Statistics and Probability Letters 186 (2022) 109470

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Maximin L_1-distance Range-fixed Level-augmented designs

Yanping Gao, Siyu Yi, Yongdao Zhou *

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

ARTICLE INFO

Article history:
Received 12 September 2021
Received in revised form 17 February 2022
Accepted 7 March 2022
Available online 14 March 2022

MSC:
62K15
62K99

Keywords:
Follow-up strategy
Maximin distance design
Upper bound

ABSTRACT

We discuss two cases of range-fixed level-augmented designs (RFLADs) under the maximin L_1-distance criterion. The upper bounds of the L_1-distance and construction methods for RFLADs are given. We also study the column-orthogonality of the constructed RFLADs.

1. Introduction

In a physical experiment, the follow-up strategy is a popularly used technique when more information is needed to reach the experimental goal after analyzing the initial design. In some cases, the levels of some of the factors may be augmented in the follow-up stages. For example, Li et al. (2016) studied the factors affecting the extraction and optimized the yields of flavonols and anthocyanins. The initial two-level design screened the important factors, and the sequential design extended one level of the important factors. Zhu et al. (2018) studied the process optimization of foam sizing for cotton yarns. Some important factors that selected by the initial two-level design were augmented into four levels to further investigate in follow-up stage. Both Li et al. (2016) and Zhu et al. (2018) augmented the levels of the factors under the constraint that the experimental domain is fixed. When the true model is unknown, space-filling designs are commonly used. Inspired by the necessity of the augmentation of the number of levels, Gao et al. (2021) discussed the range-extended and range-fixed level-augmented designs under the wrap-around L_2-discrepancy (WD, Hickernell, 1998), which is a measure for one type of space-filling properties, uniformity, of a design. Similar results can be seen in Qin et al. (2016), Yang et al. (2017), Gou et al. (2018), Yang et al. (2019), and so on.

The maximin distance criterion is another type of space-filling properties (Johnson et al., 1990), which seeks the design points over an experimental domain to maximize the minimal L_p-distance among the pairs of points. Grosso et al. (2009) proposed the iterated local search algorithms to search maximin L_2-distance Latin hypercube designs and showed that those approach is efficient in many numerical experiments. Moon et al. (2011) gave a new and well-performing "smart swap" algorithm to generate maximin Latin hypercube designs quickly under L_2-distance criterion. Wang et al. (2018) used the William transformation to construct a series of equi-distant or nearly equi-distant designs, which are

* Corresponding author.
E-mail address: ydzhou@nankai.edu.cn (Y. Zhou).

https://doi.org/10.1016/j.spl.2022.109470

maximin L_1-distance designs. Li et al. (2020) presented a construction method which constructed large maximin L_1-distance designs from small designs. Yang et al. (2021) gave some deterministic construction methods for constructing maximin L_2-distance designs in two to five dimensions based on densest packings. However, in those literatures, they did not consider the issue for constructing the range-fixed level-augmented designs (RFLADs) under the maximin L_p-distance criterion. In this paper, we consider the maximin L_1-distance RFLADs, in which the number of levels is augmented and the added points should be scattered uniformly in the whole experimental domain coupled with the initial points under the L_1-distance criterion. We mainly consider the level augmentation for the two-level initial design and the number of levels of augmentation is 1 or 2, in which the number of levels of each factor is usually not large to reduce the experimental error for physical experiment, as shown in Li et al. (2016) and Zhu et al. (2018).

The rest of the paper is organized as follows. Section 2 presents some definitions and notations. Section 3 gives a construction method for the mixed two- and three-level RFLADs. Moreover, it can be proved that when the parameters meet certain conditions, the design constructed by the method is a maximin L_1-distance RFLAD. In Section 4, the construction method and the corresponding result for the mixed two- and four-level RFLADs are presented. Some conclusions and discussions are summarized in Section 5. All the proofs are given in the Appendix.

2. Definitions and notations

Denote $D(n; q_1, \ldots, q_m)$ as a design with n runs, m factors and the kth factor taking values from $\{1, \ldots, q_k\}$ for $k = 1, \ldots, m$. If some q_i's are unequal, it is called as an asymmetrical design and denoted by $D(n; q_1^{r_1}, \ldots, q_s^{r_s})$, where $m = \sum_{i=1}^{s} r_i$; otherwise, it is called as a symmetrical design and denoted by $D(n; q^m)$. Denote all of the $D(n; q^m)$ and $D(n; q_1^{r_1}, \ldots, q_s^{r_s})$ by $\mathcal{D}(n; q^m)$ and $\mathcal{D}(n; q_1^{r_1}, \ldots, q_s^{r_s})$, respectively. For a design, if each level in each column occurs equally often, we call it as the U-type design and denote it as $U(n; q_1, \ldots, q_m)$. The U-type design $U(n; q^m)$ and $U(n; q_1^{r_1}, \ldots, q_s^{r_s})$ can be defined similarly as above, as well as $\mathcal{U}(n; q^m)$ and $\mathcal{U}(n; q_1^{r_1}, \ldots, q_s^{r_s})$. For a design $D(n; q_1, \ldots, q_m)$, we define $d_p(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{m} |x_k - y_k|^p$ as the L_p-distance of any two rows $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_m)$, where integer $p \geq 1$. Denote by $d_p(D) = \min\{d_p(\mathbf{x}, \mathbf{y}) : \mathbf{x} \neq \mathbf{y}, \mathbf{x}, \mathbf{y} \in D\}$ the L_p-distance of $D(n; q_1, \ldots, q_m)$, and by $d_{p,ave}(D)$ and $d_{p,sum}(D)$ the average and the sum of the L_p-distance among all pairs of rows in the design $D(n; q_1, \ldots, q_m)$, respectively.

