Cost-Efficient Mixed-Level Covering Designs for Testing Experiments

Yasmeen Akhtar[1] · Frederick Kin Hing Phoa[2]

Abstract

A covering design is a traditional class of experimental plans for hardware and software testing purposes. This paper presents a class of size-optimal covering designs for testing experiments with mixed-level factors. Among all the factors of different levels, one or two factors have a high number of levels while other factors form a full factorial so that all level combinations among factor pairs are "covered" at least once and appeared almost equally frequent. We use the coloring techniques for hypergraphs to construct such near-balanced mixed-level covering designs with the minimum run size.

Keywords Covering designs · Near-continuous factors · Hyperedge coloring · Mixed covering array on hypergraph

Mathematics Subject Classification 62K15

1 Introduction

In the regime of artificial intelligence and robotics, engineered systems consist of many different testing points or components, each with a different number of options, that work interactively in a variety of situations. Although one can test a specific level of a component in isolation, the origin of a system failure may come from

✉ Frederick Kin Hing Phoa
  fredphoa@stat.sinica.edu.tw

  Yasmeen Akhtar
  yasmeen@stat.sinica.edu.tw

[1] School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, USA

[2] Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

interactions among different components at specific option settings. A complete enumeration of testing interactions among all components at all possible options is infeasible especially when many components have a large number of options. However, [15] reported that testing all possible interactions between the pair components would detect around 70% of all faults. Thus, [16] suggested a pseudo-exhaustive approach that uses covering designs to test all possible interactions among every choice of t components for $2 \le t \le 6$. An $N \times k$ experimental plan with s levels is a *covering design of strength t* denoted as $CA(N, k, s, t)$ if all s^t level combinations for any t columns appear at least once. Some literature uses the term "t-covering" to describe similar design property, but we will use the term "strength t" for the rest of the paper. A covering design usually requires that all combinations of levels be "covered", but unlike an orthogonal array (OA), it does not require equal appearances of all level combinations. Because of the orthogonality requirement, the latter arrays may have a large run size whereas a covering design can be constructed with a smaller number of runs than an OA under the same number of factors but at the cost of poorer estimation of the factor effects.

When different components have a different number of options, we call it mixed-level covering design (MCD). The construction and analysis of mixed-level OAs are surveyed and detailed in [13, 24]. For reasons of run size economy or flexibility, nearly OAs [23, 25], OAs of weak strength t [12, 26], and almost OAs [18] are considered as potential alternatives to OAs. Algorithmic constructions for MCDs have been studied in [8–10]. By simply including a set of columns from an optimal size $CA(n, k', 2, 2)$ to a given MCD of strength two, [10] provided the construction for $MCD(2s, s \cdot 2^k, 2)$ with $k = 2^{s-1}$, $MCD(3s, s \cdot 2^k, 2)$ with $k = \frac{3s}{2}\binom{s}{s/2}$ if s is even, $MCD(4s, s \cdot 2^k, 2)$ with $k =$ the coefficient of x^s in $(4 + 6x + 4x^2)^s$, $MCD(9, 3^2 \cdot 2^{20}, 2)$, $MCD(12, 3 \cdot 4 \cdot 2^{243}, 2)$, and $MCD(16, 4^2 \cdot 2^{3,453}, 2)$. Some of these designs are optimal in size but they may not exhibit the near-balanced property. The constructions for optimal size MCD of strength two with four factors or five factors are presented in [21]. Algorithmic constructions for MCDs of strength two to cover every level combination at least twice are proposed in [2]. However, the generated designs are not optimal in size. A mixed covering array on the graph is another generalization of MCD in which not all pair of columns are required to have their level combinations covered. The construction of optimal size MCD on bipartite, cycles, and upper bounds on the size of mixed covering arrays on all 3-chromatic and on a large number of 4- and 5-chromatic graphs are provided in [20]. A further generalization to MCD on hypergraphs has been systematically studied in [1, 22]. We extend the results of [20] and [2] to construct optimal size MCD on the fan hypergraphs. The coloring techniques for constructing MCD on bipartite graphs from [20] and the single-vertex hyperedge hooking operations from [2] are generalized to a higher rank. The previous studies in this area mostly concentrate on the hypergraphs of rank 3 whereas the results in this paper are not limited to any such ranks. Following the definition of OA of weak strength from [26], the MCD obtained in Theorem 1 is an OA of weak strength t for $t = 1, 2$.

The purpose of this paper is to construct a class of near-balanced MCDs for a testing experiment that consists of a few near-continuous (or very-high-level) factors.

The term "near-continuous" implies that practitioners with continuous factors of interest in their experiments may consider discretizing these factors into factors with a very high number of levels; then, they can conduct their experiments using MCDs. In specific, given the factor levels $s_1 \leq s_2 \leq \cdots \leq s_r$ and a near-continuous factor attaining the values $\{1, 2, \ldots, h\}$, where $h \leq \prod_{i=1}^{r-1} s_i$, we prove the existence of a size-optimal MCD that is a full factorial design on its r discrete factors while covers all pairwise level combinations with the near-continuous factor almost equally often. Moreover, the near-continuous factor attains all its levels almost equally often (up to the difference of 1). In addition, if the near-continuous factor attains the values $\{1, 2, \ldots, h_1\}$, where $h_1 \leq \prod_{i=1}^{r-2} s_i$, then we can also include another factor attaining the values within $\{1, 2, \ldots, h_2\}$, where $h_2 \leq \min\{h, s_r\}$, without an increase in the number of runs such that all pairwise level combinations between the first $r + 1$ factors (including the first near-continuous factor) and the last factor are covered.

Starting with a full factorial design D with n runs, one could think of randomly searching for a vector v of length n with entries from a set of h symbols until it satisfies the following three properties.

1. The frequency of each symbol is either $\lfloor n/h \rfloor$ or $\lceil n/h \rceil$;
2. It covers all pairwise level combinations with each factor in D;
3. All pairwise level combinations between v and any factors in D are covered almost equally often.

This approach does not guarantee the existence of a vector with the required properties and exhaustive search is computationally infeasible. Therefore, we address this problem by providing a theoretic proof based on the existence of specific colorings in the r-uniform complete r-partite hypergraph.

This paper is organized as follows. Section 2 provides the formal definition of MCD and the necessary background for this paper along with the application of proposed designs in a software testing experiment. Section 3 contains our main results for the existence of size-optimal MCD for systems that consist of near-continuous factors. In Sect. 4, we derive results from the theory of hypergraph colorings to construct a size-optimal MCD and prove the theorems from Sect. 3. Some discussions and future directions are provided in the last section.

2 Definition and Background

2.1 Mixed-Level Covering Designs and Their Application

The *covering design*, also known as *covering array*, has been in focus during the past 30 years for its cost efficiency in implementing experiments for software testing, hardware testing, drug screening, and other areas where the interactions of multiple parameters are required to be tested. As the number of runs (test trials) is proportional to the logarithmic value of the number of factors, covering designs fulfill the desire of most applications in minimizing the experimental run

sizes or resources [7, 8, 11]. In this work, we consider covering designs where different factors can have a different number of levels, and we call it *mixed-level covering design* or MCD in short. Here is the formal definition of MCD and its balance property.

