https://doi.org/10.56093/JISAS.V78I1.8
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 78(1) 2024 63–67

On Construction of Nearly Orthogonal Latin Hypercube Designs

A. Anil Kumar[1,2], Baidya Nath Mandal[3], Rajender Parsad[2] and Sukanta Dash[2]
[1]*The Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi*
[2]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi*
[3]*ICAR-Indian Agricultural Research Institute, Gauria Karma, Jharkhand*

Received 04 September 2023; Revised 16 April 2024; Accepted 21 April 2024

SUMMARY

Orthogonal Latin hypercube designs are becoming popular in designing computer experiments. Available literature on construction of orthogonal or nearly orthogonal LHDs has one or more restriction in terms of either runs or factors. In this article, we have proposed a method of construction for obtaining nearly orthogonal Latin hypercube designs capable of accommodating flexible number of runs or factors.

Keywords: Computer experiments; Latin hypercube designs; Orthogonal or nearly orthogonal Latin hypercube designs.

1. INTRODUCTION

Computer experiments are being widely conducted in those situations where the physical experiments and the process underlying it are too expensive, time-consuming or sometimes not even possible to observe. While designing computer experiments, Latin Hypercube designs are popularly used. These designs were introduced by McKay *et al.* (1979) and said to have the one-dimensional space-filling property.

A Latin hypercube design denoted as $LHD(n, m)$ is an $n \times m$ matrix whose columns are permutations of the column vector $(1, 2, ..., n)'$. For example, an LHD with 5 runs and 3 columns (or factors) is given below where columns are permutations of levels {1,2,3,4,5}.

1	2	5
3	4	2
4	1	3
5	3	1
2	5	4

Whenever polynomial models are considered, orthogonality will be an important criterion to evaluate LHDs. For any $LHD = (l_1, \ldots, l_m)$, where l_i is the i^{th} factor or column of the LHD, $\rho_{l_i l_j}$ is defined as the correlation coefficient between l_i and l_j and is given as $\rho_{l_i l_j} = l_i' l_j / (l_i' l_i l_j' l_j)^{1/2}$. If the correlation coefficient between any two factors is zero then LHD is orthogonal LHD (OLHD) i.e. $\rho_{ij}(LHD) = 0$ for all $i \neq j$ and if $\rho_{ij}(LHD)$ is near zero then the LHD is nearly orthogonal LHD. An example of an OLHD with 7 runs and 3 factors is given below.

1	7	6
2	4	1
3	2	3
4	1	5
5	3	7
6	5	2
7	6	4

It may be checked that correlation between any two factors in the above design is zero.

Ye (1998) provided a real example of using OLHD to simulate the cooling system of an injection molding process in which there were six input variables (factors).

Corresponding author: Baidya Nath Mandal
E-mail address: mandal.stat@gmail.com

He used a 17 x 6 OLHD in the computer experiment through which he built two regression models for two reponse variables, one of them measuring quality and another measuring productivity of the system.

An OLHD with n runs and m factors is denoted as $OLHD(n,m)$ and each factor includes n uniformly spaced levels presented in its centered form i.e., levels of factors are given as $\left\{-\frac{n-1}{2},-\frac{n-3}{2},\ldots,-\frac{n-2i+1}{2},\ldots,\frac{n-3}{2},\frac{n-1}{2}\right\}$. A 1st order OLHD is an LHD with mutually orthogonal columns that ensures independent estimation of linear effects when 1st order model is considered. Similarly, an LHD is said to be a 2nd order OLHD if any two columns are orthogonal and any column is orthogonal to element wise product of any other two columns, and these designs ensure independence of estimates of main effects and any quadratic or two factor interaction effects. An example of OLHD (8,4) is given below where levels are represented in centered form.

$$D = 1/2\begin{pmatrix} 1 & 3 & 5 & 7 & -1 & -3 & -5 & -7 \\ 3 & -1 & -7 & 5 & -3 & 1 & 7 & -5 \\ 5 & 7 & -1 & -3 & -5 & -7 & 1 & 3 \\ 7 & -5 & 3 & -1 & -7 & 5 & -3 & 1 \end{pmatrix}'$$