Given an initial design $\mathbf{d}_0$, the follow-up stage may add more points inside the experimental domain to augment the number of levels of some factors. Let n_1 be the number of the additional runs, m_1 and m_2 be the number of factors which need not and need to augment the number of levels, respectively. Then, as in Gao et al. (2021), a design $D = (\mathbf{d}_0^T \;\; \mathbf{d}_1^T)^T \in \mathcal{U}(n + n_1; 2^{m_1}(2 + q)^{m_2})$ is called the range-fixed level-augmented design (RFLAD), if $\mathbf{d}_0 \in \mathcal{U}(n; 2^m)$, in which the levels $\{1, 2\}$ become $\{1, 2 + q\}$ for the m_2 level-augmented factors, and $\mathbf{d}_1 \in \mathcal{D}(n_1; 2^{m_1}(2 + q)^{m_2})$. Let $\mathcal{L}_f(n + n_1; 2^{m_1}(2 + q)^{m_2})$ denote all the RFLADs. Based on the definition of RFLAD, we focus on the cases $q = 1$ and $q = 2$ in this paper, which lead to mixed two- and three-level RFLAD $D_1 \in \mathcal{L}_f(n + n_1; 2^{m_1}3^{m_2})$ and mixed two- and four-level RFLAD $D_2 \in \mathcal{L}_f(n+n_1; 2^{m_1}4^{m_2})$. These designs are commonly used in practice, such as Li et al. (2016) and Zhu et al. (2018), and thus it is meaningful to consider these types of designs.

For a RFLAD $D = (\mathbf{d}_0^T \;\; \mathbf{d}_1^T)^T = (x_{ik})_{1\leq i\leq n+n_1, 1\leq k\leq m}$, define $d_{1,sum}(D)$ as

$$d_{1,sum}(D) = \sum_{i=1}^{n+n_1} \sum_{j=1}^{n+n_1} \sum_{k=1}^{m} |x_{ik} - x_{jk}| = d_{1,sum}(\mathbf{d}_0) + d_{1,sum}(\mathbf{d}_1) + 2 \cdot d_{1,sum}(\mathbf{d}_0, \mathbf{d}_1), \tag{1}$$

where $d_{1,sum}(\mathbf{d}_0)$ and $d_{1,sum}(\mathbf{d}_1)$ are the sums of L_1-distance among all pairs of runs in the initial design $\mathbf{d}_0$ and the added portion $\mathbf{d}_1$, respectively, and $d_{1,sum}(\mathbf{d}_0, \mathbf{d}_1)$ is the sum of the L_1-distance among all pairs of runs with one coming from $\mathbf{d}_0$ and another from $\mathbf{d}_1$. The corresponding average value of each part denoted as $\lfloor d_{1,ave}(\mathbf{d}_0)\rfloor$, $\lfloor d_{1,ave}(\mathbf{d}_1)\rfloor$ and $\lfloor d_{1,ave}(\mathbf{d}_0, \mathbf{d}_1)\rfloor$, respectively. According to the definition of $d_1(D)$, we have $d_1(D) \leq d_U(D)$, where $d_U(D) = \min\{\lfloor d_{1,ave}(\mathbf{d}_0)\rfloor, \lfloor d_{1,ave}(\mathbf{d}_1)\rfloor, \lfloor d_{1,ave}(\mathbf{d}_0, \mathbf{d}_1)\rfloor\}$. Comparing $\lfloor d_{1,ave}(D_1)\rfloor$, the common upper bound of $d_1(D_1)$ (Wang et al., 2018), we use $d_U(D)$ to measure the space-filling property of design D. The relationship between $\lfloor d_{1,ave}(D_1)\rfloor$ and $d_U(D)$ will be discussed separately in Sections 3 and 4. Then, we give the definition of the maximin L_1-distance RFLAD.

Definition 1. A level-augmented design $D \in \mathcal{L}_f(n + n_1; 2^{m_1}(2 + q)^{m_2})$ is called the maximin L_1-distance RFLAD, if its distance efficiency $d_{p,eff}(D) = d_1(D)/d_U(D)$ is equal to 1 among $\mathcal{L}_f(n + n_1; 2^{m_1}(2 + q)^{m_2})$.

In the following sections, we will give methods to construct mixed two- and three-level and mixed two- and four-level RFLADs and discuss the corresponding L_1-distance of these designs.

3. Construction of mixed two- and three-level RFLADs

In this section, we provide a general construction method for mixed two- and three-level RFLAD D_1 and prove that D_1 is a maximin L_1-distance RFLAD under some requirements. Let **1**, **2**, **3** and **4** denote the $2^k \times 1$ vectors of 1s, 2s, 3s and 4s, respectively.

Construction 1.