Definition 1 A *mixed-level covering design of strength t* $MCD(N, s_1^{k_1} s_2^{k_2} \cdots s_v^{k_v}, t)$ is an $N \times k$ matrix, where $k = k_1 + k_2 + \cdots + k_v$ is the total number of factors, in which the first k_1 columns have symbols from $\mathbb{Z}_{s_1}$, the next k_2 columns have symbols from $\mathbb{Z}_{s_2}$, and so on, with the property that in any $N \times t$ subarrays every possible t-tuple occurs at least once as a row.

When each t-tuple is covered the same number of times, then the design is called an OA. However, in a covering design, t-tuples can be covered a different number of times which makes the structure unbalanced. If the structure is highly unbalanced, then measuring the interaction effect for the estimation of variation is difficult. The equal number of observations for all combinations of factor levels is desired for variation standardization. The existence of such a design is restricted and requires the number of runs to be a multiple of the product of the number of levels for any t factors. To address this issue using several new designs like nearly OAs [25], OAs of weak strength t [12, 26], almost OAs [18] have been introduced by relaxing the requirement on an OA of strength t that all level combinations must appear equally often for any t factors. We aim to cover all tuples almost equal number of times and define the near-balance property of a design as follows.

Definition 2 A mixed-level covering design of strength t is said to be balanced if all level combinations for any t columns appear equally often. A mixed-level covering design of strength t is near-balanced if for each $1 \le t' \le t$ all level combinations for any t' columns appear as equally often as possible, that is, the difference of occurrences of level combinations does not exceed one.

A balanced covering design is essentially the same as the OA, and a near-balanced covering design of strength t is an OA of weak strength s for $s = 1, ..., t$ as defined in [26]. Unlike an OA of weak strength t defined in [12], a near-balanced covering design of strength t need not be an OA of strength $t - 1$.

Here is an example of software testing during the development of an Android application. There are a large number of configuration options for Android apps to control the behavior of the device, and these options operate across a variety of hardware and software platforms. For simplicity, we consider a scenario that an Android developer is interested in designing a test suite for an app based on the combination of five basic features: the initial status of GPS navigation (NAVIGATIONHIDDEN) during start-up, the app orientation on the device (ORIENTATION), the type of keyboard for app input (KEYBOARD), the availability of touchscreen setting for the app (TOUCHSCREEN), and a parameter UI_MODE_TYPE that describes which type of device the app is used. Table 1 lists these five parameters and their corresponding level options found in the resource configuration file for Android apps.

Table 1 Google android configuration options in the illustrative example

	Parameter name	# Options	Option values in example
X_1	NAVIGATIONHIDDEN	2	NO, YES
X_2	ORIENTATION	2	LANDSCAPE, PORTRAIT
X_3	KEYBOARD	3	12KEY, NOKEYS, QWERTY
X_4	TOUCHSCREEN	3	FINGER, NOTOUCH, STYLUS
Y	UI_MODE_TYPE	≤ 12	APPLIANCE, CAR, DESK, MASK, NORMAL, TELEVISION, UNDEFINED, VR_HEAD-SET, WATCH, ND1, ND2, ND3

In the experimental design, NAVIGATIONHIDDEN and ORIENTATION are 2-level factors X_1 and X_2, KEYBOARD and TOUCHSCREEN are 3-level factors X_3 and X_4, and Y is a near-continuous or high-level factor UI_MODE_TYPE. Note that the developer file provides only nine options, but it can be used to investigate up to 12 options, and we put ND1, ND2, ND3 to denote some new Android-based devices in the future. Since UI_MODE_TYPE has a large number of choices as compared to the other four factors, studying all level combinations of the basic features with the UI_MODE_TYPE is not cost efficient. To design an experiment for studying all combinations of basic features and the pair of each basic feature and the UI_MODE_TYPE, we need an MCD of five factors such that the first four factors form a full factorial design and the MCD is of strength 2. Figure 1 provides the design matrices of MCD$(36, 2^2 \cdot 3^2 \cdot g^1, 2)$ for $g = 9, 10, 11, 12$ obtained using Algorithm 1 from Sect. 3. These designs suggest that 36 runs are adequate to conduct an experiment that fulfills the experimental requirements for at most 12 options for UI_MODE_TYPE. It is easy to verify that the design matrix that consists of $X_1, X_2, X_3,$ and X_4 forms a full

	X_1	X_2	X_3	X_4	Y (9)	Y (10)	Y (11)	Y (12)			X_1	X_2	X_3	X_4	Y (9)	Y (10)	Y (11)	Y (12)
1	0	0	0	0	0	0	0	0		19	1	0	0	0	8	5	5	4
2	0	0	0	1	1	1	1	1		20	1	0	0	1	6	6	3	5
3	0	0	0	2	2	2	2	2		21	1	0	0	2	7	7	4	3
4	0	0	1	0	3	3	3	3		22	1	0	1	0	2	1	8	10
5	0	0	1	1	4	4	4	4		23	1	0	1	1	0	2	9	11
6	0	0	1	2	5	5	5	5		24	1	0	1	2	1	0	10	9
7	0	0	2	0	6	6	6	6		25	1	0	2	0	5	4	10	2
8	0	0	2	1	7	7	7	7		26	1	0	2	1	3	3	1	0
9	0	0	2	2	8	8	8	8		27	1	0	2	2	4	9	0	1
10	0	1	0	0	4	9	9	9		28	1	1	0	0	0	8	3	7
11	0	1	0	1	5	3	10	10		29	1	1	0	1	1	6	8	8
12	0	1	0	2	3	4	6	11		30	1	1	0	2	2	3	7	6
13	0	1	1	0	7	7	2	1		31	1	1	1	0	3	4	7	8
14	0	1	1	1	8	8	0	2		32	1	1	1	1	4	9	6	6
15	0	1	1	2	6	6	1	0		33	1	1	1	2	5	1	0	7
16	0	1	2	0	1	2	4	5		34	1	1	2	0	6	2	1	11
17	0	1	2	1	2	5	5	3		35	1	1	2	1	7	0	2	9
18	0	1	2	2	0	1	3	4		36	1	1	2	2	8	5	9	10

Fig. 1 Design matrix corresponding to a system involving one near-continuous factor

factorial design MCD($36, 2^2 \cdot 3^2, 4$). If we include the column of $Y(i)$ with $i = 9$ to the full factorial design, we have a near-balanced MCD($36, 2^2 \cdot 3^2 \cdot 9^1, 2$). Similarly, MCD($36, 2^2 \cdot 3^2 \cdot 10^1, 2$), MCD($36, 2^2 \cdot 3^2 \cdot 11^1, 2$), and MCD($36, 2^2 \cdot 3^2 \cdot 12^1, 2$) are obtained by including the columns $Y(10)$, $Y(11)$, and $Y(12)$, respectively, to the full factorial design. All three MCDs are verified as size-optimal, meaning that these MCDs use the smallest number of test trials to achieve their desired property with the indicated numbers of factors with their levels.

