Statistica Sinica **33** (2023), 1923-1942
doi:https://doi.org/10.5705/ss.202020.0317

DOUBLY COUPLED DESIGNS FOR COMPUTER EXPERIMENTS WITH BOTH QUALITATIVE AND QUANTITATIVE FACTORS

Feng Yang, C. Devon Lin, Yongdao Zhou and Yuanzhen He

Sichuan Normal University, Queen's University, Nankai University and Beijing Normal University

Abstract: Computer experiments with both qualitative and quantitative input variables occur frequently in many scientific and engineering applications. As a result, how to choose the input settings for such experiments is important for accurate statistical analysis, uncertainty quantification, and decision-making. Sliced Latin hypercube designs were the first systematic approach to address this issue. However, the cost of such designs increases with an increasing number of level combinations of the qualitative factors. To reduce the cost of the run size, marginally coupled designs have been proposed, in which the design for the quantitative factors is a sliced Latin hypercube design with respect to each qualitative factor. The drawback of such designs is that the corresponding data may not be able to capture the effects between any two (or more) qualitative and quantitative factors. To balance the run size and design efficiency, we propose a new type of design, namely doubly coupled designs. Here the design points for the quantitative factors form a sliced Latin hypercube design with respect to the level of any qualitative factor, and with respect to the level combinations of any two qualitative factors. The proposed designs have a better stratification property between the qualitative and quantitative factors compared with that of marginally coupled designs. Here, we establish the existence of the proposed designs, introduce several construction methods, and examine the properties of the resulting designs.

Key words and phrases: Completely resolvable orthogonal array, sliced Latin hypercube, stratification.

1. Introduction

Computer experiments provide an efficient way of representing real-world complex systems, and have been increasingly popular in the physical, engineering, and social sciences (Santner, Williams and Notz (2003); Fang, Li and Sudjianto (2005)). For recent works on computer experiments, refer to Chen, Santner and Dean (2018), Wang et al. (2018), Xiao and Xu (2018), Wang, Xiao and Xu

Corresponding author: Yongdao Zhou, School of Statistics and Data Science, Nankai University, Nankai District, Tianjin 300071, China. E-mail: ydzhou@nankai.edu.cn.

(2018), Huang et al. (2021), and the reference therein. One way of selecting the input settings for computer experiments is to use Latin hypercube designs (LHDs), proposed by McKay, Beckman and Conover (1979), because of the desirable feature that when projected onto any factor, the resulting design points spread out uniformly and achieve the maximum stratification. However, an LHD is not guaranteed to be space filling in two or more dimensions. As a result, several improved LHDs have been discussed, such as maxmin LHDs (Morris and Mitchell (1995); Joseph and Hung (2008); Wang, Xiao and Xu (2018)), orthogonal array-based LHDs (Tang (1993)), orthogonal LHDs (Georgiou and Efthimiou (2014); Sun and Tang (2017); Li, Liu and Tang (2020)), and strong orthogonal array-based LHDs (He and Tang (2013); Zhou and Tang (2019); Shi and Tang (2020); Wang, Yang and Liu (2021)). However, such designs can only be used when all the factors are continuous or quantitative. In some applications, qualitative factors are inevitable by nature, and play a crucial role in the study of complex systems (Rawlinson et al. (2006); Long and Bartel (2006); Joseph and Delaney (2007); Qian, Wu and Wu (2008); Hung, Joseph and Melkote (2009); Han et al. (2009); Zhou, Qian and Zhou (2011); Huang et al. (2016)). Consequently, we require designs for computer experiments that include both qualitative and quantitative factors.

The sliced Latin hypercube design (SLHD) introduced by Qian (2012) is an LHD that can be divided into several slices, each of which constitutes a smaller LHD. SLHDs maintain the maximum one-dimensional stratification for the whole design, as well as for each slice. The first systematic approach to accommodate both qualitative and quantitative factors in computer experiments uses an SLHD for the quantitative factors and a (fractional) factorial design for the qualitative factors. Here each slice for the quantitative factors corresponds to a level combination of the qualitative factors. It is evident that the run sizes of the SLHDs grow rapidly as the number of level combinations of the qualitative factors increases. That is, an SLHD may be suitable for situations in which the number of level combinations of the qualitative factors is relatively small, or the experiment is not expensive to run. Inspired by this, Deng, Hung and Lin (2015) proposed marginally coupled designs (MCDs), where the design points for the quantitative factors form an SLHD with respect to any qualitative factor. For the construction of MCDs, refer to Deng, Hung and Lin (2015), He, Lin and Sun (2017); He et al. (2017), He, Lin and Sun (2019), and Zhou, Yang and Liu (2021).

MCDs select input settings that have the desirable stratification between each qualitative factor and all quantitative factors. However, some MCDs may have poor design properties between multiple qualitative factors and all quantitative factors. Intuitively, such design properties are important when studying the interaction effects between multiple qualitative factors and quantitative factors, thus possibly affecting the accuracy of an emulator for the underlying computer simulator. Suppose that an experiment comprises three qualitative factors, namely the kind of raw materials (say, M1, M2, and M3), the shape of the raw materials (e.g., thick, medium, and thin), and the type of catalysts (C1, C2, and C3), as well as other quantitative factors. Here, it is sensible to adopt a design in which, for each kind, shape or catalyst, the associated design for the quantitative factors has a desirable space-filling property. Furthermore it would be more desirable if for each level combination of any two qualitative factors, such as (M1, thick), the corresponding design points for the quantitative factors enjoy the appealing space-filling property, which can help us to understand the effect between any two qualitative factors and the quantitative factors. In this study, we focus on designs with the appealing stratification properties between every two qualitative factors and all quantitative factors, along with all the features of MCDs. We call such designs *doubly coupled designs* (DCDs).

As in an MCD, a DCD uses an LHD for the quantitative factors. In addition, this LHD not only satisfies the constraint that for each level of any qualitative factor, the corresponding design points for the quantitative factors form an LHD, but also the constraint that for each level combination of any two qualitative factors, the corresponding design points for the quantitative factors form an LHD. In other words, for a DCD, with respect to each qualitative factor, the design for the quantitative factors is an SLHD, and with respect to any two qualitative factors, the design for the quantitative factor is also an SLHD. The concept of DCDs sounds straightforward. However, the construction procedure for DCDs is not trivial, and cannot be achieved using simple extensions of the construction for MCDs.

The rest of this paper is organized as follows. Section 2 presents the notations and definitions of the relevant designs. The theoretical results of the existence for the proposed designs are discussed in Section 3. Section 4 provides three constructions for DCDs. The last section concludes the paper. All proofs are given in the online Supplementary Material.

2. Notation and Definitions

An $n \times m$ matrix, of which the jth column has s_j levels $\{0, 1, \ldots, s_j - 1\}$, is an orthogonal array of n rows, m factors, and strength t if each of all possible level combinations occurs with the same frequency in any $n \times t$ submatrix. Such an array is denoted by $\mathrm{OA}(n, m, s_1 \cdots s_m, t)$. If some s_i are equal, we denote it by $\mathrm{OA}(n, m, s_1^{u_1} \cdots s_l^{u_l}, t)$, where $\sum_i^l u_i = m$. Furthermore, if all s_i are identical, we denote it by $\mathrm{OA}(n, m, s, t)$. An $\mathrm{OA}(n, m, s, 2)$ is called a completely resolvable orthogonal array, denoted by $\mathrm{CROA}(n, m, s, 2)$, if its rows can be divided into n/s subarrays, such that each is an $\mathrm{OA}(s, m, s, 1)$.

