JOURNAL OF COMPLEXITY 17, 86–97 (2001)
doi:10.1006/jcom.2000.0569, available online at http://www.idealibrary.com on IDEAL®

On the Isomorphism of Fractional Factorial Designs[1]

Chang-Xing Ma

Department of Statistics, Nankai University, Tianjin 300071, China

Kai-Tai Fang

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

and

Dennis K. J. Lin

Department of Management Science and Information Systems,
The Pennsylvania State University, University Park, Pennsylvania 16802

Received January 4, 2000

Two fractional factorial designs are called isomorphic if one can be obtained from the other by relabeling the factors, reordering the runs, and switching the levels of factors. To identify the isomorphism of two s-factor n-run designs is known to be an NP hard problem, when n and s increase. There is no tractable algorithm for the identification of isomorphic designs. In this paper, we propose a new algorithm based on the centered L_2-discrepancy, a measure of uniformity, for detecting the isomorphism of fractional factorial designs. It is shown that the new algorithm is highly reliable and can significantly reduce the complexity of the computation. Theoretical justification for such an algorithm is also provided. The efficiency of the new algorithm is demonstrated by using several examples that have previously been discussed by many others.

Key Words: factorial designs; Hamming distance; isomorphism; uniformity.

[1] This work was partially supported by a Hong Kong RGC grant and Science Faculty of Hong Kong Baptist University. Dennis Lin is partially supported by the *National Science Foundation* via Grant DMS-9704711 and the National Science Council of ROC via Contract NSC 87-2119-M-001-007. The corresponding author is Kai-Tai Fang.

0885-064X/01 $35.00

1. INTRODUCTION

Fractional factorial experiments have become important in all kinds of studies. Two factorial designs are called *isomorphic* if one can be obtained from the other by relabeling the factors, reordering the runs, or switching the levels of factors. Two isomorphic designs are considered to be equivalent because they share the same statistical properties in a classical ANOVA model. Therefore, it is important to identify design-isomorphism. Identifying whether two fractional designs are isomorphic has received a great deal of attention in the literature (see, for example, Clark and Dean (2000) and references therein).

Denote by $d(n, q, s)$ a factorial design of n runs and s factors each having q levels. A design $d(n, q, s)$ is usually expressed as an $n \times s$ matrix with elements $0, 1, ..., q-1$. For identifying two $d(n, q, s)$ designs, a complete search compares $n!(q!)^s s!$ designs from the definition of isomorphism. For example, it requires $13!12!2^{12} = 1.22 \times 10^{22}$ comparisons for two factorial $d(13, 2, 12)$ designs to see if these designs are isomorphic. The identification of two factorial designs is considered to be a NP hard problem when the number of runs or/and the number of factors increase.

A fractional factorial design that is constructed through defining relations among factors is called *regular*, otherwise *nonreqular*. For a q^{s-k} regular fractional factorial design D, let $A_r(D)$ be the number of factorial effects, involving r factors, which appear in the defining relation. The sequence $\{A_1(D), ..., A_s(D)\}$ is called the word-length pattern. The smallest r for $A_r(D) > 0$ is called the resolution of D. For a given (n, q, s) one searches a design $d(n, q, s)$ with highest resolution as it has less confounding. When two $d(n, q, s)$ designs D_1 and D_2 have the save level of resolution, there exists a $r > 0$ such that $A_j(D_1) = A_j(D_2) = 0$ for all $j < r$ and $A_r(D_i) > 0, i = 1, 2$. One wants to choose the design with smaller $A_r(\cdot)$. The minimum aberration criterion is based on such an idea. For a detailed discussion refer to Fries and Hunter (1980).

