
Journal of the American Statistical Association

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

# On Design Orthogonality, Maximin Distance, and Projection Uniformity for Computer Experiments

Yaping Wang, Fasheng Sun & Hongquan Xu

**To cite this article:** Yaping Wang, Fasheng Sun & Hongquan Xu (2022) On Design Orthogonality, Maximin Distance, and Projection Uniformity for Computer Experiments, Journal of the American Statistical Association, 117:537, 375-385, DOI: 10.1080/01621459.2020.1782221

**To link to this article:** https://doi.org/10.1080/01621459.2020.1782221

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2022, VOL. 117, NO. 537, 375–385: Theory and Methods
https://doi.org/10.1080/01621459.2020.1782221



# On Design Orthogonality, Maximin Distance, and Projection Uniformity for Computer Experiments

Yaping Wang*[a], Fasheng Sun*[b], and Hongquan Xu[c]

[a]KLATASDS-MOE, School of Statistics, East China Normal University, Shanghai, China; [b]KLAS and School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, China; [c]Department of Statistics, University of California, Los Angeles, CA

**ABSTRACT**
Space-filling designs are widely used in both computer and physical experiments. Column-orthogonality, maximin distance, and projection uniformity are three basic and popular space-filling criteria proposed from different perspectives, but their relationships have been rarely investigated. We show that the average squared correlation metric is a function of the pairwise $L_2$-distances between the rows only. We further explore the connection between uniform projection designs and maximin $L_1$-distance designs. Based on these connections, we develop new lower and upper bounds for column-orthogonality and projection uniformity from the perspective of distance between design points. These results not only provide new theoretical justifications for each criterion but also help in finding better space-filling designs under multiple criteria. Supplementary materials for this article are available online.

**ARTICLE HISTORY**
Received May 2019
Accepted June 2020

**KEYWORDS**
Correlation; Latin hypercube design; $L_p$-distance; Space-filling property; Uniform projection designs

## 1. Introduction

Computer experiments offer significant economic benefits for investigating complex physical systems (Santner, Williams, and Notz 2003; Fang, Li, and Sudjianto 2006). Constructing efficient designs with good space-filling properties has much importance in both computer and physical experiments; see, for example, Pronzato and Müller (2012) and Joseph (2016) for excellent surveys on space-filling designs. Latin hypercube designs (LHDs) are arguably the most widely used for computer experiments because of their maximum one-dimensional stratification properties. Bingham, Sitter, and Tang (2009) and Sun and Tang (2017a) also proposed balanced designs with flexible number of levels for many computer experiments, which include LHDs as special cases. An ordinary LHD or balanced design has uniform one-dimensional projections but still may not be space-filling because of possibly poor two and higher dimensional performance. Therefore, various optimality criteria have been proposed for design optimization and construction.

Column-orthogonality (Owen 1994; Tang 1998; Ye 1998) and maximin distance criterion (Johnson, Moore, and Ylvisaker 1990; Zhou and Xu 2014; Ba, Myers, and Brenneman 2015) are two most commonly used design criteria which have drawn much attention. Exact or near column-orthogonality can be viewed as a useful stepping stone to space-filling designs (Bingham, Sitter, and Tang 2009), but their theoretical connections with space-filling designs have not been completely revealed yet. It is one of the major objectives of this article to explain these connections. Maximin distance designs are asymptotically $D$-optimal under the ordinary kriging model as the correlations between points become weak (Johnson, Moore, and Ylvisaker 1990). A number of authors have constructed orthogonal and maximin LHDs; refer to Lin and Tang (2015) for a comprehensive review and Sun and Tang (2017a, 2017b), Xiao and Xu (2017, 2018), and Wang, Xiao, and Xu (2018) for some recent advances. On the other hand, uniform designs aim to spread the design points uniformly over the whole design space (Fang et al. 2000; Fang, Li, and Sudjianto 2006). To improve the low-dimensional projections, Sun, Wang, and Xu (2019) recently proposed and studied uniform projection designs.

It is natural to ask whether there exist some connections or equivalences among different space-filling criteria. Wang, Yang, and Xu (2018) proved that within the class of mirror-symmetric designs with $2m$ runs and $m$ factors, column-orthogonal designs are maximin $L_2$-distance designs. Sun, Wang, and Xu (2019) established a connection between the uniform projection criterion and maximin $L_1$-distance. They showed that maximin $L_1$-equidistant designs are uniform projection designs. This article aims to explore some new connections among the criteria—column-orthogonality, projection uniformity and maximin $L_1$- and $L_2$-distance. We establish some optimality results and show that these different criteria are closely related and consistent in more general cases, mainly when the designs have relatively high factor-to-run ratios. Our results provide not only new theoretical justifications for each criterion from other viewpoints, but also some insights for finding or constructing better space-filling designs with economic run size. Such designs can be used for factor screening in physical and computer experiments with a large number of factors (Butler 2005; Moon, Dean, and Santner 2012; Woods and Lewis 2016; Kleijnen 2017). For example, Moon, Dean, and Santner (2012) proposed a two-stage screening procedure for computer experiments, in which an

**CONTACT** Hongquan Xu hqxu@stat.ucla.edu Department of Statistics, University of California, Los Angeles, CA 90095.
*These authors contributed equally to this work.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JASA.

$n \times (n-1)$ preliminary design matrix $X^*$ satisfying the following three requirements is in demand for the first stage:

(P.1) The columns of $X^*$ must be uncorrelated to allow independent assessment of the effects of the different inputs; (P.2) the minimum and maximum values in each column must be 0 and 1, respectively, to prevent input values with larger ranges from having larger impacts on the response, artificially induced by the design; and (P.3) the design defined by $X^*$ should be "space-filling" to insure that all regions of the input space are explored.

The theories and designs discussed in this article agree with (P.1)–(P.3). Although the main focus of this article is space-filling designs, it is worth pointing out that our results are applicable for not only LHDs, but also general balanced designs including $s$-level fractional factorials.

The remainder of this article is unfolded as follows. Section 2 gives a brief background and presents several novel results linking column-orthogonal designs with maximin $L_2$-distance designs. Section 3 provides lower bounds of the uniform projection criterion, some connections between projection uniformity and maximin $L_1$-distance, and some general asymptotical optimality results. Section 4 studies the connection between column-orthogonal designs and uniform projection designs. Section 5 concludes this article with a general discussion. The supplementary materials contain all the proofs and some additional tables and figures.

## 2. Connection Between Column-Orthogonality and $L_2$-Distance

### 2.1. Notation, Definitions, and Preliminary Results

We use $(n, s^m)$ to denote a design of $n$ runs and $m$ factors, with each factor taking levels from the set $\mathcal{Z}_s = \{0, 1, \ldots, s-1\}$. We use an $n \times m$ matrix $D = (x_{ik})_{n\times m}$ to represent an $(n, s^m)$ design. A design $D$ is called an orthogonal array (OA) of strength $t$, denoted by $\mathrm{OA}(n, m, s, t)$, if each level combination appears equally often in any distinct $t$ columns of $D$. We call an $(n, s^m)$ design balanced if it is an OA of strength one, that is, each level appears $n/s$ times in each column. In particular, an LHD is a balanced $(n, n^m)$ design and is denoted by $\mathrm{LHD}(n, m)$. Throughout the article, we consider balanced designs only. We call a design combinatorially orthogonal if it is an OA of strength two.

Researchers have proposed various space-filling criteria and studied the corresponding optimal space-filling designs for computer experiments. Column-orthogonality (or low correlation) is a popular criterion which optimizes designs by minimizing correlations among factors (Owen 1994; Tang 1998; Ye 1998). We also call a column-orthogonal design simply as an orthogonal design. Owen (1994) proposed the mean squared correlation metric to measure the orthogonality of a design $D$, which is defined as $\rho^2(D) = 2\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}\rho_{jk}^2/[m(m-1)]$, where $\rho_{jk}$ is the sample correlation between the $j$th and $k$th columns of $D$. It is clear that $0 \le \rho^2(D) \le 1$. A design $D$ is column-orthogonal if and only if $\rho^2(D) = 0$. Obviously, a combinatorially orthogonal design must be column-orthogonal, but the converse is not necessarily true.

The maximin distance criterion is to maximize the minimum inter-site distance of a design (Johnson, Moore, and Ylvisaker 1990). For the $i$th row $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{im})$ and the $j$th row $\boldsymbol{x}_j = (x_{j1}, \ldots, x_{jm})$ of a design $D$, their $L_p$-distance is defined as $d_p(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{k=1}^{m}|x_{ik} - x_{jk}|^p$, where $p \ge 1$ is an integer. This definition is the $p$th power of the traditional $L_p$-norm, which ensures that the $L_p$-distance is additive, that is, $d_p(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{m} d_p(x_i, y_i)$, and is always an integer. The $L_p$-distance of $D$ is $d_p(D) = \min\{d_p(\boldsymbol{x}_i, \boldsymbol{x}_j), 1 \le i < j \le n\}$. A design $D$ is called a maximin $L_p$-distance design if it has the maximum $d_p(D)$ value among all competing designs. In the literature $p = 1$ and $p = 2$ are the most widely used, which correspond to the rectangular (Manhattan) distance and the Euclidean distance, respectively.

Let $\bar{d}_p = 2\sum_{1\le i<j\le n} d_p(\boldsymbol{x}_i, \boldsymbol{x}_j)/[n(n-1)]$ be the average pairwise $L_p$-distance of a balanced $(n, s^m)$ design $D$. Using the balance property, it is easy to obtain that $\bar{d}_p$ is a constant determined by $n$, $m$, and $s$. Specifically, we have

$$\bar{d}_1 = \frac{nm(s^2-1)}{3(n-1)s} \tag{1}$$

and

$$\bar{d}_2 = \frac{nm(s^2-1)}{6(n-1)}. \tag{2}$$

Based on the fact that the minimum inter-site distance cannot exceed the integer part of the average distance, Zhou and Xu (2015) derived the following upper bounds.

