Quaternary-code Designs: A Better Design Choice for Experimentations

Frederick Kin Hing Phoa[1,2]
[1] Institute of Statistical Science, Academia Sinica, Taipei, Taiwan
[2] Corresponding Author e-mail: fredphoa@stat.sinica.edu.tw

Abstracts

There was a realization that nonregular designs could be utilized in conducting efficient experiments with flexibility, run size economy, and ability to exploit interactions, but the use of such designs was not common due to the lack of structure for systematic construction. Quaternary-code (QC) design, which is a new class of nonregular designs with similar structure to regular designs, attracts numerous attention from the researchers in design of experiments in the recent years. This talk discusses the advantages on using QC designs over regular designs when an experiment in scientific research or industrial application is conducted, including how the optimal QC design is constructed when some restrictions on experimental resources are given, and how the results from these experiments are analyzed in proper manner. Some real-life examples are given in order to compare the efficiency and economy between the experiments conducted under QC designs and regular designs.

Keywords: design of experiments, nonregular designs, Quaternary-code designs, design efficiency

1. Introduction and Motivation

There is considerable scope for reducing resources used in research by designing more efficient studies. Experiments are increasingly complex, in addition to rising experimental cost and competing resources. Careful design considerations with only minor variation in traditional designs can lead to a more efficient study in terms of more precise estimates or able to estimate more effects in the study at the same cost.

In many scientific researches and investigations, the interest lies in the study of effects of many factors simultaneously. One may choose a full factorial design which is able to estimate all possible level combinations of factors, but it usually involves many unnecessary trials. To be more cost-efficient, a fractional factorial design is suggested. A good choice of fractional factorial design allows us to study many factors with relatively small run size but enables us to estimate a large number of effects.

Designs that can be constructed through defining relations among factors are called regular designs, and all other designs that do not possess this kind of defining relation are called nonregular designs. Regular designs have been popular in scientific experimentations and industrial applications in the past decades because of its structural and analytical simplicities, see Wu and Hamada (2000) and Mukerjee and Wu (2006). Nonregular designs, however, have received particular attention in the past 20 years. It is well-recognized that nonregular designs have partial aliasing structure, and thus they outperform their regular counterparts with regard to some estimation measures (e.g. resolution, projectivity, etc.), and this is a major motivating force for the current surge of interest in these designs. A comprehensive review on the development of nonregular designs is referred to Xu, Phoa and Wong (2009).

2. How Partial Aliasing Structure is Important

Comparing with regular designs, nonregular designs are more difficult to analyze because some main effects may be partially aliased with interactions. Phoa, Wong and Xu (2009) studied two examples in Toxicology, illustrating that (i) nonregular designs enjoyed the run size economy without sacrificing the estimation abilities of the designs, and (ii) nonregular designs provided additional information of the interactions through their partially aliased structure with the main effects.

Nevertheless, the complex aliasing structure is a benefit because partially aliased effects can be estimated together. A key step is to disentangle the interactions from the estimates of the main effects. As Hamada and Wu (1992) and Phoa, Xu and Wong (2009) pointed out, ignoring non-negligible interactions can lead to (i) important effects being missed, (ii) spurious effects being detected, and (iii) estimated effects having reversed signs resulting in incorrectly recommended factor levels. These potential pitfalls were demonstrated through three examples of analytical chemistry in Phoa, Wong and Xu (2009).

3. Construction Method and Design Structure

One of the major obstacles for the use of nonregular designs is the lack of simple design structure. A recent major development in nonregular two-level designs has been the use of quaternary codes for their simple construction, and the resulting two-level designs are generally called QC designs. Xu and Wong (2007) pioneered research on QC designs and reported theoretical as well as computational results.

The study of the fundamental design structure began in Phoa and Xu (2009). Some concepts and notations are recalled here. A quaternary code takes on values from $Z_4 = \{0,1,2,3\}$. Let G by an $n \times (n + p)$ generator matrix $G = (V, I_n)$, where $V = \{\vec{v}_1, \cdots, \vec{v}_p\}$ is a matrix over Z_4 that consists of p vectors of lengths n and I_n is an $n \times n$ identity matrix. All possible linear combinations of the rows in G over Z_4 form a quaternary linear code, denoted by C. Then each Z_4 entry of C is transformed into two binary codes in its binary image $D = \phi(C)$ via the Gray map, which is defined as follows:

$$\phi: \quad 0 \rightarrow (+1, +1) \quad 1 \rightarrow (+1, -1) \quad 2 \rightarrow (-1, -1) \quad 3 \rightarrow (-1, +1)$$

Note that D is a binary matrix or a two-level design with 2^{2n} runs and $2n + 2p$ factors. It is easy to verify that the identity matrix I_n generates a full $2^{2n} \times 2n$ design. Therefore, the properties of D depend on the matrix V only. Moreover, D is mostly nonregular.