Different methods of construction are available in the literature. Tang (1993) proposed orthogonal arrays-based space filling Latin hypercube designs which have better space filling properties than randomly selected Latin hypercube designs and Tang (1994) proposed method of construction of maximin Latin Hypercube designs. Such designs were based on maximization of minimum distances between the pairs of design points**.** Ye (1998) provided the method of construction of OLHD for given $n = 2^k$ or $n = 2^{k+1}$ and $m = 2^{k-2}$ for any integer k > 1. Steinberg and Lin (2006) provided construction methods for obtaining OLHD for particular 2^{2^k} run sizes for any integer k. Cioppa and Lucas (2007) proposed an algorithmic approach to obtain OLHD, given a fixed sample size in more dimensions. Lin *et al.* (2009) coupled an orthogonal array of index 1 with a small orthogonal or nearly orthogonal Latin hypercube to obtain orthogonal or nearly orthogonal Latin hypercube. Sun *et al.* (2009) proposed a method for constructing OLHDs in which all the linear terms are orthogonal not only to each other, but also to the quadratic terms. Sun *et al.* (2010) presented an approach for obtaining OLHDs with $n = r2^{c+1}$ or $r2^{c+1}+1$ runs and for $k = 2^c$ columns/factors where, r is a positive integer. Lin *et al.* (2010) proved that no OLHD exists for $4r+2$ runs, where r is any positive integer and also developed a more flexible method for constructing orthogonal or nearly orthogonal Latin hypercube designs. Dey and Sarkar (2014) extended result of Lin *et al.* (2009) and used orthogonal array of strength 2 to obtain several 1^{st} order OLHD (n, m). Efthimiou *et al.* (2015) proposed methods for constructing nearly orthogonal Latin hypercube designs with 2,4,8,12,16,20 factors having flexible run sizes. Parui *et al.* (2015) proposed construction methods for obtaining OLHDs for all permissible number of runs for 2 and 3 factors. Mandal *et al.* (2016) proposed method for obtaining OLHDs with special reference to four factors.

One can see that the available methods OLHD construction usually have restrictions either in terms of runs or factors. Among others, Bingham *et al.* (2009), Lin *et al.* (2009, 2010), Sun *et al.* (2010), Gu and Yang (2013), Wang *et al.* (2015) also obtained methods to construct nearly OLHDs. In this paper we propose a method to construct nearly OLHDs. The method of construction developed provides nearly OLHDs with flexible number of runs and factors.

This paper is organized as follows. Section 2 contains the methodology to obtain nearly OLHDs. In Section 3, a comparison is made on the designs available in the literature and the designs obtained from proposed method. In Section 4 discussion is made on obtaining nearly OLHDs.

2. METHODOLOGY

Pang *et al.* (2022) developed construction method to obtain nearly OLHD $(2^{c+1}, 2^c + 2^{c-1})$ based on the construction method available in Sun *et al.* (2009) and Algorithm 1 in Wang *et al.* (2015) and focused on maximin L_2 distance criterion to show asymptotic optimality of the design. In this section, we present construction method to obtain nearly OLHD $(r2^{c+1}, 2^c + 2^{c-1})$ based on the method given in Sun *et al.* (2010) and Algorithm 1 of Wang *et al.* (2015). Consider the following steps to first obtain OLHD $(r2^{c+1}, 2^c)$, followed by steps to obtain nearly OLHD $(r2^{c+1}, 2^c + 2^{c-1})$ from resulted OLHD.

Step 1. Let c be any positive integer. Then for $c = 1$

$$S_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \text{ and } T_1 = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$$

Step 2. For $c \geq 2$, define S_c and T_c as follow

$$S_c = \begin{pmatrix} S_{c-1} & -S^*_{c-1} \\ S_{c-1} & S^*_{c-1} \end{pmatrix},$$

$$T_c = \begin{pmatrix} T_{c-1} & -(T^*_{c-1} + 2^{c-1} S^*_{c-1}) \\ T_{c-1} + 2^{c-1} S_{c-1} & T^*_{c-1} \end{pmatrix}$$

where, * operator on an even rowed matrix results into top half of the rows multiplied by –1 and others rows left without any change.