*Lemma 1.* For a balanced $(n, s^m)$ design $D = (x_{ik})$, $d_1(D) \le \lfloor \bar{d}_1 \rfloor$ and $d_2(D) \le \lfloor \bar{d}_2 \rfloor$, where $\lfloor x \rfloor$ is the largest integer not exceeding $x$.

In particular, for LHDs the upper bounds become $d_1(D) \le \lfloor (n+1)m/3 \rfloor$ and $d_2(D) \le \lfloor n(n+1)m/6 \rfloor$. We first explore the connection between column-orthogonality and maximin distance. Here is an example.

*Example 1.* Let $n = 16, 20, 25$, and 30 and $m = 2, \ldots, n$. For each $(n, m)$ combination, we generate 100 maximin $L_2$-distance $\mathrm{LHD}(n, m)$'s by the R package SLHD (Ba, Myers, and Brenneman 2015) with default settings. Then for the 100 $\mathrm{LHD}(n, m)$'s, we calculate their $\rho^2$, $d_1$, and $d_2$ values. Table 1 summarizes the average $\rho^2$, $d_1$, and $d_2$ values and the corresponding standard errors for $n = 16$, and tables for $n = 20, 25, 30$ can be found in the supplementary materials. We see that the variability of the 100 maximin LHDs generated by the R package SLHD is relatively small in most cases. The plot of average $\rho^2$ values to the number of factors $m$ is given as the left panel in Figure 1. For comparison, we also generate 100 random $\mathrm{LHD}(n, m)$'s for each $(n, m)$ combination. The corresponding plot is given as the right panel in Figure 1.

From Figure 1 and related numerical results, we have the following observations.

(i) For random LHDs of $n$ runs, the average $\rho^2$ is about $O(1/n)$, and the curve becomes flat as $m$ grows. Actually, by Owen (1994), the expectation of $\rho^2$ for random LHDs of $n$ runs is $1/(n-1)$, and is independent of the number of columns.

**Table 1.** The average (ave) $d_1$, $d_2$, and $\rho^2$ values and the corresponding standard errors (SEs) for 100 maximin distance LHD$(n, m)$'s for $n = 16$ and $m = 2, \ldots, n$.

| | $d_1$ | | $d_2$ | | $\rho^2$ | |
|---|---|---|---|---|---|---|
| $m$ | Ave | SE | Ave | SE | Ave | SE |
| 2 | 4.96 | 0.20 | 13.00 | 0.92 | 0.0146 | 0.0202 |
| 3 | 8.80 | 0.59 | 38.65 | 2.60 | 0.0139 | 0.0112 |
| 4 | 13.30 | 0.83 | 75.08 | 4.08 | 0.0117 | 0.0063 |
| 5 | 17.97 | 0.96 | 119.19 | 5.99 | 0.0088 | 0.0037 |
| 6 | 22.84 | 1.28 | 165.43 | 6.78 | 0.0081 | 0.0029 |
| 7 | 28.36 | 1.47 | 212.84 | 8.82 | 0.0074 | 0.0026 |
| 8 | 33.80 | 1.68 | 260.75 | 10.22 | 0.0071 | 0.0026 |
| 9 | 39.30 | 1.53 | 309.26 | 10.74 | 0.0084 | 0.0023 |
| 10 | 44.67 | 1.48 | 353.92 | 10.57 | 0.0099 | 0.0021 |
| 11 | 50.05 | 1.65 | 399.45 | 11.55 | 0.0103 | 0.0021 |
| 12 | 55.19 | 1.92 | 443.20 | 11.91 | 0.0099 | 0.0021 |
| 13 | 61.05 | 2.06 | 493.29 | 10.21 | 0.0080 | 0.0020 |
| 14 | 67.08 | 1.72 | 543.84 | 12.58 | 0.0064 | 0.0024 |
| 15 | 73.22 | 1.88 | 603.19 | 15.41 | 0.0047 | 0.0025 |
| 16 | 78.18 | 1.98 | 645.23 | 16.36 | 0.0078 | 0.0020 |

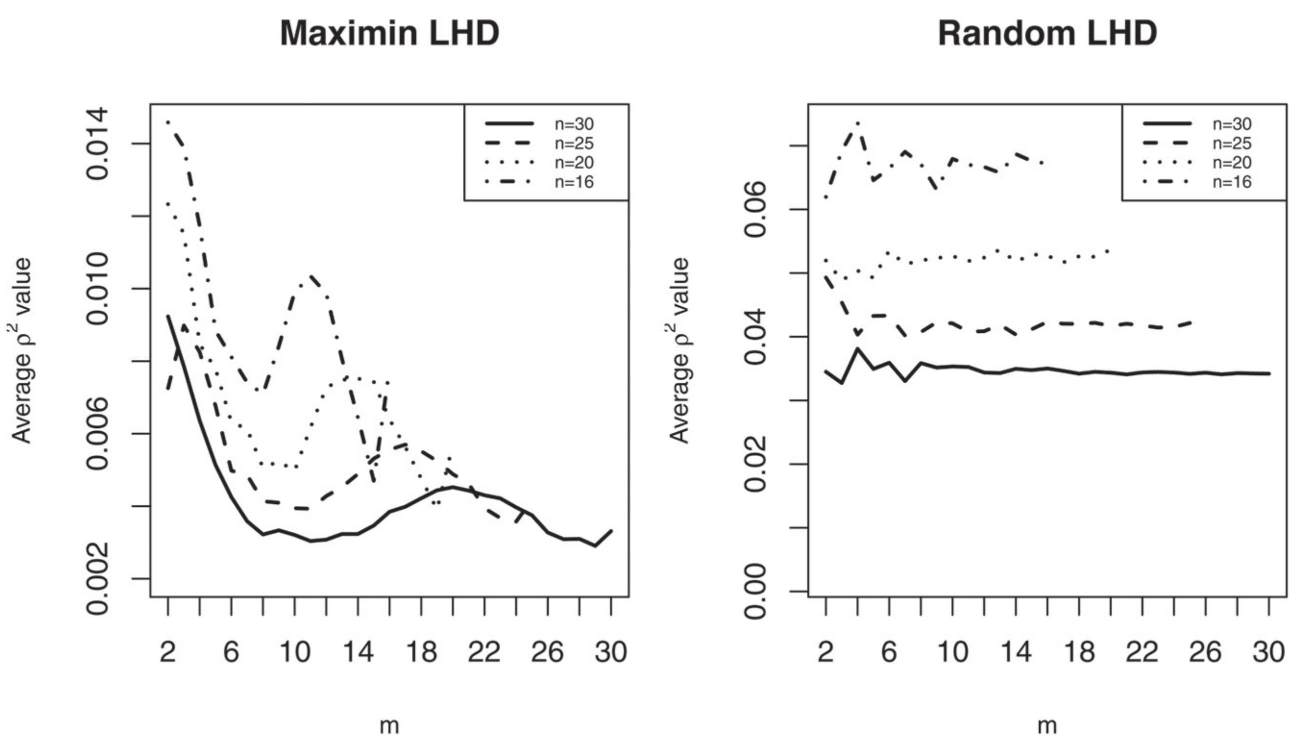


**Figure 1.** Plot of the average $\rho^2$ values of 100 maximin LHD$(n, m)$'s (left) and random LHD$(n, m)$'s (right) for $n = 16, 20, 25, 30$ and $m = 2, \ldots, n$.

(ii) In general, all the maximin distance LHDs have much smaller average squared correlations than the random LHDs. Some explanations of this phenomenon can be found in Theorems 1 and 3.

(iii) For each $n$, the average $\rho^2$ value of maximin $L_2$-distance LHDs decreases when $m$ is small (less than $n/2$). Especially, when $m$ is around $n/2$, the average $\rho^2$ value achieves a local minimum for $n = 16$ and 20. This phenomenon also exists for other $n$'s in our simulations not reported here. Example 4 and the paragraph afterward provide explanations why $m = n/2$ tends to be a local minimum point.

(iv) For each $n$, the average $\rho^2$ value of maximin $L_2$-distance LHDs decreases when $m$ is from about $2n/3$ to $n - 1$. The smallest average $\rho^2$ values are all attained when $m = n-1$. See Theorem 2, Proposition 2, and related comments and examples for some explanations.

