# Statistical Isomorphism of Three-Level Fractional Factorial Designs

Pi-Wen Tsai
Division of Biostatistics and Bioinformatics
National Health Research Institutes, Taipei, 115, Taiwan, R.O.C.
*pwtsai@nhri.org.tw*

Steven G. Gilmour
School of Mathematical Sciences, Queen Mary,
University of London, Mile End Road, London E1 4NS, UK.
*s.g.gilmour@qmul.ac.uk*

Roger Mead
School of Applied Statistics, The University of Reading,
PO Box 240, Earley Gate, Reading RG6 6FN, UK.
*roger.mead@tesco.net*

**Abstract**

From a statistician's standpoint, the interesting kind of isomorphism for fractional factorial designs depends on the statistical application. Combinatorially isomorphic fractional factorial designs may have different statistical properties when factors are quantitative. This idea is illustrated by using Latin squares of order 3 to obtain fractions of the $3^3$ factorial design in 18 runs.

*Key words:* efficiency; optimal design; orthogonal array.

## 1 Introduction

Two fractional factorial designs are called *combinatorially isomorphic* if one can be obtained from the other by permutations of the experimental units (rows), factor labels (columns) and levels of the factors. To identify the combinatorial isomorphism of two fractional factorial designs is known to be an NP hard problem, i.e. the number of steps required to solve it is not bounded by a polynomial in the size of the input, here the number of factors multiplied by the number of experimental units. Clark and Dean [5] proposed using Hamming distances to detect distinct designs for two-level experiments. Ma, Fang and Lin [7] proposed a necessary criterion based on the uniformity of distances between points in the design [2] for detecting non-isomorphic designs. Cheng and Ye [4] used indicator functions as a tool to classify geometrically non-isomorphic designs. We now emphasise that the statistically interesting kind of isomorphism depends on the statistical application. Two combinatorially isomorphic three-level designs may have different statistical properties when factors are quantitative. We use the simple example of three-level designs to demonstrate our idea.

## 2 Latin squares and fractional factorials

A *Latin square* of order $n$ is an $n \times n$ array with entries from a set of $n$ symbols, such that each symbol in the set occurs exactly once in each row and exactly once in each column. For example, assuming that the symbols in the square are $\{0, 1, 2\}$, the square

$$\begin{array}{ccc} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{array} \tag{1}$$

is a Latin square of order 3. Clearly, if we permute the rows, the columns or the symbols of a Latin square, the result is still a Latin square.

A Latin square can be represented as an $n^2 \times 3$ array. Let S be the set of $n^2$ triples of the form $(i, j, k)$ where the symbol in row $i$ and column $j$ of the square is $k$. Labelling the rows $\{0, 1, 2\}$ and the columns likewise, the Latin square in

(1) can be represented by the following nine triples:

$$000 \quad 012 \quad 021 \quad 101 \quad 110 \quad 122 \quad 202 \quad 211 \quad 220 \tag{2}$$

When written with each triple on a separate line of the page, this becomes an orthogonal array of strength 2 for three factors, each with 3 levels.

This orthogonal array can also be obtained by using the usual *defining contrast subgroups* procedure for the construction of regular $3^{3-1}$ fractional factorial designs. The procedure is to ensure the resulting design contains exactly nine treatment combinations that satisfy the relationship (which is called the *defining relation* of the design)

$$\mathrm{x}_1 + \alpha_2\mathrm{x}_2 + \alpha_3\mathrm{x}_3 = a \text{ (mod 3)}, \tag{3}$$

where $\mathrm{x}_i$ is the level of the $i$th factor in a particular treatment combination and $\mathrm{x}_i = 0$ (low level), 1 (middle level) or 2 (high level), $\alpha_i = 1$ or 2 and $a = 0$, 1 or 2. The orthogonal array obtained from (2) is formed when we select $\mathrm{x}_1 + 2\mathrm{x}_2 + 2\mathrm{x}_3 = 0$ (mod 3) as the defining relation.

It is easily verified that there are 12 distinct sets of $\alpha$s and $a$s corresponding to the 12 regular $3^{3-1}$ fractional factorials. These 12 designs are said to be *combinatorially isomorphic* since one can be obtained from any other by relabelling the factors, reordering the runs and relabelling the levels of factors. Each of these designs can be represented by a $3 \times 3$ Latin square and these Latin squares are said to be *isotopic* since one can be obtained from another by permuting rows, columns and symbols.

However, when the factors are quantitative, with equally spaced levels, and we are interested in checking linearity of response, we use linear and quadratic contrasts, with coefficients $(-1, 0, 1)$ and $(1/2, -1, 1/2)$ respectively, to decompose the main effects of factors. Assume now that the three levels are denoted by $-1$, 0 and $+1$ for the low, middle and high levels of a factor. Then designs obtained from one another by some relabellings of levels of factors, namely switching 0 with $+1$ or 0 with $-1$ or cycling $(-1, 0, +1)$ to either $(0, +1, -1)$ or $(+1, -1, 0)$, are said to be in different *design families* as defined by [9], or *geometrically non-isomorphic* as defined by [3]. Only one design in a design family is needed, since the other members of the family have the same properties. For this classification, the 12 regular $3^{3-1}$ designs are divided into two design families: one consists of Latin squares with a centre point (0,0,0); the other consists of Latin squares without the centre point. These are denoted by LS_wc and LS_wt respectively.