Step 3. Let $H_c = \left(T_c - \frac{S_c}{2} \right)$. For any positive integer c and r and $B_{r.2^c * 2^c} = \left(\left(H_c^1 \right)', \left(H_c^2 \right)', \ldots, \left(H_c^r \right)' \right)'$, where H_c^i is calculated as $H_c^i = H_c + (i-1)(2^c.S_c)$, where $i = 1, 2 \ldots r$.

Step 4. Then OLHD $(r2^{c+1}, 2^c)$ is given as

D $(r2^{c+1}, 2^c) = \begin{pmatrix} B_{r2^c * 2^c} \\ -B_{r2^c * 2^c} \end{pmatrix}$ and denote it by L.

Step 5. Let OLHD $(r2^c, 2^{c-1})$ be denoted as L_0, then E_1 and F_1 is given as, $E_1 = 2L_0 - J/2$, $F_1 = 2L_0 + J/2$, where J is a unit matrix of imension ($(r2^c * 2^{c-1})$.

Step 6. Let, $K = \left(E_1', F_1' \right)'$. Then nearly OLHD say P of dimensions $(r2^{c+1}, 2^c + 2^{c-1})$ is given as P $(r2^{c+1}, 2^c + 2^{c-1}) = (L, K)$.

Example 1. For $r = 3$ and $c = 2$, consider obtaining a nearly OLHD (24,6).

Firstly, we obtain T_2 from S_1 and T_1.

$$T_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & -4 & 3 \\ 3 & 4 & -1 & -2 \\ 4 & -3 & 2 & -1 \end{pmatrix}$$

Then, $H_2 = \left(T_2 - \frac{S_2}{2} \right) = \begin{pmatrix} 0.5 & 1.5 & 2.5 & 3.5 \\ 1.5 & -0.5 & -3.5 & 2.5 \\ 2.5 & 3.5 & -0.5 & -1.5 \\ 3.5 & -2.5 & 1.5 & -0.5 \end{pmatrix}$.

Now for $r = 2$ and $c = 2$, $B_{2.2^2 * 2^2} = \left(\left(H_2^1 \right)', \left(H_2^2 \right)' \right)'$ is obtained.

$H_2^1 = H_2 + (1-1)(2^2.S_2)$ (for $i = 1$). Therefore,

$$H_2^1 = \begin{pmatrix} 0.5 & 1.5 & 2.5 & 3.5 \\ 1.5 & -0.5 & -3.5 & 2.5 \\ 2.5 & 3.5 & -0.5 & -1.5 \\ 3.5 & -2.5 & 1.5 & -0.5 \end{pmatrix}.$$

Similarly, $H_2^2 = \begin{pmatrix} 4.5 & 5.5 & 6.5 & 7.5 \\ 5.5 & -4.5 & -7.5 & 6.5 \\ 6.5 & 7.5 & -4.5 & -5.5 \\ 7.5 & -6.5 & 5.5 & -4.5 \end{pmatrix}$ and

$H_2^3 = \begin{pmatrix} 8.5 & 9.5 & 10.5 & 11.5 \\ 9.5 & -8.5 & -11.5 & 10.5 \\ 10.5 & 11.5 & -8.5 & -9.5 \\ 11.5 & -10.5 & 9.5 & -8.5 \end{pmatrix}$ is obtained. Now OLHD $(2.2^{2+1}, 2^2)$ is given as OLHD $(16,4) = \begin{pmatrix} B_{3.2^2, 2^2} \\ -B_{3.2^2, 2^2} \end{pmatrix} =$

$$B_{12,4} = \begin{pmatrix} 0.5 & 1.5 & 2.5 & 3.5 & 4.5 & 5.5 & 6.5 & 7.5 & 8.5 & 9.5 & 10.5 & 11.5 \\ 1.5 & -0.5 & -3.5 & 2.5 & 5.5 & -4.5 & -7.5 & 6.5 & 9.5 & -8.5 & -11.5 & -10.5 \\ 2.5 & 3.5 & -0.5 & -1.5 & 6.5 & 7.5 & -4.5 & -5.5 & 10.5 & 11.5 & -8.5 & -9.5 \\ 3.5 & -2.5 & 1.5 & -0.5 & 7.5 & -6.5 & 5.5 & -4.5 & 11.5 & -10.5 & 9.5 & -8.5 \end{pmatrix}'.$$