A Latin hypercube of n rows and m factors, denoted by $\mathrm{LH}(n, m)$, is an $n \times m$ matrix, each column of which is a permutation of the n equally spaced levels, say $\{0, 1, \ldots, n-1\}$. Given a Latin hypercube $L = (l_{ij})$, a random LHD $D = (d_{ij})$ can be generated by $d_{ij} = (l_{ij} + u_{ij})/n$, where u_{ij} is a random number from (0,1). An LHD possesses the property that each of the n equally spaced intervals has exactly one design point. A random LHD may not be space filling in two or more dimensional projections. Orthogonal array-based Latin hypercubes, introduced by Owen (1992) and Tang (1993), resolve this issue, and guarantee the same grid stratification in low-dimensional projections as that of the original orthogonal array. We review the construction method here. Assume an $\mathrm{OA}(n, m, s, t)$ exists. For each column of the orthogonal array, replace the n/s positions of level i by a random permutation of $\{i(n/s), i(n/s)+1, \ldots, (i+1)(n/s)-1\}$, for $i = 0, 1, \ldots, s-1$. The resulting design is an $\mathrm{LH}(n, m)$. Throughout this paper, we refer to this method as the level-expansion method. Conversely, an array can be obtained by replacing $\{i(n/s), i(n/s)+1, \ldots, (i+1)(n/s)-1\}$ with the integer i, for $i = 0, \ldots, s-1$, which we refer to as the level-collapsion method.

Let $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ be the n-run designs for q qualitative factors and p quantitative factors, respectively, and denote $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$. A design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ is called a *marginally coupled design* if $\boldsymbol{D}_2$ is a Latin hypercube and the rows in $\boldsymbol{D}_2$ corresponding to each level of each factor in $\boldsymbol{D}_1$ form an LHD.

MCDs possess the appealing stratification property between each qualitative and all quantitative factors. We extend the concept of MCDs, and introduce a general notion, namely *ω-way coupled designs*, which have a stronger stratification property between the two types of factors.

Definition 1. An n-run design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ with q s-level qualitative factors and p quantitative factors is called a ω-way coupled design if it satisfies the following: (i) $\boldsymbol{D}_2$ is an $\mathrm{LH}(n, p)$; and (ii) the rows in $\boldsymbol{D}_2$ corresponding to each level combination of any l factors in $\boldsymbol{D}_1$ form an LHD, for $l = 1, \ldots, \omega$.

Clearly, a ω-way coupled design is also an l-way coupled design for any $l < \omega$. In addition, a one-way coupled design is exactly an MCD. In this study, we focus on a two-way coupled design called a DCD. We denote such a design by $\mathrm{DCD}(n, s^q, p)$. We concentrate on DCDs in which $\boldsymbol{D}_1$ is $\mathrm{OA}(n, q, s, 2)$.

Table 1. Design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ in Example 1.

$\boldsymbol{D}_1^T$	0	1	0	1	0	1	0	1
	0	1	1	0	0	1	1	0
$\boldsymbol{D}_2^T$	1	0	6	7	4	5	3	2
	0	4	2	6	5	1	7	3
	0	4	6	2	5	1	3	7
	1	0	2	3	4	5	6	7

Example 1 provides a DCD and its visualization.

Example 1. Consider the design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ in Table 1. Let $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ be the two qualitative factors, and $\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3$, and $\boldsymbol{d}_4$ be the four quantitative factors.

Figures 1(a), (b), and (c) display the design points for the first two quantitative factors $\boldsymbol{d}_1$ versus $\boldsymbol{d}_2$ with respect to the level combinations of $(\boldsymbol{z}_1, \boldsymbol{z}_2)$, the levels of $\boldsymbol{z}_1$, and the levels of $\boldsymbol{z}_2$, respectively. From Figure 1(a), it is apparent that all eight points form an LHD, whereas the points of $\boldsymbol{D}_2$ corresponding to each of the four level combinations of $(\boldsymbol{z}_1, \boldsymbol{z}_2)$ are LHDs with two levels. Figures 1(b) and 1(c) reveal that the points in $\boldsymbol{D}_2$ corresponding to each level of $\boldsymbol{z}_1$ or $\boldsymbol{z}_2$ form an LHD. The plots for other quantitative dimensions are similar, so we omit them to save space. From Definition 1, this is a DCD$(8, 2^2, 4)$. Clearly, this DCD has a better stratification property between the qualitative and quantitative factors than that of an MCD, because the design points for the quantitative factors in an MCD may not enjoy the maximum one-dimensional projection uniformity with respect to each level combination of any two qualitative factors, as shown in Figure 1(a).

3. Existence of DCDs

This section focuses on the properties of DCDs, and establishes the existence of a DCD(n, s^q, p) that quantifies all the characteristics of the sub-designs $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ in a DCD.

For ease of expression, we introduce some additional notations. For $\boldsymbol{D}_1 = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_q)$ and $\boldsymbol{D}_2 = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p)$ in a DCD, we define $\widetilde{\boldsymbol{D}}_2$ and $\widetilde{\widetilde{\boldsymbol{D}}}_2$ as

$$\widetilde{\boldsymbol{D}}_2 = \left\lfloor \frac{\boldsymbol{D}_2}{s} \right\rfloor = (\tilde{\boldsymbol{d}}_1, \ldots, \tilde{\boldsymbol{d}}_p) \text{ and } \widetilde{\widetilde{\boldsymbol{D}}}_2 = \left\lfloor \frac{\widetilde{\boldsymbol{D}}_2}{s} \right\rfloor = (\tilde{\tilde{\boldsymbol{d}}}_1, \ldots, \tilde{\tilde{\boldsymbol{d}}}_p), \tag{3.1}$$

where $\lfloor a \rfloor$ represents the largest integer not exceeding a. Because $\boldsymbol{D}_2$ is an LH(n, p), we have that $\widetilde{\boldsymbol{D}}_2$ is an OA$(n, p, n/s, 1)$ and $\widetilde{\widetilde{\boldsymbol{D}}}_2$ is an OA$(n, p, n/s^2, 1)$.

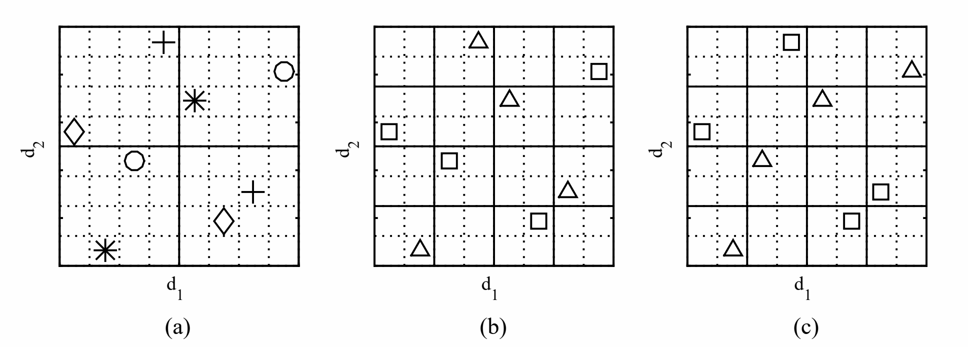

Figure 1. Scatterplots of $\boldsymbol{d}_1$ versus $\boldsymbol{d}_2$ in Example 1: (a) the points represented by $*, +, \circ$, and $\Diamond$ correspond to the level combinations (0, 0), (0, 1), (1, 0), and (1, 1) of the factors $(\boldsymbol{z}_1, \boldsymbol{z}_2)$; (b) the points marked by $\triangle$ and $\square$ correspond to the levels 0 and 1 of $\boldsymbol{z}_1$; (c) the points represented by $\triangle$ and $\square$ correspond to the levels 0 and 1 of $\boldsymbol{z}_2$.

Conversely, $\boldsymbol{D}_2$ can be obtained from $\widetilde{\boldsymbol{D}}_2$ using the level-expansion method.

Theorem 1 provides the necessary and sufficient conditions on both $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ to ensure that a DCD exists.

Theorem 1. *Suppose* $\boldsymbol{D}_1 = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_q)$ *is an* $OA(n, q, s, 2)$ *and* $\boldsymbol{D}_2 = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p)$ *is an* $LH(n, p)$. *The design* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ *is a* $DCD(n, s^q, p)$ *if and only if*

(a) $(\boldsymbol{z}_i, \tilde{\boldsymbol{d}}_k)$ *is an* $OA(n, 2, s\,(n/s)\,, 2)$, *for any* $1 \le i \le q, 1 \le k \le p$; *and*

(b) $(\boldsymbol{z}_i, \boldsymbol{z}_j, \tilde{\tilde{\boldsymbol{d}}}_k)$ *is an* $OA(n, 3, s^2\left(n/s^2\right), 3)$, *for any* $1 \le i \ne j \le q, 1 \le k \le p$.