A necessary condition for two regular factorial designs to be isomorphic is that they have identical word-length pattern. Draper and Mitchell (1968) gave two $L_{512}(2^{12})$ orthogonal designs which have identical word-length patterns, but are not isomorphic. Here, $L_n(q^s)$ denotes an orthogonal array with n runs and s columns. Draper and Mitchell (1970) gave a more sensitive criterion for isomorphism, called "letter pattern comparison," and tabulated 1024-run designs of resolution 6. Let a_{ij} be the number of words of length j in which letter i appears in a regular design D and $A = (a_{ij})$ be the letter pattern matrix of D. They conjectured that two designs D and D' are isomorphic if and only if $A = PA'$, where P is a permutation matrix. Obviously, two designs having identical letter pattern matrices necessarily have identical word-length patterns. Chen and Lin (1991) gave two nonisomorphic designs 2^{31-15} with identical letter pattern matrices and thus showed that the criterion "letter pattern matrix" is not sufficient for design isomorphism. Note that both the word-length and letter pattern matrix are not easy to calculate and can be applied only to regular factorial designs.

Recently, Clark and Dean (2000), denoted by [CD00] for convenience, gave a sufficient and necessary condition for isomorphism of designs. Let $H=(d_{ij})$ be the *Hamming distance matrix* of a design D, where d_{ij} is the *Hamming distance* of the ith and jth runs of D and is defined as the number of levels of the factors where they differ. This clever method is invariant under the permutations of levels, but the complexity here makes the calculation intractable. For example, it may require $12!12!12 = 2.75 \times 10^{18}$ comparisons for two non-isomorphic $d(13, 2, 12)$ designs.

In this paper we propose a necessary criterion for detecting non-isomorphic (regular and nonregular) factorial designs based on uniformity, a criterion that is crucial in space-filling designs for computer experiments (Bates *et al.* (1996)) and in uniform designs (Fang and Wang (1994) and Fang *et al.* (2000)). The centered L_2-discrepancy proposed by Hickernell (1998) is employed as the measure of uniformity in this study and is introduced in Section 2. An algorithm for detecting the isomorphism of two-level $d(n, 2, s)$ designs is also proposed there. Section 3 applies the proposed algorithm to several examples that were discussed by others. Section 4 discusses the extension to higher level designs, and an example is given for illustration. The conclusion and further discussion are given in Section 5.

2. ISOMORPHISM OF TWO-LEVEL DESIGNS

Recall that Clark and Dean (2000) algorithm is mainly based on the following lemma. This will be called as the *HD-method* for convenience. Also note that the HD-method requires to find the permutation $\{c_1, ..., c_p\}$ and the permutation matrix R.

LEMMA 1. *Let D_1 and D_2 be two $d(n, q, s)$ designs. Then D_1 and D_2 are isomorphic if and only if there exist an $n \times n$ permutation R and a permutation $\{c_1, ..., c_s\}$ of $\{1, ..., s\}$ such that for $p = 1, ..., s$*

$$H_{D_1}^{\{1, ..., p\}} = R H_{D_2}^{\{c_1, ..., c_p\}} R',$$

where $H_D^{\{c_1, ..., c_p\}}$ is the Hamming distance matrix of the design formed by columns $\{c_1, ..., c_p\}$ of design D.

A design $d(n, q, s)$ can be viewed as n points in the unit cube $C^s = [0, 1)^s$, after proper coding, e.g., the q levels $(0, 1, ..., q-1) \Rightarrow (0.5/q, 1.5/q, ..., (q-0.5)/q)$. Let $\mathcal{P} = \{u_1, ..., u_n\}$ be a set of n points in C^s. Many criteria have been proposed for the measures of uniformity. We shall concentrate on the centered L_2-discrepancy (CD_2 for short) in our study here. The CD_2 has some nice and unique properties, such as it is invariant under reordering the runs, relabeling coordinates and reflections of the points about any plane passing through the center of the unit cube and parallel to its faces. Furthermore, the CD_2 captures uniformity over the unit cube as well as the uniformity over all projected subdimensions. Hickernell (1998) gave an analytical formula for the CD_2 as

$$(CD_2(\mathcal{P}))^2 = \left(\frac{13}{12}\right)^s - \frac{2}{n} \sum_{k=1}^{n} \prod_{j=1}^{s} \left(1 + \frac{1}{2}|u_{kj} - 0.5| - \frac{1}{2}|u_{kj} - 0.5|^2\right)$$
$$+ \frac{2^s}{n^2} \sum_{k=1}^{n} \sum_{j=1}^{n} \prod_{i=1}^{s} \left[1 + \frac{1}{2}|u_{ki} - 0.5| + \frac{1}{2}|u_{ji} - 0.5| - \frac{1}{2}|u_{ki} - u_{ji}|\right], \tag{1}$$

where $\mathbf{u}_k = (u_{k1}, ..., u_{ks})'$. Recently, Fang and Mukerjee (2000) show the relationship between uniformity and aberration for two-level regular fractions. Ma *et al.* (1999) found links between uniformity and orthogonality for some factorials.