Step 1. Given a two-level design $X \in \mathcal{U}(2^k; 2^{2^k-1})$, the initial design $\mathbf{d}_0 = \begin{pmatrix} \mathbf{1} & X & X \\ \mathbf{2} & X & X_f \end{pmatrix}$ is a two-level design with 2^{k+1} runs and $2^{k+1}-1$ columns, where X_f is the foldover design of X, i.e., the level 1 in X convert to the level 2 in X_f, and the level 2 in X convert to the level 1 in X_f;

Step 2. Change **2** in the first column of $\mathbf{d}_0$ into **3** to obtain $\mathbf{d}'_0$;

Step 3. Add the additional portion $\mathbf{d}_1 = \begin{pmatrix} \mathbf{2} & X_f & X \end{pmatrix}$ to the design $\mathbf{d}'_0$, $D_1 = \begin{pmatrix} \mathbf{d}'^T_0 & \mathbf{d}^T_1 \end{pmatrix}^T$ is the resulting mixed two- and three-level RFLAD.

In the Construction Method 1, the number of the added runs is the minimal value such that the resulting RFLAD is a U-type design. It is reasonable to consider this case for saving cost. To derive the requirements for constructing a maximin RFLAD under L_1-distance in Construction Method 1, we first give a lemma to calculate the sum of L_1-distance among all pairs of runs in the RFLAD D_1.

Lemma 1. *For any initial design $\mathbf{d}_0 \in \mathcal{U}(n; 2^m)$ and the additional portion $\mathbf{d}_1 \in \mathcal{D}(n_1, 2^{m_1}3^{m_2})$, we have*

$$(1)\ d_{1,sum}(\mathbf{d}_0) = n^2\, m, \quad (2)\ d_{1,sum}(\mathbf{d}_1) = n_1^2 m_1, \quad (3)\ d_{1,sum}(\mathbf{d}_0, \mathbf{d}_1) = nn_1\, m.$$

The proof of Lemma 1 is given in the Appendix. From Lemma 1, it is easy to obtain that for any $D_1 \in \mathcal{L}_f(n+n_1; 2^{m_1}3^{m_2})$, $d_U(D_1) = \min\{\lfloor nm/(n-1)\rfloor, \lfloor n_1 m_1/(n_1-1)\rfloor, \lfloor m\rfloor\}$, it is used to measure the space-filling property of RFLAD D_1. Moreover, by Definition 1, if a design $D_1 \in \mathcal{L}_f(n+n_1; 2^{m_1}3^{m_2})$ satisfies $d_1(D_1) = d_U(D_1)$, then D_1 is the mixed two- and three-level maximin L_1-distance RFLAD. The following theorem gives the conditions for being a maximin L_1-distance RFLAD.

Theorem 1. *If the design $X \in \mathcal{U}(2^k; 2^{2^k-1})$ is a two-level saturated design, then the initial design $\mathbf{d}_0$ that is constructed by Construction Method 1 is a maximin design under L_1-distance, and the resulting mixed two- and three-level RFLAD $D_1 \in \mathcal{L}_f(3\cdot 2^k; 2^{2^{k+1}-2}3^1)$ is a maximin L_1-distance RFLAD with the L_1-distance being $2^{k+1}-3$.*

Remark 1. Based on the settings of n and m in Theorem 1, we have $d_U(D_1) \le \lfloor d_{1,ave}(D_1)\rfloor$.

The proof of Theorem 1 and Remark 1 are given in the Appendix. By Remark 1, it is more reasonable to use $d_U(D_1)$ than $\lfloor d_{1,ave}(D_1)\rfloor$ to measure the space filling of D_1. Theorem 1 requires that the design X is a two-level saturated design. According to Fang et al. (2005), the initial design $\mathbf{d}_0$ constructed by Step 1 of Construction Method 1 is a maximin L_1-distance design. By the proof of Theorem 1, we can also obtain D_1 is a mixed two- and three-level maximin L_1-distance RFLAD. To illustrate the usefulness of Construction Method 1, we give an example as follows.

Example 1. According to Step 1, the initial design $\mathbf{d}_0$ formed based on $X \in \mathcal{U}(4; 2^3)$ is a two-level saturated design with eight runs and seven columns. By Step 2, we change the initial design $\mathbf{d}_0$ to $\mathbf{d}'_0$ whose last six columns are the same as $\mathbf{d}_0$ and the first column is $(1, 1, 1, 1, 3, 3, 3, 3)^T$. Then, we add the additional portion $\mathbf{d}_1$ by Step 3, where X and $\mathbf{d}_1$ are as follows:

$$X = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix}, \quad \mathbf{d}_1 = \begin{pmatrix} 2 & 2 & 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 2 & 2 & 2 & 1 & 1 \\ 2 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}.$$

Then $D_1 = (\mathbf{d}'_0\ \ \mathbf{d}_1)^T$ is the resulting mixed two- and three-level maximin RFLAD with the L_1-distance being 5.

4. Construction of mixed two- and four-level RFLADs

Adding more points to the interior of the experimental domain to extend the levels of some factors from two to four is also worthy studying. As similar as Section 3, we first give a construction method for mixed two- and four-level RFLAD D_2, and then prove that D_2 is a maximin L_1-distance RFLAD under some requirements.

Construction 2.