Zhang et al. (2011) suggested a mathematical expression of Gray map via a trigonometric approach. It reduces the calculation of design properties to a mathematical problem that involves some manipulations of trigonometric identities. The details are skipped due to the technical complications.

The above construction always leads to QC designs with 2 to the even power runs and even number of factors. In order to obtain QC designs with other types of dimensions, e.g. designs with odd number of factors, additional techniques are developed. Column deletion (Phoa, Mukerjee and Xu 2012) reduces the QC designs with 2^{2n} runs and $2n + 2p$ factors to those with 2^{2n} runs and $2n + 2p - 1$ factors. One needs to select a column among all $2n + 2p$ columns and deletes it. Column branching (Phoa and Xu 2009), on the other hands, reduces the QC designs to those with 2^{2n-1} runs and $2n + 2p - 1$ factors. Ones first selects a column as the branching column, then half number of rows with the same symbol ($+$ or $-$) in the branching column are deleted. Finally the branching column itself is deleted

4. Design Properties and Estimation Ability

Phoa (2012) generalized the characters of design properties into simple matrix formulations $K = CF$ and $A = BF$. In the first equation, K is a set of k-equation, which is similar to the wordlength in regular designs, describing the number of factors in a set such that the entrywise total product of these factors is not all equal to 0. F is a vector recording the frequencies of entry combinations appearing in V of generator matrix G and C is a weighing matrix, describing different weighing of each frequency towards different k-equations. Similar formulation is given in the second equation, where A is a set of a-equation describing the aliasing structure of the design and B is a weighing matrix describing different weighing of each frequency towards different a-equations.

In the first formulation, each class of QC design has its unique C and F, and thus leads to a unique K for design characterization. For example, given a $(1/4)^{\text{th}}$-fraction QC designs, which is a nonregular design possessing one-fourth number of rows of its corresponding full factorial designs, only two k-equations exist according to Phoa and Xu (2009): $K = (k_1, k_2)'$, $F = (f_0, f_1, f_2, f_3)'$ and

$$C = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$

A different combination of F may lead to a QC design that has different design properties. Some combinations of F are good and they result in QC designs with very good estimation abilities, while some are bad that even main effects are entangled with some important two-factor interactions. A careful choice of F is important and suggestions of these F are given in Phoa and Xu (2009) for $(1/4)^{\text{th}}$-fractions, Phoa, Mukerjee and Xu (2012) for $(1/8)^{\text{th}}$- and $(1/16)^{\text{th}}$-fractions, and Phoa, Chen and Lin (2013+) for $(1/64)^{\text{th}}$-fractions.

5. How Good the Quaternary-code Designs Are

There are many criteria to justify how good a design is. Among them, Phoa and Xu (2009) and its continuing works focus on three main criteria: Generalized resolution (Deng and Tang 1999), generalized minimum aberration (Tang and Deng 1999) and projectivity (Box and Tyssedal 1996). For illustration purpose, we consider only generalized resolution here.

Generalized resolution is a measure of the goodness of a design from the viewpoint of estimation power. More technically, when a design with m factors is considered, some of these m factors can be grouped into some nonorthogonal sets, or called words. The integer part of the generalized resolution counts the number of factors in the smallest set among all sets, and the decimal part describes how aliased these factors in this set are. Therefore, it is obvious that the higher the resolution, the higher the estimation power and the better the designs.

Figure 1 shows the comparison between the best QC designs and the best regular designs of the same size in terms of generalized resolution. In specific, each box in the figure represents the comparison result between two best designs. A red box means that the best QC design possesses higher generalized resolution than the best regular design in that dimension and a blue box means the reverse. A green box means that they have the same generalized resolution. If the box is white, the QC design in such dimension has not been discussed in Phoa and Xu (2009), Phoa, Mukerjee and Xu (2012) and Phoa, Chen and Lin (2013+) and they are either not published yet or still under investigation. A black box means a design in such dimension is not valid in the construction of QC designs. Some of them are half-fractions or full factorials and others contain repeated rows.

