Metrika (2022) 85:405–418
https://doi.org/10.1007/s00184-021-00833-2

Asymptotically optimal maximin distance Latin hypercube designs

Tonghui Pang[1] · Yan Wang[1] · Jian-Feng Yang[2]

Received: 28 July 2020 / Accepted: 5 July 2021 / Published online: 20 July 2021

Abstract
Maximin distance designs and orthogonal designs have become increasingly popular in computer and physical experiments. The construction of such designs is challenging, especially under the maximin distance criterion. This paper studies a class of Latin hypercube designs by calculating the minimum distances between design points. We derive a general formula for the minimum intersite distance of this kind of design. The row pairs with the minimum intersite distance are also specified. The results show that such kind of Latin hypercube design is asymptotically optimal under both the maximin distance criterion and the orthogonality criterion.

Keywords Computer experiment · Maximin distance design · Space-filling design · Orthogonality

1 Introduction

Designs of computer experiments have received extensive attention in the past few decades. More and more scientists are trying to conduct computer simulation experiments to better understand complex physical phenomena. Latin hypercube designs (LHDs), proposed by McKay et al. (1979), have been popularly used for computer experiments because of their uniform coverage of each individual factor (Fang et al. (2006); Santner et al. (2003)). An $n \times k$ LHD, denoted by LHD(n, k), is a matrix of n rows and k columns each being a permutation of n equally-spaced levels. Hereafter, whenever we say a column, it refers to the one with n equally-spaced levels. In this paper, the n levels are taken to be $\{-(n-1)/2, -(n-3)/2, \ldots, (n-3)/2, (n-1)/2\}$. For an LHD$(n, k)$ $L = (l'_1, l'_2, \ldots, l'_n)'$, the L_2-distance between two distinct rows $l_i =$

✉ Jian-Feng Yang
jfyang@nankai.edu.cn

[1] School of Mathematics, Tianjin University, Tianjin 300350, China

[2] School of Statistics and Data Science, LPMC & KLMDASR, Nankai University, Tianjin 300071, China

$(l_{i1}, \ldots, l_{ik})$ and $l_j = (l_{j1}, \ldots, l_{jk})$ in L is defined to be $d_2(l_i, l_j) = \sum_{r=1}^{k} |l_{ir} - l_{jr}|^2$, which is the square of the Euclidean distance between l_i and l_j. The L_2-distance of L is defined to be $d_2(L) = \min\{d_2(l_i, l_j) : 1 \leq i \neq j \leq n\}$.

The maximin distance criterion, introduced by Johnson et al. (1990), is a significant criterion for choosing LHDs. A maximin distance design maximizes the $d_2(L)$ value so that no two design points are too close. Then maximin distance designs achieve the space-filling property in the full-dimensional space that cannot be guaranteed by random LHDs. A maximin distance design is asymptotically D-optimal under the Gaussian process model by Johnson et al. (1990). Tang (1994) constructed LHDs based on full factorial designs without replicates and demonstrated that if the latter is a maximin distance design, so is the corresponding LHD. Van Dam et al. (2007) constructed exact maximin L_2-distance LHDs with run sizes no more than 70 by the branch-and-bound algorithm. Zhou and Xu (2015) constructed the maximin L_1- and L_2-distance LHDs based on good lattice point sets and derived the upper bound of L_2-distance of a general U-type design. Ba et al. (2015) proposed a new construction approach of sliced LHDs (SLHDs) and such designs are approximate maximin distance LHDs. Joseph et al. (2015) proposed maximum projection (MaxPro) designs that maximize space-filling properties on projections to all subsets of factors. Lin and Kang (2016) provided an algebraic construction for maximin distance LHDs. Sun and Tang (2017) introduced a general method of constructing orthogonal designs that are space-filling in low-dimensional projections. Wang et al. (2018) showed the equivalence between orthogonality and maximin L_2-distance for a special class of LHDs. Some other studies on maximin distance designs include Morris and Mitchell (1995), Ye et al. (2000), Jin et al. (2005), Liefvendahl and Stocki (2006), Joseph and Hung (2008), Grosso et al. (2009), Van Dam et al. (2009), Viana et al. (2010), Zhu et al. (2012), Xiao and Xu (2018), and He (2019), among others. Although there is already rich literature for obtaining maximin distance designs, the construction of such designs is still difficult. For most design parameters, maximin distance LHDs can only be obtained by stochastic algorithms, which are inefficient and difficult to ensure that these designs can reach the largest minimum distance. For other space-filling designs, readers are referred to Lin and Tang (2015) for an excellent review.

In this paper, we revisit a special class of LHDs by calculating their minimum distances. The results show that these designs are asymptotically optimal maximin distance LHDs. As the sizes of the designs increase, such designs show better characteristics in terms of the maximin distance criterion. Wang et al. (2015) demonstrated that such designs are nearly orthogonal with a low correlation. So these designs are good in terms of both the orthogonality criterion and the maximin distance criterion.

This paper is organized as follows. Section 2 introduces the construction method of a kind of LHD(2^{c+1}, $2^c + 2^{c-1}$) based on existing results in Sun et al. (2009) and Wang et al. (2015), and gives some properties of the fold-over design M_c of Sun et al. (2009). Section 3 introduces a criterion for assessing the maximin distance of LHDs proposed by Zhou and Xu (2015) and derives the L_2-distance of such kind of LHD. The row pairs with the minimum intersite distance are also given. Section 4 contains some comparisons with the existing designs. Section 5 concludes with some remarks. All proofs are deferred to the Appendix.