Step 1. Given a two-level design $X \in \mathcal{U}(2^k; 2^{2^k-1})$, the initial design $\mathbf{d}_{00} = \begin{pmatrix} X & X \end{pmatrix}$ is the supersaturated design with 2^k runs and $2^{k+1}-2$ columns;

Step 2. Change the levels 1, 2 in any lth ($l = 1, 2, \ldots, 2^k-1$) column of X to levels 1, 4, respectively, the changed design is denoted as X_1. Then we obtain $\mathbf{d}_0 = \begin{pmatrix} X & X_1 \end{pmatrix}$;

Step 3. Change the levels 1,2 in the lth column of X_f to the levels 2,3, respectively, denoted as X_2. Then the additional portion $\mathbf{d}_1 = \begin{pmatrix} X & X_2 \end{pmatrix}$, and $D_2 = \begin{pmatrix} \mathbf{d}^T_0 & \mathbf{d}^T_1 \end{pmatrix}^T$ is the resulting mixed two- and four-level RFLAD.

Table 1
The two- and four-level RFLADs D_2, D_2', D_2''.

	$D_2(8, 2^5 4^1)$		$D_2'(8, 2^4 4^2)$		$D_2''(8, 2^4 4^1)$
$\mathbf{d}_0$	$\begin{pmatrix} 1 & 1 & 2 & 1 & 1 & 4 \\ 1 & 2 & 1 & 1 & 2 & 1 \\ 2 & 1 & 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 4 \end{pmatrix}$	$\mathbf{d}_0'$	$\begin{pmatrix} 1 & 1 & 4 & 1 & 1 & 4 \\ 1 & 2 & 1 & 1 & 2 & 1 \\ 2 & 1 & 1 & 2 & 1 & 1 \\ 2 & 2 & 4 & 2 & 2 & 4 \end{pmatrix}$	$\mathbf{d}_0''$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 4 \\ 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 4 \end{pmatrix}$
$\mathbf{d}_1$	$\begin{pmatrix} 1 & 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 1 & 2 & 1 & 3 \\ 2 & 1 & 1 & 1 & 2 & 3 \\ 2 & 2 & 2 & 1 & 1 & 2 \end{pmatrix}$	$\mathbf{d}_1'$	$\begin{pmatrix} 1 & 1 & 3 & 2 & 2 & 2 \\ 1 & 2 & 2 & 2 & 1 & 3 \\ 2 & 1 & 2 & 1 & 2 & 3 \\ 2 & 2 & 3 & 1 & 1 & 2 \end{pmatrix}$	$\mathbf{d}_1''$	$\begin{pmatrix} 1 & 1 & 2 & 2 & 2 \\ 1 & 2 & 2 & 1 & 3 \\ 2 & 1 & 1 & 2 & 3 \\ 2 & 2 & 1 & 1 & 2 \end{pmatrix}$

The meaning of X_f is the same as that in Construction Method 1. The number of added runs in Construction Method 2 is also the smallest value to satisfy the requirements of U-type structure and cost saving. In the following, we give a lemma to calculate the sum of L_1-distance among all pairs of runs in the RFLAD D_2.

Lemma 2. *For any initial design $\mathbf{d}_0 \in \mathcal{U}(n; 2^m)$ and the additional design $\mathbf{d}_1 \in \mathcal{D}(n_1, 2^{m_1}4^{m_2})$, we have*

$$(1)\ d_{1,sum}(\mathbf{d}_0) = 3n^2m/2,\ \ (2)\ d_{1,sum}(\mathbf{d}_1) = n_1^2(6m_1 + 5m_2)/4,\ (3)\ d_{1,sum}(\mathbf{d}_0, \mathbf{d}_1) = 3nn_1m/2.$$

The proof of Lemma 2 is similar to Lemma 1 and is omitted. In addition, apart from the distance property, the column-orthogonality is another important issue. Based on Lemma 2, we obtain the condition for ensuring the maximinixity under the L_1-distance of D_2 and give some discussions about the column-orthogonality as follows.

Theorem 2. *Let the design $X \in \mathcal{U}(2^k; 2^{2^k-1})$ be a two-level saturated design. Denote $\begin{pmatrix} X^T & X^T \end{pmatrix}^T$ as D_{21} and $\begin{pmatrix} X_1^T & X_2^T \end{pmatrix}^T$ as D_{22}, where X_1 and X_2 are defined in Steps 2 and 3 in Construction Method 2. Suppose that the lth column of D_{22} is the level-augmented column.*

(1) The initial design $\mathbf{d}_0$ constructed by Construction Method 2 is a maximin design under L_1-distance and the mixed two- and four-level RFLAD $D_2 = (D_{21}\ D_{22}) \in \mathcal{L}_f(2^{k+1}; 2^{2^{k+1}-3}4^1)$ is also a maximin L_1-distance RFLAD with the L_1-distance being $3 \cdot 2^k - 4$. All the columns of D_2 are orthogonal to each other, except for the lth column of D_{21} and the lth column of D_{22}.

(2) If the lth column of D_{21} is also augmented, the resulting design $D_2' \in \mathcal{L}_f(2^{k+1}; 2^{2^{k+1}-4}4^2)$ is the mixed two- and four-level maximin L_1-distance RFLAD with two level-augmented factors and the L_1-distance of D_2' is $3 \cdot 2^k - 4$. All the columns of D_2' are orthogonal to each other, except for the lth column of D_{21} and the lth column of D_{22}.