For a $d(n, q, s)$ design D with levels $0, 1, ..., q-1$, when we calculate its CD_2-value we always assume to map its q levels into $1/2q, 3/2q, ..., (2q-1)/2q$. This understanding is useful in links among uniformity, Hamming distance, distance distribution and weight distribution of a design D.

THEOREM 1. *For a two-level $d(n, 2, s)$ design D, we have*

$$CD_2^2(D) = \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \frac{1}{n^2}\left(\frac{5}{4}\right)^s \left(n + 2\sum_{i=1}^{n} \sum_{j=1}^{i-1} \left(\frac{4}{5}\right)^{d_H(\mathbf{u}_i, \mathbf{u}_j)}\right), \tag{2}$$

where $\mathbf{u}_i$, $i = 1, ..., n$ are n runs of D and $d_H(\mathbf{u}_i, \mathbf{u}_j)$ is the Hamming distance between $\mathbf{u}_i$ and $\mathbf{u}_j$.

Proof. Note that the two levels are chosen as $1/4$ and $3/4$. The second term on the right hand side of (1) becomes $2(\frac{35}{32})^s$. Consequently,

$$[1 + \tfrac{1}{2}|u_{ki} - 0.5| + \tfrac{1}{2}|u_{ji} - 0.5| - \tfrac{1}{2}|u_{ki} - u_{ji}|] = \begin{cases} 5/4 & \text{if } u_{ki} = u_{ji}, \\ 1 & \text{otherwise.} \end{cases}$$

Therefore, we have

$$\prod_{i=1}^{s}\left[1+\frac{1}{2}|u_{ki}-0.5|+\frac{1}{2}|u_{ji}-0.5|-\frac{1}{2}|u_{ki}-u_{ji}|\right]=\left(\frac{5}{4}\right)^{s-d_H(\mathbf{u}_i,\,\mathbf{u}_j)}.$$

The theorem follows from (1).

For any $d(n, q, s)$ design D, let $E_i(D)$ be $\frac{1}{n}$ times the number of pairs of two runs whose Hamming distance to be i, i.e.,

$$E_i(D)=\frac{1}{n}\ \#\{(\mathbf{c},\mathbf{d})\mid \mathbf{c},\mathbf{d}\in D,\ d_H(\mathbf{c},\mathbf{d})=i\},$$

where $d_H(\mathbf{c}, \mathbf{d})$ is the Hamming distance between two runs $\mathbf{c}$ and $\mathbf{d}$. The sequence $\{E_0(D), ..., E_s(D)\}$ is referred to as the *distance distribution* of D. From Theorem 1, we can establish a link between the distance distribution and CD_2-value of a two-level design.

THEOREM 2. *For a $d(n, 2, s)$ design D, we have*

$$CD_2^2(D)=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\frac{1}{n}\left(\frac{5}{4}\right)^s\sum_{k=0}^{s}E_k(D)\left(\frac{4}{5}\right)^k. \tag{3}$$

Comparing Eqs. (2) and (3), although both cost $O(sn^2)$, but the former needs to compute $n(n-1)/2$ powers of $\frac{4}{5}$ while the latter needs only s powers of $\frac{4}{5}$. The latter reduces the complexity of computation. In fact, if D is a regular two-level design, the complexity can be further reduced into $O(ns)$. Let $D=\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ be a regular factorial design with two levels 0 and 1. The weight distribution of D is defined by

$$W_k(D)=\frac{1}{n}\ \#\Big\{\mathbf{x}_i\ \Big|\ \sum_{j=1}^{s}x_{ij}=k\Big\}.$$