2 The construction method

This section will first give the construction method of the LHD($2^{c+1}, 2^c + 2^{c-1}$) based on the construction method in Sun et al. (2009) and Algorithm 1 in Wang et al. (2015). It should be noted that Sun et al. (2009) constructed the LHD($2^{c+1}, 2^c$) with fold-over structure and Wang et al. (2015) expanded a general fold-over LHD to accommodate more factors. Both Sun et al. (2009) and Wang et al. (2015) only focused on the orthogonality criterion. Our construction method below constructs the LHD($2^{c+1}, 2^c + 2^{c-1}$) by first iteratively adopting the LHDs constructed in Sun et al. (2009) and then following Algorithm 1 in Wang et al. (2015). However, we focus on the maximin distance criterion for the resulting designs.

Construction 1

Step 1. For $c = 1$, let

$$S_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T_1 = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

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

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

where the operator $*$ works for any matrix with an even number of rows by multiplying the entries in the top half of the matrix by -1 and leaving those in the bottom half unchanged.

Step 3. The $2^{c+1} \times 2^c$ matrix M_c can be obtained as

$$M_c = ((T_c - S_c/2)', -(T_c - S_c/2)')', \tag{2}$$

where A' denotes the transpose of A.

Step 4. Take $E_{c-1} = 2M_{c-1} - J/2$ and $F_{c-1} = 2M_{c-1} + J/2$, where J is a $2^c \times 2^{c-1}$ matrix with all elements unity.

Step 5. Take $H_c = (E'_{c-1}, F'_{c-1})'$.

Step 6. Take $P_c = (M_c, H_c)$.

For any LHD L, the sum of the elements in any column is zero, so we define the correlation coefficient of any two columns $l_{(s)}$ and $l_{(t)}$ of the design as $\rho_{st}(L) = l'_{(s)} l_{(t)} / (l'_{(s)} l_{(s)} l'_{(t)} l_{(t)})^{1/2}$. An LHD is called orthogonal if the correlation coefficient between any two distinct columns is zero. A frequently used measure of orthogonality is $\rho_M(L) = \max\{|\rho_{st}(L)|, s \neq t, l_{(s)}, l_{(t)} \in L\}$. We call L nearly orthogonal if $\rho_M(L)$ is nearly zero. The following lemma shows some useful properties of the T_c, M_c, H_c and P_c in Construction 1 (Sun et al. 2009; Wang et al. 2015).

Lemma 1

(i) The T_c in (1) consists of rows and columns of permutations of the 2^c elements $1, \ldots, 2^c$, up to sign changes.
(ii) The matrix M_c in (2) is an orthogonal LHD$(2^{c+1}, 2^c)$.
(iii) The matrices H_c and P_c are nearly orthogonal LHD$(2^{c+1}, 2^{c-1})$ and nearly orthogonal LHD$(2^{c+1}, 2^c + 2^{c-1})$, respectively.

Then, we will further introduce several properties of M_c, which will be used in Sect. 3. According to the construction process of M_c, the following properties are obvious. Therefore, the proofs are omitted here.

Property 1 For the M_c in (2) ($c > 1$), all elements in the first row are positive, all elements in the $(2^c + 1)$th row are negative, and for other rows, half of the elements are positive and the other half are negative.

Property 2 For the M_c in (2) ($c > 1$), we have that

(i) the first row has the form of $(1, 3, \ldots, 2^{c+1} - 1)/2$;
(ii) in the $(2^{c-1} + 1)$th row,
  (a) the first half has the form of $(2^c + 1, 2^c + 3, \ldots, 2^{c+1} - 1)/2$; and
  (b) the second half has the form of $(-1, -3, \ldots, -(2^c - 1))/2$;
(iii) the $(2^c + 1)$th row has the form of $(-1, -3, \ldots, -(2^{c+1} - 1))/2$; and
(iv) in the $(2^c + 2^{c-1} + 1)$th row,
  (a) the first half has the form of $(-(2^c + 1), -(2^c + 3), \ldots, -(2^{c+1} - 1))/2$; and
  (b) the second half has the form of $(1, 3, \ldots, 2^c - 1)/2$.

Example 1 For $c = 2$, the orthogonal LHD(8, 4) M_2 constructed by Construction 1 is

$$M_2 = \begin{pmatrix} 0.5 & 1.5 & 2.5 & 3.5 \\ 1.5 & -0.5 & -3.5 & 2.5 \\ 2.5 & 3.5 & -0.5 & -1.5 \\ 3.5 & -2.5 & 1.5 & -0.5 \\ -0.5 & -1.5 & -2.5 & -3.5 \\ -1.5 & 0.5 & 3.5 & -2.5 \\ -2.5 & -3.5 & 0.5 & 1.5 \\ -3.5 & 2.5 & -1.5 & 0.5 \end{pmatrix}. \tag{3}$$

It is easy to check that Properties 1 and 2 hold.

3 Calculation of the L_2-distances

This section studies the L_2-distance of the LHD$(2^{c+1}, 2^c + 2^{c-1})$ P_c generated by Construction 1. The result shows that the design P_c is an asymptotically optimal maximin distance LHD. Note from Lemma 1 that, the design P_c is a nearly orthogonal LHD. Therefore, such kind of design is shown to be asymptotically optimal under both the maximin distance criterion and the orthogonality criterion. Hereafter, we denote $M_c = (m_{is})_{2^{c+1}\times 2^c} = (m_1', m_2', \ldots, m_{2^{c+1}}')'$, $H_c = (h_{is})_{2^{c+1}\times 2^{c-1}} = (h_1', h_2', \ldots, h_{2^{c+1}}')'$ and $P_c = (p_1', p_2', \ldots, p_{2^{c+1}}')'$ for any given c.