2.2 Some Backgrounds on the Hyperedge Coloring

Our construction of an MCD is related to hypergraph and its coloring scheme. We associate an r-uniform r-partite hypergraph $H = (V_1 \cup \cdots \cup V_r, E)$ to a given set of r factors $\{f_1, \ldots, f_r\}$ in a design by defining a class of vertices V_i for each factor f_i. Each class has a vertex for each level of the associated factor and each trial is represented as a hyperedge containing the vertices for the level combinations covered in that trial. Then, we use a coloring of the hyperedges in H to assign the levels for a new factor included in the given design. Any two trials test the same level for the new factor when the corresponding two hyperedges have the same color. A further detail on this is provided in Sect. 3. Here, we briefly introduce some terminologies and concepts in hypergraphs that are useful in this work.

A *hypergraph H* consists of a pair of sets (V, E), where V is a finite set of vertices and E is a finite family of nonempty subsets of V called hyperedges. The H is called *r-uniform* hypergraph if every hyperedge contains precisely r vertices. Let $H = (V, E)$ be a hypergraph, where $E = \{e_1, e_2, \ldots, e_m\}$. For a set $J \subset \{1, 2, ..., m\}$, the *partial hypergraph* generated by J is the hypergraph $(V', \{e_i : i \in J\})$, where $V' = \bigcup_{i \in j} e_i$. For a set $A \subset V$, the *sub-hypergraph* H_A induced by A is defined as $H_A = (A, \{e_i \cap A : 1 \le i \le m, e_i \cap A \ne \emptyset\})$. The *r-uniform complete r-partite hypergraph* $K^r_{n_1, n_2, \ldots, n_r}$ is an r-uniform hypergraph with vertex set V decomposed into r disjoint sets (called sides) $V_1, V_2, \ldots, V_r$, where $|V_i| = n_i$ for $1 \le i \le r$ such that every choice of r vertices, one from each side, is a hyperedge. The *degree* $d_H(v)$ of a vertex v is the size of the family $H(v) = \{e : e \in E \text{ and } v \in e\}$.

A *k-coloring* (weak coloring) of the hyperedges in $H = (V, E)$ is a k-partition $(S_1, S_2, \ldots, S_k)$ of E such that for every vertex v with $d_H(v) > 1$, $H(v)$ has at least two hyperedges of different colors. A *good k-coloring* of the hyperedges is a k-coloring such that for every vertex v, $H(v)$ contains the largest possible number of different colors (taking account of the value of k), namely $\min\{d_H(v), k\}$. An *equitable k-coloring* of the hyperedges is a k-coloring such that for every vertex v, all the colors appear the same number of times (or to within 1, if k does not divide $d_H(v)$) in $H(v)$, that is,

$$\left\lfloor \frac{d_H(v)}{k} \right\rfloor \le |H(v) \cap S_i| \le \left\lceil \frac{d_H(v)}{k} \right\rceil \quad i = 1, 2, \ldots, k.$$

Hence, for every k, an equitable k-coloring is always a good coloring. A k-coloring of hyperedges is *uniform* if the number of hyperedge of the same color is always the same (or to within 1, if k does not divide $|E|$), that is,

$$\left\lfloor \frac{|E|}{k} \right\rfloor \le |S_i| \le \left\lceil \frac{|E|}{k} \right\rceil \quad i = 1, 2, \ldots, k.$$

A *hyperedge cover* is a subset of E such that each vertex of the hypergraph is contained in at least one hyperedge in that subset. Thus, for every $k \le \min_{v \in V} d_H(v)$, each color class of a good k-coloring of hyperedges is a hyperedge cover of H. A k-coloring $(S_1, S_2, \ldots, S_k)$ of hyperedges defines a *factorization* of hypergraph into k factors $H^1, H^2, \ldots, H^k$, where $H^i = (V, S_i)$. A factor H^i is an *f*-factor if H^i is an *f*-regular hypergraph and hypergraph is *f-factorizable* if there exists a factorization such that each factor is an *f*-factor. A factorization of the complete uniform hypergraph is provided in [3]. For details and the description of undefined terms used in this article, we refer to [6].

2.3 Covering Design on Hypergraph

When conducting an experiment or performing a test, it may be the case that certain factors are known to not interact or the interaction between them is not important or influential. Then, there is no need to cover all level combinations between such factors. Thus, only a set of parameters that jointly affect the response must be considered for testing. The covering designs on hypergraphs provide the design matrix to perform testing for such system. A *weighted hypergraph* with a positive weight assigned to each vertex is used to represent the interacting factors and their associated number of factor levels.

Let $H = (V, E)$ be a weighted hypergraph with k vertices and weights $s_1 \le s_2 \le \cdots \le s_k$ and let N be a positive integer. A *mixed-level covering design on H*, denoted as $\mathrm{MCD}(N, H, \prod_{i=1}^{k} s_i)$, is an $N \times k$ array with the following properties:

1. column i corresponds to vertex $v_i \in V$ with weight s_i;
2. the entries in column i are from $\mathbb{Z}_{s_i}$;
3. if $e = \{v_1, v_2, \ldots, v_t\}$ is a hyperedge in E, the columns correspond to vertices $v_1, v_2, \ldots, v_t$ contain all possible ordered t-tuples at least once as a row.

The other names that have been used for mixed-level covering designs on hypergraphs are *mixed covering arrays on hypergraphs* and *variable strength mixed covering arrays*. For details, we refer to [1]. The *product weight $PW(H)$* of a weighted hypergraph is defined as

$$PW(H) = \max \left\{ \prod_{v_i \in e} w_H(v_i) \ : \ e \in E \right\},$$

where $w_H(v_i)$ is the weight of vertex v_i. An MCD on a hypergraph with size $PW(H)$ is known to be *size optimal*. A *near-balanced covering design on H* is a covering design on H with the properties:

1. all levels for any columns appear as equally often as possible, that is, the difference of occurrences of levels does not exceed one;
2. all level combinations for any pair of columns correspond to the vertices in a hyperedge appear as equally often as possible, that is, the difference of occurrences of level combinations does not exceed one.

The columns of MCD are closely related to the *qualitatively independent partitions* introduced in [19]. We generalize this definition for the mixed-level case as follows. Let k_1, k_2, and n be positive integers with $n \geq k_1 k_2$. Let $\mathcal{A}$ be a k_1-partition and $\mathcal{B}$ be a k_2-partition of an n-set. Assume $\mathcal{A} = \{A_1, A_2, \ldots, A_{k_1}\}$ and $\mathcal{B} = \{B_1, B_2, \ldots, B_{k_2}\}$. The partitions $\mathcal{A}$ and $\mathcal{B}$ are *qualitatively independent* if $A_i \cap B_j \neq \varnothing$ for all i and j. It is easy to observe that for each pair of factors in an MCD with N runs, the two partitions of an N-set defined by the observations corresponding to the same factor level are qualitatively independent partitions.

3 Existence of Mixed-Level Covering Designs

In this section, we state two theorems on the existence of size-optimal MCDs and provide an algorithm to construct such designs.