Condition (a) of Theorem 1 is the necessary and sufficient condition for $(\boldsymbol{D}_1, \boldsymbol{D}_2)$ to be an MCD; see He, Lin and Sun (2017). Condition (b) states that for an MCD to be a DCD, $(\boldsymbol{z}_i, \boldsymbol{z}_j, \tilde{\tilde{\boldsymbol{d}}}_k)$ must be a full factorial design.

In addition, note that Conditions (a) and (b) are independent; that is, if a design satisfies Condition (a), it may not meet Condition (b), and vice versa. We give two designs to illustrate this point. Let $\boldsymbol{D}^{(a)} = (\boldsymbol{D}_1, \boldsymbol{D}_2^{(a)})$ and $\boldsymbol{D}^{(b)} = (\boldsymbol{D}_1, \boldsymbol{D}_2^{(b)})$, where $\boldsymbol{D}_1$ is from Table 1, $\boldsymbol{D}_2^{(a)} = ((1, 0, 6, 7, 3, 2, 4, 5)^T, (0, 4, 2, 6, 5, 1, 7, 3)^T)$, and $\boldsymbol{D}_2^{(b)} = \left((6, 0, 1, 4, 3, 5, 7, 2)^T,\ (2, 4, 0, 5, 7, 1, 6, 3)^T\right)$. Clearly, $\boldsymbol{D}^{(a)}$ meets Condition (a), but not (b), whereas $\boldsymbol{D}^{(b)}$ satisfies Condition (b), but not (a).

Remark 1. In Theorem 1, Condition (a) indicates that $(\boldsymbol{D}_1, \tilde{\boldsymbol{d}}_k)$ is an $\mathrm{OA}(n, q + 1, s^q\,(n/s)\,, 2)$. In addition, Condition (b) implies that $(\boldsymbol{D}_1, \tilde{\tilde{\boldsymbol{d}}}_k)$ is an $\mathrm{OA}(n, q + 1, s^q(n/s^2), 2)$.

We now revisit Example 1 to apply Theorem 1.

Table 2. The $\boldsymbol{D}_1$, $\widetilde{\boldsymbol{D}}_2$ and $\widetilde{\widetilde{\boldsymbol{D}}}_2$ in Example 2.

$\boldsymbol{D}_1$		$\widetilde{\boldsymbol{D}}_2$				$\widetilde{\widetilde{\boldsymbol{D}}}_2$			
0	0	0	0	0	0	0	0	0	0
1	1	0	2	2	0	0	1	1	0
0	1	3	1	3	1	1	0	1	0
1	0	3	3	1	1	1	1	0	0
0	0	2	2	2	2	1	1	1	1
1	1	2	0	0	2	1	0	0	1
0	1	1	3	1	3	0	1	0	1
1	0	1	1	3	3	0	0	1	1

Example 2. (Example 1 continued) For the given $\boldsymbol{D}_2$, we can obtain $\widetilde{\boldsymbol{D}}_2$ and $\widetilde{\widetilde{\boldsymbol{D}}}_2$ using (3.1). We display these two designs and $\boldsymbol{D}_1$ in Table 2. Here $(\boldsymbol{z}_i, \tilde{\boldsymbol{d}}_k)$ is an OA$(8, 2, 2 \times 4, 2)$ and $(\boldsymbol{z}_i, \boldsymbol{z}_j, \tilde{\tilde{\boldsymbol{d}}}_k)$ is an OA$(8, 3, 2, 3)$, for any $1 \le i \ne j \le 2$ and $1 \le k \le 4$. According to Theorem 1, the design $\boldsymbol{D}$ in Example 1 is a DCD$(8, 2^2, 4)$.

Theorem 1 establishes the existence of DCDs in terms of the relationship between the individual columns in $\boldsymbol{D}_1$ and $\tilde{\boldsymbol{d}}_k$, and the relationship between any pair of columns in $\boldsymbol{D}_1$ and $\tilde{\tilde{\boldsymbol{d}}}_k$. Interestingly, we can also give the existence of DCDs in terms of the design property of the entire design $\boldsymbol{D}_1$, which shows the required structure of $\boldsymbol{D}_1$ in a DCD. The precise result is presented in Theorem 2.

Theorem 2. *A DCD(n, s^q, p) exists if and only if $\boldsymbol{D}_1$ can be partitioned into n/s^2 CROA$(s^2, q, s, 2)$.*

Theorem 2 presents the requirement on $\boldsymbol{D}_1$ in a DCD. In the construction of a DCD, the $\boldsymbol{D}_1$ required by Theorem 2 is the cornerstone. For the given design parameters, only when the expected $\boldsymbol{D}_1$ exists can we construct the corresponding $\boldsymbol{D}_2$ such that $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ is a DCD.

As an example of Theorem 2, see $\boldsymbol{D}_1$ in Table 1. The first four rows and the last four rows of $\boldsymbol{D}_1$ are CROA$(4, 2, 2, 2)$. The sufficiency of the proof provides a procedure for constructing $\boldsymbol{D}_2$. The detailed process is shown in Construction 1 of Section 4.

Theorem 3 states the existence of a DCD in terms of the relationship between the columns of $\boldsymbol{D}_1$ and the columns of two relevant arrays, denoted as $\boldsymbol{B}$ and $\boldsymbol{C}$, respectively.

Theorem 3. *Suppose $\boldsymbol{D}_1$ is an OA$(n, q, s, 2)$ and $\boldsymbol{D}_2$ is an LH(n, p). The design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ is a DCD(n, s^q, p) if and only if there exist two arrays, $\boldsymbol{B} =$*

$OA(n,p,n/s^2,1)$ and $\boldsymbol{C} = OA(n,p,s,1)$, such that for any $1 \le i \ne j \le q$ and $1 \le k \le p$, both $(\boldsymbol{z}_i, \boldsymbol{z}_j, \boldsymbol{b}_k)$ and $(\boldsymbol{z}_i, \boldsymbol{c}_k, \boldsymbol{b}_k)$ are $OA(n,3,s^2\left(n/s^2\right),3)$, where $\boldsymbol{z}_i$ is the ith column of $\boldsymbol{D}_1$, $\boldsymbol{b}_k$ and $\boldsymbol{c}_k$ are the kth columns of $\boldsymbol{B}$ and $\boldsymbol{C}$, respectively, and $\widetilde{\boldsymbol{D}}_2$ in (3.1) can be written as $\widetilde{\boldsymbol{D}}_2 = s\boldsymbol{B} + \boldsymbol{C}$.

Remark 2. The condition $\widetilde{\boldsymbol{D}}_2 = s\boldsymbol{B} + \boldsymbol{C}$ in Theorem 3 implies $\widetilde{\widetilde{\boldsymbol{D}}}_2 = \boldsymbol{B}$, which further implies that the space-filling property of $\boldsymbol{D}_2$ relies heavily on that of $\boldsymbol{B}$, and is slightly affected by $\boldsymbol{C}$. If $\boldsymbol{B}$ has a better stratification property, so does $\boldsymbol{D}_2$. For example, if $\boldsymbol{B}$ is an $\mathrm{OA}(n,p,n/s^2,2)$ instead of an $\mathrm{OA}(n,p,n/s^2,1)$, $\boldsymbol{D}_2$ achieves stratifications on $(n/s^2) \times (n/s^2)$ grids for any two quantitative factors.

Theorems 1, 2, and 3 provide the necessary and sufficient conditions for a DCD to exist. These conditions are essentially the same, but are described in different ways for different purposes and usages. Theorem 3 reveals that to construct a DCD, we need to find the $\boldsymbol{D}_1$, $\boldsymbol{B}$, and $\boldsymbol{C}$ that satisfy the conditions. The next section provides three ways of providing such $\boldsymbol{D}_1$, $\boldsymbol{B}$, and $\boldsymbol{C}$.