The following lemma provides the upper bound of the L_2-distance for a general LHD, which was proposed by Zhou and Xu (2015). Although this upper bound is unreachable in many cases, it is still often used to evaluate whether a design has a good L_2-distance.

Lemma 2 *For any LHD(n, k) L, we have $d_2(L) \le \lfloor n(n+1)k/6 \rfloor$, where $\lfloor x \rfloor$ represents the integer part of x.*

In order to quantitatively measure the performance of an LHD under the maximin distance criterion, we define the L_2-distance efficiency of an LHD(n, k) L as

$$d_{eff}(L) = d_2(L)/\lfloor n(n+1)k/6 \rfloor. \tag{4}$$

We call an LHD L to be an asymptotically optimal maximin distance LHD if $\lim_{n,k\to+\infty} d_{eff}(L) = 1$.

Before elaborating on the main results of this section, it is necessary to restate the following lemma proposed by Wang et al. (2018). This lemma illustrates the L_2-distance of any two distinct rows in the design M_c.

Lemma 3 *For the M_c in (2) ($c \ge 1$), we have*

$$d_2(m_i, m_j) = \begin{cases} \frac{1}{3} \times (2^{3c+2} - 2^c), & \text{if } |i-j| = 2^c; \\ \frac{1}{6} \times (2^{3c+2} - 2^c), & \text{otherwise.} \end{cases}$$

For convenience, we define a function ϕ of two distinct rows $x_i = (x_{i1}, \ldots, x_{ik})$ and $x_j = (x_{j1}, \ldots, x_{jk})$ in an $n \times k$ matrix X by $\phi(x_i, x_j) = \sum_{r=1}^{k}(x_{ir} - x_{jr})$. The following lemma will deduce the maximum ϕ value and the second-largest ϕ value for M_c. This lemma will play an important role in calculating the L_2-distance between two distinct rows in H_c.

Lemma 4 *For the design M_c in (2) ($c > 1$), we have*

- (i) *the maximum ϕ value of any two distinct rows is 2^{2c}, which can be achieved only at the ϕ value of the first row and the $(2^c + 1)$th row; and*
- (ii) *the second-largest ϕ value is $(3/4) \times 2^{2c}$, which can be achieved only at $\phi(m_1, m_{2^c+2^{c-1}+1})$ and $\phi(m_{2^{c-1}+1}, m_{2^c+1})$.*

In Example 1, by calculating the ϕ value of any two rows in M_2 in (3), it is easy to see that the maximum ϕ value is $\phi(m_1, m_5) = 2^4 = 16$ and the second-largest ϕ value is $\phi(m_1, m_7) = \phi(m_3, m_5) = (3/4) \times 2^4 = 12$.

In order to calculate the L_2-distance of the design $P_c = (M_c, H_c)$, we need to know the minimum distance and the second-smallest distance between any two distinct rows of H_c. The results are stated as the following theorem.

Theorem 1 *For the matrix H_c ($c \ge 2$), we have*

(i) *the minimum L_2-distance between any two distinct rows is 2^{c-1}, which can be achieved if and only if the row pair h_i and h_j satisfies $|i-j| = 2^c$; and*
(ii) *the second-smallest L_2-distance is*

$$\frac{1}{3} \times 2^{3c} - \frac{3}{4} \times 2^{2c} + \frac{1}{6} \times 2^c,$$

which can be achieved only at $d_2(h_1, h_{2^c+2^{c-1}+2^{c-2}+1})$ and $d_2(h_{2^{c-2}+1}, h_{2^c+2^{c-1}+1})$.

Example 2 For the design M_2 in Example 1, after the transformation of Steps 4 and 5 of Construction 1, we can get a 16×4 matrix H_3 as below

$$\begin{pmatrix} 0.5 & 2.5 & 4.5 & 6.5 & -1.5 & -3.5 & -5.5 & -7.5 & 1.5 & 3.5 & 5.5 & 7.5 & -0.5 & -2.5 & -4.5 & -6.5 \\ 2.5 & -1.5 & 6.5 & -5.5 & -3.5 & 0.5 & -7.5 & 4.5 & 3.5 & -0.5 & 7.5 & -4.5 & -2.5 & 1.5 & -6.5 & 5.5 \\ 4.5 & -7.5 & -1.5 & 2.5 & -5.5 & 6.5 & 0.5 & -3.5 & 5.5 & -6.5 & -0.5 & 3.5 & -4.5 & 7.5 & 1.5 & -2.5 \\ 6.5 & 4.5 & -3.5 & -1.5 & -7.5 & -5.5 & 2.5 & 0.5 & 7.5 & 5.5 & -2.5 & -0.5 & -6.5 & -4.5 & 3.5 & 1.5 \end{pmatrix}'.$$

The minimum L_2-distance between two distinct rows in H_3 is 4, which is $d_2(h_i, h_j)$ with $|i-j| = 8$. The second-smallest L_2-distance is $d_2(h_1, h_{15}) = d_2(h_3, h_{13}) = 124$. All L_2-distances between other row pairs in H_3 are larger than 124.

The main theorem of the paper is now ready to be stated. This theorem gives a general formula of the L_2-distance of the design P_c in Construction 1.