Therefore, OLHD $(24,4) = \begin{pmatrix} B_{12,4} \\ -B_{12,4} \end{pmatrix} =$

Runs	X_1	X_2	X_3	X_4	**Runs**	X_1	X_2	X_3	X_4
1	0.5	1.5	2.5	3.5	**13**	-0.5	-1.5	-2.5	-3.5
2	1.5	-0.5	3.5	-2.5	**14**	-1.5	0.5	-3.5	2.5
3	2.5	-3.5	-0.5	1.5	**15**	-2.5	3.5	0.5	-1.5
4	3.5	2.5	-1.5	-0.5	**16**	-3.5	-2.5	1.5	0.5
5	4.5	5.5	6.5	7.5	**17**	-4.5	-5.5	-6.5	-7.5
6	5.5	-4.5	7.5	-6.5	**18**	-5.5	4.5	-7.5	6.5
7	6.5	-7.5	-4.5	5.5	**19**	-6.5	7.5	4.5	-5.5
8	7.5	6.5	-5.5	-4.5	**20**	-7.5	-6.5	5.5	4.5
9	8.5	9.5	10.5	11.5	**21**	-8.5	-9.5	-10.5	-11.5
10	9.5	-8.5	11.5	-10.5	**22**	-9.5	8.5	-11.5	10.5
11	10.5	-11.5	-8.5	9.5	**23**	-10.5	11.5	8.5	-9.5
12	11.5	10.5	-9.5	-8.5	**24**	-11.5	-10.5	9.5	8.5

Now, let OLHD $(r2^c * 2^{c-1})$ i.e., OLHD (12,2)
$$=\begin{pmatrix} 0.5 & 1.5 & 2.5 & 3.5 & 4.5 & 5.5 & -0.5 & -1.5 & -2.5 & -3.5 & -4.5 & -5.5 \\ 1.5 & -0.5 & 3.5 & -2.5 & 5.5 & -4.5 & -1.5 & 0.5 & -3.5 & 2.5 & -5.5 & 4.5 \end{pmatrix},$$
be considered as L_0. Then, E_1 and F_1 are obtained,
$$E_1=\begin{pmatrix} 1.5 & 3.5 & 5.5 & 7.5 & 9.5 & 11.5 & -0.5 & -2.5 & -4.5 & -6.5 & -8.5 & -10.5 \\ 3.5 & -0.5 & 7.5 & -4.5 & 11.5 & -8.5 & -2.5 & 1.5 & -6.5 & 5.5 & -10.5 & 9.5 \end{pmatrix},$$
and
$$F_1=\begin{pmatrix} 0.5 & 2.5 & 4.5 & 6.5 & 8.5 & 10.5 & -1.5 & -3.5 & -5.5 & -7.5 & -9.5 & -11.5 \\ 2.5 & -1.5 & 6.5 & -5.5 & 10.5 & -9.5 & -3.5 & 0.5 & -7.5 & 4.5 & -11.5 & 8.5 \end{pmatrix},$$
and K is given as,

$K=\left(E_1',F_1'\right)'$. Therefore, nearly OLHD $(r.2^{c+1}, 2^c + 2^{c-1})$ is given as $P=(L',K')'$.

The above obtained design P is a nearly orthogonal Latin hypercube design of dimension $(r.2^{c+1}, 2^c + 2^{c-1})\, i.e. (24,6)$. The correlation coefficient found between the column pairs (1,5), (1,6), (3,5), (3,6), (4,5), (4,6) and (5,6) is 0.06206, 0.06206, 0.010435, 0.010435, -0.005217, 0.005217 and 0.005217, respectively. For other column pairs, the correlation coefficient is zero. The maximum absolute correlation is 0.06206 found between column/factor pairs (1,5) and (1,6).