Theorem 1 *Let $s_1 \leq s_2 \leq \cdots \leq s_r$ be positive integers. Then, for any $s \leq \prod_{i=1}^{r-1} s_i$, there exists a size-optimal near-balanced $\mathrm{MCD}(\prod_{i=1}^{r} s_i, s_1 s_2 \cdots s_r s, 2)$ such that the projection on the first r factors is a full factorial design.*

The proof of Theorem 1 is given in the next section. This theorem states the existence of a size-optimal MCD with $r+1$ factors, where the first r factors form a full factorial design and when any factor from the first r factors is paired with the last factor, all level combinations between these two factors covered at least once and an almost equal number of times. Moreover, first r factors are uniform, and the last factor is almost uniform on the respective level set. Given a set of r factors and factor levels, we associate an r-uniform complete r-partite hypergraph $(V_1 \cup \ldots \cup V_r, E)$ to it by defining a vertex $v_x^i \in V_i$ for each level x of the ith factor and adding a hyperedge $\{v_x^1, v_y^2 \ldots, v_z^r\}$ for every factor-level combination $(x, y, \ldots, z)$. Following the algorithm outlines the construction of a near-balanced mixed-level covering design described in Theorem 1.

Algorithm 1 Near-balanced MCD via hyperedge coloring

Input: The number of factors r and factor levels $s_1 \leq \ldots \leq s_r$, and a positive integer $s \leq \prod_{i=1}^{r-1} s_i$.
Step 1. Construct the r-uniform complete r-partite hypergraph $K_{s_1,\ldots,s_r}^r$ associated with the given number of factors and factor levels.
Step 2. Find an equitable, uniform s-coloring $(F_1, \ldots, F_s)$ of hyperedges in $K_{s_1,\ldots,s_r}^r$.
Step 3. Construct an MCD with $r+1$ factors and $s_1, \ldots, s_r, s$ levels and $\prod_{i=1}^r s_i$ runs by adding a run $(x, y, \ldots, z, i)$ for each hyperedge $\{v_x^1, v_y^2, \ldots, v_z^r\}$ in F_i for $i = 1, \ldots, s$.
Output: A near-balanced MCD with $r+1$ factors and $s_1, \ldots, s_r, s$ levels.

Example 1 Consider the testing of an Android app described in Sect. 2. There are five basic features and let the feature UI_MODE_TYPE needs to be investigated for ten choices. To study all combinations of five basic features, we start with the construction of $K_{2,2,3,3}^4$ in Step 1. The hyperedges in $K_{2,2,3,3}^4$ represent the runs in a full factorial design for four factors with levels 2, 2, 3, and 3, respectively. Figure 2 shows an equitable and uniform 10-coloring $(F_0, F_1, \ldots, F_9)$ of the hyperedges in $K_{2,2,3,3}^4$. Every hyperedge in $K_{2,2,3,3}^4$ has four vertices and it is represented as a line joining these vertices. For example, the hyperedge $\{v_0^1, v_0^2, v_0^3, v_0^4\}$ is represented as a dotted line joining the vertices v_0^1, v_0^2, v_0^3, and v_0^4. Each color class F_i can be seen as a factor of $K_{2,2,3,3}^4$. Next, for each hyperedge $\{v_w^1, v_x^2, v_y^3, v_z^4\}$ in F_i, we include a run (w, x, y, z, i) to construct a MCD($36, 2^2 \cdot 3^2 \cdot 10, 2$). For example, the hyperedges $\{v_0^1, v_0^2, v_0^3, v_1^4\}, \{v_0^1, v_1^2, v_2^3, v_2^4\}, \{v_1^1, v_0^2, v_0^3, v_0^4\}$, and $\{v_1^1, v_1^2, v_1^3, v_2^4\}$ all have red color, so we include $(0, 0, 0, 1, 1), (0, 1, 2, 2, 1), (1, 0, 1, 0, 1)$, and $(1, 1, 1, 2, 1)$ as rows in the design matrix. The resulting design has factors X_1, X_2, X_3, X_4, and $Y(10)$, where the first four factors form a full factorial design as shown in Fig. 1. Since the coloring is uniform, each F_i has an almost equal number of hyperedges (either 3 or 4). Therefore, each level in $Y(10)$ appears either 3 or 4 times. Moreover, being an equitable coloring, each vertex is contained in an almost equal number of hyperedges (either 1 or 2) in each F_i and hence each level combination between X_i and $Y(10)$ is covered either once or twice. Thus, the design shown in Fig. 1 is a near-balanced design.

Derived from Theorem 1, we state the existence of an optimal size, near-balanced MCD on a specific hypergraph structure defined below.

Definition 3 The *fan* of rank r is a hypergraph F_r having r edges of cardinality two and one hyperedge of cardinality r, arranged as in Fig. 3.

Corollary 1 *Let F_r be a weighted fan of rank r shown in Fig. 3 and $s_1 \leq s_2 \leq \cdots \leq s_r$ be the weights of vertices v_i, where $i = 1, 2, \ldots, r$, and s be the weight of u. If $s \leq \prod_{i=1}^{r-1} s_i$, then there exists a size-optimal near-balanced MCD on F_r.*

Consider the design D described in Theorem 1. We consider the associations between the ith column of D and the vertex v_i in F_r for $1 \leq i \leq r$ and between the last column of D and the vertex u in F_r. Since the first r columns of D form a full
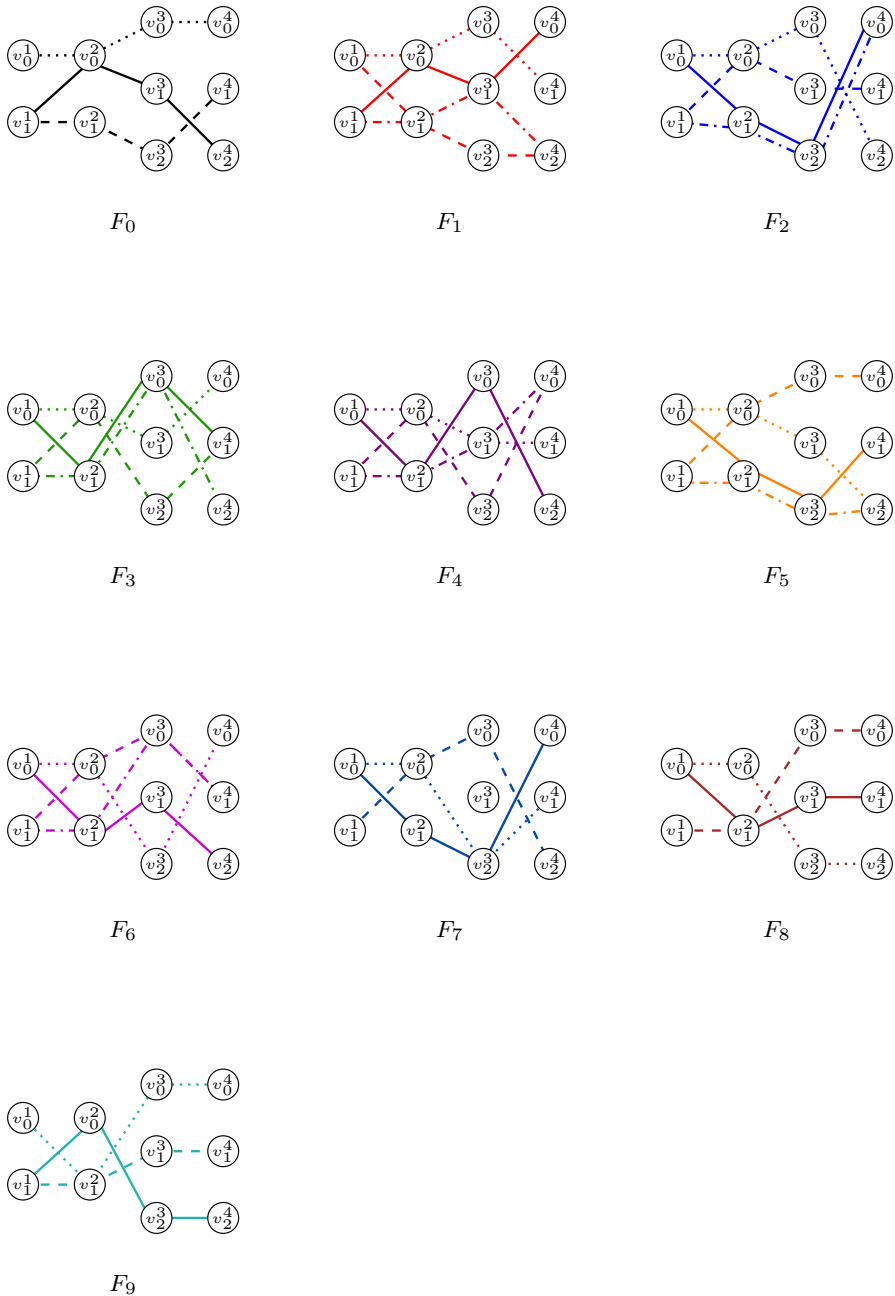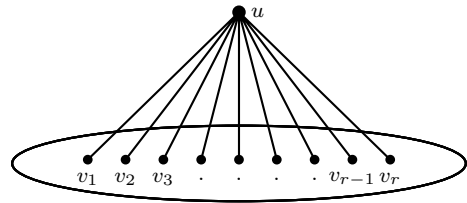
Fig. 2 A factorization of $K_{2,2,3,3}^4$ obtained via an equitable and uniform 10-coloring of the hyperedges in $K_{2,2,3,3}^4$