There are several other criteria considering low-dimensional projection properties. The minimum average reciprocal distance (minARD) criterion, proposed by Draguljić, Santner, and Dean (2012), aims to minimize

$$\mathrm{ARD}(D) = \left\{ \frac{1}{\binom{n}{2} \sum_{q \in J} \binom{m}{q}} \sum_{q \in J} \sum_{|\mathcal{U}|=q} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left( \frac{q^{1/2}}{d_2^{1/2}(\boldsymbol{x}_{i,\mathcal{U}}, \boldsymbol{x}_{j,\mathcal{U}})} \right)^{\lambda} \right\}^{1/\lambda}$$

for an $n \times m$ $D = (x_{ik})$, where $\lambda \geq 1$ is a prespecified real number, $J$ and $\mathcal{U}$ are subsets of $\{1, 2, \ldots, m\}$, $|\mathcal{U}|$ stands for the cardinality of $\mathcal{U}$, and $\boldsymbol{x}_{i,\mathcal{U}}$ and $\boldsymbol{x}_{j,\mathcal{U}}$ are, respectively, the $i$th and $j$th runs of $D$ projected onto dimensions indexed by the elements of $\mathcal{U}$. The maximum projection (MaxPro) criterion, proposed

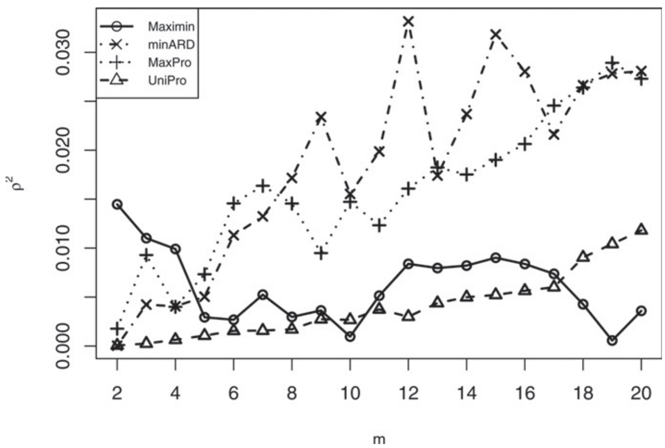


**Figure 2.** Plot of the $\rho^2$ values for maximin, minARD, MaxPro, and uniform projection (UniPro) LHD$(n,m)$'s for $n=20$ and $m=2,\ldots,n$.

by Joseph, Gul, and Ba (2015), aims to minimize

$$\psi(D)=\left\{\frac{1}{\binom{n}{2}}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{1}{\prod_{k=1}^{m}(x_{ik}-x_{jk})^2}\right\}^{1/m}.$$

Both ARD and $\psi$ consider the distance between points in low-dimensional projections. The uniform projection criterion proposed by Sun, Wang, and Xu (2019) aims to minimize the two-dimensional discrepancy $\phi(D)$ defined in (8); see Section 3.

*Example 2.* We compare four types of space-filling designs. For $n=20$ and each $m=2,\ldots,n$, we construct a maximin LHD, a minARD LHD, a MaxPro LHD, and a uniform projection LHD. The maximin LHD$(n,m)$ is chosen as the best design from the 100 LHD$(n,m)$'s generated in Example 1. The MaxPro LHD$(n,m)$ is chosen as the best design by running MaxProLHD function in the R package MaxPro (Joseph, Gul, and Ba 2015) for 100 times with default settings. The minARD LHDs (using $\lambda=1$ and $J=\{1,2\}$) and uniform projection LHDs are similarly obtained via a threshold accepting algorithm (see Fang et al. 2000). Figure 2 shows the plot of $\rho^2$ values of these LHDs against $m$. Compared with random LHDs in Figure 1, all four types of LHDs have much smaller $\rho^2$ values. As $m$ increases, maximin distance and uniform projection LHDs outperform minARD and MaxPro LHDs. It is worth noting that maximin distance LHDs have small correlations without explicit consideration of low-dimensional projection properties. However, maximin-distance LHDs do not perform as well as the other three types of LHDs under the projection distance and projection uniformity criteria; see the supplementary materials.

Joseph and Hung (2008) pointed out that orthogonal LHDs may not be space-filling in terms of distance and vice versa. This is true when the number of columns is small compared to the number of runs. For a fixed run size, from Example 1 we see that maximin $L_2$-distance LHDs tend to become "more" orthogonal as the factor-to-run ratio approaches to one. Interestingly, the average $\rho^2$ values are rather small when $m$ is around $n/2$ and $n-1$. Wang, Yang, and Xu (2018) provided an explanation for the special case $n=2m$ by showing that there exists a strong relationship between column-orthogonality and maximin $L_2$-distance for mirror-symmetric $(n,s^m)$ designs with $n=2m$. We will explore the connection between the two criteria in a general situation.

### 2.2. Some Theoretical Results

The following lemma shows that the average $L_1$- and $L_2$-distances between a fixed row and all other rows in a design is determined by the $L_2$-distance between the fixed row and the center point. This lemma is very useful in proving the upcoming theorems.

*Lemma 2.* Let $D=(x_{ik})_{n\times m}$ be a balanced $(n,s^m)$ design. For any $i=1,\ldots,n$,

$$\begin{aligned}\sum_{j=1}^{n}d_1(\boldsymbol{x}_i,\boldsymbol{x}_j)&=\tfrac{nm(s^2-1)}{4s}+\tfrac{n}{s}d_2(\boldsymbol{x}_i,s_0),\\ \sum_{j=1}^{n}d_2(\boldsymbol{x}_i,\boldsymbol{x}_j)&=\tfrac{nm(s^2-1)}{12}+nd_2(\boldsymbol{x}_i,s_0),\end{aligned}\tag{3}$$

where $x_i$ is the $i$th row of $D$, $s_0=(s-1)/2$, and $d_2(\boldsymbol{x}_i,s_0)=\sum_{k=1}^{m}(x_{ik}-s_0)^2$.

Now we give an analytical expression of the average squared correlation $\rho^2(D)$ in terms of the pairwise $L_2$-distances between design points of $D$.

*Theorem 1.* For a balanced $(n,s^m)$ design $D=(x_{ik})$, we have

$$\rho^2(D)=\frac{h(D)}{n^2m(m-1)(s^2-1)^2/36}+1,\tag{4}$$

where

$$h(D)=\sum_{i=1}^{n}\sum_{j=1}^{n}d_2^2(\boldsymbol{x}_i,\boldsymbol{x}_j)-\frac{2}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}d_2(\boldsymbol{x}_i,\boldsymbol{x}_j)\right)^2.$$

Theorem 1 establishes a novel link between the relationship of rows and columns of the design. This link provides new insights on the design orthogonality from the distance perspective. Based on this, we obtain the following lower bound for $\rho^2(D)$.

*Theorem 2.* For a balanced $(n,s^m)$ design $D$,

$$\begin{aligned}\rho^2(D)\geq\rho^2_{\mathrm{LB}}&=\max\left\{\frac{m+1-n}{(n-1)(m-1)},0\right\}\\ &=\begin{cases}0 & \text{if } m\leq n-1,\\ \frac{m+1-n}{(n-1)(m-1)} & \text{if } m>n-1.\end{cases}\end{aligned}$$

Furthermore, when $m\geq n-1$, the lower bound is achieved if and only if $D$ is an equidistant design under the $L_2$-distance.

There are a series of similar results in the literature of supersaturated designs when $m\geq n-1$. For a balanced $(n,2^m)$ design $D$ with two levels taken as $\{-1,1\}$, the $E(s^2)$ criterion is defined to be $2\sum_{j=1}^{m-1}\sum_{k=j+1}^{m}s_{jk}^2/[m(m-1)]$, where $s_{jk}$ is the inner product of the $j$th and $k$th columns of $D$ (Booth and Cox 1962). Obviously, for a two-level design $D$, $E(s^2)=n^2\rho^2(D)$. Nguyen (1996) and Tang and Wu (1997) independently showed $E(s^2)\geq n^2(m+1-n)/[(n-1)(m-1)]$. Cheng (1997) gave an alternative proof by considering the inner products of the rows instead of the columns. Butler (2005) generalized the $E(s^2)$ criterion to supersaturated LHDs and proved the same lower bound. Theorem 2 extends these results and incorporates two-level designs and LHDs as special cases. As pointed out by a referee, Theorem 2 can be proved by using the same techniques in Tang and Wu (1997), Nguyen (1996), Cheng (1997), and Butler (2005). Our proof is based on Theorem 1 and provides a new space-filling perspective for general $s$-level designs. Furthermore, Theorem 2 reveals that, when $m \geq n-1$, $\rho^2(D) = (m+1-n)/[(n-1)(m-1)]$ is equivalent to the geometrical property that all of the design points in $D$ have the same $L_2$-distance from each other.

It is well known that there do not exist column-orthogonal designs when $m \geq n$, which can also be seen from Theorem 2 since $\rho^2(D) \geq (m+1-n)/[(n-1)(m-1)] > 0$. On the other hand, from the proof of Theorem 2, if an $L_2$-equidistant balanced $(n, s^m)$ design $D$ exists, $\rho^2(D)$ must equal $(m+1-n)/[(n-1)(m-1)]$, which is nonnegative only if $m \geq n-1$. Therefore, we have the following proposition.

*Proposition 1.* There do not exist $L_2$-equidistant balanced $(n, s^m)$ designs when $m < n-1$.

For saturated $(n, s^m)$ designs in the sense of $m = n-1$, there is an exact equivalence between column-orthogonality and maximin $L_2$-distance.

*Proposition 2.* Let $D$ be a balanced $(n, s^{n-1})$ design. Then $D$ is column-orthogonal if and only if it is a maximin equidistant design under the $L_2$-distance.

*Example 3.* Let $s = 2$, $m = n-1$, and $D$ be any $(n, s^m)$ design. Then from Proposition 2, $D$ is column-orthogonal if and only if it is an equidistant design under the $L_2$-distance. Since for $s = 2$, any $L_p$-distance is equivalent to the Hamming distance, that is, the number of positions where two rows differ, $D$ must be an equidistant design under the Hamming distance. Such a design is well known as a saturated two-level OA$(n, n-1, 2, 2)$.