3. COMPARISON WITH DESIGNS AVAILABLE IN LITERATURE

Table 1 presents some nearly orthogonal LHDs constructed using proposed methodology and its comparison with some of the existing methods of constructing nearly OLHDs. In the first column, run sizes of the designs has been given and corresponding maximum number of factors under existing methods and proposed method has been given in other columns.

Table 1. Comparison of maximum number of factors of nearly orthogonal Latin hypercube designs for some given run sizes (< 200)

Run size	LMT	GY	GE	WYLL	PWY	Proposed method
8	-	6	-	-	6	6
16	-	12	-	15	12	12
24	-	-	-	23	-	6
32	-	24	-	31	24	24
40	-	-	-	-	-	6
48	-	-	-	47	-	12
56	-	-	-	24	-	6
64	48	48	-	63	48	48
72	-	-	-	-	-	6
80	-	-	-	-	-	12
88	-	-	-	-	-	6
96	-	-	24	71	-	24
104	-	-	-	-	-	6
112	-	-	-	-	-	12
120	-	-	-	-		6
128	-	96	-	96	96	96
136	-	-	-	-	-	6
144	-	-	-	-	-	12
152	-	-	-	-	-	6
160	-	-	-	36	-	24
168	-	-	-	-	-	6
176	-	-	-	-	-	6
184	-	-	-	-	-	6
192	-	-	-	72	-	48
200	-	-	-	48	-	6

Note: LMT = Lin, Mukerjee and Tang (2009); GY = Gu and Yang (2012): GE = Georgiou and Efthimiou (2015); WYLL = Wang, Yang, Lin and Liu (2015); PWY = Pang, Wang and Yang (2022).

n	X_1	X_2	X_3	X_4	X_5	X_6	***n***	X_1	X_2	X_3	X_4	X_5	X_6
1	0.5	1.5	2.5	3.5	1.5	3.5	**13**	-0.5	-1.5	-2.5	-3.5	0.5	2.5
2	1.5	-0.5	3.5	-2.5	3.5	-0.5	**14**	-1.5	0.5	-3.5	2.5	2.5	-1.5
3	2.5	-3.5	-0.5	1.5	5.5	7.5	**15**	-2.5	3.5	0.5	-1.5	4.5	6.5
4	3.5	2.5	-1.5	-0.5	7.5	-4.5	**16**	-3.5	-2.5	1.5	0.5	6.5	-5.5
5	4.5	5.5	6.5	7.5	9.5	11.5	**17**	-4.5	-5.5	-6.5	-7.5	8.5	10.5
6	5.5	-4.5	7.5	-6.5	11.5	-8.5	**18**	-5.5	4.5	-7.5	6.5	10.5	-9.5
7	6.5	-7.5	-4.5	5.5	-0.5	-2.5	**19**	-6.5	7.5	4.5	-5.5	-1.5	-3.5
8	7.5	6.5	-5.5	-4.5	-2.5	1.5	**20**	-7.5	-6.5	5.5	4.5	-3.5	0.5
9	8.5	9.5	10.5	11.5	-4.5	-6.5	**21**	-8.5	-9.5	-10.5	-11.5	-5.5	-7.5
10	9.5	-8.5	11.5	-10.5	-6.5	5.5	**22**	-9.5	8.5	-11.5	10.5	-7.5	4.5
11	10.5	-11.5	-8.5	9.5	-8.5	-10.5	**23**	-10.5	11.5	8.5	-9.5	-9.5	-11.5
12	11.5	10.5	-9.5	-8.5	-10.5	9.5	**24**	-11.5	-10.5	9.5	8.5	-11.5	8.5

It can be seen that proposed method can construct nearly OLHDs for many more number of runs than the existing methods for given number of columns.