THEOREM 3. *Let $D=\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ be a regular factorial design with two levels* 0 *and* 1. *We have*

$$CD_2^2(D)=\left(\frac{13}{12}\right)^s-2\left(\frac{35}{32}\right)^s+\frac{1}{n}\left(\frac{5}{4}\right)^s\sum_{k=0}^{s}W_k(D)\left(\frac{4}{5}\right)^k. \tag{4}$$

Proof. Because runs of a regular design D must form a linear subspace of the full design 2^s over $GF(2)$, $(\mathbf{x}_i-\mathbf{x}_j)$ (mod 2) is also a run of D, i.e., $\{(\mathbf{x}_i-\mathbf{x}_j) \text{ (mod 2)} \mid j=1, ..., n\}$ are all the designs, D, for any $i=1, ..., n$. The proof is completed by Eq. (3).

To compute $E_k(D)$, one needs to calculate $n(n-1)/2$ distances, but to find $W_k(D)$, one needs only n summations for the runs.

For two isomorphic $d(n, 2, s)$ designs D_1 and D_2 they have the same set of Hamming distances, the same sequences of E_k, and thus an identical CD_2-value. Furthermore, they have the same CD_2-value distribution for each of their projection designs. For given k $(1 \leqslant k < s)$ there are $\binom{s}{k}$ sub-designs for a design $D(n, 2, s)$. The distribution of CD_2-values of these sub-designs is called the k-dimensional CD_2-value distribution of D and we denote it by $F_k(D)$. The following necessary condition is thus obtained.

Uniformity Criterion for Isomorphism (*UCI*) *of Two-Level Designs.* The necessary conditions for two $d(n, 2, s)$ designs D_1 and D_2 to be isomorphic are a) they have the same CD_2-value; b) they have the same distribution $F_k(D_1) = F_k(D_2)$ for $1 \leqslant k < s$.

Based on the UCI we propose the following algorithm, called NIU algorithm, for detecting isomorphic $d(n, 2, s)$ designs. Let D_1 and D_2 be two $d(n, 2, s)$ designs.

NIU ALGORITHM.

Step 1. Comparing $CD_2(D_1)$ and $CD_2(D_2)$, if $CD_2(D_1) \neq CD_2(D_2)$ we conclude D_1 and D_2 are not isomorphic and terminate the process, otherwise go to Step 2.

Step 2. Let $[x]$ be the integer part of a positive number x. For $k = 1, s-1, 2, s-2, ..., [s/2], s-[s/2]$, comparing $F_k(D_1)$ and $F_k(D_2)$, if $F_k(D_1) \neq F_k(D_2)$ we conclude D_1 and D_2 are not isomorphic and terminate the process, otherwise this step goes to the next k-value.

We next apply the NIU algorithm to several examples that have been studied by others. As will be seen, the NIU algorithm efficiently detects the non-isomorphism of designs, typically at Step 1 or Step 2. Note that the NIU algorithm needs $O(n^2 s 2^s)$ operations to compare 2^{s+1} CD_2-values in the worst case. This is polynomial in n and exponential in s. This is a significant improvement over the complete search (which takes $n!s!2^s$ comparisons and is superexponential in both n and s) and the HD-method (which requires $s(s!)^2$ comparisons and each comparison required $O(n!)$ operations in the worst case).

3. EXAMPLES OF TWO-LEVEL DESIGNS

In this section we apply the NIU algorithm to several designs that have been studied in the literature and show that the new algorithm is very useful. The calculation was carried on a PC-computer with double precision.

EXAMPLE 1. Consider two $d(7, 2, 6)$ designs D_1 and D_2. Their treatment matrices are given below. All their treatments are the same except the sixth treatment combination.

$$D_1 = \begin{pmatrix} 0&0&0&1&1&1 \\ 0&0&1&1&0&0 \\ 0&1&0&0&1&0 \\ 0&1&1&0&0&1 \\ 1&0&0&0&0&1 \\ 1&0&1&1&1&0 \\ 1&1&0&1&0&1 \end{pmatrix}, \quad \text{and} \quad D_2 = \begin{pmatrix} 0&0&0&1&1&1 \\ 0&0&1&1&0&0 \\ 0&1&0&0&1&0 \\ 0&1&1&0&0&1 \\ 1&0&0&0&0&1 \\ 0&1&0&0&0&1 \\ 1&1&0&1&0&1 \end{pmatrix}.$$

The designs D_1 and D_2 are non-isomorphic because $CD_2^2(D_1) = 0.2792$ while $CD_2^2(D_2) = 0.4245$. The process terminates at the first step.