Theorem 2 *For any given $c \geq 2$, the L_2-distance of the LHD($2^{c+1}, 2^c + 2^{c-1}$) P_c is*

$$2^{3c} - \frac{3}{4} \times 2^{2c}, \tag{5}$$

which can be achieved only at $d_2(p_1, p_{2^c+2^{c-1}+2^{c-2}+1})$ and $d_2(p_{2^{c-2}+1}, p_{2^c+2^{c-1}+1})$.

Example 3 For $c = 3$, Construction 1 can generate an LHD(16, 12) P_3 as below

$$\left(\begin{array}{cccccccccccccccc} 0.5 & 1.5 & 2.5 & 3.5 & 4.5 & 5.5 & 6.5 & 7.5 & -0.5 & -1.5 & -2.5 & -3.5 & -4.5 & -5.5 & -6.5 & -7.5 \\ 1.5 & -0.5 & 3.5 & -2.5 & 5.5 & -4.5 & 7.5 & -6.5 & -1.5 & 0.5 & -3.5 & 2.5 & -5.5 & 4.5 & -7.5 & 6.5 \\ 2.5 & -3.5 & -0.5 & 1.5 & 6.5 & -7.5 & -4.5 & 5.5 & -2.5 & 3.5 & 0.5 & -1.5 & -6.5 & 7.5 & 4.5 & -5.5 \\ 3.5 & 2.5 & -1.5 & -0.5 & 7.5 & 6.5 & -5.5 & -4.5 & -3.5 & -2.5 & 1.5 & 0.5 & -7.5 & -6.5 & 5.5 & 4.5 \\ 4.5 & 5.5 & -6.5 & -7.5 & -0.5 & -1.5 & 2.5 & 3.5 & -4.5 & -5.5 & 6.5 & 7.5 & 0.5 & 1.5 & -2.5 & -3.5 \\ 5.5 & -4.5 & -7.5 & 6.5 & -1.5 & 0.5 & 3.5 & -2.5 & -5.5 & 4.5 & 7.5 & -6.5 & 1.5 & -0.5 & -3.5 & 2.5 \\ 6.5 & -7.5 & 4.5 & -5.5 & -2.5 & 3.5 & -0.5 & 1.5 & -6.5 & 7.5 & -4.5 & 5.5 & 2.5 & -3.5 & 0.5 & -1.5 \\ 7.5 & 6.5 & 5.5 & 4.5 & -3.5 & -2.5 & -1.5 & -0.5 & -7.5 & -6.5 & -5.5 & -4.5 & 3.5 & 2.5 & 1.5 & 0.5 \\ \hdashline 0.5 & 2.5 & 4.5 & 6.5 & -1.5 & -3.5 & -5.5 & -7.5 & 1.5 & 3.5 & 5.5 & 7.5 & -0.5 & -2.5 & -4.5 & -6.5 \\ 2.5 & -1.5 & 6.5 & -5.5 & -3.5 & 0.5 & -7.5 & 4.5 & 3.5 & -0.5 & 7.5 & -4.5 & -2.5 & 1.5 & -6.5 & 5.5 \\ 4.5 & -7.5 & -1.5 & 2.5 & -5.5 & 6.5 & 0.5 & -3.5 & 5.5 & -6.5 & -0.5 & 3.5 & -4.5 & 7.5 & 1.5 & -2.5 \\ 6.5 & 4.5 & -3.5 & -1.5 & -7.5 & -5.5 & 2.5 & 0.5 & 7.5 & 5.5 & -2.5 & -0.5 & -6.5 & -4.5 & 3.5 & 1.5 \end{array}\right)'.$$

The L_2-distance of P_3 is $2^9 - \frac{3}{4} \times 2^6 = 464$, which occurs at the L_2-distance between the first and 15th rows and the L_2-distance between the 3rd and 13th rows.

According to Theorem 2, we can easily get the L_2-distance efficiency of the design P_c, which is stated as the following theorem.

Table 1 L_2-distance of the LHD for some given orders

c	order	P_c	UB	SLHD		
				Maximum	Median	Minimum
2	8×6	52	72	62	56	51
3	16×12	464	544	466	446	399
4	32×24	3904	4224	3740	3598	3378
5	64×48	32000	33280	29656	29064	27836
6	128×96	259072	264192	238488	235582.5	229246

Theorem 3 *For the design P_c, the L_2-distance efficiency is*

$$d_{eff}(P_c) = 1 - \frac{5}{2^{c+2} + 2},$$

which tends to one as c increases infinitely.

The above theorem shows that the design P_c is asymptotically optimal under maximin L_2-distance criterion.

4 Some comparisons

In this section, we make some comparisons for our designs, SLHDs from Ba et al. (2015) and MaxPro LHDs by Joseph et al. (2015).

The R package *SLHD* for implementing the SLHD algorithm, proposed by Ba et al. (2015), can be downloaded from http://cran.r-project.org/. If the number of slices is one, the design produced by the package *SLHD* is an LHD. The comparison results are reported in Table 1 for $c = 2, \ldots, 6$, where the column "order" refers to the run size and column size for each c, the column "P_c" shows the L_2-distance of the P_c in Construction 1, and the column "UB" is the upper bound in Lemma 2. For SLHDs, we ran the *maximinSLHD* command with default settings 100 times and computed the L_2-distance for each design. Table 1 lists their maximum, median and minimum values.

The following observations can be seen from Table 1.

(i) For $c = 2$ and 3, the L_2-distance of P_c is smaller than that of the SLHD.
(ii) It is difficult to reach the upper bound in Lemma 2 for both P_c and SLHD.
(iii) As the parameter c increases, the design P_c performs better than the SLHD in terms of maximin L_2-distance.