Before we move on to the construction of DCDs, we consider a theoretically and practically important topic in the study of DCDs, that is, the maximum number of s-level qualitative factors that an n-run DCD can entertain. The following corollary gives the upper bound of the qualitative factors in a DCD.

Corollary 1. *If a DCD with $\boldsymbol{D}_1$ being an $OA(n,q,s,2)$ exists, then $q \leqslant s$.*

The proof of Corollary 1 is straightforward by Theorem 2 and Lemma 1 of Deng, Hung and Lin (2015), and is thus omitted. This corollary shows that the number of qualitative factors in a DCD cannot exceed s. Although the result seems restrictive, it is still practical. There are applications in the literature in which the number of qualitative factors is no more than the number of qualitative levels. For example, Phadke (1989) considered a router bit experiment with two qualitative four-level factors and seven quantitative factors. Moreover, when s is a prime power, there always exists a $\mathrm{CROA}(s^2,s,s,2)$ by deleting one column from the saturated $\mathrm{OA}(s^2,s+1,s,2)$. By stacking n/s^2 such CROAs to obtain the desired $\boldsymbol{D}_1$ in Theorem 2, the number of the qualitative factors of the resulting design $\boldsymbol{D}_1$ reaches the upper bound, s.

4. Construction of DCDs

When constructing DCDs, the computational search approach is often infeasible. This section presents three constructions for generating various DCDs. The two methods in Subsection 4.1 construct DCDs by using the permutation approach, which can accommodate a large number of quantitative factors. Subsection 4.2 provides DCDs with the guaranteed projection space-filling properties on the quantitative factors, whereas the number of quantitative factors in DCDs may be relatively limited. The constructions use orthogonal arrays of strength two or three, which are readily available in, for example, Hedayat, Sloane and Stufken (1999) and the design catalogues in Sloane (2014).

4.1. Constructions of design $\boldsymbol{D}_2$ using permutations

In this subsection, we give two procedures based on permutations to construct DCDs with a large number of quantitative factors.

Let $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_\lambda$ be OA$(s^2, q+1, s, 2)$. Without loss of generality, assume the last column of every $\boldsymbol{A}_i$ is $(\mathbf{0}_s^T, \mathbf{1}_s^T, \ldots, (\mathbf{s}-\mathbf{1})_s^T)^T$, where $\boldsymbol{y}_s$ represents a column vector of length s, with all elements being y. The λ orthogonal arrays are used to generate $\boldsymbol{D}_1$ in this subsection.

One construction procedure of $\boldsymbol{D}_1$, $\boldsymbol{B}$, and $\boldsymbol{C}$ in Theorem 3 works as follows, using the idea of the proof of Theorem 2.

Construction 1.

Step 1. *Obtain the array* $\boldsymbol{D}_1$ *by deleting the last column of* $(\boldsymbol{A}_1^T, \ldots, \boldsymbol{A}_\lambda^T)^T$.

Step 2. *Let* $\boldsymbol{b}_k = \boldsymbol{v}_k \otimes \mathbf{1}_{s^2}$, *where* $\boldsymbol{v}_k$ *is a random permutation of* $(0, 1, \ldots, \lambda-1)^T$, *for* $1 \le k \le p$, *and denote* $\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p)$.

Step 3. *Let* $\boldsymbol{c}_k = ((\boldsymbol{w}_{k1} \otimes \mathbf{1}_s)^T, \ldots, (\boldsymbol{w}_{k\lambda} \otimes \mathbf{1}_s)^T)^T$, *where* $\boldsymbol{w}_{kj}$ *is a random permutation of* $(0, 1, \ldots, s-1)^T$, *for* $1 \le k \le p$ *and* $1 \le j \le \lambda$, *and denote* $\boldsymbol{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$.

Step 4. *Let* $\widetilde{\boldsymbol{D}}_2 = s\boldsymbol{B} + \boldsymbol{C}$, *and obtain* $\boldsymbol{D}_2 = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p)$ *from* $\widetilde{\boldsymbol{D}}_2$ *using the level-expansion method. Denote* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$.

Proposition 1. *The design* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ *generated by Construction* 1 *is a DCD*$(\lambda s^2, s^q, p)$.

The proof is straightforward, and thus we omit it. In Construction 1, Step 1 constructs $\boldsymbol{D}_1$ to meet the requirement in Theorem 2. The orthogonal arrays $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_\lambda$ are OA$(s^2, q+1, s, 2)$. Note that $\boldsymbol{A}_i$ can be either the same or different (isomorphic or non-isomorphic), respectively. However, using different $\boldsymbol{A}_i$ is more desirable for generating $\boldsymbol{D}_1$ with higher strength. The proposed procedure produces $(s!)^{\lambda s} \cdot (s!)^\lambda \cdot \lambda!$ different quantitative columns in Proposition 1; that is, Construction 1 provides DCDs with a considerable number of quantitative factors. For $2 \leq s \leq 11$ and a positive integer λ, Construction 1 can produce the DCDs with λs^2 runs, q qualitative factors, and p quantitative factors, where $q = s$ for a prime power s, $q = 2,\ 3$ for $s = 6,\ 10$, respectively, and $p \leq (s!)^{\lambda s} \cdot (s!)^{\lambda} \cdot \lambda!$. Details are given in Table 1 of the online Supplementary Material.

Example 3 illustrates Construction 1. To save space, we set $p = 3$.

Example 3. Suppose we aim to construct a DCD$(27, 3^3, 3)$; that is, $s = 3, \lambda = 3, q = 3$, and $p = 3$. We use the following three OA$(9, 4, 3, 2)$ below:

$$\boldsymbol{A}_1 = \left[\begin{array}{cccc} 0&0&0&0 \\ 1&1&2&0 \\ 2&2&1&0 \\ \hdashline 0&2&2&1 \\ 1&0&1&1 \\ 2&1&0&1 \\ \hdashline 0&1&1&2 \\ 1&2&0&2 \\ 2&0&2&2 \end{array}\right], \quad \boldsymbol{A}_2 = \left[\begin{array}{cccc} 0&0&1&0 \\ 1&1&0&0 \\ 2&2&2&0 \\ \hdashline 0&2&0&1 \\ 1&0&2&1 \\ 2&1&1&1 \\ \hdashline 0&1&2&2 \\ 1&2&1&2 \\ 2&0&0&2 \end{array}\right], \quad \boldsymbol{A}_3 = \left[\begin{array}{cccc} 0&0&2&0 \\ 1&1&1&0 \\ 2&2&0&0 \\ \hdashline 0&2&1&1 \\ 1&0&0&1 \\ 2&1&2&1 \\ \hdashline 0&1&0&2 \\ 1&2&2&2 \\ 2&0&1&2 \end{array}\right]. \tag{4.1}$$

In Step 1, stack $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$ by row, and delete the last column of the resulting design to obtain $\boldsymbol{D}_1$, which is an OA$(27, 3, 3, 2)$, and can be divided into three CROA$(9, 3, 3, 2)$. In Steps 2 and 3, let $\boldsymbol{v}_1 = (1, 2, 0)^T, \boldsymbol{v}_2 = (0, 2, 1)^T, \boldsymbol{v}_3 = (1, 0, 2)^T$, $\boldsymbol{w}_{11} = (0, 1, 2)^T$, $\boldsymbol{w}_{12} = (1, 0, 2)^T$, $\boldsymbol{w}_{13} = (0, 2, 1)^T$, $\boldsymbol{w}_{21} = (1, 2, 0)^T$, $\boldsymbol{w}_{22} = (1, 0, 2)^T$, $\boldsymbol{w}_{23} = (0, 1, 2)^T$, $\boldsymbol{w}_{31} = (2, 0, 1)^T$, $\boldsymbol{w}_{32} = (0, 1, 2)^T$, and $\boldsymbol{w}_{33} = (1, 0, 2)^T$. Then by Construction 1, we obtain

$$\boldsymbol{B} = \begin{pmatrix} 1&1&1&1&1&1&1&1&1&2&2&2&2&2&2&2&2&2&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&2&2&2&2&2&2&2&2&2&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&2&2&2&2&2&2&2&2&2 \end{pmatrix}^T$$

and

$$\boldsymbol{C} = \begin{pmatrix} 0&0&0&1&1&1&2&2&2&1&1&1&0&0&0&2&2&2&0&0&0&2&2&2&1&1&1 \\ 1&1&1&2&2&2&0&0&0&1&1&1&0&0&0&2&2&2&0&0&0&1&1&1&2&2&2 \\ 2&2&2&0&0&0&1&1&1&0&0&0&1&1&1&2&2&2&1&1&1&0&0&0&2&2&2 \end{pmatrix}^T.$$

In Step 4, let $\widetilde{\boldsymbol{D}}_2 = 3\boldsymbol{B} + \boldsymbol{C}$ and obtain $\boldsymbol{D}_2$. According to Proposition 1, the resulting design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ is a DCD$(27, 3^3, 3)$, shown in Table 3. Moreover, the generated $\boldsymbol{D}_1$ is of strength three. The number of qualitative factors in this example achieves the upper bound in Corollary 1.

