in a way that s is a prime number while m is even. The focus of this study is to consider s=7 to produce 8 input variables. We set U to be 0.5 to initialize the program in order to achieve the desired OALHDs.

7. CONCLUSION

This study presents the construction of OALHDs from Bush Construction Type I. There are several techniques and criteria available for the construction of space-filling designs. These include the use of difference matrices, Galois fields and orthogonal arrays among other existing techniques. The construction of OALHDs has been made easier by simplifying the rigorous mathematics involved into a computer program that runs in a twinkling of an eye. Our focus is to utilize the design to develop a borehole computer experiment for future research.The borehole model only comprises a univariate output and 8 input variables.

REFERENCES

[BB52] **R. C. Bose, K. A. Bush** - *Orthogonal Arrays of Strength Two and Three*, Annals of Mathematical Statistics, 23, 508–524, 1952.

[HSS99] **A. S. Hedayat, N. J. A. Sloane, J. Stufken** - *Orthogonal Arrays: Theory and Applications*, Springer-Verlag, New York, 1999.

[KO96] **J. R. Koehler, A. B. Owen** - *Computer Experiments: Handbook of Statistics*, 13, 261–308, 1996.

[LBK03] **S. Leary, A. Bhaskar, A. Keane** - *Optimal Orthogonal Array-Based Latin Hypercubes*, J. Applied Statist. 30, 585-598, 2003.

[MBC79] **M. D. McKay, R. J. Beckman, W. J. Conover** - *A comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code.* Technometrics, 21, 239 -245, 1979.

[MM95] **M. D. Morris, T. J. Mitchell** - *Exploratory Designs for Computer Experiments*, Journal of Statistical Planning and Inference, 43, 381-402, 1995.

[OYA14a] **K. A. Osuolale, W. B. Yahya, B. L. Adeleke** - *An Algorithm for Constructing Space-Filling Designs for Hadamard Matrices of Orders 4λ and $8\ \lambda$*, Proceedings of the 49^{th} Annual Conference of the Science Association of Nigeria, 27^{th} April-1^{st} May, 2014.

[OYA14b] **K. A. Osuolale, W. B. Yahya, B. L. Adeleke** - *Performing a Simple Pendulum Experiment as a Demonstrative Example for Computer Experiments*, Annals. Computer Science Series, Tome 12, Fasc. 2, 2014.

[OYA14c] **K. A. Osuolale, W. B. Yahya, B. L. Adeleke** - *Construction of Space-Filling Designs for Three Input Variables Computer Experiments*, WASET, International Journal of Computer, Control, Quantum and Information Engineering, 8(9), 1624-1628, 2014.

[Owe92a] **A. B. Owen** - *A Central Limit Theorem for Latin Hypercube Sampling*, J. Royal Statist. Soc. Ser. B 54, 541–551, 1992.

[Owe92b] **A. B. Owen** - *Orthogonal Arrays for Computer Experiments*, Integration and Visualization, Statistica Sinica, 2, 439-452, 1992.

[Pat54] **H. D. Patterson** - *The Error of Lattice Sampling*, Journal of Royal Statistical Society, Series B, 16,140 -149, 1954.

[Rao46] **C. R. Rao** - *Hypercube of Strength 'd' Leading to Confounded Designs in Factorial Experiments*, Bull. Calcutta Math. Soc., 38, 67-78, 1946.

[Rao47] **C. R. Rao** - *Factorial Experiments Derivable from Combinatorial Arrangements of Arrays*, Journal of the Royal Statistical Society (Suppl.), 9, 128-139, 1947.

[SWN14] **T. J. Santner, B. J. Williams, W. I. Notz** - *The Design and Analysis of Computer Experiments*, New York, Springer, 2014.

[Str03] **S. Strogatz** - *The Real Scientific Hero of 1953*, New York Times, March 4, Editorial/Op-Ed, 2003.

[Tan91] **B. Tang** - *Orthogonal Array-Based Latin Hypercubes*, IIQP Research Report, RR 91-10, 1991.

[Tan93] **B. Tang** - *Orthogonal Array-Based Latin Hypercubes*, Journal of the American Statistical Association, 88, 1392–1397, 1993.

[Tan94] **B. Tang** - *A Theorem for Selecting OA-Based Latin Hypercubes Using A Distance Criterion*, Comm. Statist., A Theory and Methods, 23, 2047-2058, 1994.

[WH00] **C. F. J. Wu, M. Hamada** - *Experiments: Planning, Analysis and Parameter Design Optimization*, Wiley, New York, 2000.