EXAMPLE 2. The three $d(12, 2, 5)$ designs d_1, d_2 and d_3 considered in Example 3.1 of [CD00] have different squared CD-values 0.1893, 0.1882, and 0.1682. So they are non-isomorphic to each other.

EXAMPLE 3. The NIU algorithm is applied to the three $d(18, 2, 17)$ designs, $d_9, d_{6,3}$ and $d_{3,3,3}$, discussed by Cohn (1994) (also given in [CD00] as Examples 2.1 and 3.1). First of all, all three designs produce an identical squared CD_2-value of 2.952. We next calculate their k-dimensional CD_2-value distributions. For $k = 1$, the $F_1(D)$ distribution is identical for all three designs: 9 projection designs have $CD_2^2 = 0.0208$ and 8 designs have $CD_2^2 = 0.0224$; the $F_{16}(D)$ distribution is also identical: $CD_2^2 = 2.3877$ (twice), $CD_2^2 = 2.3891$ (6 times), $CD_2^2 = 2.3958$ (6 times), and $CD_2^2 = 2.3965$ (3 times). For $k = 2$, the $F_2(D)$ distribution is identical for all three designs: $CD_2^2 = 0.0469$ (36 times), $CD_2^2 = 0.0484$ (72 times), and $CD_2^2 = 0.0503$ (28 times). However, the $F_{15}(D)$ distribution of $d_{3,3,3}$ is different from the other two designs. For $k = 3$, we found that the $F_3(D)$ distributions are different between designs d_9 and $d_{6,3}$. Finally, we conclude that the three designs are not isomorphic to each other.

EXAMPLE 4. Consider the two $L_{512}(2^{12})$ designs, D_1 and D_2, given in Draper and Mitchell (1968, as designs 3.4 and 3.5 in Table I). They showed that these two designs are not isomorphic but share an identical word-length pattern. When applying NIU to these two designs, an identical CD-value of 0.8609440 was obtained. Obviously, for $k = 1, 2$, all the k-projection designs of D_1 and D_2 are isomorphic, because both designs are orthogonal designs. However, $F_{11}(D_1) \neq F_{11}(D_2)$. We thus conclude that they are not isomorphic. Specifically, $F_{11}(D_1)$ has $CD_2^2 = 0.7057277$ (12 times); while $F_{11}(D_2)$ has $CD_2^2 = 0.7057209$ (3 times), $CD_2^2 = 0.7057277$

TABLE I

The Distribution F_{12} for the Five Designs

D_1 CD_2^2	Freq.	D_2 CD_2^2	Freq.	D_3 CD_2^2	Freq.	D_4 CD_2^2	Freq.	D_5 CD_2^2	Freq.
0.97930	35	0.97930	19	0.97930	11	0.97930	7	0.97930	7
0.98407	420	0.98288	64	0.98288	96	0.98288	112	0.98288	112
		0.98407	372	0.98407	348	0.98407	336	0.98407	336

(6 times) and $CD_2^2 = 0.7057345$ (3 times). Note that it is not feasible to apply the HD-method for such a large design.

EXAMPLE 5. It is well known that the numbers of non-isomorphic Hadamard matrices are 1, 1, 1, 5, 3, and 60 respectively for orders $n = 4$, 8, 12, 16, 20 and 24. We have successfully applied the NIU to all six of these cases. The details for $n = 16$ are given here.