(3) If the lth column of D_{21} is deleted, the resulting design D_2'' is a maximin L_1-distance RFLAD with the L_1-distance being $3 \cdot 2^k - 5$ and all the columns in $D_2'' \in \mathcal{L}_f(2^{k+1}; 2^{2^{k+1}-4}4^1)$ are orthogonal to each other.

Remark 2. Based on the settings of n and m in Theorem 2, we have $d_U(D_2) \le \lfloor d_{1,ave}(D_2) \rfloor$, $d_U(D_2') \le \lfloor d_{1,ave}(D_2') \rfloor$, and $d_U(D_2'') \le \lfloor d_{1,ave}(D_2'') \rfloor$.

The proof of Theorem 2 is given in the Appendix. The proof of Remark 2 is similar to Remark 1 and is omitted. Similarly, by Remark 2, it is also reasonable to use $d_U(D_2)$, $d_U(D_2')$ and $d_U(D_2'')$ to measure the space-filling property of D_2, D_2' and D_2'', respectively.

Next, we give an example to illustrate the usefulness of Construction Method 2.

Example 2. According to Step 1 of Construction Method 2, the initial design $\mathbf{d}_{00}$ is formed by design $X \in \mathcal{U}(4; 3)$, where the X is the same as used in Example 1. By Step 2, we change the initial design $\mathbf{d}_{00}$ to $\mathbf{d}_0$ whose first five columns are the same as $\mathbf{d}_{00}$, and the last column is $(4, 1, 1, 4)^T$, see Table 1. Then, we add an additional portion $\mathbf{d}_1 \in \mathcal{D}(4; 2^5 4^1)$ by Step 3, which is listed in Table 1. Then $D_2 = (\mathbf{d}_0^T\ \ \mathbf{d}_1^T)^T$ is the mixed two- and four-level RFLAD with one level-augmented factor. By Theorem 2(1), the third column and the sixth column of D_2 are not orthogonal. Then, we change the third column of $\mathbf{d}_0$ to $(4, 1, 1, 4)^T$ to construct $\mathbf{d}_0'$ and use an additional portion $\mathbf{d}_1' \in \mathcal{D}(4; 2^4 4^2)$. Then $D_2' = \begin{pmatrix} \mathbf{d}_0'^T & \mathbf{d}_1'^T \end{pmatrix}^T$ is the mixed two- and four-level RFLAD with two level-augmented factors. According to Theorem 2(3), we delete the third column of D_2 to obtain D_2'' and it is the mixed two- and four-level RFLAD in which all the columns are orthogonal to each other.

5. Discussion and conclusion

In this paper, we discuss the construction methods for the mixed two- and three-level and the mixed two- and four-level maximin L_1-distance RFLADs. For the latter one, from Theorem 2, it is found that only one pair of columns in the design is not orthogonal. If the experimenter pays more attention to the orthogonality of RFLADs, one column of such pair can be deleted and then all the columns in the resulting design are orthogonal to each other.

We only focus on RFLADs with minimum number of the added runs to ensure their U-type structure in this paper. We may increase the number of the added runs more flexibly for further investigation. Moreover, for constructing general

RFLADs, the numerical algorithms, such as the threshold accepting algorithm, can be considered. In addition, topics of high-level sequential designs and the multi-stage maximin L_1-distance sequential designs of computer experiment are also worth studying, which will be studied in the future.

Acknowledgments

The authors would like to thank the editor, the associate editor and the two referees for their valuable comments. This work was supported by National Natural Science Foundation of China (11871288 and 12131001), Natural Science Foundation of Tianjin, China (19JCZDJC31100) and KLMDASR, China. The authorship is listed in alphabetical order.

Appendix

Proof Lemma 1. From the structure of $\mathbf{d}_0$ and $\mathbf{d}_1$, we have,

(1) $d_{1,sum}(\mathbf{d}_0) = (n/2 \cdot n/2 \cdot |3-1| \cdot m_2) \times 2 + [n/2 \cdot n/2 \cdot (|2-1| \times 2) \cdot m_1] \times 2 = n^2 m$;

(2) $d_{1,sum}(\mathbf{d}_1) = 2 \cdot (n_1/2 \cdot n_1/2 \cdot |2-1| \times 2 \cdot m_1) = n_1^2 m_1$;

(3) $d_{1,sum}(\mathbf{d}_0, \mathbf{d}_1) = 2 \cdot (n/2 \cdot n_1/2 \cdot |2-1| \times 2) \cdot m_1 + 2 \cdot (n/2 \cdot n_1/2 \cdot |1-2|) \cdot m_2 = nn_1 m$. The proof is completed. □