A refinement of Theorem 2 for LHDs or balanced designs under some parameters can be obtained by considering the following results. Lin et al. (2010) showed that an orthogonal LHD$(n, m)$ with $n \geq 4$ exists if and only if $n \neq 4k+2$ for any integer $k \geq 0$. Karunanayaka and Tang (2018) further showed that an orthogonal balanced $(n, s^m)$ design does not exist if $n/s$ is odd and $s = 4k+2$ for some integer $k \geq 0$. Actually, for an LHD$(n, m)$ with $n = 4k+2 > 0$, Lin (2008) obtained in her PhD thesis that $\rho^2(D) \geq 36/[n^2(n^2-1)^2]$. Using the same technique, it is not difficult to show that for a balanced $(n, s^m)$ design $D$, if $s = 4k+2$ for some integer $k \geq 0$ and $n/s$ is odd, then $\rho^2(D) \geq 36/[n^2(s^2-1)^2]$ and, with the combination of Theorem 2,

$$\rho^2(D) \geq \max\left\{\frac{m+1-n}{(n-1)(m-1)}, \frac{36}{n^2(s^2-1)^2}\right\}.$$

Equidistant designs with $d_2(D) = \bar{d}_2$ are maximin distance designs and optimal under the $\rho^2(D)$ criterion by Theorem 2. However, in many cases equidistant designs do not exist and the lower bound in Theorem 2 is not achievable. Recall that for an $(n, s^m)$ design $D$, the average pairwise $L_p$-distance, $\bar{d}_p$, is a constant. Let

$$V_p(D) = \sum_{1\leq i<j\leq n} \left(d_p(\boldsymbol{x}_i, \boldsymbol{x}_j) - \bar{d}_p\right)^2 \tag{5}$$

be the variation of all pairwise $L_p$-distances in $D$. It is reasonable to expect that a good space-filling design, such as a maximin $L_p$-distance design, should minimize $V_p(D)$ as much as possible. The smaller the $V_p(D)$ value is, the better the design is. The next theorem shows that $\rho^2(D)$ is controlled by $V_2(D)$.

*Theorem 3.* Let $D$ be a balanced $(n, s^m)$ design and $V_2(D)$ be defined in (5) with $p = 2$. We have

$$\rho^2(D) \leq \frac{m+1-n}{(n-1)(m-1)} + \frac{72V_2(D)}{n^2 m(m-1)(s^2-1)^2}. \tag{6}$$

The equality holds if and only if $d_2(\boldsymbol{x}_i, s_0) = m(s^2-1)/12$ for any $1 \leq i \leq n$, where $s_0 = (s-1)/2$ and $d_2(\boldsymbol{x}_i, s_0) = \sum_{k=1}^m (x_{ik} - s_0)^2$.

*Example 4.* A regular cross-polytope in $\mathbb{R}^m$ is the convex hull of $m$ mutually perpendicular line segments of equal length, intersecting at the midpoint of each of them. An $(n, s^m)$ design $D$ is called mirror-symmetric if for any point $(x_1, \ldots, x_m)$ in $D$, $(s-1-x_1, \ldots, s-1-x_m)$ is also a point in $D$. Let $D$ be any maximin mirror-symmetric $(n, s^m)$ design with $n = 2m$ and $d_2(D) = m(s^2-1)/6$. Wang, Yang, and Xu (2018) proved that the geometric structure of $D$ is unique—the $2m$ points are the vertices of a cross-polytope in $\mathbb{R}^m$. All the $n(n-1)/2 = 2m^2 - m$ pairwise $L_2$-distances of $D$ only take two values, with $(2m^2-2m)$ $d_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$'s equal to $m(s^2-1)/6$ and the remaining $m$ $d_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$'s equal to $m(s^2-1)/3$. By (2), we have $\bar{d}_2 = m^2(s^2-1)/(6m-3)$. Therefore, the variation of all $d_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$'s is $V_2(D) = m^3(m-1)\left(s^2-1\right)^2/(36m-18)$. Then by Theorem 3, we obtain $\rho^2(D) \leq 0$, which means that $\rho^2(D) = 0$, that is, $D$ is orthogonal.

Since maximin distance designs under the $L_2$-distance tend to have small $V_2(D)$, Theorem 3 explains why the average $\rho^2(D)$ values are small for maximin distance LHDs in Example 1(ii). When $n = 2m$ is an even number such that an orthogonal mirror-symmetric LHD$(n, m)$ $D$ in Example 4 exists, for example, $n \geq 4$ being any power of 2 (Sun, Liu, and Lin 2009), we have $d_2(D)/\bar{d}_2 = 1 - 1/n$. Wang, Yang, and Xu (2018) conjectured that such $D$ also has maximin $L_2$-distance among all possible LHD$(n, m)$'s. We believe this is true at least for small $n$, for example, $n = 8, 16, 24$, etc., based on numerical search. Thus, for such $n$ and $m = n/2$ searching for a maximin $L_2$-distance LHD tends to yield a nearly orthogonal LHD as observed in Example 1 (iii). When $m$ is close to $n-1$, maximin distance designs under the $L_2$-distance are close to be $L_2$-equidistant, hence, Theorem 2 and Proposition 2 explain why the average $\rho^2(D)$ values become smaller as $m$ approaches $n-1$ and why the minimum average $\rho^2(D)$ values are all attained when $m = n-1$ in Example 1(iv). Sun, Wang, and Xu (2019) compared the projection properties of four $19 \times 18$ designs and observed a phenomenon that maximin $L_2$-distance LHDs tend to have small correlations in all projected dimensions. The above theoretical results also explain this phenomenon.

In view of Theorems 2 and 3, we naturally expect that a nearly $L_2$-equidistant design should have a small $\rho^2(D)$ value. For an $(n, s^m)$ design $D$, suppose $\max_{1\le i<j\le n} |d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) - \bar{d}_2| \le \delta$ for some $\delta > 0$. It is clear that $V_2(D) \le n(n-1)\delta^2/2$. Then by Theorem 3, $\rho^2(D)$ is also controlled by $\delta$. We summarize this as a corollary.

*Corollary 1.* Let $D$ be a balanced $(n, s^m)$ design and $\bar{d}_2$ be defined in (2). We have

$$\rho^2(D) \le \frac{mn(\delta/\bar{d}_2)^2 + m + 1 - n}{(n-1)(m-1)}, \tag{7}$$

where $\delta = \max_{1\le i<j\le n} |d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) - \bar{d}_2|$.

Theorem 2 and Corollary 1 together give

$$\max\left\{\frac{m+1-n}{(n-1)(m-1)}, 0\right\} \le \rho^2(D) \le \frac{mn(\delta/\bar{d}_2)^2 + m + 1 - n}{(n-1)(m-1)}.$$

In particular, $\delta = 0$ implies $\rho^2(D) = (m+1-n)/[(n-1)(m-1)]$. A nearly $L_2$-equidistant design usually has a small ratio of $\delta/\bar{d}_2$. Hence, Corollary 1 implies that a nearly $L_2$-equidistant design does have a small $\rho^2(D)$ value. For example, suppose if $m = n-1$ and $\delta/\bar{d}_2 \le 0.1$, then $\rho^2(D) \le 0.01mn/\{(n-1)(m-1)\} = 0.01n/(n-2)$. The next two examples illustrate the value of Corollary 1.

*Example 5.* Let

$$D = \begin{pmatrix} 3 & 1 & 0 & 2 & 1 & 0 & 0 & 2 & 3 & 3 & 1 & 2 \\ 2 & 3 & 1 & 0 & 3 & 1 & 1 & 0 & 2 & 2 & 3 & 0 \\ 3 & 1 & 2 & 1 & 0 & 2 & 1 & 0 & 0 & 2 & 3 & 3 \\ 2 & 3 & 0 & 3 & 1 & 0 & 3 & 1 & 1 & 0 & 2 & 2 \\ 3 & 3 & 3 & 1 & 2 & 1 & 0 & 2 & 1 & 0 & 0 & 2 \\ 2 & 2 & 2 & 3 & 0 & 3 & 1 & 0 & 3 & 1 & 1 & 0 \\ 3 & 0 & 2 & 3 & 3 & 1 & 2 & 1 & 0 & 2 & 1 & 0 \\ 2 & 1 & 0 & 2 & 2 & 3 & 0 & 3 & 1 & 0 & 3 & 1 \\ 3 & 1 & 0 & 0 & 2 & 3 & 3 & 1 & 2 & 1 & 0 & 2 \\ 2 & 3 & 1 & 1 & 0 & 2 & 2 & 3 & 0 & 3 & 1 & 0 \end{pmatrix}^{\mathrm{T}}$$

be the balanced $(12, 4^{10})$ design from Example 1 of Sun, Pang, and Liu (2011), with its original levels mapped to 0, 1, 2, 3 by $-3 \to 0$, $-1 \to 1$, $1 \to 2$, $3 \to 3$. This design has $d_2(D) = 25$ and $\max_{1\le i<j\le n} d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) = 30$. By (2) we obtain $\bar{d}_2 = 27$, hence, $\delta = 3$ and $D$ is nearly maximin $L_2$-equidistant with $\delta/\bar{d}_2 = 1/9$. From Corollary 1, we know that $\rho^2(D) \le 4.8634 \times 10^{-3}$. In fact, $D$ is orthogonal.

*Example 6.* Using Theorem 4 of Wang, Xiao, and Xu (2018), we can construct an LHD(9, 9)

$$D = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 6 & 8 & 9 & 7 & 5 & 3 & 1 \\ 3 & 6 & 9 & 7 & 4 & 1 & 2 & 5 & 8 \\ 4 & 8 & 7 & 3 & 1 & 5 & 9 & 6 & 2 \\ 5 & 9 & 4 & 1 & 6 & 8 & 3 & 2 & 7 \\ 6 & 7 & 1 & 5 & 8 & 2 & 4 & 9 & 3 \\ 7 & 5 & 2 & 9 & 3 & 4 & 8 & 1 & 6 \\ 8 & 3 & 5 & 6 & 2 & 9 & 1 & 7 & 4 \\ 9 & 1 & 8 & 2 & 7 & 3 & 6 & 4 & 5 \end{pmatrix}.$$

This design has $d_2(D) = 126$ and $\max_{1\le i<j\le n} d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) = 140$. By (2) we obtain $\bar{d}_2 = 135$, hence, $\delta = 9$ and $D$ is nearly maximin $L_2$-equidistant with $\delta/\bar{d}_2 = 1/15$. From Corollary 1, we know that $\rho^2(D) \le 0.02125$. In fact, $D$ is nearly orthogonal with $\rho^2(D) = 0.01763$.