4. DISCUSSION

LHDs are widely used in designing computer experiments. LHD with its orthogonal factors help in independent estimation of effect of individual factors on the output of the model. In this regard, construction of methods to obtain an orthogonal LHDs is gaining importance and because of difficulties in developing methods to obtain OLHDs, methods are developed for obtaining nearly OLHDs. In this paper a method to obtain nearly OLHD is developed based on the construction methods in Sun *et al.* (2010) and Wang *et al.* (2015). One restriction of the proposed method is that the run sizes can only be multiples of two. Therefore, future scope of the research would be to develop methods for other run size and/or factor combination. One can also study L_2 distance efficiency of the designs and see whether design is asymptotically optimal or not under maximin L_2 distance criterion and can extend results of Wang *et al.* (2018) to show any connection between maximin distance designs and orthogonal designs.

ACKNOWLEDGEMENTS

The authors are thankful to the anonymous reviewers for their noteworthy suggestions leading to the improvement in the manuscript.

REFRENCES

Bingham, D., Sitter, R.R., and Tang, B. (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika*, **96**(1), 51-65.

Cioppa, T. M., and Lucas, T.W. (2007). Efficient nearly orthogonal and space filling Latin hypercubes, *Technometrics,* **49**(1), 45-55.

Dey, A. and Sarkar, D. (2016). A note on the construction of orthogonal Latin hypercube designs, *Journal of Combinatorial Designs,* **24**(3), 105-111.

Efthimiou, I., Georgiou, S. D., and Liu, M. Q. (2015). Construction of nearly orthogonal Latin hypercube designs*. Metrika*, **78**, 45-57.

Gu, L. and Yang, J.F. (2013). Construction of nearly orthogonal Latin hypercube designs. *Metrika,* **76**, 819-830.

Lin, C.D., Mukerjee, R., and Tang, B. (2009). Construction of orthogonal and nearly orthogonal Latin hypercubes, *Biometrika,* **96**(1), 243-247.

Lin, C.D., Bingham, D., Sitter, R.R. and Tang, B. (2010). A new and flexible method for constructing designs for computer experiments. *The Annals of Statistics,* **38**(3), 1460-1477.

Mandal, B.N., Dash, S., Parui, S., and Parsad, R. (2016). Orthogonal Latin hypercube designs with special reference to four factors. *Statistics & Probability Letters*, **119,** 181-185.

McKay, M.D., Beckman, R.J., and Conover, W.J. (1979). Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, **21**(2), 239-245.

Pang, T., Wang, Y., and Yang, J. F. (2022). Asymptotically optimal maximin distance Latin hypercube designs. *Metrika*, **85**(4), 405-418.

Parui, S., Mandal, B.N., Parsad, R., and Dash, S. (2018). Orthogonal Latin hypercube designs for three columns. *Utilitas Mathematica*. **108**, 149-158

Steinberg, D.M., and Lin, D.K. (2006). A construction method for orthogonal Latin hypercube designs, *Biometrika*, **93**(2), 279-288.

Sun, F., Liu, M.Q. and Lin, D.K. (2009). Construction of orthogonal Latin hypercube designs. *Biometrika*, **96**, 971-974.

Sun, F., Liu, M.Q. and Lin, D.K. (2010). Construction of orthogonal Latin hypercube designs with flexible run sizes. *Journal of Statistical Planning and Inference*, **140**(11), 3236-3242.

Tang, B. (1993). Orthogonal array-based Latin hypercubes. *Journal of the American Statistical Association,* **88**, 1392-1397.

Tang, B. (1994). A theorem for selecting OA-based Latin hypercubes using a distance criterion. *Communications in Statistics: Theory and Methods*, **23**, 2047-2058.

Wang, L., Yang, J.F., Lin, D.K. and Liu, M.Q. (2015). Nearly orthogonal Latin hypercube designs for many design columns. *Statistica Sinica*, 1599-1612.

Wang, Y., Yang, J. and Xu, H. (2018). On the connection between maximin distance designs and orthogonal designs. *Biometrika*, **105**(2), 471-477.

Ye, K.Q. (1998). Orthogonal column Latin hypercubes and their application in computer experiments. *Journal of the American Statistical Association*, **93**(444), 1430-1439.