Fig. 3 The fan of rank r : F_r

factorial design, it is a covering design of strength r on the hyperedge of cardinality r. Being an MCD($\prod_{i=1}^{r} s_i, s_1 s_2 \cdots s_r s, 2$), all possible ordered tuples for each pair of columns appear at least once as a row in the respective sub-design of D. Since the projection on any $t \leq 2$ factors is near balanced, D is a near-balanced MCD on F_r. Moreover, the equality of the product weight of F_r and $\prod_{i=1}^{r} s_i$ implies that the array D is a size-optimal design.

Theorem 2 *Let $s_1 \leq s_2 \leq \cdots \leq s_r$ be positive integers. Then, for any $s \leq \prod_{i=1}^{r-2} s_i$ and $s' \leq \min\{s, s_r\}$, there exists a size-optimal MCD($\prod_{i=1}^{r} s_i, s_1 s_2 \cdots s_r ss', 2$) such that the projection on first r factors is a full factorial design.*

The proof of Theorem 2 is given in the next section. Similar to Theorem 1, this theorem states that the existence of an MCD with $r + 2$ factors of which the first r factors form a full factorial design. When any factor from the first r factors is paired with any one factor of the last two factors, all level combinations between these two factors are covered at least once. Moreover, all level combinations between the last two factors also exist at least once. Derived from Theorem 2, we state the existence of an optimal MCD on another specific fan-type hypergraph structure in the following corollary.

Corollary 2 *Let H be a weighted hypergraph shown in Fig. 4 and $s_1 \leq s_2 \leq \cdots \leq s_r$ be the weights of the vertices $v_1, v_2, \ldots, v_r$, respectively. Let s, s' be the weights of u and w, respectively. If $s \leq \prod_{i=1}^{r-2} s_i$ and $s' \leq \min\{s, s_r\}$, then there exists an optimal MCD on H.*

Consider the design D described in Theorem 2. The ith column in D corresponds to the vertex v_i for $1 \leq i \leq r$, and the last two columns in D correspond to the vertices u and w, respectively. Then, D forms an MCD on H. Since $PW(H) = \prod_{i=1}^{r} s_i$, D is size-optimal.

Fig. 4 A weighted hypergraph H
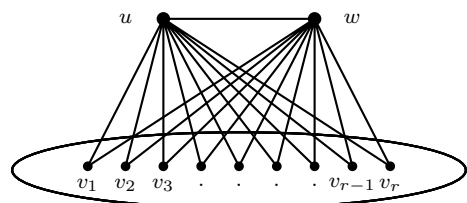
The following is an example of an optimal size MCD($24, 2^3 \cdot 3^2 \cdot 4, 2$) obtained using the construction in Theorem 2 such that the projection on the first four factors is a full factorial design (Fig. 5).

4 Proofs

Some additional lemmas are required from graph theory to prove two main theorems, and we summarize them in the first subsection below.

4.1 Coloring the Hyperedges in $K^r_{n_1,n_2,\ldots,n_r}$

The uniform and equitable k-coloring of the hyperedges in the r-uniform complete r-partite hypergraph $K^r_{n_1,n_2,\ldots,n_r}$ has been studied in [4, 5]. The following theorem for such hypergraphs where multiple hyperedges are admissible is proved in [4].

Lemma 1 *For every $k \geq 2$, the hyperedges of the r-uniform complete r-partite hypergraph $K^r_{n_1,n_2,\ldots,n_r}$ admit an equitable k-coloring that is uniform.*

In this section, we study the good coloring of hyperedges in the $K^r_{n_1,n_2,\ldots,n_r}$ and prove the existence of two distinct good colorings that are also qualitatively independent partitions. When $n_i(\geq 2)$ are pairwise coprime, we prove the existence of an equitable k-coloring of hyperedges, where k is coprime with each n_i, that is also uniform, and the corresponding factors exhibit a complete bipartite graph between each pair of sides in the induced sub-hypergraph.

	v_1 (2)	v_2 (2)	v_3 (2)	v_4 (3)	u (4)	w (3)
1	0	0	0	0	0	0
2	0	0	0	1	1	0
3	0	0	0	2	2	0
4	0	0	1	0	1	1
5	0	0	1	1	2	1
6	0	0	1	2	3	2
7	0	1	0	0	2	2
8	0	1	0	1	3	1
9	0	1	0	2	0	1
10	0	1	1	0	3	0
11	0	1	1	1	0	2
12	0	1	1	2	1	2

	v_1 (2)	v_2 (2)	v_3 (2)	v_4 (3)	u (4)	w (3)
13	1	0	0	0	3	0
14	1	0	0	1	0	2
15	1	0	0	2	1	2
16	1	0	1	0	2	2
17	1	0	1	1	3	1
18	1	0	1	2	0	1
19	1	1	0	0	1	1
20	1	1	0	1	2	1
21	1	1	0	2	3	2
22	1	1	1	0	0	0
23	1	1	1	1	1	0
24	1	1	1	2	2	0

Fig. 5 A size-optimal MCD($24, 2^3 \cdot 3^2 \cdot 4, 2$) on a fan-type hypergraph

Lemma 2 *Let $n_1 \leq n_2 \leq \cdots \leq n_r$ and $h \leq \prod_{i=1}^{r-2} n_i$ be positive integers. Then, for any positive integers $k \leq \min\{h, n_r\}$, the hyperedges of the r-uniform complete r-partite hypergraph $H = K_{n_1,n_2,\ldots,n_r}^r$ admit a good k-coloring $\mathcal{A} = \{A_0, A_1, \ldots, A_{k-1}\}$ and a good h-coloring $\mathcal{B} = \{B_0, B_1, \ldots, B_{h-1}\}$ such that $\mathcal{A}$ and $\mathcal{B}$ are qualitatively independent partitions.*

Proof Let $v_0^r, v_1^r, \ldots, v_{n_r-1}^r$ be the vertices in the side V_r of H. Let G be the partial hypergraph of H, generated by $E \setminus \cup_{i=0}^{k-1} H(v_i^r)$. It is easy to observe that if G is non-empty then it is the r-uniform complete r-partite hypergraph $K_{n_1,n_2,\ldots,(n_r-k)}^r$. Using Lemma 1, the hyperedges in G admit an equitable k-coloring $\mathcal{S} = \{S_0, S_1, \ldots, S_{k-1}\}$ and an equitable h-coloring $\mathcal{T} = \{T_0, T_1, \ldots, T_{h-1}\}$ that are also uniform.