Table 3. $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ in Example 3.

	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
$\boldsymbol{D}_1^T$	0	1	2	2	0	1	1	2	0	0	1	2	2	0	1	1	2	0	0	1	2	2	0	1	1	2	0
	0	2	1	2	1	0	1	0	2	1	0	2	0	2	1	2	1	0	2	1	0	1	0	2	0	2	1
	9	10	11	13	14	12	15	16	17	22	23	21	19	18	20	24	25	26	2	0	1	7	8	6	4	5	3
$\boldsymbol{D}_2^T$	3	5	4	6	7	8	0	1	2	21	22	23	19	20	18	26	24	25	11	10	9	13	14	12	16	15	17
	16	17	15	10	11	9	12	13	14	1	2	0	4	5	3	8	7	6	21	22	23	19	20	18	24	25	26

Remark 3. If an $\mathrm{OA}(\lambda s^2, q, s, t)$ with $t \geqslant 3$ can be partitioned into λ $\mathrm{CROA}(s^2, q, s, 2)$, a DCD with $\boldsymbol{D}_1$ of strength $t \geqslant 3$ can be constructed, such as the $\boldsymbol{D}_1$ in Example 3.

We now introduce the second method for constructing the required arrays $\boldsymbol{D}_1$, $\boldsymbol{B}$, and $\boldsymbol{C}$ in Theorem 3, based on the permutation method.

Construction 2.

Step 1. *Obtain the array* $\boldsymbol{D}_1$ *by deleting the last column of* $\mathbf{1}_\lambda \otimes \boldsymbol{A}_1$.

Step 2. *Let* $\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p)$, *where* $\{b_{i,k}, b_{i+s^2,k}, \ldots, b_{i+(\lambda-1)s^2,k}\}$ *is a random permutation of* $\{0, 1 \ldots, \lambda - 1\}$ *and* $b_{i,k}$ *is the ith entry of* $\boldsymbol{b}_k$, *for* $1 \leq i \leq s^2$ *and* $1 \leq k \leq p$.

Step 3. *Let* $\boldsymbol{c}_k = \mathbf{1}_\lambda \otimes (\boldsymbol{w}_k \otimes \mathbf{1}_s)$, *where* $\boldsymbol{w}_k$ *is a random permutation of* $(0, 1, \ldots, s-1)^T$, *for* $1 \leq k \leq p$, *and denote* $\boldsymbol{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$.

Step 4. *Let* $\widetilde{\boldsymbol{D}}_2 = s\boldsymbol{B} + \boldsymbol{C}$, *and obtain* $\boldsymbol{D}_2 = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p)$ *from* $\widetilde{\boldsymbol{D}}_2$ *using the level-expansion method. Denote* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$.

Proposition 2. *The design* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ *produced by Construction* 2 *is a* $DCD(\lambda s^2, s^q, p)$.

In Proposition 2, DCDs with at most $(s!)^{\lambda s} \cdot s! \cdot (\lambda!)^{s^2}$ distinct quantitative columns can be generated, which indicates that Construction 2 can also construct DCDs containing a large number of quantitative factors.

Example 4 below provides an illustration of Construction 2.

Example 4. Generate a $\mathrm{DCD}(27, 3^3, 3)$ and choose $\boldsymbol{A}_1$ shown in Example 3. In Step 1, delete the last column of $\mathbf{1}_\lambda \otimes \boldsymbol{A}_1$ to obtain $\boldsymbol{D}_1$. In Step 2, let

$$\boldsymbol{B} = \begin{pmatrix} 2&0&1&2&0&1&2&0&1&0&1&2&0&1&2&0&1&2&1&2&0&1&2&0&1&2&0 \\ 2&1&0&0&2&1&1&0&2&0&2&1&1&0&2&2&1&0&1&0&2&2&1&0&0&2&1 \\ 0&2&1&2&1&0&1&0&2&1&0&2&0&2&1&2&1&0&2&1&0&1&0&2&0&2&1 \end{pmatrix}^T.$$

Table 4. $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ in Example 4.

	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
$\boldsymbol{D}_1^T$	0	1	2	2	0	1	1	2	0	0	1	2	2	0	1	1	2	0	0	1	2	2	0	1	1	2	0
	0	2	1	2	1	0	1	0	2	0	2	1	2	1	0	1	0	2	0	2	1	2	1	0	1	0	2
	19	1	9	22	3	13	25	8	17	0	11	18	5	14	21	6	16	24	10	20	2	12	23	4	15	26	7
$\boldsymbol{D}_2^T$	23	12	4	7	26	17	10	1	19	3	22	14	16	8	25	18	9	0	13	5	21	24	15	6	2	20	11
	8	26	17	20	10	2	13	5	21	16	7	25	1	19	11	22	14	3	24	15	6	9	0	18	4	23	12

One can easily check that $\{b_{i,k}, b_{i+9,k}, b_{i+18,k}\}$ is a permutation of $\{0, 1, 2\}$, for $1 \leq i \leq 9$ and $1 \leq k \leq 3$. In Step 3, let $\boldsymbol{w}_1 = (0, 1, 2)^T, \boldsymbol{w}_2 = (1, 2, 0)^T$, and $\boldsymbol{w}_3 = (2, 0, 1)^T$, yielding

$$\boldsymbol{C} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}^T.$$

The obtained design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ is a DCD, as shown in Table 4.

In practice, for a predetermined p, an optimal $\boldsymbol{D}_2$ according to some optimization criteria (e.g., the maximin distance, uniform discrepancies, etc.) can be found by ranking all possible candidate designs or by using the greedy search algorithms, such as the simulated annealing or threshold accepting algorithms, if the number of candidate designs is exceedingly large (Morris and Mitchell (1995); Ba, Myers and Brenneman (2015)).

4.2. Constructions for a better space-filling property on quantitative factors

This subsection provides a construction method that uses one array we call $\boldsymbol{A}$ to provide $\boldsymbol{D}_1$ and $\boldsymbol{C}$ required in Theorem 3. That is, the new construction involves only two arrays, $\boldsymbol{A}$ and $\boldsymbol{B}$. Two specific cases of the construction are provided to produce the required $\boldsymbol{A}$ and $\boldsymbol{B}$, where the resulting DCDs may share some extra high-dimensional space-filling properties between the quantitative factors. Suppose $\boldsymbol{A}$ is an $\text{OA}(n, q+1, s, 2)$ and $\boldsymbol{B}$ is an $\text{OA}(n, p, n/s^2, 1)$. Construction 3 works as follows.

Construction 3.

Step 1. *Randomly choose q columns from $\boldsymbol{A}$ to obtain $\boldsymbol{D}_1$. Denote the remaining column of $\boldsymbol{A}$ by $\boldsymbol{a}^*$.*

Step 2. *Let $\boldsymbol{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$, where $\boldsymbol{c}_k$ is obtained by permuting the levels of $\boldsymbol{a}^*$,* *for any* $1 \le k \le p$.