The MaxPro designs maximize space-filling properties on projections to all subsets of factors. Our designs and SLHDs guarantee a good space-filling property in the full-dimensional space. From Fig. 1 in Joseph et al. (2015), it can be seen that the performance of MaxPro designs is not as good as SLHDs in the full-dimensional space, so it is not as good as our designs.

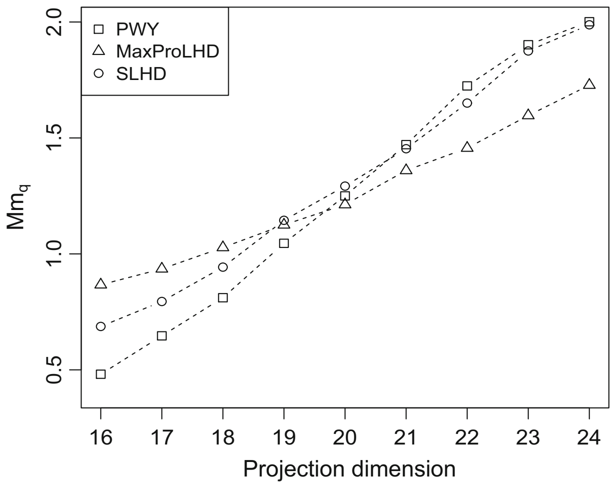

Fig. 1 Plot of Mm_q against q for our design (PWY), MaxPro LHD and SLHD

To study the projection properties, we projected the 32×24 designs onto 16 to 24 dimensions to compare our design with SLHD and MaxPro LHD. The SLHD we used is the best design in Table 1, which has an L_2-distance up to 3740. The MaxPro LHD is generated by the *MaxProLHD* function in the R package *MaxPro*. The comparison results are shown in Figs. 1, 2 and 3 . Figure 1 is the results for the maximin criterion Mm_q (larger the better) proposed by Joseph et al. (2015). Figures 2 and 3 are the results for the L_2-efficiency (larger the better) defined in (4), where for each case of dimensions, Fig. 2 compares the best one from all possible projections while Fig. 3 compares the means for all possible projections. From Fig. 1, it can be seen that, under the Mm_q criterion, the MaxPro LHD performs the best when projected onto smaller dimensions, while our design performs the best when projected onto larger dimensions. From Figs. 2 and 3 , we can see that, under the L_2-distance efficiency criterion, our design performs the best on all projections of 16–24 dimensions. For designs of size 16×12, similar comparison results can be found, so we omit the corresponding plots here.

5 Concluding remarks

In this work, by combining the constructions in Sun et al. (2009) and Wang et al. (2015) together, we found a new class of Latin hypercube designs (LHDs). Such LHDs, having 2^{c+1} runs and $2^c + 2^{c-1}$ factors for any $c > 1$, are shown to be asymptotically optimal in terms of the maximin L_2-distance criterion. Furthermore, the row pairs that achieve the minimum L_2-distance have been specified. In addition, we note that such LHDs are also nearly orthogonal from Wang et al. (2015). Therefore, the resulting LHDs perform well in terms of both the maximin L_2-distance criterion and the orthogonality criterion.

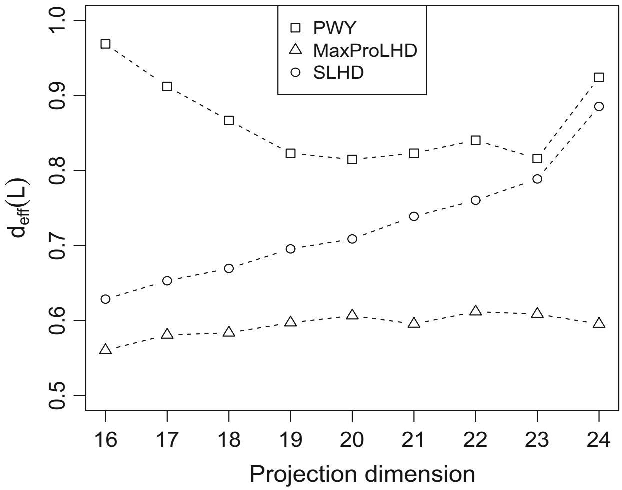

Fig. 2 Plot of the best $d_{eff}(L)$'s for our design (PWY), MaxPro LHD and SLHD projected onto q dimensions

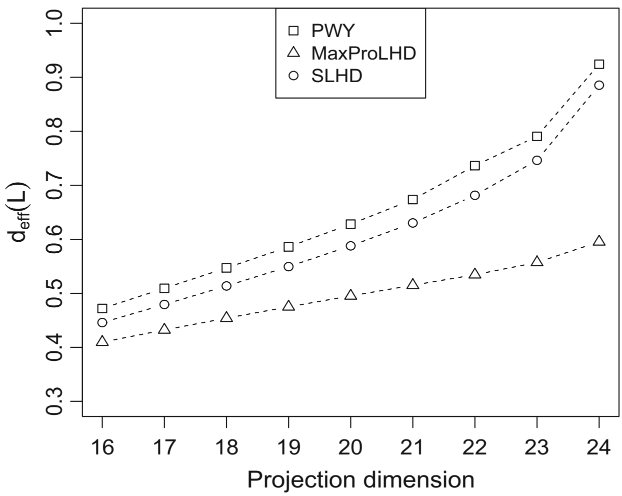

Fig. 3 Plot of the mean $d_{eff}(L)$'s for our design (PWY), MaxPro LHD and SLHD projected onto q dimensions

Wang et al. (2018) showed some connection between maximin distance designs and orthogonal designs when the column size is one half of the run size. Our work extended the results of Wang et al. (2018) and implied the coincidence, in some sense, when the column size is three quarters of the run size. One restriction of the proposed method is that the run size can only be a power of two. For other sizes of runs and/or columns, how to obtain maximin distance designs by construction method? Do they still have some connection with orthogonality? These are issues worth future study.