Hall (1961) found that there exist exactly five non-isomorphic groups of Hadamard matrices of order 16. After deleting the column of 1's from each of the matrices after normalizing, the remaining matrices are denoted by D_j, $j = 1, ..., 5$, respectively. All five designs have the same squared CD-values of 1.8988504 and identical distributions of F_{14} and F_{13} when projected into 14 and 13 sub-dimensions. However, the different F_{12}-distributions given in Table I indicate the non-isomorphism of five designs, except designs D_4 and D_5. In the next step, we found that $F_{11}(D_4) \neq F_{11}(D_5)$, as shown in Fig. 1, and thus concluded the non-isomorphism of five designs.

EXAMPLE 6. Chen and Lin (1991, p. 97, Table 1) gave two non-isomorphic $L_{32768}(2^{31})$ designs with the same letter pattern matrix. It requires a large amount of computation to detect their non-isomorphism

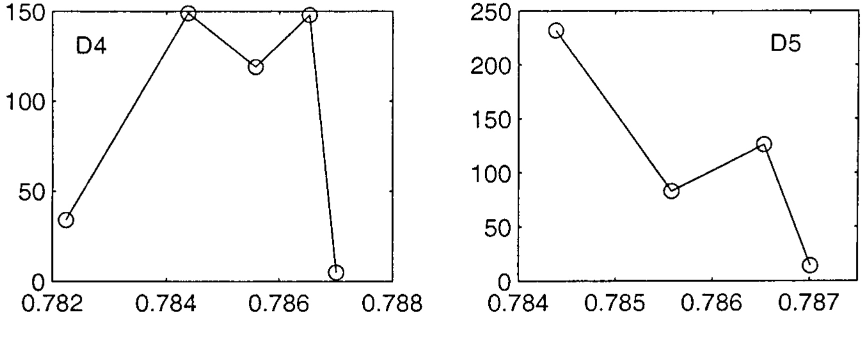

FIG. 1. Plots of Distribution F_{11} for D_4 and D_5.

by any other existing algorithms, including the HD-method. The non-isomorphism, however, can be easily detected by the NIU algorithm as follows:

(1) two designs have the same $CD_2^2 = 18.30959$;

(2) all the 30-dimensional projection designs have the same $CD_2^2 = 15.86654$;

(3) all the 29-dimensional projection designs have the same $CD_2^2 = 13.73364$;

(4) the F_{28}-distributions of the two designs are different as follows:

Design (a)		Design (b)	
CD_2^2	Freq.	CD_2^2	Freq.
11.872776	155	11.872791	155
11.872806	4340	11.872796	465
		11.872801	930
		11.872806	1550
		11.872811	1395

The conclusion that two designs are not isomorphic follows by implementing only four steps of the algorithm.

4. FRACTIONAL FACTORIAL DESIGNS OF HIGHER LEVELS

In this section we consider the problem of detecting non-isomorphic designs for high-level factorial designs. Let D be a $d(n, q, s)$ design and $E_k(D)$'s be its distance distribution defined in Section 2. Denote

$$B_a(D) = \sum_{i=1}^{s} E_i(D)\, a^i$$

as the distance enumerator of D (Roman, 1992, p. 226). For a two-level design D we have from (3)

$$(Cd_2(D))^2 = \frac{1}{n}\left(\frac{5}{4}\right)^s B_{4/5}(D) - 2\left(\frac{35}{32}\right)^s + \left(\frac{13}{12}\right)^s$$

which provides a link between the distance enumerator and uniformity. In fact, the UCI is equivalent to the measure $B_{4/5}(D)$ for two-level designs. This measure can naturally be used for high-level factorial designs. Given k $(1 \leqslant k \leqslant s)$, the distribution of B_a-values over all k-dimensional projection subdesigns is denoted by $F_{B_{a,k}}(D)$. We now can have an NIU version for the high-level designs. As the parameter a is a pre-determined value, we omit a from the notation for simplicity.

NIU ALGORITHM FOR HIGH-LEVEL DESIGNS.

Step 1. Comparing $B(D_1)$ and $B(D_2)$, if $B(D_1) \neq B_2(D_2)$, we conclude D_1 and D_2 are not isomorphic and terminate the process. Otherwise go to Step 2.