*Remark 1.* We point out that for $D = (x_{ik})$ in Example 6, some of the two-dimensional projections have obvious patterns that are not space-filling. This phenomenon occurs in many specialized constructions, especially when the factor-to-run ratio is large, for example, the good-lattice-point-based methods in Example 6 and Section 3.2, and some existing constructions for column-orthogonal designs. Such phenomenon can be somewhat mitigated by adding $u_{ik}/n$ (or $u_{ik}/s$ for an $s$-level design) to each $x_{ik}$ in practice, where $u_{ik}$'s are independent random variables from uniform (0, 1).

## 3. Connection Between Projection Uniformity and Maximin $L_1$-Distance

### 3.1. Projection Uniformity and Lower Bounds

Uniformity is another favorable space-filling criterion which aims to scatter points as uniformly as possible over the whole design space by minimizing certain discrepancy metric. The most widely used discrepancies include the centered $L_2$-discrepancy, the wrap-around $L_2$-discrepancy and the discrete discrepancy (see, e.g., Fang, Li, and Sudjianto 2006). Sun, Wang, and Xu (2019) considered the design's projection uniformity under the centered $L_2$-discrepancy measure. For an $(n, s^m)$ design $D = (x_{ik})$, its (squared) centered $L_2$-discrepancy (CD) is defined as

$$\mathrm{CD}(D) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\left(1 + \frac{1}{2}|z_{ik}| + \frac{1}{2}|z_{jk}| - \frac{1}{2}|z_{ik} - z_{jk}|\right) - \frac{2}{n}\sum_{i=1}^{n}\prod_{k=1}^{m}\left(1 + \frac{1}{2}|z_{ik}| - \frac{1}{2}|z_{ik}|^2\right) + \left(\frac{13}{12}\right)^m,$$

where $z_{ik} = (2x_{ik} - s + 1)/(2s)$.

By focusing on two-dimensional projection uniformity based on the CD criterion, Sun, Wang, and Xu (2019) proposed the uniform projection criterion, which is defined as

$$\phi(D) = \frac{2}{m(m-1)}\sum_{|\mathcal{U}|=2}\mathrm{CD}(D_{\mathcal{U}}), \tag{8}$$

where $\mathcal{U} \subset \{1, 2, \ldots, m\}$, $|\mathcal{U}|$ stands for the cardinality of $\mathcal{U}$ and $D_{\mathcal{U}}$ is the projected design of $D$ onto dimensions indexed by the elements of $\mathcal{U}$. A design is called a uniform projection design if it attains the minimum $\phi(D)$ value. Note that for a balanced design, the criterion (8) is equivalent to the $CL_{2,t}$ criterion with $t = 2$ proposed in Ma, Fang, and Lin (2003).

According to Theorem 1 of Sun, Wang, and Xu (2019) and their Section 4's numerical results, a uniform projection design $D$ tends to have small $\phi(D_{\mathcal{U}})$ values for all projections. We can define other uniform projection criterion using other discrepancies, but we restrict our attention to the original definition (8) due to the associated good theoretical properties given below. The following two lemmas from Sun, Wang, and Xu (2019) connect the projection uniformity with the pairwise $L_1$-distance between design points.

*Lemma 3.* For a balanced $(n, s^m)$ design $D$,

$$\phi(D) = \frac{g(D)}{4n^2m(m-1)s^2} + C(m, s), \tag{9}$$

where

$$g(D)=\sum_{i=1}^{n}\sum_{j=1}^{n} d_1^2(\boldsymbol{x}_i,\boldsymbol{x}_j)-\frac{2}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{n} d_1(\boldsymbol{x}_i,\boldsymbol{x}_j)\right)^2$$

and

$$C(m,s)=\frac{4(5m-2)s^4+30(3m-5)s^2+15m+33}{720(m-1)s^4}+\frac{1+(-1)^s}{64s^4}.$$

*Lemma 4.* For a balanced $(n,s^m)$ design $D$, $\phi(D)\geq\phi_{\mathrm{LB1}}$, where

$$\phi_{\mathrm{LB1}}=\frac{5m(4s^4+2(13n-17)s^2-n+5)-(n-1)(8s^4+150s^2-33)}{720(n-1)(m-1)s^4}+\frac{1+(-1)^s}{64s^4}. \quad (10)$$

The bound is achieved if and only if $D$ is an equidistant design under the $L_1$-distance.

The lower bound (10) can be viewed as an analogy of the saturated and supersaturated bound $(m+1-n)/[(n-1)(m-1)]$ for $\rho^2(D)$. The lower bound is effective when the factor-to-run ratio $m/n$ is large, but less useful when the factor-to-run ratio is small. For example, for an LHD(20, 3), the bound in (10) is $-0.002957$, which is negative. Now we give an improved lower bound of $\phi(D)$ which incorporates the small factor-to-run ratio case into consideration.

*Theorem 4.* For a balanced $(n,s^m)$ design $D$, $\phi(D)\geq\phi_{\mathrm{LB}}=\max\{\phi_{\mathrm{LB1}},\phi_{\mathrm{LB2}}\}$, where $\phi_{\mathrm{LB1}}$ is defined in (10) and

$$\phi_{\mathrm{LB2}}=\frac{26s^2-1}{144s^4}+\frac{1+(-1)^s}{64s^4}. \quad (11)$$

Furthermore, the lower bound (11) is achieved if and only if $D$ is combinatorially orthogonal.

Theorem 4 improves Lemma 4 and avoids the lower bound being negative. The lower bound (11) is from the perspective of combinatorial orthogonality. It shows that an OA$(n,m,s,2)$, if it exists, has the best two-dimensional projection uniformity $\phi(D)$. Let us examine when the lower bound $\phi_{\mathrm{LB2}}$ is better than $\phi_{\mathrm{LB1}}$ in more detail. Comparing (10) and (11), we have

$$\phi_{\mathrm{LB1}}-\phi_{\mathrm{LB2}}=\frac{\left(s^2-1\right)\left(s^2(5m-2n+2)-5m-7n+7\right)}{180(m-1)(n-1)s^4}.$$

Therefore, $\phi_{\mathrm{LB1}}\leq\phi_{\mathrm{LB2}}$ if and only if

$$m\leq\frac{2s^2+7}{5s^2-5}(n-1). \quad (12)$$

From (12), we see that when $s=2$, the lower bound $\phi_{\mathrm{LB2}}$ is better than $\phi_{\mathrm{LB1}}$ if and only if $m<n-1$, which is the same condition when the lower bound $(m+1-n)/[(n-1)(m-1)]$ for $\rho^2(D)$ is negative in Theorem 2. However, as $s$ increases, the ratio $(2s^2+7)/(5s^2-5)$ decreases and condition (12) differs from that for $\rho^2(D)$. For example, when $s=3$ and 4, condition (12) becomes $m\leq 5(n-1)/8$ and $m\leq 39(n-1)/75$, respectively. In the extreme case of LHDs with $s=n$, $\phi_{\mathrm{LB1}}\leq\phi_{\mathrm{LB2}}$ if and only if $m\leq(2n^2+7)/(5n+5)$, which is approximately equivalent to $m/(n-1)\leq 2/5$ for medium or large $n$.

*Example 7.* Sun, Wang, and Xu (2019) constructed a uniform projection design $D$ via algorithm in their Example 1. This design is an LHD(25, 3) with $\phi(D)=5.279\times10^{-4}$. The original lower bound established in Sun, Wang, and Xu (2019) is $\phi_{\mathrm{LB1}}=-3.558\times10^{-3}<0$. Since $m/(n-1)=1/8<2/5$, by (12), $\phi_{\mathrm{LB2}}$ is better than $\phi_{\mathrm{LB1}}$. The improved lower bound is $\phi_{\mathrm{LB}}=\phi_{\mathrm{LB2}}=3.135\times10^{-4}$, which is of the same order as $\phi(D)$.

### 3.2. Projection Uniformity and Maximin $L_1$-Distance

This subsection presents some further results on the relationship between $\phi(D)$ and $L_1$-distance parallel to the results regarding $\rho^2(D)$. We also establish asymptotical optimality of several classes of LHDs.

*Theorem 5.* Let $D$ be a balanced $(n,s^m)$ design and $V_1(D)$ be defined in (5) with $p=1$. We have

$$\phi(D)\leq\phi_{\mathrm{LB1}}+\frac{V_1(D)}{2n^2m(m-1)s^2}, \quad (13)$$

where $\phi_{\mathrm{LB1}}$ is defined in (10). The equality holds if and only if $d_2(\boldsymbol{x}_i,s_0)=m(s^2-1)/12$ for any $1\leq i\leq n$, where $s_0=(s-1)/2$ and $d_2(\boldsymbol{x}_i,s_0)=\sum_{k=1}^{m}(x_{ik}-s_0)^2$.

Maximin $L_1$-distance designs tend to have small $V_1(D)$; thus, by Theorem 5, they tend to have small $\phi(D)$ values and good two-dimensional projection uniformity. We have the following corollary, parallel to Corollary 1.