Now, let G_i be the partial hypergraph of H generated by $H(v_i^r)$, where $i = 0, 1, \ldots, k-1$. As each G_i is the r-uniform complete r-partite hypergraph $K_{n_1,n_2,\ldots,n_{r-1},1}^r$ using Lemma 1, there exists an equitable h-coloring $\mathcal{P} = \{P_0, P_1, \ldots, P_{h-1}\}$ of the hyperedges in G_i. Since for each vertex v, $d_{G_i}(v) \geq \prod_{i=1}^{r-2} n_i$, the equitable coloring implies that each color class is a hyperedge cover of G_i.

Now, define $R_{ij} = P_{(i+j \bmod k)}$, where $0 \leq i \neq j \leq k-1$, and $R_{ii} = E \setminus \underset{0 \leq i \neq j \leq k-1}{\cup} R_{ij}$. Let $\mathcal{R}_i = \{R_{i0}, R_{i1}, \ldots, R_{i(k-1)}\}$ be the k-coloring of hyperedges in G_i, where $0 \leq i \leq k-1$. As $k \leq h$, for each vertex v, each color class R_j in this k-coloring contains at least one hyperedge that contains v and hence it is a good k-coloring. Take $A_j = S_j \cup (\cup_{i=0}^{k-1} R_{ij})$, where $j = 0, 1, \ldots, k-1$. Then, $\mathcal{A} = \{A_0, A_1, \ldots, A_{k-1}\}$ is a k-coloring of the hyperedges in H. Since $\mathcal{S}$ is an equitable k-coloring and each $\mathcal{R}_i$ is a good k-coloring, the coloring $\mathcal{A}$ is also a good k-coloring.

Now, take $B_j = T_j \cup P_j$, where $j = 0, 1, \ldots, h-1$. Then, $\mathcal{B} = \{B_0, B_1, \ldots, B_{h-1}\}$ is an h-coloring of the hyperedges in H, and since both $\mathcal{T}$ and $\mathcal{P}$ are equitable h-colorings; the coloring $\mathcal{B}$ is a good h-coloring of H. Moreover, for each pair of color classes A_j and B_l, where $A_j \in \mathcal{A}$ and $B_l \in \mathcal{B}$

$$A_j \cap B_l \supseteq (\cup_{i=0}^{k-1} R_{ij}) \cap P_l$$
$$\supseteq (\cup_{i=0}^{h-1} P_i) \cap P_l$$
$$\supseteq P_l.$$

Thus, $\mathcal{A}$ and $\mathcal{B}$ are qualitatively independent partitions. $\square$

Lemma 3 *Let $n_1 < n_2 < \ldots < n_r$ and $k < n_1 n_2 \cdots n_{r-2}$ be positive integers that are pairwise coprime. Then, the hyperedges of the r-uniform complete r-partite hypergraph $H = K_{n_1,n_2,\ldots,n_r}^r$ admit an equitable k-coloring that is uniform and in each corresponding factor H^l, where $1 \leq l \leq k$; for every pair of sides (V_i, V_j), where $1 \leq i < j \leq r$, the induced sub-hypergraph $H_{V_i \cup V_j}^l$ is a complete bipartite graph.*

Proof Let $v_0^i, v_1^i, \ldots, v_{n_i-1}^i$ be the vertices in the side V_i, where $1 \leq i \leq r$, and E be the set of hyperedges in H. For $N = \prod_{i=1}^r n_i$, we define the following bijective map from $\mathbb{Z}_N$ to the set of hyperedges in H.

$$\varphi : \mathbb{Z}_N \to E$$
$$s \mapsto e = \{v_{(s \mod n_1)}^1, v_{(s \mod n_2)}^2, \ldots, v_{(s \mod n_r)}^r\}$$

Since H does not have multiple hyperedges, the map φ is well defined and the bijective property follows from the Chinese Remainder Theorem. For $0 \leq s \leq N - 1$, we label the hyperedge e as e_s, where $s = \varphi^{-1}(e)$. For $l = 0, 1, \ldots, k-1$, let $S_l = \{e_j : j \equiv l \mod k\}$ be the subset of E. Then, $(S_0, S_1, \ldots, S_{k-1})$ defines a partition of E. To prove that this partition is a coloring of hyperedges in H, we show that for each vertex v in H, $H(v) \not\subset S_l$, where $l = 0, \ldots, k-1$. Consider the hyperedges e_j and e_{j+n_i} in $H(v_j^i)$. Since n_i and k are coprime, $j \not\equiv (j + n_i) \mod k$, the hyperedges e_j and e_{j+n_i} belong to two different color classes. Thus, $H(v_j^i)$ has at least two hyperedges of different colors. The uniform k-coloring follows from the fact that $|S_l| = \lceil \frac{N}{k} \rceil$, where $0 \leq l \leq N - 1 - k\lfloor \frac{N}{k} \rfloor$, and $|S_l| = \lfloor \frac{N}{k} \rfloor$, where $N - k\lfloor \frac{N}{k} \rfloor \leq l \leq k - 1$. For an arbitrary vertex v_j^i and any S_l, using the Chinese Remainder Theorem the pair of congruences $x \equiv j \mod n_i$ and $x \equiv l \mod k$ has a unique solution modulo $n_i k$. Thus, $|H(v_j^i) \cap S_l|$ is either $\lfloor \frac{N}{n_i k} \rfloor$ or $\lceil \frac{N}{n_i k} \rceil$. Hence, it is an equitable k-coloring of E. Moreover, for any pair of vertices $v_a^i \in V_i$ and $v_b^j \in V_j$, where $i \neq j$ and any S_l, again using the Chinese Remainder Theorem, the system of congruences $x \equiv a \mod n_i, x \equiv b \mod n_j$, and $x \equiv l \mod k$ has a unique solution modulo $n_i n_j k$. Thus, $|H(v_a^i) \cap H(v_b^j) \cap S_l|$ is either $\lfloor \frac{N}{n_i n_j k} \rfloor$ or $\lceil \frac{N}{n_i n_j k} \rceil$ that is at least 1. Hence, the induced sub-hypergraph $H_{V_i \cup V_j}^l$ is a complete bipartite graph. $\square$