Step 2. For $k = 1, s-1, 2, s-2, \ldots, [s/2], s-[s/2]$, compare $F_{B_k}(D_1)$ and $F_{B_k}(D_2)$. If $F_{B_k}(D_1) \neq F_{B_k}(D_2)$, we conclude D_1 and D_2 are not isomorphic and terminate the process, otherwise this step goes to the next k-value.

For a simple illustration, consider the four $L_{18}(3^7)$ in Table II, where Design (a) is from Masuyama (1957), Design (c) is from Fang *et al.* (2000), and Design (b) is from http://www.research.att.com/ ~ njas/oadir/. Taking

TABLE II

Four $L_{18}(3^7)$ Designs

No.	(a)	(b)	(c)	(d)
1	1 1 1 1 1 1 1	1 1 1 1 1 1 1	1 2 1 1 3 2 1	3 1 2 1 2 1 2
2	1 2 2 2 2 2 2	1 1 2 3 2 3 2	1 3 1 1 2 1 2	1 1 2 3 1 3 1
3	1 3 3 3 3 3 3	1 2 1 3 3 2 2	1 1 2 3 2 2 3	2 2 2 3 2 1 3
4	2 1 1 2 2 3 3	1 2 3 2 1 3 3	1 3 2 3 1 3 1	1 2 3 1 1 2 2
5	2 2 2 3 3 1 1	1 3 2 2 3 1 3	1 1 3 2 3 3 2	3 1 3 1 2 3 3
6	2 3 3 1 1 2 2	1 3 3 1 2 2 2	1 2 3 2 1 1 3	1 2 1 1 2 1 1
7	3 1 2 1 3 2 3	2 1 1 2 3 3 2	2 1 1 3 1 1 2	1 3 2 2 2 3 2
8	3 2 3 2 1 3 1	2 1 3 3 1 2 3	2 2 1 3 3 3 3	2 3 1 1 3 3 1
9	3 3 1 3 2 1 2	2 2 2 2 2 2 1	2 2 2 2 2 2 2	2 1 3 2 2 1 1
10	1 1 3 3 2 2 1	2 2 3 1 3 1 1	2 3 2 2 3 1 1	3 3 1 3 2 2 2
11	1 2 1 1 3 3 2	2 3 1 3 2 1 3	2 1 3 1 1 2 1	3 2 2 2 3 2 1
12	1 3 2 2 1 1 3	2 3 2 1 1 3 2	2 3 3 1 2 3 3	1 3 3 2 3 1 3
13	2 1 2 3 1 3 2	3 1 2 1 3 2 3	3 1 1 2 2 3 1	2 2 3 3 3 3 2
14	2 2 3 1 2 1 3	3 1 3 2 2 1 2	3 3 1 2 1 2 3	3 3 3 3 1 1 1
15	2 3 1 2 3 2 1	3 2 1 1 2 3 3	3 1 2 1 3 1 3	1 1 1 3 3 2 3
16	3 1 3 2 3 1 2	3 2 2 3 1 1 2	3 2 2 1 1 3 2	2 1 1 2 1 1 2
17	3 2 1 3 1 2 3	3 3 1 2 1 2 1	3 2 3 3 2 1 1	3 2 1 2 1 3 3
18	3 3 2 1 2 3 1	3 3 3 3 3 3 1	3 3 3 3 3 2 2	2 3 2 1 1 2 3

$a = 4/5$, for example, the four designs have the same distance enumerator 6.685248. However, the distributions of distance enumerator of all 6-dimensional projection designs are different as indicated below. Therefore, we conclude that Designs (a), (c), and (d) are non-isomorphic. Note that an exhaustive comparison indicates that Designs (a) and (b) are indeed isomorphic.

$B_{(a)}$	Freq.	$B_{(b)}$	Freq.	$B_{(c)}$	Freq.	$B_{(d)}$	Freq.
7.6683	1	7.6683	1	7.6719	1	7.6737	3
7.6765	6	7.6765	6	7.6747	2	7.6765	4
				7.6765	4		

5. DISCUSSION

The uniformity criterion proposed in this paper is useful for detecting design isomorphism. The HD-method gives a necessary and sufficient link between the isomorphism and the Hamming distance matrices of two designs. As a matter of fact, $CD_2(D)$ is a function of the Hamming distance matrix of D. Equations (3) and (4) can significantly reduce the computation efforts. This makes the NIU algorithm a powerful tool, as clearly seen from Examples 1–7.