Step 3. *Let* $\widetilde{\boldsymbol{D}}_2 = s\boldsymbol{B} + \boldsymbol{C}$, *and obtain* $\boldsymbol{D}_2 = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p)$ *from* $\widetilde{\boldsymbol{D}}_2$ *using the level-expansion method. Denote* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$.

Theorem 4. *Suppose* $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{q+1})$ *is an* $OA(n, q+1, s, 2)$ *and* $\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p)$ *is an* $OA(n, p, n/s^2, 1)$. *If* $(\boldsymbol{a}_i, \boldsymbol{a}_j, \boldsymbol{b}_k)$ *is an* $OA(n, 3, s^2\left(n/s^2\right), 3)$, *for any* $1 \le i \ne j \le q+1$ *and* $1 \le k \le p$, *then* $\boldsymbol{D}$ *obtained using Construction* 3 *is a* $DCD(n, s^q, p)$.

Theorem 4 tells us that to construct a DCD(n, s^q, p), the most important task is to find the two required arrays, $\boldsymbol{A}$ and $\boldsymbol{B}$. Under the condition of Theorem 4, it can be verified that the three arrays $\boldsymbol{D}_1$, $\boldsymbol{B}$, and $\boldsymbol{C}$ in Construction 3 meet the conditions in Theorem 3. Hence, the obtained $\boldsymbol{D}$ of Construction 3 is a DCD. As such Theorem 4 can be viewed as a special case of Theorem 3.

Next, we present an example in which we apply Construction 3.

Example 5. Suppose we want to construct a DCD$(8, 2^2, 4)$. Let

$$\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) = \begin{pmatrix} 0\,0\,0 \\ 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \\ 0\,0\,0 \\ 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \end{pmatrix} \text{ and } \boldsymbol{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4) = \begin{pmatrix} 0\,0\,0\,0 \\ 0\,1\,1\,0 \\ 1\,0\,1\,0 \\ 1\,1\,0\,0 \\ 1\,1\,1\,1 \\ 1\,0\,0\,1 \\ 0\,1\,0\,1 \\ 0\,0\,1\,1 \end{pmatrix}.$$

It can be checked that $\boldsymbol{A}$ is an OA$(8, 3, 2, 2)$, $\boldsymbol{B}$ is an OA$(8, 4, 2, 2)$, and $(\boldsymbol{a}_i, \boldsymbol{a}_j, \boldsymbol{b}_k)$ is an OA$(8, 3, 2, 3)$, for $1 \le i \ne j \le 3, 1 \le k \le 4$. Thus, $\boldsymbol{A}$ and $\boldsymbol{B}$ satisfy the requirements in Theorem 4. In Step 1, select $\boldsymbol{a}_2, \boldsymbol{a}_3$ of $\boldsymbol{A}$ to be $\boldsymbol{D}_1$, and denote $\boldsymbol{a}^* = \boldsymbol{a}_1$. In Step 2, each column of $\boldsymbol{C}$ is generated by $\boldsymbol{a}_1$ using the level permutation. Without loss of generality, let $\boldsymbol{c}_k = \boldsymbol{a}_1$, for $1 \le k \le 4$. In Step 3, we obtain the corresponding $\widetilde{\boldsymbol{D}}_2$ and apply the level-expansion method. The resulting design $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ is shown in Example 1. In addition, $\boldsymbol{B}$ is an orthogonal array of strength two. Therefore, the resulting $\boldsymbol{D}_2$ achieves stratifications on 2×2 grids for any two quantitative factors, which we verify in Figure 1.

We now present two cases to generate the required arrays $\boldsymbol{A}$ and $\boldsymbol{B}$ in Theorem 4, which can produce DCDs with s^3 and s^u runs, respectively, for $u \geqslant 3$.

Furthermore, Case 1 is suitable for any $s \geqslant 2$, whereas Case 2 works for any prime power s. In both cases, the resulting DCDs enjoy extra two-dimensional space-filling properties between the quantitative factors.

Case 1. *Let* $\boldsymbol{G}$ *be an* $OA(s^3, m, s, 3)$. *Split the columns of* $\boldsymbol{G}$ *randomly into two arrays,* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*, where* $\boldsymbol{A}$ *has* $q+1$ *columns and* $\boldsymbol{B}$ *has* p *columns, and* $m = p + q + 1$.

The orthogonal array $\boldsymbol{G}$ of strength three in Case 1 can be taken directly from, for example, Sloane (2014).

Proposition 3. *The* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ *constructed using Construction* 3 *by using* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *in Case* 1 *is a DCD*(s^3, s^q, p)*, where* $q + p = m - 1$. *Furthermore, we have*

(a) $\boldsymbol{D}_1$ *is an* $OA(s^3, q, s, t)$*, where* $t = q$ *if* $q < 3$*, and* $t = 3$ *if* $q \geqslant 3$*;*

(b) $(\tilde{\boldsymbol{d}}_k, \tilde{\boldsymbol{d}}_{k'})$*, for any* $1 \leq k \neq k' \leq p$*, achieves a stratification on* $s^2 \times s$ *and* $s \times s^2$ *grids; and*

(c) $\tilde{\tilde{\boldsymbol{D}}}_2$ *is an* $OA(s^3, p, s, t)$*, where* $t = p$ *if* $p < 3$*, and* $t = 3$ *if* $p \geqslant 3$.

Parts (b) and (c) in Proposition 3 mean that $\boldsymbol{D}_2$ enjoys the two-dimensional and three-dimensional space-filling properties. For $s \leq 10$, the sum of the number of qualitative and quantitative factors of the DCDs produced by Case 1 of Construction 3 is no more than three for $s = 2, 3, 6, 10$, five for $s = 4, 5$, seven for $s = 7$, and nine for $s = 8, 9$. Table 2 of the online Supplementary Material shows the details.

We now introduce Case 2, which is based on regular fractional factorial designs (Wu and Hamada (2009)). For any prime power s and any integer $u \geq 3$, let $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_u$ be independent columns of length s^u with entries from $GF(s)$, the Galois field of order s.

Case 2.

Step 1. *Let*

$$\begin{aligned} \boldsymbol{A} &= \{\boldsymbol{\xi}_1 + \mu_2 \boldsymbol{\xi}_2 \mid \mu_2 \in GF(s)\} \cup \{\boldsymbol{\xi}_2\} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{s+1}), \\ \boldsymbol{R}_v &= \{\boldsymbol{\xi}_1 + \mu_2 \boldsymbol{\xi}_2 + \mu_{v+2} \boldsymbol{\xi}_{v+2} \mid \mu_2 \in GF(s), \mu_{v+2} \in GF(s) \backslash \{0\}\} \\ &\quad \cup \{\boldsymbol{\xi}_2 + \mu_{v+2} \boldsymbol{\xi}_{v+2} \mid \mu_{v+2} \in GF(s) \backslash \{0\}\} \\ &\quad \cup \{\boldsymbol{\xi}_{v+2}\} = (\boldsymbol{r}_{v,1}, \ldots, \boldsymbol{r}_{v,s^2}), \end{aligned} \tag{4.2}$$

where $\boldsymbol{r}_{v,f}$ *is a column vector of length* s^u*, for* $1 \leq v \leq u-2$ *and* $1 \leq f \leq s^2$.

Step 2. For any $1 \leq f \leq s^2$, let

$$\boldsymbol{B}_f = (\boldsymbol{r}_{1,f}, \ldots, \boldsymbol{r}_{u-2,f})\boldsymbol{T}, \tag{4.3}$$

where

$$\boldsymbol{T} = \begin{pmatrix} s^{u-3} & 1 & \cdots & s^{u-5} & s^{u-4} \\ s^{u-4} & s^{u-3} & \cdots & s^{u-6} & s^{u-5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s & s^2 & \cdots & s^{u-3} & 1 \\ 1 & s & \cdots & s^{u-4} & s^{u-3} \end{pmatrix} = (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{u-2}).$$

There are $u-2$ columns in each $\boldsymbol{B}_f$.

Step 3. Let $\boldsymbol{B} = (\boldsymbol{B}_1, \ldots, \boldsymbol{B}_{s^2}) = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{(u-2)s^2})$.