*Corollary 2.* Let $D$ be a balanced $(n,s^m)$ design and $\bar{d}_1$ be defined in (1). We have

$$\phi(D)\leq\phi_{\mathrm{LB1}}+\frac{nm(s^2-1)^2(\delta/\bar{d}_1)^2}{36(n-1)(m-1)s^4}, \quad (14)$$

where $\delta=\max_{1\leq i<j\leq n}|d_1(\boldsymbol{x}_i,\boldsymbol{x}_j)-\bar{d}_1|$.

The next lemma gives an upper bound of $\phi(D)$ for balanced designs.

*Lemma 5.* For a balanced $(n,s^m)$ design $D$, $\phi(D)\leq\phi_{\mathrm{UB}}$, where

$$\phi_{\mathrm{UB}}=\frac{(10m-8)s^4+(140m-150)s^2-25m+33}{720(m-1)s^4}+\frac{(-1)^s+1}{64s^4}. \quad (15)$$

We define the relative $\phi$-efficiency of an $(n,s^m)$ design $D$ as

$$\phi_{\mathrm{RE}}(D)=\frac{\phi_{\mathrm{UB}}-\phi(D)}{\phi_{\mathrm{UB}}-\phi_{\mathrm{LB}}}, \quad (16)$$

where $\phi_{\mathrm{LB}}=\max\{\phi_{\mathrm{LB1}},\phi_{\mathrm{LB2}}\}$ is defined in Theorem 4, and $\phi_{\mathrm{UB}}$ is given in (15). The larger the $\phi_{\mathrm{RE}}(D)$ value is, the better the projection uniformity is. Obviously, $\phi_{\mathrm{RE}}(D)=1$ if and only if $\phi(D)=\phi_{\mathrm{LB}}$, that is, $D$ is either an OA$(n,m,s,2)$ or a maximin $L_1$-equidistant design.

In the sequel, we investigate the projection uniformity of several classes of maximin LHDs constructed via good lattice point (GLP) designs. Let $(h_1,\ldots,h_m)$ be a set of integers coprime to $n$. A GLP design $D=(x_{ik})$ is defined by $x_{ik}=i\times h_k$

(mod $n$) for $i = 1, \ldots, n$ and $k = 1, \ldots, m$. The GLP design $D$ is an $n \times m$ LHD. Let $\varphi(n)$ be the Euler function, that is, the number of positive integers coprime to $n$ and less than $n$. We can construct an $n \times m$ GLP design for any $m \leq \varphi(n)$. In particular, when $n$ is a prime, $\varphi(n) = n - 1$. For $b \in \mathcal{Z}_n$, let $D_b = D + b = (x_{ik} + b) \pmod n$ be a linear permutation of the GLP design $D$. Let $W : \mathcal{Z}_n \to \mathcal{Z}_n$ be the Williams transformation (Williams 1949), where

$$W(x) = \begin{cases} 2x, & 0 \leq x < n/2 \\ 2(n-x) - 1, & n/2 \leq x < n. \end{cases}$$

The following general construction procedure was first proposed by Wang, Xiao, and Xu (2018) for constructing maximin $L_1$-distance LHDs.

1. Given an integer $n$, generate an $n \times m$ GLP design $D$, where $m \leq \varphi(n)$.
2. For $b = 0, \ldots, n-1$, generate $D_b = D + b \pmod n$ and $E_b = W(D_b)$.
3. Find the best $E_b$ under the given optimality criterion.

We consider two cases. First consider the case $n = p$ and $m = n - 1$, where $p$ is an odd prime. Denote $E_{b^*}$ the maximin $L_1$-distance LHD and $E_{b^{**}}$ the uniform projection LHD by the general construction procedure, where $b^*$ and $b^{**} \in \mathcal{Z}_n$. Recall that $\bar{d}_1 = (n+1)m/3$ for an LHD. Theorem 2 in Wang, Xiao, and Xu (2018) states that $d_1(E_{b^*})$ almost attains the upper bound in Lemma 1 with $d_1(E_{b^*})/\bar{d}_1 \geq 1 - 2/\sqrt{3(n^2-1)}$. The following proposition shows that either $b^* = b^{**}$ or the maximin $L_1$-distance LHD $E_{b^*}$ has the second smallest $\phi$ value.

*Proposition 3.* Let $n = p$ be an odd prime, $m = n - 1$, $c_0 = \lfloor \sqrt{(n^2-1)/12} \rfloor$ and $E_{b^*}$ be the design maximizing $d_1(E_b)$.

(i) If $c_0 < \sqrt{n^2/12 - 11/36} - 2/3$ or $c_0 \geq \sqrt{(n^2-4)/12} - 1/2$, then $b^* = b^{**}$, that is, $E_{b^*}$ is also the uniform projection design among all $E_b$'s, $b \in \mathcal{Z}_n$ and

$$\phi(E_{b^*}) = \begin{cases} \phi_{\text{LB1}} + \{(c_0+1)^2 - (n^2 - 1)/12\}^2/\{(n-2)n^4\}, & c_0 < \sqrt{n^2/12 - 11/36} - 2/3, \\ \phi_{\text{LB1}} + \{c_0^2 - (n^2 - 1)/12\}^2/\{(n-2)n^4\}, & c_0 \geq \sqrt{(n^2-4)/12} - 1/2, \end{cases}$$

where $\phi_{\text{LB1}} = (12n^3 + 154n^2 - 12n - 29)/(720n^4)$ is defined in (10) with $m = n - 1$ and $s = n$.

(ii) Otherwise, $E_{b^*}$ has the second smallest $\phi$ value among all $E_b$'s, $b \in \mathcal{Z}_n$, and $\phi(E_{b^*}) = \phi_{\text{LB1}} + \{c_0^2 - (n^2-1)/12\}^2/\{(n-2)n^4\}$.

(iii) Both $\phi_{\text{RE}}(E_{b^*})$ and $\phi_{\text{RE}}(E_{b^{**}})$ are of order $1 - O(1/n^3)$ and converge to one as $n$ increases.

The following proposition shows that the uniform projection design also has the largest or second largest $L_1$-distance.

*Proposition 4.* Let $n = p$ be an odd prime, $n \geq 5$, and $E_{b^{**}}$ be the design minimizing $\phi(E_b)$. Then $E_{b^{**}}$ is either the maximin $L_1$-distance design or has the second largest $d_1(D)$ value among all $E_b$'s, $b \in \mathcal{Z}_n$.

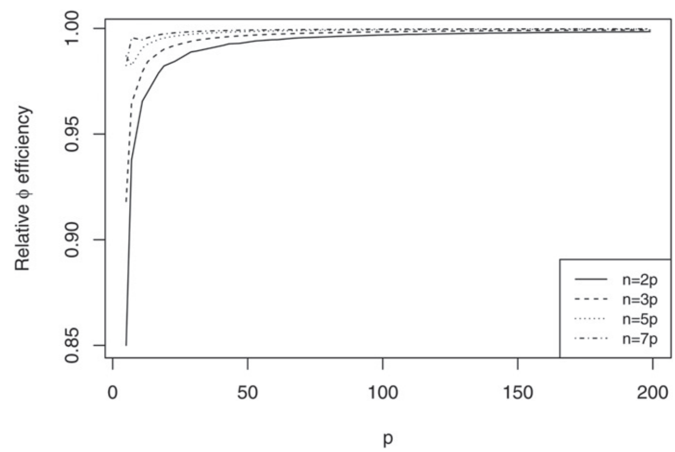


**Figure 3.** Plot of the $\phi_{\text{RE}}(E_b)$ values for $n = 2p, 3p, 5p, 7p$, $p \leq 200$ and $b = \lfloor n(1 + 1/\sqrt{3})/4 \rfloor$.

Now we consider the second case with $n = kp$ where $k$ is a prime and $p$ is an odd prime. Let $m = \varphi(n)$ and $b = \lfloor n(1 + 1/\sqrt{3})/4 \rfloor$. The Williams transformed GLP design $E_b = W(D_b) = W(D + b) \pmod n$ is an LHD$(n, m)$. Although $E_b$ may not be the best design for $b \in \mathcal{Z}_n$, Wang, Xiao, and Xu (2018) proved that when $k = 2$, $E_b$ is asymptotically maximin $L_1$-distance with $d_1(E_b)/\bar{d}_1 = 1 - O(1/n)$. For $k = 3, 5, 7$, Wang, Xiao, and Xu (2018) showed numerically that as $n > 100$, $d_1(E_b)/\bar{d}_1$ converges to 1 for $n = 7p$ and are greater than 0.95 for $n = 3p, 5p$.

Figure 3 shows the $\phi_{\text{RE}}(E_b)$ values for $n = 2p, 3p, 5p, 7p$ and $p \leq 200$. We see that the $\phi_{\text{RE}}(E_b)$ values are all larger than 0.85 for $k = 2$ and larger than 0.9 for $k = 3, 5, 7$. When $p$ increases, the relative $\phi$-efficiencies approach to one quickly in all four cases. These LHDs are all nearly optimal under the criteria of orthogonality, projection uniformity and maximin $L_1$-distance.

The following theorem shows that $\phi_{\text{RE}}(E_b)$ convergences to one when $n = 2p$.