Proof Theorem 1. Denote $\mathbf{d}_{01} = \begin{pmatrix} 1 & X & X \end{pmatrix} = (\mathbf{x}_1, \ldots, \mathbf{x}_{2^k})^T$, $\mathbf{d}_{02} = \begin{pmatrix} 2 & X & X_f \end{pmatrix} = (\mathbf{y}_1, \ldots, \mathbf{y}_{2^k})^T$, where $\mathbf{x}_i$ and $\mathbf{y}_i$ are the ith rows in $\mathbf{d}_{01}$ and $\mathbf{d}_{02}$, respectively. Since $X \in \mathcal{U}(2^k; 2^{2^k-1})$ is a two-level saturated design, both X and X_f are the Hamming-equidistant designs with the coincidence number λ_0, the number of places where two rows take the same value, of any two rows, being $2^{k-1}-1$. Denote $\lambda(\mathbf{x}, \mathbf{y})$ as the coincidence number for any two rows $\mathbf{x}, \mathbf{y}$ in design. Thus, $\lambda(\mathbf{x}_i, \mathbf{x}_j) = 2^k - 1$, $i \neq j = 1, \ldots, 2^k$, where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the ith row and the jth row in $\mathbf{d}_{01}$. Similarly, $\lambda(\mathbf{y}_i, \mathbf{y}_j) = 2^k - 1$, $i \neq j = 1, \ldots, 2^k$, where $\mathbf{y}_i$ and $\mathbf{y}_j$ are the ith row and the jth row in $\mathbf{d}_{02}$. Thus, both $\mathbf{d}_{01}$ and $\mathbf{d}_{02}$ are the Hamming-equidistant designs with the coincidence number of any two rows being $2^k - 1$. For the ith row in $\mathbf{d}_{01}$ and the jth row in $\mathbf{d}_{02}$, $\lambda(\mathbf{x}_i, \mathbf{y}_j) = 2^k - 1$. Hence, the initial design $\mathbf{d}_0$ is a Hamming-equidistant design with the coincidence number being $2^k - 1$, and $d_1(\mathbf{d}_0) = 2^{k+1}$. With similar argument, we can obtain that $\mathbf{d}_1$ is also a Hamming-equidistant design and its L_1-distance is 2^{k+1}.

In addition, $\mathbf{d}_1 = \begin{pmatrix} 2 & X_f & X \end{pmatrix} = (\mathbf{z}_1, \ldots, \mathbf{z}_{2^k})^T$, where $\mathbf{z}_l$ is the lth row in $\mathbf{d}_1$. We have $\lambda(\mathbf{x}_i, \mathbf{z}_l) = 2^k - 2$, $i = 1, \ldots, 2^k$, $l = 1, \ldots, 2^k$, where $\mathbf{x}_i$ and $\mathbf{z}_l$ are the ith row in $\mathbf{d}_{01}$ and the lth row in $\mathbf{d}_1$ respectively, and the L_1-distance of the $\mathbf{x}_i$ and $\mathbf{z}_l$ is $2^{k+1}+1$. Similarly, $\lambda(\mathbf{y}_j, \mathbf{z}_l) = 0$ or 2^k, $j = 1, \ldots, 2^k$, $l = 1, \ldots, 2^k$, where $\mathbf{y}_j$ and $\mathbf{z}_l$ are the jth row in $\mathbf{d}_{02}$ and the lth row in $\mathbf{d}_1$ respectively, then the corresponding L_1-distance of the $\mathbf{y}_j$ and $\mathbf{z}_l$ is $2^{k+2}-4$ or $2^{k+1}-3$.

According to Lemma 1, we have

$$\lfloor d_{1,ave}(\mathbf{d}_0) \rfloor = 2^{k+1},\ \lfloor d_{1,ave}(\mathbf{d}_1) \rfloor = 2^{k+1},\ \lfloor d_{1,ave}(\mathbf{d}_0, \mathbf{d}_1) \rfloor = 2^{k+1} - 1,$$

which implies that $d_U(D_1) = 2^{k+1} - 1$. Under the maximin L_1-distance criterion, the added portion $\mathbf{d}_1$ should be scattered uniformly in the whole experimental domain coupled with the initial points, so the requirement that the design $\mathbf{d}_1$ should be a Hamming-equidistant design, and the form of $\mathbf{d}_1$ constructed by Construction Method 1 meets this requirement. Under the form of $\mathbf{d}_1$ and conditions in Theorem 1, we have $d_1(D_1) = 2^{k+1} - 3$, which is not much different from the tight upper bound $d_U(D_1)$. However, the value of $2^{k+1} - 3$ is the largest value for $d_1(D_1)$ among all the designs in $\mathcal{L}_f(n+n_1; 2^{m_1}3^{m_2})$ given the initial design $\mathbf{d}_0$. Thus, the design D_1 constructed by the Construction Method 1 is the mixed two- and three-level maximin L_1-distance RFLAD. The proof is complete. □

Proof Remark 1. For a RFLAD $D_1 = (\mathbf{d}_0^T \;\; \mathbf{d}_1^T)^T \in \mathcal{L}_f(n+n_1; 2^{m_1}3^{m_2})$, according to Lemma 1 and Theorem 1, if $n = 2n_1$, $m = 2^{k+1} - 1$ and $m_1 = 2^{k+1} - 2$, $d_U(D_1) = \lfloor d_{1,ave}(\mathbf{d}_0, \mathbf{d}_1) \rfloor = \lfloor m \rfloor$. We can also have

$$\lfloor d_{1,ave}(D_1) \rfloor = \lfloor \frac{d_{1,sum}(\mathbf{d}_0) + d_{1,sum}(\mathbf{d}_1) + 2d_{1,sum}(\mathbf{d}_0, \mathbf{d}_1)}{(3n/2)(3n/2-1)} \rfloor = \lfloor \frac{8nm + nm_1}{3(3n-2)} \rfloor,$$

then $d_{1,ave}(D_1) - d_{1,ave}(\mathbf{d}_0, \mathbf{d}_1) > 0$. The proof is completed. □