# 3 Example for three factors

One of the frequently used three-level orthogonal arrays in industrial experiments is the $L_{18}(3^7)$ orthogonal array, an 18-run design for seven three-level factors, described in, for example, [6]. There are 7 different designs for three factors when we project $L_{18}(3^7)$ onto sets of three factors, as described in [10] and [3]. These designs can be obtained by putting two $3 \times 3$ Latin squares together and have either 18, 15 or 9 distinct points.

When we put together two $3 \times 3$ Latin squares which have no points in common, we have two designs with 18 distinct points, namely designs D1 and D2 in Table 1. These two designs, although they are combinatorially isomorphic, have different statistical properties and are separated as two different design families. Design D2 is formed by putting two mutually orthogonal $3 \times 3$ Latin squares together.

When we put together two $3 \times 3$ Latin squares which have points in common, we either have a design with three repeated points, i.e. one of designs D3, D4, D5 and D6 in Table 1, or a design with nine repeated points, D7 and D8 in Table 1. Table 2 summarises the properties of these designs. The difference between D3 and D4 is that the former has repeated points corresponding to the middle level, 0, of one factor, whereas the latter has repeated points corresponding to $+1$ or $-1$ of one factor. Design D8 is obtained by using a LS_wt twice and cannot be obtained by projection from $L_{18}(3^7)$. Wang and Wu [10] found that there are

Table 1: Plans for designs for three 3-level factors in 18 runs

| D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 | D11 | D12 | D13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| −−− | −−− | −−− | −−− | −−− | −−− | −−− | −−− | −−− | −−− | −−− | −−− | −−− |
| −−+ | −−0 | −−+ | −−− | −−− | −−0 | −−− | −−− | −−0 | −−0 | −−0 | −−− | −−− |
| −0− | −0− | −0− | −00 | −00 | −0− | −0+ | −00 | −0− | −0− | −0− | −00 | −00 |
| −00 | −0+ | −00 | −00 | −0+ | −0+ | −0+ | −00 | −0+ | −0+ | −0+ | −0+ | −0+ |
| −+0 | −+0 | −+0 | −++ | −+0 | −+0 | −+0 | −++ | −+0 | −+0 | −+0 | −+0 | −+0 |
| −++ | −++ | −++ | −++ | −++ | −++ | −+0 | −++ | −++ | −++ | −++ | −++ | −++ |
| 0−0 | 0−− | 0−0 | 0−0 | 0−+ | 0−+ | 0−+ | 0−+ | 0−0 | 0−+ | 0−− | 0−0 | 0−0 |
| 0−+ | 0−+ | 0−0 | 0−− | 0−+ | 0−+ | 0−+ | 0−+ | 0−+ | 0−+ | 0−+ | 0−+ | 0−+ |
| 00− | 000 | 00+ | 00+ | 00− | 000 | 000 | 00− | 00− | 00− | 000 | 00− | 00− |
| 00+ | 00+ | 00+ | 00− | 000 | 000 | 000 | 00− | 00+ | 000 | 000 | 00+ | 000 |
| 0+− | 0+− | 0+− | 0++ | 0+− | 0+− | 0+− | 0+0 | 0+− | 0+− | 0+− | 0+− | 0+− |
| 0+0 | 0+0 | 0+− | 0+0 | 0+0 | 0+− | 0+− | 0+0 | 0+0 | 0+0 | 0++ | 0+0 | 0++ |
| +−− | +−0 | +−− | +−0 | +−0 | +−− | +−0 | +−0 | +−− | +−− | +−0 | +−0 | +−0 |
| +−0 | +−+ | +−+ | +−+ | +−0 | +−0 | +−0 | +−0 | +−+ | +−0 | +−+ | +−+ | +−+ |
| +00 | +0− | +0− | +0− | +0− | +0− | +0− | +0+ | +00 | +00 | +0− | +0− | +0− |
| +0+ | +00 | +00 | +0+ | +0+ | +0+ | +0− | +0+ | +00 | +0+ | +0+ | +00 | +0+ |
| ++− | ++− | ++0 | ++− | ++− | ++0 | +++ | ++− | ++− | ++− | ++− | ++− | ++− |
| +++ | +++ | +++ | ++0 | +++ | +++ | +++ | ++− | +++ | +++ | ++0 | +++ | ++0 |

three types of projected three-factor designs and named them L18_3.1, L18_3.2 and L18_3.3 respectively, as in Table 2. Cheng and Ye [4], who permuted the levels of the projected designs to generate some additional 18-run designs, also found these eight 18-run three-factor design families.