Acknowledgements We are grateful to the editor, the associate editor and two anonymous referees for their insightful comments and suggestions, which have led to significant improvements. This work was supported by the National Natural Science Foundation of China (Grant Nos. 11871033, 11871052 and 11971204), the Natural Science Foundation of Tianjin (20JCYBJC01050) and the Fundamental Research Funds for the Central Universities (63211090). The authorship is listed in alphabetic order.

Appendix

A.1. Proof of Lemma 4 First, based on Property 2, we can easily obtain the ϕ values for the following three row pairs

$$\begin{aligned}
\phi(m_1, m_{2^c+1}) &= \sum_{r=1}^{2^c}(m_{1,r} - m_{2^c+1,r}) \\
&= 2 \times (1 + 3 + \cdots + 2^{c+1} - 1)/2 = 2^{2c}, \\
\phi(m_1, m_{2^c+2^{c-1}+1}) & \\
&= (1 + 3 + \cdots + (2^{c+1}-1))/2 - (1 + 3 \\
&\quad + \cdots + (2^c-1) - (2^c+1) - (2^c+3) - \cdots - (2^{c+1}-1))/2 \\
&= 2 \times ((2^c + 1) + (2^c + 3) + \cdots + (2^{c+1} - 1))/2 = \frac{3}{4} \times 2^{2c},
\end{aligned}$$

and

$$\begin{aligned}
\phi(m_{2^{c-1}+1}, m_{2^c+1}) & \\
&= \sum_{r=1}^{2^c}(m_{2^{c-1}+1,r} - m_{2^c+1,r}) = \sum_{r=1}^{2^c}((-m_{2^c+2^{c-1}+1,r}) - (-m_{1,r})) \\
&= \sum_{r=1}^{2^c}(m_{1,r} - m_{2^c+2^{c-1}+1,r}) \\
&= \phi(m_1, m_{2^c+2^{c-1}+1}) = \frac{3}{4} \times 2^{2c}.
\end{aligned}$$

Next, we confirm that the value $\phi(m_1, m_{2^c+1})$ is the maximum ϕ value and the values $\phi(m_1, m_{2^c+2^{c-1}+1})$, $\phi(m_{2^{c-1}+1}, m_{2^c+1})$ are the second-largest ϕ value. According to Properties 1 and 2 , it is easy to see that the maximum ϕ value is the ϕ value of the first row and the $(2^c + 1)$th row. We will focus on the second-largest ϕ value.

For any $i, j \neq 1, 2^c + 1$, since the row vector m_i has a half of elements positive and others negative, so half of the 2^{c+1} elements of the sum formula $\phi(m_i, m_j)$ are negative. Then, if possible, the maximum value of $\phi(m_i, m_j)$ is

$$\begin{aligned}
&2 \times (((2^c + 1) + (2^c + 3) + \cdots + (2^{c+1} - 1))/2 - (1 + 3 + \cdots + (2^c - 1))/2) \\
&\quad = 2^c \times 2^{c-1} = \frac{1}{2} \times 2^{2c} < \frac{3}{4} \times 2^{2c}.
\end{aligned}$$

By Property 1, we can know the sum formula $\phi(m_1, m_i)$ has $2^c + 2^{c-1}$ positive elements. For the values $\phi(m_{2^c+1}, m_i)$, $\phi(m_i, m_1)$, and $\phi(m_i, m_{2^c+1})$, we have a similar conclusion. By the (i) and (iii) of Property 2, we have

$$\begin{aligned}
&\max\{\phi(m_1, m_i), \phi(m_{2^c+1}, m_i), \phi(m_i, m_1), \phi(m_i, m_{2^c+1})\}\\
&\quad\le ((1+3+\cdots+(2^{c+1}-1))+((2^c+1)+(2^c+3)+\cdots\\
&\qquad+(2^{c+1}-1))-(1+3+\cdots+(2^c-1)))/2\\
&\quad= \frac{3}{4}\times 2^{2c},
\end{aligned}$$

where the equality achieves at $\phi(m_1, m_{2^c+2^{c-1}+1})$ and $\phi(m_{2^{c-1}+1}, m_{2^c+1})$. Also, we note that

$$\phi(m_{2^c+1}, m_1) = -\phi(m_1, m_{2^c+1}) = -2^{2c} < \frac{3}{4}\times 2^{2c}.$$

Summarizing the above results, we can see that the second-largest ϕ value can only be achieved at $\phi(m_1, m_{2^c+2^{c-1}+1})$ and $\phi(m_{2^{c-1}+1}, m_{2^c+1})$. This completes the proof. □

A.2. Proof of Theorem 1 For $c = 2$, it is easy to show that the minimum L_2-distance between any two distinct rows in H_2 is $d_2(h_i, h_j) = 2$, where $|i-j| = 2^c$; and the second-smallest L_2-distance is $d_2(h_1, h_8) = d_2(h_2, h_7) = \frac{1}{3}\times 2^6 - \frac{3}{4}\times 2^4 + \frac{1}{6}\times 2^2 = 10$. In the following, we only consider the case of $c > 2$.

For convenience, we always assume $i < j$ and denote $\mathscr{A} = \{1, 2, \ldots, 2^c\}$ and $\mathscr{B} = \{2^c+1, 2^c+2, \ldots, 2^{c+1}\}$. We calculate the L_2-distances between any two distinct rows in H_c by dividing i and j into the following four cases.