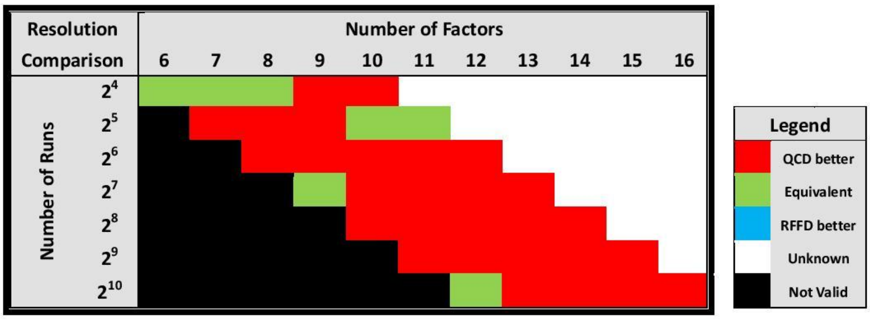

Figure 1. Resolution comparison between the best QC designs and regular designs.

The comparison in Figure 1 clearly shows that QC designs possess equivalent, or otherwise better, estimation power than the regular designs of the same size. Although there are some QC designs in larger dimensions that have lower generalized resolution due to the construction restrictions of QC designs, some additional twists in the construction immediately improves their generalized resolution. In addition, this comparison ends at 16 factors and 1024 runs in Figure 1. The comparisons in larger dimensions or in other criteria (e.g. minimum aberration, projectivity, etc.) can be found in the literature.

6. Conclusions

This paper briefly reviews how a nonregular design can be better choice than regular design when a scientific experiment or an industrial application is conducted. Although nonregular design possesses some good design properties, it is not as popular as regular designs do due to its lack of structure. Therefore, this paper introduces a new class of nonregular designs called Quaternary-code (QC) designs, which possesses a simple design structure constructed via quaternary linear codes. The design structure and design properties are briefly introduced thereafter, showing how these QC designs are systematically constructed and theoretically investigated. A resolution comparison to regular designs illustrates that the QC designs possess better estimation ability than the regular designs of the same size.

Acknowledgements

This work is supported by National Science Council of Taiwan ROC Grants 100-2118-M-001-002-MY2.

References

Box, G.E.P. and Tyssedal, J. (1996) “Projective properties of certain orthogonal arrays,” *Biometrika*, 83, 950-955.

Deng, L.Y. and Tang, B. (1999) “Generalized resolution and minimum aberration criteria for Plackett-Burman and other nonregular factorial designs,” *Statistica, Sinica*, 9, 1071-1082.

Hamada, M. and Wu, C.F.J. (1992) “Analysis of designed experiments with complex aliasing,” *Journal of Quality Technology*, 24, 130-137.

Mukerjee, R. and Wu, C.F.J. (2006) *A Modern Theory of Factorial Designs*, Springer, New York.

Phoa, F.K.H. (2012) “A code arithmetic approach for Quaternary code designs and its application to $(1/64)^{th}$-fractions,” *Annals of Statistics*, 40, 3161-3175.

Phoa, F.K.H., Chen, H.W. and Lin, S.C. (2013+) “One-thirty-second- and one-sixty-fourth-fraction Quaternary code designs with high resolution,” in review.

Phoa, F.K.H., Mukerjee, R. and Xu, H. (2012) “One-eighth- and one-sixteenth-fraction Quaternary code designs with high resolution,” *Journal of Statistical Planning and Inferences*, 142, 1073-1080.

Phoa, F.K.H., Wong, W.K. and Xu, H. (2009) “The need of considering the interactions in the analysis of screening designs,” *Journal of Chemometrics*, 23, 545-553.

Phoa, F.K.H. and Xu, H. (2009) “A class of quarter-fractioned nonregular designs constructed by Quaternary codes,” *Annals of Statistics*, 37, 2561-2581.

Phoa, F.K.H., Xu, H. and Wong, W.K. (2009) “The use of nonregular fractional factorial designs in combination toxicity studies,” *Food and Chemical Toxicology*, 47, 2183-2188.

Tang, B. and Deng, L.Y. (1999) “Minimum G_2-aberration for nonregular fractional factorial designs,” *Annals of Statistics*, 27, 1914-1926.

Wu, C.F.J. and Hamada, M. (2000) *Experiments: Planning, Analysis and Parameter Design Optimization*, Wiley, New York.

Xu, H., Phoa, F.K.H. and Wong, W.K. (2009) “Recent developments in nonregular

fractional factorial designs," *Statistics Surveys*, 3, 18-46.
Xu, H. and Wong, A. (2009) "Two-level nonregular designs from quaternary linear codes," *Statistica Sinica*, 17, 1191-1213.
Zhang, R., Phoa, F.K.H., Mukerjee, R. and Xu, H. (2011) "A trigonometric approach to Quaternary code designs with application to one-eighth and one-sixteenth fractions," *Annals of Statistics*, 39, 931-955.