*Theorem 6.* Let $p \geq 7$ be an odd prime. Let $n = 2p$, $m = p - 1$ and $b = \lfloor n(1 + 1/\sqrt{3})/4 \rfloor$. As $n$ increases, $\phi_{\text{RE}}(E_b)$ converges to one with

$$\phi_{\text{RE}}(E_b) \geq 1 - \frac{\left(11 - 6\sqrt{3}\right)n^4 + 2\left(9\sqrt{3} - 14\right)n^3 - 4\left(5\sqrt{3} - 7\right)n^2 + 8\left(2\sqrt{3} - 1\right)n + 16}{(n-2)^2 n \left(n^2 - 1\right)} = 1 - O(1/n).$$

## 4. Connection Between Projection Uniformity and Column-Orthogonality

In this section, we study the relationship between projection uniformity and column-orthogonality. Both criteria aim to optimize the two-dimensional projections of the design. The results in Sections 2 and 3 indicate that these two criteria are quite similar from the distance perspective. The expression of $h(D)$ in Theorem 1 and the expression of $g(D)$ in Lemma 3 are identical except that the former is with respect to $L_2$-distances while the latter is with respect to $L_1$-distances.

In the proof of Theorem 1, we derive an auxiliary expression of $\rho^2(D)$, Equation (S2) in the supplementary materials,

$$\rho^2(D) = \frac{\sum_{i=1}^n \sum_{j=1}^n d_2^2(\boldsymbol{x}_i, \boldsymbol{x}_j) - 2n \sum_{i=1}^n d_2^2(\boldsymbol{x}_i, s_0)}{n^2 m(m-1)(s^2-1)^2/36} - \frac{m+2}{2(m-1)}.$$

By combining (3) in Lemma 2 and Theorem 2 of Sun, Wang, and Xu (2019), we can also represent $\phi(D)$ as

$$\phi(D) = \frac{\sum_{i=1}^n \sum_{j=1}^n s^2 d_1^2(\boldsymbol{x}_i, \boldsymbol{x}_j) - 2n \sum_{i=1}^n d_2^2(\boldsymbol{x}_i, s_0)}{4m(m-1)n^2 s^4} + C_1(n, m, s),$$

where $C_1(n, m, s)$ is a constant related to $n, m, s$ only. Intuitively, for any $(n, s^m)$ design $D = (x_{ik})$, the terms $\sum_{i=1}^n \sum_{j=1}^n d_2^2(\boldsymbol{x}_i, \boldsymbol{x}_j)$ and $\sum_{i=1}^n \sum_{j=1}^n d_1^2(\boldsymbol{x}_i, \boldsymbol{x}_j)$ should be highly correlated. In fact, they are equivalent to $V_2(D)$ and $V_1(D)$, respectively. Theorem 3 shows that $\rho^2(D)$ is controlled by $V_2(D)$ and Theorem 5 shows that $\phi(D)$ is controlled by $V_1(D)$. A design with a small $V_2(D)$ value tends to have a small $V_1(D)$ value and vice versa. This can also be observed from the following inequalities

$$d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{k=1}^m (x_{ik} - x_{jk})^2 \leq (s-1) \sum_{k=1}^m |x_{ik} - x_{jk}| = (s-1) d_1(\boldsymbol{x}_i, \boldsymbol{x}_j)$$

and

$$m\, d_2(\boldsymbol{x}_i, \boldsymbol{x}_j) = m \sum_{k=1}^m (x_{ik} - x_{jk})^2 \geq \left( \sum_{k=1}^m |x_{ik} - x_{jk}| \right)^2 = d_1^2(\boldsymbol{x}_i, \boldsymbol{x}_j),$$

which states that $L_1$- and $L_2$-distances are bounded by each other. Therefore, designs with small $\phi(D)$ values tend to have small correlations between columns and vice versa.

*Example 8.* To study the relationship between column-orthogonality and projection uniformity numerically, we randomly generate 100 LHDs with $n = 19$ runs and $m$ factors for $m = 6, 12, 18$. We compute the $\rho^2(D)$ and $\phi(D)$ values of the 100 designs. For comparison, we also compute the $\mathrm{CD}(D)$ values.

Figure 4 shows the scatterplots of $\phi$ values (multiplied by 1000) and CD values against $\rho^2$ values, where the numbers in parentheses are the corresponding correlation coefficients. The $\phi(D)$ and $\rho^2(D)$ values have strong correlations (about 0.9) for each $m$. The correlations between the $\phi(D)$ and $\mathrm{CD}(D)$ values are strong for small $m$ and become weaker for larger $m$. This phenomenon also exists for other LHDs and $(n, s^m)$ designs we simulated. The criteria $\phi(D)$ and $\rho^2(D)$ are highly consistent while $\phi(D)$ and $\mathrm{CD}(D)$ are not consistent for designs with large factor-to-run ratio.

In particular, for two-level designs, since $L_1$- and $L_2$-distances are both equivalent to the Hamming distance, the criteria $\phi(D)$ and $\rho^2(D)$ are equivalent, and they both measure a design's combinatorial orthogonality. We summarize this in the following corollary.

*Corollary 3.* For a balanced $(n, 2^m)$ design $D$,

$$\phi(D) = \frac{1}{64}\rho^2(D) + \frac{215}{4608}.$$

As a result, orthogonality and projection uniformity are exactly equivalent for two-level designs.

## 5. Discussion

We have shown some connections among the three criteria—column-orthogonality, projection uniformity, and maximin distance. Column-orthogonality and projection uniformity focus on two-dimensional projections of designs which are useful in experiments where only a subset of input variables are active. The maximin distance criterion focuses on the separation of design points in the full-dimensional space. The connections among these criteria reveal some interesting relationships between rows and columns of a design. These results not only help justify one criterion from each other, but also provide new insights on design constructions; see Section 3.2 for example.

Optimal or nearly optimal designs given in our examples are (nearly) saturated or supersaturated, which have high factor-to-run ratios. They are economic to explore experiments with many factors. In many physical or computer experiments, there are a large number of factors at an early stage and often only a few of them are important (Moon, Dean, and Santner 2012; Woods and Lewis 2016). Our designs are suitable for screening factors in such experiments since they have large distance, low correlations, and good projection properties, so that they should perform well under various modeling and data analysis strategies. It would be meaningful to investigate how to use the connections among different space-filling criteria for constructing space-filling designs with $m$ much smaller than $n$. Theorems 3 and 5 indicate that designs with small $L_1$- or $L_2$-distance variance tend to have small correlation or good projection uniformity. Conversely, maximin distance designs may be constructed among the class of column-orthogonal designs or uniform projection designs when $m$ is much smaller than $n$. One difficulty of obtaining maximin distance designs for small $m/n$ ratio is that the upper bounds of $L_1$- or $L_2$-distance given in Lemma 1 are too loose. Equations (4) and (9) might be helpful in establishing better maximin distance bounds in such cases. We will further study this problem in the future.

Throughout the article, we assume that the design is balanced, that is, an OA of strength one. Two problems are worthy of further study. The first problem is to relax the balance and discrete-level constraints and explore the relationships among maximin distance, column-orthogonality, and projection uniformity for a general design in the $m$-dimensional cube. The main results of this article no longer holds and a general connection is difficult, but in some special cases we can find good designs with certain geometrical structures and study their space-filling properties, for example, the cross-polytope

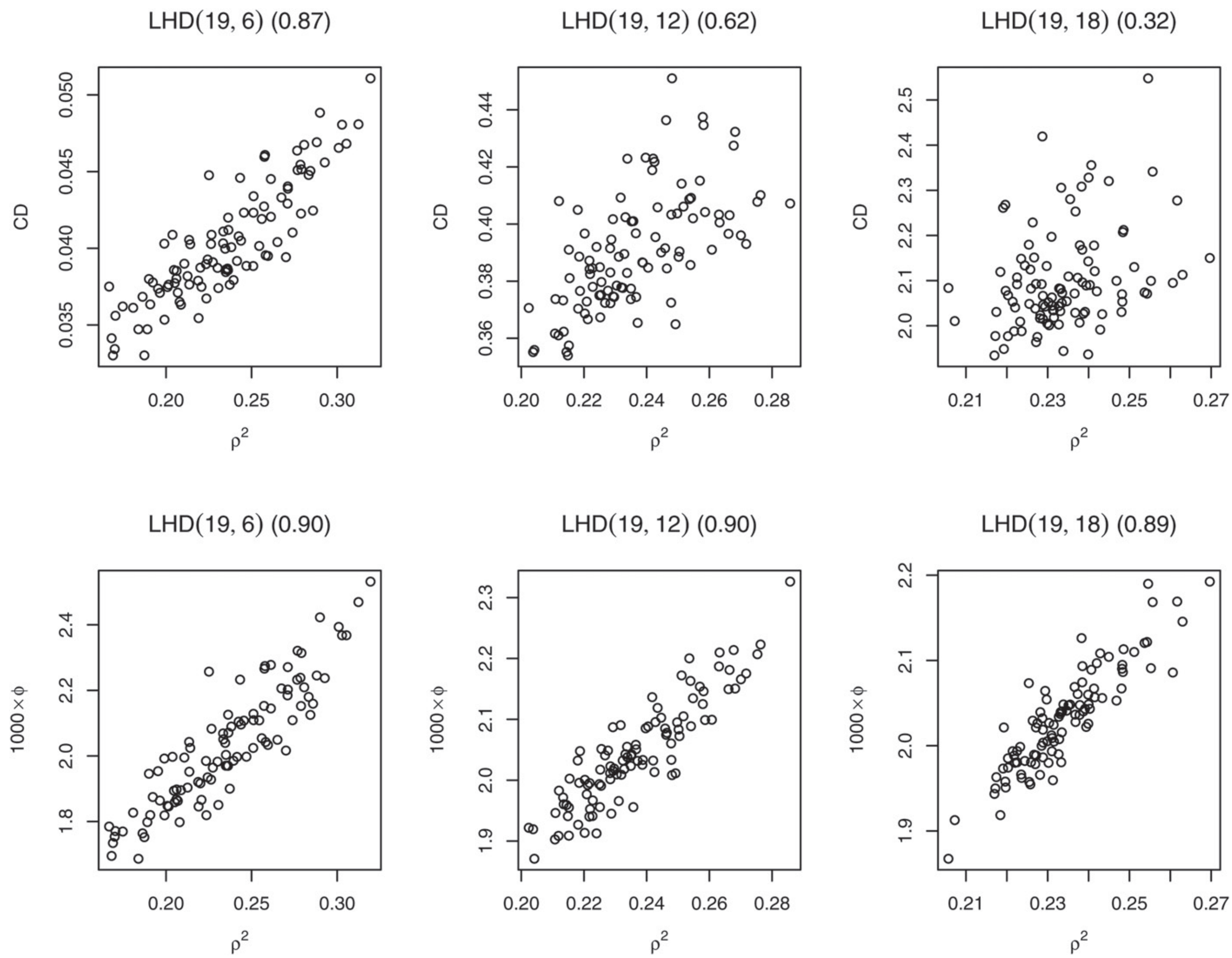


**Figure 4.** Scatterplots of CD and $\phi$ (multiplied by 1000) against $\rho^2$, with corresponding correlation coefficients in parentheses.