Proof Theorem 2. (1) Denote $\mathbf{d}_0 = \begin{pmatrix} X & X_1 \end{pmatrix} = (\mathbf{x}_1, \ldots, \mathbf{x}_{2^k})^T$, $\mathbf{d}_1 = \begin{pmatrix} X & X_2 \end{pmatrix} = (\mathbf{y}_1, \ldots, \mathbf{y}_{2^k})^T$, where $\mathbf{x}_i$ and $\mathbf{y}_i$ are the ith rows in $\mathbf{d}_0$ and $\mathbf{d}_1$, respectively. According to Lemma 2, we can obtain that $d_U(D_2) = 3 \cdot 2^k - 3$. Since $X \in \mathcal{U}(2^k; 2^{2^k-1})$ is a two-level saturated design, it is also a Hamming-equidistant design with the coincidence number of any two rows being $2^{k-1}-1$, denoted by λ_0, and the design $\mathbf{d}_{00}$ in Construction Method 2 is a supersaturated design constructed by Lin (1993). Following (Fang et al., 2003), $\mathbf{d}_{00}$ is a two-level Hamming-equidistant design with the coincidence number of any two rows being $2\lambda_0$, so it is a maximin L_1-distance design with the L_1-distance being 2^k. According to Step 2, $\mathbf{d}_0$ is also a Hamming-equidistant design, in which the coincidence number of any two rows is the same as that of $\mathbf{d}_{00}$, therefore, $d_1(\mathbf{d}_0) = 3 \cdot 2^k$. By calculation, the coincidence number of any two different rows in $\mathbf{d}_1$ is $2^k - 2$, then the L_1-distance of the ith row and jth row in $\mathbf{d}_1$ is $3 \cdot 2^k$ or $3 \cdot 2^k - 2$, so $d_1(\mathbf{d}_1) = 3 \cdot 2^k - 2$. And if $\mathbf{x}_i \in \mathbf{d}_0$ and $\mathbf{y}_j \in \mathbf{d}_1$, $i = 1, 2, \ldots, 2^k$, $j = 1, 2, \ldots, 2^k$, we have $d_1(\mathbf{d}_0, \mathbf{d}_1) = 3 \cdot 2^k - 4$. In summary, the L_1-distance of any two different rows in

D_2 is $3\cdot2^k-4$, which is not much different from the tight upper bound $d_U(D_2)$. The form of the added portion that listed in Step 3 of the Construction Method 2 can guarantee the space-filling property of the added portion $\mathbf{d}_1$ and couple with the design D_2, so we get that $3\cdot2^k-4$ is the largest value for the L_1-distance among all the designs in $\mathcal{L}_f(2^{k+1}; 2^{2^{k+1}-3}4^1)$. Therefore, under the conditions in Theorem 2(1), the design D_2 constructed by the Construction Method 2 is the mixed two- and four-level maximin L_1-distance RFLAD.

We normalize $D_2 = (x_{ij})_{1\le i\le\cdots\le2^{k+1}, 1\le j\le\cdots\le2^{k+1}-2}$ by $x_{ij}-\bar{x}_j$, where $\bar{x}_j$ is the mean of the jth column in D_2. Denote $\mathbf{x}^r$ and $\mathbf{x}^k$ as the rth column and kth column of the design. For $\mathbf{x}^r, \mathbf{x}^k \in X$, since X is a two-level saturated design, so we have $\langle\mathbf{x}^r, \mathbf{x}^k\rangle = 0, 1 \le r \ne k \le 2^k-1$, where $\langle\cdot,\cdot\rangle$ represents the internal product of the two vectors. So, for $\mathbf{x}^{r_{21}}, \mathbf{x}^{k_{21}} \in D_{21}$, we have $\langle\mathbf{x}^{r_{21}}, \mathbf{x}^{k_{21}}\rangle = 0, 1 \le r_{21} \ne k_{21} \le 2^k-1$. Similarly, for $\mathbf{x}^{r_1}, \mathbf{x}^{k_1} \in X_1$, we have $\langle\mathbf{x}^{r_1}, \mathbf{x}^{k_1}\rangle = 0, 1 \le r_1 \ne k_1 \le 2^k-1$; for $\mathbf{x}^{r_2}, \mathbf{x}^{k_2} \in X_2$, we have $\langle\mathbf{x}^{r_2}, \mathbf{x}^{k_2}\rangle = 0, 1 \le r_2 \ne k_2 \le 2^k-1$. So, for $\mathbf{x}^{r_{22}}, \mathbf{x}^{k_{22}} \in D_{22}$, we have $\langle\mathbf{x}^{r_{22}}, \mathbf{x}^{k_{22}}\rangle = 0, 1 \le r_{22} \ne k_{22} \le 2^k-1$. If $\mathbf{x}^{l_{21}} \in D_{21}$ is the lth column of level augmentation, then for $\mathbf{x}^{l_{22}} \in D_{22}$, we have $\langle\mathbf{x}^{l_{21}}, \mathbf{x}^{l_{22}}\rangle \ne 0$, where $1 \le l_{21} = l_{22} \le 2^k-1$; if $\mathbf{x}^{r_{21}} \in D_{21}$ and $\mathbf{x}^{k_{22}} \in D_{22}$, we have $\langle\mathbf{x}^{r_{21}}, \mathbf{x}^{k_{22}}\rangle = 0$ when $1 \le r_{21} \ne k_{22} \le 2^k-1$. Therefore, all the columns of D_2 are orthogonal to each other, except for the lth column of D_{21} and the lth column of D_{22}.