Clearly, $\boldsymbol{A}$ consists of the independent columns $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ and all possible interactions of these two columns, and thus $\boldsymbol{A}$ has $s+1$ columns. The column vectors in $\boldsymbol{R}_v$ must involve $\boldsymbol{\xi}_{v+2}$ and may contain $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$. Lemma 1 summarizes the design properties of $\boldsymbol{A}, \boldsymbol{R}_1, \ldots, \boldsymbol{R}_{u-2}$. The proof is straightforward, and is thus omitted.

Lemma 1. *For $\boldsymbol{A}, \boldsymbol{R}_1, \ldots, \boldsymbol{R}_{u-2}$ in Case 2, we have,*

(a) $\boldsymbol{A}$ is an $OA(s^u, s+1, s, 2)$;

(b) $(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_{u-2})$ is an $OA(s^u, (u-2)s^2, s, 2)$;

(c) $(\boldsymbol{a}_i, \boldsymbol{a}_j, \boldsymbol{r}_{1,f}, \ldots, \boldsymbol{r}_{u-2,f})$ is an $OA(s^u, u, s, u)$, for any $1 \leq f \leq s^2$ and $1 \leq i \neq j \leq s+1$; and

(d) $(\boldsymbol{r}_{1,f}, \ldots, \boldsymbol{r}_{u-2,f}, \boldsymbol{r}_{v,l})$ is an $OA(s^u, u-1, s, u-1)$, for any $1 \leq v \leq u-2$ and $1 \leq f \neq l \leq s^2$.

Lemma 1(c) means that taking two distinct columns from $\boldsymbol{A}$ and one column from each $\boldsymbol{R}_v$, for $v = 1, \ldots, u-2$, the resulting u columns form an s-level orthogonal array of s^u runs and strength u, that is, a full factorial design of s levels and u columns. Similarly, Lemma 1(d) implies that the array of $u-1$ columns, consisting of two distinct columns of $\boldsymbol{R}_v$ and one column of each of the remaining $u-3$ arrays $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_{v-1}, \boldsymbol{R}_{v+1}, \ldots, \boldsymbol{R}_{u-2}$, is an orthogonal array of strength $u-1$, that is, a full factorial design of s levels and $u-1$ columns.

From Lemma 1 and Case 2, we have the following result.

Lemma 2. *For $\boldsymbol{A}$ and $\boldsymbol{B}$ in Case 2, we have* *(a)* $(\boldsymbol{a}_i, \boldsymbol{a}_j, \boldsymbol{b}_k)$ *is an* $OA(s^u, 3, s^2(s^{u-2}), 3)$*, for any* $1 \leq i \neq j \leq s+1$ *and* $1 \leq k \leq (u-2)s^2$*; and*

(b) $\boldsymbol{A}$ *is an* $\mathrm{OA}(s^u, s+1, s, 2)$ *and* $\boldsymbol{B}$ *is an* $\mathrm{OA}(s^u, (u-2)s^2, s^{u-2}, 1)$*.*

Lemma 2 states that $\boldsymbol{A}$ and $\boldsymbol{B}$ in Case 2 are the required arrays in Theorem 4. For $s = 2$ and $u = 3$, the two arrays $\boldsymbol{A}$ and $\boldsymbol{B}$ in Case 2 are shown in Example 5.

Proposition 4. *For any prime power* s *and any integer* $u \geqslant 3$*,* $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$ *constructed from Construction* 3 *by using* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *in Case* 2 *is a* $DCD(s^u, s^s, (u-2)s^2)$*, with* $\boldsymbol{D}_1$ *being an* $OA(s^u, s, s, 2)$ *and* $\boldsymbol{D}_2$ *being an* $LH(s^u, (u-2)s^2)$*. In addition,* $\boldsymbol{D}_2$ *has the following properties:*

(a) if $\lfloor (i-1)/(u-2) \rfloor = \lfloor (i'-1)/(u-2) \rfloor$*,* $\tilde{\tilde{\boldsymbol{d}}}_i$ *and* $\tilde{\tilde{\boldsymbol{d}}}_{i'}$ *achieve* $s \times s$ *grids stratification; and*

(b) if $\lfloor (i-1)/(u-2) \rfloor \neq \lfloor (i'-1)/(u-2) \rfloor$*,* $\tilde{\tilde{\boldsymbol{d}}}_i$ *and* $\tilde{\tilde{\boldsymbol{d}}}_{i'}$ *achieve* $s^{u-2} \times s$ *and* $s \times s^{u-2}$ *grids stratification.*

Obviously, the number of qualitative factors for the DCDs in Proposition 4 is s, which reaches the upper bound in Corollary 1, and the number of quantitative factors is $(u-2)s^2$.

5. Conclusion

In this paper, we have proposed ω-way coupled designs, with $\omega \geqslant 2$, for computer experiments involving both qualitative and quantitative factors. We focus on the properties and constructions of the two-way coupled designs, which we call DCDs. Similarly to MCDs, such designs are an economical alternative to SLHDs. In contract to MCDs, they require that for each level combination of every two qualitative factors, the corresponding design points for the quantitative factors must form an LHD. This additional requirement leads to the result that given the same run size, DCDs can accommodate fewer qualitative factors than MCDs can. In addition, DCDs are equipped with better stratification properties between the qualitative and quantitative factors than those of MCDs.

When the design for the qualitative factors $\boldsymbol{D}_1$ is an $\mathrm{OA}(n, q, s, 2)$, we provide the necessary and sufficient conditions for the existence of a DCD, and give a tight upper bound for the number of qualitative factors. Three construction methods are provided, which are different, but related. In particular, Constructions 1 and 2 are both based on the idea of permutations, but they generate $\boldsymbol{D}_1, \boldsymbol{B}$, and $\boldsymbol{C}$ in different ways. More specifically, Step 1 of Construction 2 uses λ identical orthogonal arrays, whereas $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_\lambda$ used in Step 1 of Construction 1 can be identical, isomorphic, or non-isomorphic. Step 2 of Construction 1 is a special case of Step 2 of Construction 2 when $\{b_{i,k}, b_{i+s^2,k}, \ldots, b_{i+(\lambda-1)s^2,k}\}$ is the same random permutation of $\{0, 1, \ldots, \lambda-1\}$, for $1 \leq i \leq s^2$. Step 3 of Construction 2 is a special case of Step 3 of Construction 1 when ω_{kj} is the same permutation of $\{0, 1, \ldots, \lambda-1\}$, for $1 \leq j \leq \lambda$. Construction 3 differs from Constructions 1 and 2 in that it uses an array $\boldsymbol{A}$ to provide $\boldsymbol{D}_1$ and $\boldsymbol{C}$. Thus, the building blocks of Construction 3 are the arrays $\boldsymbol{A}$ and $\boldsymbol{B}$ that meet the conditions in Theorem 4. Two cases of such $\boldsymbol{A}$ and $\boldsymbol{B}$ are given. Because $\boldsymbol{B}$ in Case 1 and $\boldsymbol{R}$ in Case 2 are orthogonal arrays, the $\boldsymbol{D}_2$ of the DCDs produced by Construction 3 involving Cases 1 and 2 are orthogonal array-based Latin hypercubes. Constructions 1 and 2 can accommodate a larger number of quantitative factors than Construction 3 can, but their limitation is that we cannot ensure the space-filling properties of the designs for the quantitative factors. On the other hand, the $\boldsymbol{D}_2$ constructed using Construction 3 with Case 1 and Case 2 can guarantee some desirable stratification on the grids for the quantitative factors. However, the number of quantitative factors of the resulting designs may be relatively limited. Because all the constructions are algebraic, they incur a low cost in terms of computing time. For practical use, we list examples of the DCDs provided by the proposed construction methods in the online Supplementary Material.