structure studied in Wang, Yang, and Xu (2018). The second problem is to investigate the relationships among space-filling criteria for OAs of strength two or higher, OA-based LHDs (Tang 1993; Xiao and Xu 2018) and strong OAs (He and Tang 2013, 2014; He, Cheng, and Tang 2018).

## Supplementary Materials

The online supplementary materials contain the proofs of Lemmas 2 and 5, Theorems 1–6, and Propositions 3 and 4, and some additional tables and figures.


## Acknowledgments

We are grateful to the editor, an associate editor, and three reviewers for their insightful comments and constructive suggestions.

## Funding

Wang was supported by NSFC (grant nos. 11971004 and 71931004) and Shanghai Sailing Program 19YF1412800, and Sun was supported by NSFC (grant nos. 11971098 and 11771220), Fundamental Research Funds for the Central Universities, and Jilin Provincial Science and Technology Development Plan Funded Project (grant no. 20180520026JH).


## References


Ba, S., Myers, W. R., and Brenneman, W. A. (2015), "Optimal Sliced Latin Hypercube Designs," *Technometrics*, 57, 479–487. [375,376]

Bingham, D., Sitter, R. R., and Tang, B. (2009), "Orthogonal and Nearly Orthogonal Designs for Computer Experiments," *Biometrika*, 96, 51–65. [375]

Booth, K. H. V., and Cox, D. R. (1962), "Some Systematic Supersaturated Designs," *Technometrics*, 4, 489–495. [378]

Butler, N. A. (2005), "Supersaturated Latin Hypercube Designs," *Communications in Statistics—Theory and Methods*, 34, 417–428. [375,379]

Cheng, C. S. (1997), "$E(s^2)$-Optimal Supersaturated Designs," *Statistica Sinica*, 7, 929–939. [378,379]

Draguljić, D., Santner, T. J., and Dean, A. M. (2012), "Non-Collapsing Space-Filling Designs for Bounded Nonrectangular Regions," *Technometrics*, 54, 169–178. [377]

Fang, K. T., Li, R., and Sudjianto, A. (2006), *Design and Modeling for Computer Experiments*, New York: CRC Press. [375,380]

Fang, K. T., Lin, D. K. J., Winker, P., and Zhang, Y. (2000), "Uniform Design: Theory and Application," *Technometrics*, 42, 237–248. [375,378]

He, Y., Cheng, C. S., and Tang, B. (2018), "Strong Orthogonal Arrays of Strength Two Plus," *The Annals of Statistics*, 46, 457–468. [384]

He, Y., and Tang, B. (2013), "Strong Orthogonal Arrays and Associated Latin Hypercubes for Computer Experiments," *Biometrika*, 100, 254–260. [384]

——— (2014), "A Characterization of Strong Orthogonal Arrays of Strength Three," *The Annals of Statistics*, 42, 1347–1360. [384]

Johnson, M. E., Moore, L. M., and Ylvisaker, D. (1990), "Minimax and Maximin Distance Designs," *Journal of Statistical Planning and Inference*, 26, 131–148. [375,376]

Joseph, V. R. (2016), "Space-Filling Designs for Computer Experiments: A Review," *Quality Engineering*, 28, 28–35. [375]

Joseph, V. R., Gul, E., and Ba, S. (2015), "Maximum Projection Designs for Computer Experiments," *Biometrika*, 102, 371–380. [378]

Joseph, V. R., and Hung, Y. (2008), "Orthogonal-Maximin Latin Hypercube Designs," *Statistica Sinica*, 18, 171–186. [378]

Karunanayaka, R. C., and Tang, B. (2018), "On the Existence and Constructions of Orthogonal Designs," *Australian and New Zealand Journal of Statistics*, 60, 471–480. [379]

Kleijnen, J. P. (2017), "Design and Analysis of Simulation Experiments: Tutorial," in *Advances in Modeling and Simulation*, eds. A. Tolk, J. Fowler, G. Shao, and E. Yucesan, New York: Springer, pp. 135–158. [375]

Lin, C. D. (2008), "New Developments in Designs for Computer Experiments and Physical Experiments," Ph.D. thesis, Simon Fraser University. [379]

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, 1460–1477. [379]

Lin, C. D., and Tang, B. (2015), "Latin Hypercubes and Space-Filling Designs," in *Handbook of Design and Analysis of Experiments*, eds. A. Dean, M. Morris, J. Stufken, and D. Bingham, New York: Chapman and Hall/CRC, pp. 593–625. [375]

Ma, C. X., Fang, K. T., and Lin, D. K. J. (2003), "A Note on Uniformity and Orthogonality," *Journal of Statistical Planning and Inference*, 113, 323–334. [380]

Moon, H., Dean, A., and Santner, T. (2012), "Two-Stage Sensitivity-Based Group Screening in Computer Experiments," *Technometrics*, 54, 376–387. [375,383]

Nguyen, N. K. (1996), "An Algorithmic Approach to Constructing Supersaturated Designs," *Technometrics*, 38, 69–73. [378,379]

Owen, A. B. (1994), "Controlling Correlations in Latin Hypercube Samples," *Journal of the American Statistical Association*, 89, 1517–1522. [375,376]

Pronzato, L., and Müller, W. G. (2012), "Design of Computer Experiments: Space Filling and Beyond," *Statistics and Computing*, 22, 681–701. [375]

Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, New York: Springer. [375]

Sun, F. S., Liu, M. Q., and Lin, D. K. J. (2009), "Construction of Orthogonal Latin Hypercube Designs," *Biometrika*, 96, 971–974. [379]

Sun, F. S., Pang, F., and Liu, M. (2011), "Construction of Column-Orthogonal Designs for Computer Experiments," *Science China Mathematics*, 54, 2683–2692. [380]

Sun, F. S., and Tang, B. (2017a), "A Method of Constructing Space-Filling Orthogonal Designs," *Journal of the American Statistical Association*, 112, 683–689. [375]

——— (2017b), "A General Rotation Method for Orthogonal Latin Hypercubes," *Biometrika*, 104, 465–472. [375]

Sun, F. S., Wang, Y., and Xu, H. (2019), "Uniform Projection Designs," *The Annals of Statistics*, 47, 641–661. [375,378,379,380,381,383]

Tang, B. (1993), "Orthogonal Array-Based Latin Hypercubes," *Journal of the American Statistical Association*, 88, 1392–1397. [384]

——— (1998), "Selecting Latin Hypercubes Using Correlation Criteria," *Statistica Sinica*, 8, 965–977. [375,376]

Tang, B., and Wu, C. F. J. (1997), "A Method for Constructing Supersaturated Designs and Its $E(s^2)$ Optimality," *The Canadian Journal of Statistics*, 25, 191–201. [378,379]

Wang, L., Xiao, Q., and Xu, H. (2018), "Optimal Maximin $L_1$-Distance Latin Hypercube Designs Based on Good Lattice Point Designs," *The Annals of Statistics*, 46, 3741–3766. [375,380,382]

Wang, Y., Yang, J., and Xu, H. (2018), "On the Connection Between Maximin Distance Designs and Orthogonal Designs," *Biometrika*, 105, 471–477. [375,378,379,384]

Williams, E. J. (1949), "Experimental Designs Balanced for the Estimation of Residual Effects of Treatments," *Australian Journal of Scientific Research*, 2, 149–168. [382]

Woods, D. C., and Lewis, S. M. (2016), "Design of Experiments for Screening," in *Handbook of Uncertainty Quantification*, eds. R. Ghanem, D. Higdon, and H. Owhadi, New York: Springer, pp. 1143–1185. [375,383]

Xiao, Q., and Xu, H. (2017), "Construction of Maximin Distance Latin Squares and Related Latin Hypercube Designs," *Biometrika*, 104, 455–464. [375]

——— (2018), "Construction of Maximin Distance Designs via Level Permutation and Expansion," *Statistica Sinica*, 28, 1395–1414. [375,384]

Ye, K. Q. (1998), "Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments," *Journal of the American Statistical Association*, 93, 1430–1439. [375,376]

Zhou, Y. D., and Xu, H. (2014), "Space-Filling Fractional Factorial Designs," *Journal of the American Statistical Association*, 109, 1134–1144. [375]

——— (2015), "Space-Filling Properties of Good Lattice Point Sets," *Biometrika*, 102, 959–966. [376]