(2) If the levels of the lth column of D_{21} is also augmented, then the design D_2' is the RFLAD with two columns augmenting levels. As similar as the proof in Theorem 2(1), the $d_1(\mathbf{d}_0)$, $d_1(\mathbf{d}_1)$ and $d_1(\mathbf{d}_0, \mathbf{d}_1)$ of D_2' need to be calculated separately. The calculation process is similar to Theorem 2(1) and is omitted here. By calculation, it can be obtained that $d_1(D_2') = 3\cdot2^k-4$, and we can also obtain $d_U(D_2') = 3\cdot2^k-3$. The L_1-distance of any two different rows in D_2' is not much different from the upper bound $d_U(D_2')$. Moreover, we also obtain that $3\cdot2^k-4$ is the largest value for the L_1-distance among all the designs in $\mathcal{L}_f(n+n_1; 2^{m_1}4^{m_2})$ because of the form of the added portion that described in Theorem 2(2) can guarantee the space-filling property of the added portion $\mathbf{d}_1'$ and couple with the design $\mathbf{d}_0'$, therefore D_2' is a maximin L_1-distance RFLAD with two factors of level augmentation. Similar to the proof in (1), the lth column of D_{21} is not orthogonal to the lth column of D_{22}. But except for these two columns, other columns in D_2' are column-orthogonal to each other.

(3) By calculation, we can obtain that both $d_U(D_2'')$ and $d_1(D_2'')$ equal to $3\cdot2^k-5$. Thus, D_2'' is the mixed two- and four-level maximin L_1-distance RFLAD. From the method to get D_2'', it is natural to know that all the columns in D_2'' are orthogonal to each other. The proof is completed. □

References

Fang, K.T., Lin, D.K.J., Liu, M.Q., 2003. Optimal mixed-level supersaturated design. Metrika 58, 279–291.
Fang, K.T., Tang, Y., Yin, J., 2005. Lower bounds for wrap-around L_2-discrepancy and constructions of symmetrical uniform designs. J. Complexity 21, 757–771.
Gao, Y.P., Yi, S.Y., Zhou, Y.D., 2021. Level-augmented uniform designs. Stat. Pap. Online http://dx.doi.org/10.1007/s00362-021-01247-y.
Gou, T.X., Qin, H., Kashinath, C., 2018. Efficient asymmetrical extended designs under wrap-around L_2-discrepancy. J. Syst. Sci. Complex 31, 1391–1404.
Grosso, A., Jamali, A.R.M.J.U., Locatelli, M., 2009. Finding maximin latin hypercube designs by iterated local search heuristics. Eur. J. Oper. Res. 197, 541–547.
Hickernell, F.J., 1998. Lattice rules: How well do they measure up? In: Hellekalek, P., Larcher, G. (Eds.), Random and Quasi-Random Point Sets. Springer-Verlag, pp. 106–166.
Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin distance designs. J. Statist. Plann. Inference 26, 131–148.
Li, X., Chen, F., Li, S., Jia, J., Gu, H., Yang, L., 2016. An efficient homogenate-microwave-assisted extraction of flavonols and anthocyanins from blackcurrant marc: optimization using combination of plackett-burman design and box-behnken design. Ind. Crops. Prod. 94, 834–847.
Li, W.L., Liu, M.Q., Tang, B., 2020. A method of constructing maximin distance designs. Biometrika 108 (4), 845–855.
Lin, D.K.J., 1993. A new class of supersaturated designs. Technometrics 35, 28–31.
Moon, H., Dean, A., Santner, T., 2011. Algorithms for generating maximin latin hypercube and orthogonal designs. J. Stat. Theory. Pract. 5, 81–98.
Qin, H., Gou, T.X., Chatterjee, K., 2016. A new class of two-level optimal extended designs. J. Korean Stat. Soc. 45, 168–175.
Wang, L., Xiao, Q., Xu, H., 2018. Optimal maximin L_1-distance latin hypercube designs based on good lattice point designs. Ann. Stat. 46, 3741–3766.
Yang, L.Q., Zhou, Y.D., Liu, M.Q., 2021. Maximin distance designs based on densest packings. Metrika 84, 615–634.
Yang, F., Zhou, Y.D., Zhang, X.R., 2017. Augmented uniform designs. J. Statist. Plann. Inference 182, 61–73.
Yang, F., Zhou, Y.D., Zhang, A.J., 2019. Mixed-level column augmented uniform designs. J. Complexity 53, 23–39.
Zhu, B., Liu, J., Gao, W., 2018. Optimization of operational parameters of foam sizing process for cotton yarns based on plackett-burman experiment design. AUTEX. Res. J. 18, 61–66.