The needed arrays $\boldsymbol{D}_1$, $\boldsymbol{B}$, and $\boldsymbol{C}$ in Theorem 3 can be constructed using other approaches in the future. The methods used to generate the two arrays $\boldsymbol{A}$ and $\boldsymbol{B}$ required by Theorem 4 are not limited to the two cases given in this paper, and additional pairs of $\boldsymbol{A}$ and $\boldsymbol{B}$ can be considered. Furthermore, we can consider the DCD $\boldsymbol{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2)$, with $\boldsymbol{D}_1$ being a mixed-level orthogonal array of strength t or possessing some good space-filling properties. Further investigation of the space-filling property of $\boldsymbol{D}_1$ is an important topic. Zhou and Xu (2014) studied the space-filling property of orthogonal arrays under two commonly used space-filling measures, namely, the discrepancy and maximin distance. Because of the required relationship between the columns in $\boldsymbol{D}_1$ and the columns in $\boldsymbol{D}_2$, this would require additional effort to explore the theoretical space-filling property of $\boldsymbol{D}_1$ in a DCD. Another possible direction is to construct DCDs in which $\boldsymbol{D}_2$ has high-dimensional space-filling properties, such as three or four dimensions. In addition, an interesting, but challenging direction is to construct ω-way coupled designs with $\omega > 2$. The construction of such designs is not trivial, and cannot be easily extended. We hope to study this and report the results in future work.

Supplementary Material

The supplementary material gives all the proofs, and the tables of some possible DCDs produced by the proposed constructions.

Acknowledgments

Part of this research was conducted when the first author visited the Department of Mathematics and Statistics at Queen's University. Yang was supported by the China Scholarship Council, National Natural Science Foundation of China (12101435), and Research Foundation of Sichuan Normal University (pt12101435, ky20200919). Zhou was supported by the National Natural Science Foundation of China (11871288) and Natural Science Foundation of Tianjin (19JCZDJC31100). Lin was supported by the Natural Sciences and Engineering Research Council of Canada. The authors would also like to thank Professor Jianfeng Yang for his valuable suggestions on this topic.

References

Ba, S., Myers, W. R and Brenneman, W. A. (2015). Optimal sliced Latin hypercube designs. *Technometrics* **57**, 479–487.

Chen, P. H., Santner, T. J. and Dean, A. M. (2018). Sequential pareto minimization of physical systems using calibrated computer simulators. *Statistica Sinica* **28**, 671–692.

Deng, X., Hung, Y. and Lin, C. D. (2015). Design for computer experiments with qualitative and quantitative factors. *Statistica Sinica* **25**, 1567–1581.

Fang, K. T., Li, R. and Sudjianto, A. (2005). *Design and Modeling for Computer Experiments.* Chapman and Hall/CRC, New York.

Georgiou, S. D. and Efthimiou, I. (2014). Some classes of orthogonal Latin hypercube designs. *Statistica Sinica* **24**, 101–120.

Han, G., Santner, T. J., Notz, W. I. and Bartel, D. L. (2009). Prediction for computer experiments having quantitative and qualitative input variables. *Technometrics* **51**, 27–288.

He, Y., Lin, C. D. and Sun, F. (2017). On the construction of marginally coupled designs. *Statistica Sinica* **27**, 665–683.

He, Y., Lin, C. D. and Sun, F. (2019). Construction of marginally coupled designs by subspace theory. *Bernoulli* **25**, 2163–2182.

He, Y., Lin, C. D., Sun, F. and Lv, B. J. (2017). Marginally coupled designs for two-level qualitative factors. *Journal of Statistical Planning and Inference* **187**, 103–108.

He, Y. and Tang, B. (2013). Strong orthogonal arrays and associated Latin hypercubes for computer experiments. *Biometrika* **100**, 254–260.

Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Application.* Springer, New York.

Huang, H., Lin, D. K. J., Liu, M. Q. and Yang, J. (2016). Computer experiments with both qualitative and quantitative variables. *Technometrics* **58**, 495–507.

Huang, H., Yu, H., Liu, M. Q. and Wu, D.(2021). Construction of uniform designs and complex-structured uniform designs via partitionable t-designs *Statistica Sinica* **31**, 1689–1706.

Hung, Y., Joseph, V. R. and Melkote, S. N. (2009). Design and analysis of computer experiments with branching and nested factors. *Technometrics* **51**, 354–365.

Joseph, V. R. and Delaney, J. D. (2007). Functionally induced priors for the analysis of experiments. *Technometrics* **49**, 1–11.

Joseph, V. R. and Hung, Y. (2008). Orthogonal-maximin Latin hypercube designs. *Statistica Sinica* **18**, 171–186.

Li, W., Liu, M. Q and Tang, B. (2020). A method of constructing maximin distance designs. *Biometrika 108*, 845–855.

Long, J. P. and Bartel, D. L. (2006). Surgical variables affect the mechanics of a hip resurfacing system. *Clinical Orthopaedics and Related Research* **453**, 115–122.

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

Morris, M. D. and Mitchell, T. J. (1995). Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference* **43**, 381–402.

Owen, A. B. (1992). Orthogonal arrays for computer experiments, integration and visuazation. *Statistica Sinica* **2**, 439–452.

Phadke, M. S. (1989). *Quality Engineering Using Robust Design.* Prentice-Hall, Englewood Cliffs.

Qian, P. Z. G. (2012). Sliced Latin hypercube designs. *Journal of the American Statistical Association* **107**, 393–399.

Qian, P. Z. G., Wu, H. and Wu, C. F. J. (2008). Gaussian process models for computer experiments with qualitative and quantitative factors. *Technometrics* **50**, 383–396.

Rawlinson, J. J., Furman, B. D., Li, S., Wright, T. M. and Bartel, D. L. (2006). Retrieval, experimental and computational assessment of the performance of total knee replacements. *Journal of Orthopaedic Research Official Publication of the Orthopaedic Research Society* **24**, 1384–1394.

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

Shi, C. and Tang, B. (2020). Construction results for strong orthogonal arrays of strength three. *Bernoulli* **26**, 418–431.

Sloane, N. J. A. (2014). A Library of Orthogonal Arrays. Web:`http://neilsloane.com/oadir/`.

Sun, F. and Tang, B. (2017). A general rotation method for orthogonal Latin hypercubes. *Biometrika* **104**, 465–472.

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

Wang, L., Sun, F., Lin, D. K. J. and Liu, M. Q. (2018). Construction of orthogonal symmetric Latin hypercube designs. *Statistica Sinica* **28**, 1503–1520.

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**, 3741C66.

Wang, C., Yang, J. and Liu, M. Q. (2021). Construction of strong group-orthogonal arrays. *Statistica Sinica* **32**, 1225–1243

Wu, C. F. J. and Hamada, M. (2009). *Experiments: Planning, Analysis, and Optimization.* John Wiley & Sons, Inc, New Jersey.

Xiao, Q. and Xu, H. (2018). Construction of maximin distance designs via level permutation and expansion. *Statistica Sinica* **28**, 1395–1414.

Zhou, Q., Qian, P. Z. G. and Zhou, S. (2011). A simple approach to emulation for computer models with qualitative and quantitative factors. *Technometrics* **53**, 266–273.

Zhou, W., Yang, J. and Liu, M. Q. (2021). Construction of orthogonal marginally coupled designs. *Statistical Papers* **62**, 1795–1820.

Zhou, Y. and Tang, B. (2019). Column-orthogonal strong orthogonal arrays of strength two plus and three minus. *Biometrika* **106**, 997–1004.

Zhou, Y. D. and Xu, H. (2014). Space-filling fractional factorial designs. *Journal of the American Statistical Association* **109**, 1134–1144.

Feng Yang

School of Mathematical Sciences & Laurent Mathematics Center, Sichuan Normal University, Chengdu 610066, China.

E-mail: yangfeng@sicnu.edu.cn

C. Devon Lin

Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada.

E-mail: devon.lin@queensu.ca

Yongdao Zhou

School of Statistics and Data Science, Nankai University, Tianjin 300071, China.

E-mail: ydzhou@nankai.edu.cn

Yuanzhen He

School of Statistics, Beijing Normal University, Beijing 100875, China.

E-mail: heyuanzhen@bnu.edu.cn

(Received August 2020; accepted November 2021)