However, by using the design search procedure of [9], we generated five other three-factor designs. They each have only one repeated point and cannot be formed either by putting two Latin squares together or by projection from the

Table 2: The relationship between the 18-run designs and Latin squares

| Design | Two Latin squares | #Distinct runs | |
|---|---|---|---|
| D1 | LS_wt+LS_wt | 18 | L18_3.1 |
| D2 | LS_wt+LS_wc | 18 | |
| D3 | LS_wt+LS_wt | 15 | L18_3.2 |
| D4 | LS_wt+LS_wt | 15 | |
| D5 | LS_wt+LS_wc | 15 | |
| D6 | LS_wc+LS_wc | 15 | |
| D7 | LS_wc twice | 9 | L18_3.3 |
| D8 | LS_wt twice | 9 | |

Table 3: A $(3 \times 3)/2$ Semi-Latin square

| | $-$ | 0 | $+$ |
|---|---|---|---|
| $-$ | 1 4 | 2 6 | 3 5 |
| 0 | 3 6 | 1 5 | 2 4 |
| $+$ | 2 5 | 3 4 | 1 6 |

Table 4: Correspondence between entries in semi-Latin square and labels in 3-level designs

| Label | D9 | D10 | D11 | D12 | D13 |
|---|---|---|---|---|---|
| $-$ | 3,4 | 5,6 | 4,6 | 1,4 | 1,4 |
| 0 | 5,6 | 4,5 | 1,5 | 2,3 | 5,6 |
| $+$ | 5,6 | 3,6 | 2,3 | 5,6 | 2,3 |

$L_{18}(3^7)$ array. They are labelled D9, D10, D11, D12 and D13 in Table 1. In fact, these new designs can be obtained from the $(3 \times 3)/2$ semi-Latin square labelled (b) in [8], shown in Table 3. Labelling the rows $(-, 0, +)$ and the columns similarly, as shown, these designs are obtained by labelling the numbers $(-, 0, +)$, as shown in Table 4. Designs D1-D6 can be obtained in a similar way from this semi-Latin square, whereas designs D7 and D8 can be obtained from the other $(3 \times 3)/2$ semi-Latin square, labelled (a) in [8], namely the inflated Latin square.

Studying the linear and quadratic contrasts of the factors for these 13 designs, we find that the linear and quadratic effects of a factor are sometimes correlated with interactions not involving that factor, and that the interactions are correlated with each other. These designs have different efficiencies for fitting the models that contain some interactions as well as main effects. Table 5 gives values of the $A_s$ criterion, i.e. the mean of the variances of the parameter estimates, excluding the intercept, for fitting a second-order model; see [1] for a fuller description of this and related optimality criteria. It shows that designs that are not formed by putting two Latin squares together, namely those labelled D9, D10, D11, D12 and D13, do not necessarily give less efficient parameter esti-

Table 5: Properties of designs for three factors in 18 runs

| Design | $A_s$ | Design | $A_s$ |
|---|---|---|---|
| D1 | 0.1334 | D9 | 0.1158 |
| D2 | 0.1187 | D10 | 0.1204 |
| D3 | 0.1121 | D11 | 0.1529 |
| D4 | 0.1688 | D12 | 0.1394 |
| D5 | 0.1741 | D13 | 0.1289 |
| D6 | 0.1420 | | |
| D7 | $\infty$ | | |
| D8 | $\infty$ | | |

mates than those obtained from putting two Latin squares together, even for the second-order model.

## Acknowledgements

The authors thank an anonymous referee for asking whether our new designs could be obtained from semi-Latin squares and the Managing Editor for many helpful suggestions on the presentation.

## References

[1] A. C. Atkinson and A. N. Donev (1992) *Optimum Experimental Designs*. Oxford: Oxford University Press.

[2] R. A. Bates, E. Buck, E. Riccomagno and H. P. Wynn (1996) Experimental design and observation for large systems (with discussion). *Journal of the Royal Statistical Society, Series B*, **58**, 77–111.

[3] S. W. Cheng and C. F. J. Wu (2001) Factor screening and response surface exploration. *Statistica Sinica*, **11**, 553–604.

[4] S. W. Cheng and K. Q. Ye (2004) Geometric isomorphism and minimum aberration for factorial designs with quantitative factors. *Annals of Statistics*, to appear.

[5] J. B. Clark and A. M. Dean (2001) Equivalence of fractional factorial designs. *Statistica Sinica*, **11**, 537–547.

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

[7] C. X. Ma, K. T. Fang and D. K. J. Lin (2001) On the isomorphism of fractional factorial designs. *Journal of Complexity*, **17**, 86–97.

[8] D. A. Preece and G. H. Freeman (1983) Semi Latin squares and related designs. *Journal of the Royal Statistical Society, Series B*, **45**, 267–277.

[9] P. W. Tsai, S. G. Gilmour and R. Mead (2000) Projective three-level main effects designs robust to model uncertainty. *Biometrika*, **87**, 467–475.

[10] J. C. Wang and C. F. J. Wu (1995) A hidden projection property of Plackett-Burman and related designs. *Statistica Sinica*, **5**, 235–250.