(i) If $i, j \in \mathscr{A}$ or $i, j \in \mathscr{B}$, and $|i-j| = 2^{c-1}$, by Lemma 3, we have

$$d_2(h_i, h_j) = \sum_{r=1}^{2^{c-1}} |h_{ir} - h_{jr}|^2 = 4\times\frac{2^{3(c-1)+2} - 2^{c-1}}{3} = \frac{2}{3}\times 2^{3c} - \frac{2}{3}\times 2^c.$$

(ii) If $i, j \in \mathscr{A}$ or $i, j \in \mathscr{B}$, and $|i-j| \neq 2^{c-1}$, by Lemma 3, we have

$$d_2(h_i, h_j) = 4\times\frac{2^{3(c-1)+2} - 2^{c-1}}{6} = \frac{1}{3}\times 2^{3c} - \frac{1}{3}\times 2^c.$$

(iii) If $i \in \mathscr{A}$, $j \in \mathscr{B}$ and $|i-j| = 2^c$, we have

$$d_2(h_i, h_j) = \sum_{r=1}^{2^{c-1}} |h_{ir} - h_{jr}|^2 = \sum_{r=1}^{2^{c-1}} \left|(2m_{ir} - \frac{1}{2}) - (2m_{ir} + \frac{1}{2})\right|^2 = 2^{c-1}.$$

(iv) If $i \in \mathscr{A}$, $j \in \mathscr{B}$ and $|i-j| \neq 2^c$, we have

$$\begin{aligned} d_2(h_i, h_j) &= \sum_{r=1}^{2^{c-1}} |h_{ir} - h_{jr}|^2 \\ &= \sum_{r=1}^{2^{c-1}} \left| (2m_{ir} - \frac{1}{2}) - (2m_{j-2^c,r} + \frac{1}{2}) \right|^2 \\ &= 4d_2(m_i, m_{j-2^c}) - 4\phi(m_i, m_{j-2^c}) + 2^{c-1}. \end{aligned}$$

If $|i-(j-2^c)| = 2^{c-1}$, then we have

$$\begin{aligned} d_2(h_i, h_j) =&\ 4 \times \frac{2^{3(c-1)+2} - 2^{c-1}}{3} - 4\phi(m_i, m_{j-2^c}) + 2^{c-1} \\ \geq&\ 4 \times \frac{2^{3(c-1)+2} - 2^{c-1}}{3} - 4\phi(m_1, m_{1+2^{c-1}}) + 2^{c-1} \\ =&\ 4 \times \frac{2^{3(c-1)+2} - 2^{c-1}}{3} - 4 \times 2^{2(c-1)} + 2^{c-1} \\ =&\ \frac{2}{3} \times 2^{3c} - 2^{2c} - \frac{1}{6} \times 2^c. \end{aligned} \tag{6}$$

The equality (6) holds if and only if $i = 1$ and $j = 1 + 2^{c-1} + 2^c$.
If $|i-(j-2^c)| \neq 2^{c-1}$, then we have

$$\begin{aligned} d_2(h_i, h_j) =&\ 4 \times \frac{2^{3(c-1)+2} - 2^{c-1}}{6} - 4\phi(m_i, m_{j-2^c}) + 2^{c-1} \\ \geq&\ 4 \times \frac{2^{3(c-1)+2} - 2^{c-1}}{6} - 4 \times \frac{3}{4} \times 2^{2(c-1)} + 2^{c-1} \\ =&\ \frac{1}{3} \times 2^{3c} - \frac{3}{4} \times 2^{2c} + \frac{1}{6} \times 2^c. \end{aligned} \tag{7}$$

The equality (7) holds if and only if one of the following two conditions holds: (a) $i = 1$ and $j = 2^c + 2^{c-1} + 2^{c-2} + 1$; and (b) $i = 2^{c-2} + 1$ and $j = 2^c + 2^{c-1} + 1$.

By summarizing the above results, when $c \geq 2$, we can see that the minimum L_2-distance between any two distinct rows in H_c is $d_2(h_i, h_j) = 2^{c-1}$, where $|i-j| = 2^c$; and the second-smallest L_2-distance is $d_2(h_1, h_{2^c+2^{c-1}+2^{c-2}+1}) = d_2(h_{2^{c-2}+1}, h_{2^c+2^{c-1}+1}) = \frac{1}{3} \times 2^{3c} - \frac{3}{4} \times 2^{2c} + \frac{1}{6} \times 2^c$. All other L_2-distances between row pairs are larger than $\frac{1}{3} \times 2^{3c} - \frac{3}{4} \times 2^{2c} + \frac{1}{6} \times 2^c$. This completes the proof. □

A.3. Proof of Theorem 2 By Lemma 3, we know that, for any two rows of M_c, their L_2-distance is

$$d_2(m_i, m_j) = \begin{cases} \frac{1}{3} \times (2^{3c+2} - 2^c), & \text{if } |i-j| = 2^c; \\ \frac{1}{6} \times (2^{3c+2} - 2^c), & \text{otherwise.} \end{cases}$$

By Theorem 1, we know that the minimum L_2-distance between two distinct rows in H_c is 2^{c-1}. This value is the L_2-distance between h_i and h_j with $|i-j|=2^c$. So when $i, j \in \{1, 2, \ldots, 2^{c+1}\}$ and $|i-j|=2^c$, the L_2-distance between two rows p_i and p_j is

$$d_2(p_i, p_j) = \frac{1}{3} \times (2^{3c+2} - 2^c) + 2^{c-1} = \frac{4}{3} \times 2^{3c} + \frac{1}{6} \times 2^c.$$