4.2 Proof of Theorem 1

Proof To construct a covering design MCD($\prod_{i=1}^r s_i, s_1 s_2 \cdots s_r s, 2$), we consider the r-uniform complete r-partite hypergraph $H = K_{s_1, s_2, \ldots, s_r}^r = (V_1 \cup V_2 \cup \cdots \cup V_r, E)$. Let $v_0^i, v_1^i, \ldots, v_{s_i-1}^i$ be the vertices in the side V_i, where $1 \leq i \leq r$. There are $N = \prod_{i=1}^r s_i$ hyperedges in E. Using Lemma 1, there exists an equitable s-coloring $(S_0, S_1, \ldots, S_{s-1})$ of the hyperedges in E that is also uniform. Since $s \leq \prod_{i=1}^{r-1} s_i$, for each vertex $v \in V_i$ and each color class S_k,

$$1 \leq \left\lfloor \frac{\prod_{j=1, j \neq i}^r s_j}{s} \right\rfloor \leq |H(v) \cap S_k| \leq \left\lceil \frac{\prod_{j=1, j \neq i}^r s_j}{s} \right\rceil. \tag{1}$$

For each hyperedge $e_j = \{v_{a_1}^1, v_{a_2}^2, \ldots, v_{a_r}^r\} \in S_k$, where $0 \leq k \leq s-1$, $0 \leq a_i \leq s_i - 1$, and $i = 1, 2, \ldots, r$, we include the factor-level combination $r_j = (a_1, a_2, \ldots, a_r, k)$ as a row to construct an $N \times (r+1)$ array D. Since H is the r-uniform complete r-partite hypergraph, for every level combination $(a_1, a_2, \ldots, a_r) \in \mathbb{Z}_{s_1} \times \mathbb{Z}_{s_2} \times \cdots \times \mathbb{Z}_{s_r}$, there exists a hyperedge $\{v_{a_1}^1, v_{a_2}^2, \ldots, v_{a_r}^r\}$. Therefore, the first r columns in D form a full factorial design.

Next, for each pair $a \in \mathbb{Z}_{s_i}$ and $b \in \mathbb{Z}_s$ using Eq. 1, there exists at least one hyperedge e_j in S_b that contains the vertex v_a^i. Then, $D(j, i) = a$ and $D(j, r+1) = b$. Thus, the projection on the factors i and $r+1$ is a full factorial design and D is a covering design MCD($\prod_{i=1}^r s_i, s_1 s_2 \cdots s_r s, 2$). Moreover, as H is a simple r-uniform complete r-partite hypergraph, the projection on any two factors among the first r factors is a balanced design. Since $(S_0, S_1, \ldots, S_{s-1})$ is an equitable s-coloring, the projection on any two factors involving the last one is a near-balanced design. $\qquad\square$

Note: If $s_1 < s_2 < \ldots < s_r$ and $s < \prod_{i=1}^{r-2} s_i$ are positive integers that are pairwise coprime. Then, Lemma 3 provides an $MCD(\prod_{i=1}^r s_i, s_1 s_2 \cdots s_r s, 3)$ such that the projection on first r factors is a full factorial design.

4.3 Proof of Theorem 2

Proof Let H be the r-uniform complete r-partite hypergraph $K_{s_1, s_2, \ldots, s_r}^r = (V_1 \cup V_2 \cup \cdots \cup V_r, E)$ and $v_0^i, v_1^i, \ldots, v_{s_i-1}^i$ be the vertices in the side V_i, where $1 \leq i \leq r$. There are $N = \prod_{i=1}^r s_i$ hyperedges in E. Using Lemma 2, there exist a good s-coloring $\mathcal{A} = (A_0, A_1, \ldots, A_{s-1})$ and a good s'-coloring $\mathcal{B} = (B_0, B_1, \ldots, B_{s'-1})$ such that $\mathcal{A}$ and $\mathcal{B}$ are qualitatively independent partitions of E. Since $s' \leq s \leq \prod_{i=1}^{r-2} s_i < d_H(v)$, where v is an arbitrary vertex of H, and both the coloring are good coloring of hyperedges, for each vertex v and each color class A_k,

$$|H(v) \cap A_k| \geq 1. \tag{2}$$

Similarly, for each color class B_l,

$$|H(v) \cap B_l| \geq 1. \tag{3}$$

To construct a covering design MCD($\prod_{i=1}^r s_i, s_1 s_2 \cdots s_r s s', 2$), for each hyperedge $e_j = \{v_{a_1}^1, v_{a_2}^2, \ldots, v_{a_r}^r\} \in A_k \cap B_l$, where $0 \leq k \leq s-1$, $0 \leq l \leq s'-1$, $0 \leq a_i \leq s_i - 1$, and $i = 1, 2, \ldots, r$; we include the factor-level combination $r_j = (a_1, a_2, \ldots, a_r, k, l)$ as a row to construct an $N \times (r+2)$ array D. As H is the r-uniform complete r-partite hypergraph, for every level combination $(a_1, a_2, \ldots, a_r) \in \mathbb{Z}_{s_1} \times \mathbb{Z}_{s_2} \times \cdots \times \mathbb{Z}_{s_r}$ there exists a hyperedge $\{v_{a_1}^1, v_{a_2}^2, \ldots, v_{a_r}^r\}$. Therefore, the array D contains a row having $(a_1, a_2, \ldots, a_r)$ on the first r columns. Thus, the first r factors in D form a full factorial design. Next, for each pair $a \in \mathbb{Z}_{s_i}$ and $b \in \mathbb{Z}_s$, using Eq. 2, there exists at least one hyperedge e_j in A_b that contains the vertex v_a^i. Then, $D(j, i) = a$ and $A(j, r+1) = b$. Thus, the projection on each factor i

and $r + 1$ is a full factorial design. Similarly, using Eq. 3, for each $c \in \mathbb{Z}_{s'}$, there exists at least one hyperedge e_j in B_c that contains the vertex v_a^i and hence $D(j, i) = a$ and $D(j, r + 2) = c$. Thus, the projection on the factors i and $r + 2$ is a full factorial design. Now, for each pair $b \in \mathbb{Z}_s$ and $c \in \mathbb{Z}_{s'}$ being the qualitatively independent partitions, there exists at least one hyperedge $e_j \in A_b \cap B_c$. Thus, $D(j, r + 1) = b$ and $D(j, r + 2) = c$. Hence, the projection on the factors $r + 1$ and $r + 2$ is a full factorial design and D is a covering design MCD($\prod_{i=1}^{r} s_i, s_1 s_2 \cdots s_r s s', 2$). $\qquad\square$

5 Discussion

This work introduces a class of mixed-level coverage designs (MCD), for experiments that have the following requirements:

1. The experiment consists of a few near-continuous or high-level factors and many low-level factors of interest.
2. All pairwise level combinations between the high-level factors and the low-level factors are required to be covered at least once in the experiment.
3. All level combinations among low-level factors are required to be covered in the experiment.

In specific, using a graph-theoretic approach, we provide the existence results (Theorems 1 and 2) for MCDs with one or two high-level factors, together with the condition on their near-balance properties.