For the research on projection properties of a given fractional factorial design, we need to classify all its projection designs. For example, Lin and Draper (1992) studied the Plackett–Burman $L_{12}(2^{11})$ design and its projection designs. For the five-dimensional case they have classified 462 projection designs into two non-isomorphic groups. This is computationally-intensive work, using definition of the isomorphism directly. Moreover, many optimality criteria such as D-optimality, A-optimality, were used for the classification of projection designs. Unfortunately, there is no analytic link between isomorphism and D-optimality (or A-optimality), and many non-isomorphic designs result in an identical information matrix, and thus optimalities based on the information matrix, such as D- or A-optimality, are inappropriate in detecting the design isomorphism. The NIU algorithm proposed here provides a very efficient and meaningful way for classifying projection designs.

The UCI is only a necessary condition for design isomorphism. We conjecture that UCI is also a sufficient condition, but fail to prove it at this stage. If the conjecture is true, the computing complexity of comparing two $d(n, q, s)$ is polynomial in n and q and exponential in s. Nonetheless, it is probably the most efficient algorithm for detecting non-isomorphism of designs.

REFERENCES

1. R. A. Bates, R. J. Buck, E. Riccomagno, and H. P. Wynn, Experimental design and observation for large systems, *J. Roy. Statist. Soc. Ser. B* **58** (1996), 77–94.
2. J. Chen and D. K. J. Lin, On the identity relationship of 2^{k-p} designs, *J. Statist. Plann. Inference* **28** (1991), 95–98.
3. J. B, Clark and A. M. Dean, Equivalence of fractional factorial designs, *Statist. Sinica*, in press.
4. J. H. E., Cohn, On the number of *d*-optimal designs, *J. Combin. Theory* **66** (1994), 214–225.
5. N. R. Draper and T. J. Mitchell, Construction of the set of 256-run designs of resolution $\geqslant 5$ and set of even 512-run designs of resolution $\geqslant 6$ with special reference to the unique saturated designs, *Ann. Math. Statist.* **39** (1968), 246–255.
6. N. R. Draper and T. J. Mitchell, Construction of a set of 512-run designs of resolution $\geqslant 5$ and a set of even 1024-run designs of resolution $\geqslant 6$, *Ann. Math. Statist.* **41** (1970), 876–887.
7. K. T. Fang and Y. Wang, "Number-Theoretic Methods in Statistics," Chapman & Hall, London, 1994.
8. K. T. Fang, D. K. J. Winker, and P. Zhang, Uniform design: Theory and applications, *Technometrics* **42** (2000), 237–248.
9. K. T. Fang and R. Mukerjee, Connection between uniformity and aberration in regular fractions of two-level factorials, *Biometrika* **87** (2000), 193–198.
10. A. Fries and W. G. Hunter, Minimum Aberration 2^{k-p} designs, *Technometrics* **22** (1980), 601–608.
11. F. J. Hickernell, A generalized discrepency and quadrature error bound, *Math. Comp.* **67** (1998), 299–322.
12. M. Hall, Jr., "Hadamard Matrices of Order 16," Research Summary, No. 36-10, Vol. 1, pp. 21–26, Jet Propulsion Lab, Pasadena, CA, 1961.
13. D. K. J. Lin and N. R. Draper, Projection properties of Plackett and Burman designs, *Technometrics* **34** (1992), 423–428.
14. C. X. Ma, K. T. Fang, and D. K. J. Lin, A note on uniformity and orthogonality, submitted for publication.
15. M. Masuyama, On difference sets for constructing orthogonal arrays of index two and strength two, *Rep. Statist. Appl. Res. Un. Japan. Sci. Engrs.* **5** (1957), 27–34.
16. R. L. Plackett and J. P. Burman, The design of optimum multifactorial experiments, *Biometrika* **34** (1946), 305–325.
17. S. Roman, "Coding and Information Theory," Springer-Verlag, New York, 1992.