By Theorem 1, we know that the second-smallest L_2-distance between two distinct rows of H_c is $d_2(h_1, h_{2^c+2^{c-1}+2^{c-2}+1}) = d_2(h_{2^{c-2}+1}, h_{2^c+2^{c-1}+1}) = \frac{1}{3} \times 2^{3c} - \frac{3}{4} \times 2^{2c} + \frac{1}{6} \times 2^c$. So when $i, j \in \{1, 2, \ldots, 2^{c+1}\}$ and $|i-j| \neq 2^c$, the minimum L_2-distance between two rows p_i and p_j is

$$\begin{aligned} d_2(p_1, p_{2^c+2^{c-1}+2^{c-2}+1}) &= d_2(p_{2^{c-2}+1}, p_{2^c+2^{c-1}+1}) \\ &= \frac{1}{6} \times (2^{3c+2} - 2^c) + \frac{1}{3} \times 2^{3c} - \frac{3}{4} \times 2^{2c} + \frac{1}{6} \times 2^c \\ &= 2^{3c} - \frac{3}{4} \times 2^{2c}. \end{aligned}$$

By summarizing above results, we know that the L_2-distance of the design P_c is $2^{3c} - \frac{3}{4} \times 2^{2c}$, which is either $d_2(p_1, p_{2^c+2^{c-1}+2^{c-2}+1})$ or $d_2(p_{2^{c-2}+1}, p_{2^c+2^{c-1}+1})$. This completes the proof. □

A.4. Proof of Theorem 3 According to Theorem 2, the L_2-distance efficiency of the design P_c is

$$d_{eff}(P_c) = \frac{2^{3c} - \frac{3}{4} \times 2^{2c}}{\frac{1}{6} \times 2^{c+1}(2^{c+1}+1)(2^c + 2^{c-1})} = 1 - \frac{5}{2^{c+2}+2},$$

which clearly tends to one as c increases infinitely. □

References

Ba S, Myers WR, Brenneman WA (2015) Optimal sliced Latin hypercube designs. Technometrics 57(4):479–487

Fang K-T, Li R, Sudjianto A (2006) Design and modeling for computer experiments. CRC Press, Boca Raton

Grosso A, Jamali A, Locatelli M (2009) Finding maximin Latin hypercube designs by iterated local search heuristics. Eur J Oper Res 197(2):541–547

He X (2019) Interleaved lattice-based maximin distance designs. Biometrika 106(2):453–464

Jin R, Chen W, Sudjianto A (2005) An efficient algorithm for constructing optimal design of computer experiments. J Stat Plan Inference 134(1):268–287

Johnson ME, Moore LM, Ylvisaker D (1990) Minimax and maximin distance designs. J Stat Plan Inference 26(2):131–148

Joseph VR, Hung Y (2008) Orthogonal-maximin Latin hypercube designs. Stat Sin 18:171–186

Joseph VR, Gul E, Ba S (2015) Maximum projection designs for computer experiments. Biometrika 102(2):371–380

Liefvendahl M, Stocki R (2006) A study on algorithms for optimization of Latin hypercubes. J Stat Plann Inference 136(9):3231–3247

Lin CD, Kang L (2016) A general construction for space-filling Latin hypercubes. Stat Sin 26:675–690

Lin CD, Tang B (2015) Latin hypercubes and space-filling designs. In: Dean A, Morris J, Stufken J, Bingham D (eds) Handbook of design and analysis of experiments. Chapman and Hall/CRC, New York, pp 593–625

McKay MD, Beckman RJ, Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2):239–245

Morris MD, Mitchell TJ (1995) Exploratory designs for computational experiments. J Stat Plan Inference 43(3):381–402

Santner TJ, Williams BJ, Notz WI (2003) The design and analysis of computer experiments. Springer, NewYork

Sun F, Liu M-Q, Lin DK (2009) Construction of orthogonal Latin hypercube designs. Biometrika 96(4):971–974

Sun F, Tang B (2017) A method of constructing space-filling orthogonal designs. J Am Stat Ass 112(518):683–689

Tang B (1994) A theorem for selecting oa-based Latin hypercubes using a distance criterion. Commun Stat Theory Methods 23(7):2047–2058

Van Dam ER, Husslage B, Den Hertog D, Melissen H (2007) Maximin Latin hypercube designs in two dimensions. Oper Res 55(1):158–169

Van Dam ER, Rennen G, Husslage B (2009) Bounds for maximin Latin hypercube designs. Oper Res 57(3):595–608

Viana FA, Venter G, Balabanov V (2010) An algorithm for fast optimal Latin hypercube design of experiments. Int J Numer Meth Eng 82(2):135–156

Wang L, Yang J-F, Lin DK, Liu M-Q (2015) Nearly orthogonal Latin hypercube designs for many design columns. Stat Sin 25(4):1599–1612

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

Xiao Q, Xu H (2018) Construction of maximin distance designs via level permutation and expansion. Stat Sin 28(3):1395–1414

Ye QK, Li W, Sudjianto A (2000) Algorithmic construction of optimal symmetric Latin hypercube designs. J Stat Plan Inference 90(1):145–159

Zhou Y, Xu H (2015) Space-filling properties of good lattice point sets. Biometrika 102(4):959–966

Zhu H, Liu L, Long T, Peng L (2012) A novel algorithm of maximin Latin hypercube design using successive local enumeration. Eng Optim 44(5):551–564