Traditionally, the covering designs are used in hardware and software testing experiments with a large number of testing points. The use of MCD to these testing experiments is important as it allows a sequential testing procedure when the experiments performed at a different time. We consider the following scenario. One conducts a testing experiment in the first phase and finds out 12 different testing-point orientations for the optimization of a system. In the second phase, four additional testing points (two 2-level and two 3-level) are added to the system, and it is known that the interactions among these four additional testing points require detailed investigation. Then, one may consider conducting this experiment using a 36-run MCD where the last factor is 12-level, and the level setting of this MCD is given in Sect. 2.1. In general, the high-level factors can be treated as some optimal settings in the previous phrase while the low-level factors (those form a full factorial design) are the new testing points in the current phrase.

The MCD has a lot of other applications. For example, it can be used in a choice experiment. A discrete choice experiment (DCE) is a quantitative technique for eliciting individual preferences. It is an essential tool for conjoint analysis in economics, marketing, health care, and many other research areas. A comprehensive study on DCE is referred to [17]. In the framework of MCD, the high-level factors are the choice sets consisting of hard-to-adjust factors, while the low-level factors are factors that are easy or inexpensive to adjust. This experimental plan is applied to some practical problems like a study of the dish

orientation on the restaurants' dinner valued menu, where the high-level factor is the choice of the main dish, and the low-level factors include the side dishes and drinks.

Although we always define in the MCD the low-level factors as those form a full factorial design and the high-level factors are the rest, the whole framework works well even one chooses to put a low-level factor in the position of "high" level. For example, one may consider a 40-run MCD with four "low"-level factors (three of them are 2-level and one of them is 5-level) and one "high"-level factor with only three levels. Unless it is strictly required in the experiment, the assignment of low-level factors to form a full factorial design is always the most cost efficient. Moreover, the whole framework also works well when one chooses to put important factors, regardless of low- or high-levels, in the position of full factorials and a few less-important factors in the original position of "high" level factors. In other words, our construction of MCD is useful in designing an experiment involving a group of important factors and one or two factors that are not as important as the aforementioned group of factors. The full factorial design enables to analyze the important factors in detail while the other less-important factors cover all factor-level combinations with the other factors (including the important ones).

The first extension from the current framework of MCD is to accommodate an additional number of high-level factors. For example, for an MCD with three high-level factors, one needs to ensure that the pair formed from one of the three high-level factors and any one of the low-level factors consists of all level combinations, and all three pairwise combinations of high-level factors consist of all level combinations too. From a graph-theoretic perspective, it is a fan-type structure where the high-level vertices form a triad rather than a dyad or just a vertex. To a better understanding and compare covering designs with the same parameters, several graphical methods have been introduced in [14]. The coverage evaluation plot, coverage evaluation scatterplot matrix, and the correlation-based r-plot, etc. are proposed to visualize different properties of a covering design. It would be interesting to use these graphical methods for evaluating the MCDs obtained in this paper. In particular, we would like to investigate the performance of our MCDs based on the t-way coverage evaluation plot when $t \geq 3$.

The second extension is to consider a further cost reduction in the experiment, or a further reduction in the run size of an MCD. It is achievable if we consider forming a fractional factorial design rather than a full factorial design among low-level factors. However, it then introduces the problem of aliasing among these factors, no matter if it is full aliasing or partial aliasing, which hinders the analysis of experiments. The trade-off between design estimation capacity and the experimental cost efficiency always appears when a fractionated rather than a full combination of factor levels is used, and a further study is under investigation.

References

1. Akhtar Y, Maity S (2017) Mixed covering arrays on 3-uniform hypergraphs. Discrete Appl Math 232:8–22
2. Ang N (2015) Generating Mixed-Level Covering Arrays with $\lambda = 2$ and Test Prioritization, M.Sc. thesis, Arizona State University
3. Baranyai Z (1975) On the factorization of the complete uniform hypergraph. In: Infinite and Finite Sets, Coll. Math. J. Bolyai, 10, Bolyai J. Mat. Tirsulat, Budapest, North Holland, Amsterdam, pp 91–108
4. Baranyai Z (1978) La coloration des arêtes des hypergraphes complets II. In: Problèmes Combinatoires et Théorie des Graphes, Coll. C.N.R.S. Orsay 1976, (Bermond, Fournier, Las Vergnas, Sotteau), C.N.R.S., Paris, pp 19–22
5. Baranyai Z (1979) The edge-coloring of complete hypergraphs I. J Comb Theory Ser B 26:276–294
6. Berge C (1989) Hypergraphs-combinatorics of finite sets, vol 45. Elsevier Science Publishers B.V, North-Holland
7. Bryce RC, Colbourn CJ (2009) A density-based greedy algorithm for higher strength covering arrays. Softw Test Verif Reliab 19(1):37–53
8. Cohen DM, Dalal SR, Fredman ML, Patton GC (1997) The AETG system: an approach to testing based on combinatorial design. IEEE Trans Softw Eng 23(7):437–444
9. Cohen MB, Gibbons PB, Mugridge WB, Colbourn CJ (2003) Constructing test suites for interaction testing. In: 25th international conference on software engineering, proceedings, Portland, OR, USA, 38-48
10. Dalal SR, Mallows CL (1998) Factor-covering designs for testing software. Technometrics 40(3):234–243
11. Godbole AP, Skipper DE, Sunley RA (1996) t-covering arrays: upper bounds and Poisson approximations. Comb Probab Comput 5:105–118
12. Gromping U, Xu H (2014) Generalized resolution for orthogonal arrays. Ann Stat 42:918–939
13. Hedayat AS, Sloane NJA, Stufken J (1999) Orthogonal arrays, theory and applications, springer series in statistics. Springer, New York
14. Kim Y, Jang DH, Anderson-Cook CM (2016) Graphical methods for evaluating covering arrays. Qual Reliab Eng Int 32:1467–1481
15. Kuhn DR, Reily M (2002) An investigation of the applicability of design of experiments to software testing. In: Proceedings of 27th annual NASA Goddard/IEEE software engineering workshop, pp 91-96
16. Kuhn DR, Kacker R, Lei Y (2013) Introduction to combinatorial testing. CRC Press, Boca Raton
17. Li W, Nachtsheim CJ, Wang K, Reul R, Albrecht M (2013) Conjoint analysis and discrete choice experiments for quality improvement. J Qual Technol 45:74–99
18. Lin YL, Phoa FKH, Kao MH (2017) Optimal design of fMRI experiments using circulant (almost-) orthogonal arrays. Ann Stat 45:2483–2510
19. Meagher K (2005) Covering arrays on graphs: qualitative independence graphs and extremal set partition theory, Ph.D. thesis, University of Ottawa
20. Meagher K, Moura L, Zekaoui L (2007) Mixed covering arrays on graphs. J Comb Des 15:393–404
21. Moura L, Stardom J, Stevens B, Williams A (2003) Covering arrays with mixed alphabet sizes. J Comb Des 11:413–432
22. Raaphorst S (2013) Variable strength covering arrays, Ph.D. thesis, University of Ottawa
23. Wang J, Wu CFJ (1992) Nearly orthogonal arrays with mixed levels and small runs. Technometrics 34(4):409–422
24. Wu CFJ, Hamada MS (2009) Experiments: planning, analysis and optimization, 2nd edn. Wiley, New Jersey
25. Xu H (2002) An algorithm for constructing orthogonal and nearly-orthogonal arrays with mixed levels and small runs. Technometrics 44:356–368
26. Xu H (2003) Minimum moment aberration for nonregular designs and supersaturated designs. Stat Sin 13:691–708