Comprehensive description of main effects and interactions in Plackett Burman design

Yoshiki OISHI*1, Shu YAMADA1

1. Keio University • Hiyoshi 3-14-1, Yokohama, 223-8522

* yoshiki61216@keio.jp

Abstract:
Plackett Burman designs are used to estimate main effects because the columns assigned main effects are mutually orthogonal. However, correlations between estimates of a main effect column and a two-factor interaction column are not orthogonal in general. There are variations in the magnitude of correlation depending on the size of Plackett Burman design. For example, correlation coefficients between a main effect and a two-factor interaction are either 0 or ±0.33 in Plackett Burman design N=12. However, most of them are ±0.2 and some of them are 0.6 in Plackett Burman design N=20. When a correlation coefficient between main effects and two-factor interactions is significantly large, analysis results may be depart from the true status. Therefore, in this study, we comprehensively describe correlation coefficients between a main effect and a two-factor interaction for all designs proposed by the Plackett and Burman designs such that N=8,...,100. In addition, for designs that have large correlation coefficients, they are classified according to how to construct a design table and examined in detail. We describe certain rules in combination of a main effect and a two-factor interaction with large correlation coefficient. We also show how to construct designs that avoids large correlation based on the rules.

Keywords
triplets; main effect; two-factor interaction; correlation; design table

1. Introduction

In this study, we comprehensively describe correlation coefficients between of a main effect column and a two-factor interaction column in the Plackett Burman design. In experimental design, confounding always occurs in general when an experiment is performed using a certain number of factors. As shown in Hamada and Wu (1992), Lin (1993), Iida (1994) and Yamada (2004), Plackett Burman designs have a property that main effects do not confound each other, while main effects and two-factor interactions are partially confounded. In the followings, three-factor or more factors interactions will not be discussed, so we will write them as interactions simply. At this time, if interactions cannot be neglected, this confounding may cause an analysis result to be depart from the true status. Oishi and Yamada (2018) introduces an example that the analysis results are stable in N=12, but those become unstable in N=20 because of correlations. In N=12, correlation coefficients between a main effect column and an interaction column are all constant. On the other hand, in N=20, some of the correlations are large. When a large impact of interaction are appeared in a column with high correlation with main effect columns, analysis results will be depart from the true status. It implies that analysis results differ due to magnitudes and correlations between a main effect column and an interaction column. Also, the correlation coefficients differ depending on the runs size, say N, of design. In this paper, we examine magnitudes and variability of correlation coefficient in all design shown in Plackett and Burman (1946). In addition, we discuss a method with avoiding large correlation coefficients between a main effect column and an interaction column. The examined designs are shown in the original work by Plackett and Burman (1946). Table 1 summarizes N=12 Plackett Burman design that is very popular because of some good properties.

After comprehensively describing the correlation coefficients in Plackett Burman design, we propose a constructing method with avoiding large correlation coefficients for the design in N=8, ..., 100. Then we classify into four groups focusing on the method of constructing the design table. The proposed solution for each group is...
introduced for each group that can reduce the possibility that analysis results will differ from the true status. The proposed methods improve the usage of design tables, so it can be substituted without burden. Plackett and Burman (1946) provides orthogonal designs for \( N = 8, \ldots, 100 \) except 92 by showing a representative row for each \( N \). The designs are derived by circulating the representative row. While Plackett and Burman designs are popular when \( N \) is multiple of four but not \( 2^k \), we also consider the designs of \( N = 2^k \), such that 8, 16, 32, 64, because similar approach can be applied to improve the property.

Previous research

Wu (1993) uses Plackett Burman design \( N = 12 \) and 20 for constructing supersaturated designs. Supersaturated designs are one of the class of fractional factorial designs, where the number of factors is larger than the number of experiments. Wu (1993) focuses on the magnitude of correlations of paired columns from main effect columns and interaction columns and constructing supersaturated designs to avoid particularly large correlations. It is obtained that there are 48 pairs whose correlation coefficient is equal to 0.6 by a numerical search. On the other hand, Oishi and Yamada (2018) has focused on the correlation between a main effect column and interaction column in Plackett Burman design \( N = 20 \). It is obtained that there are 48 pairs whose correlation coefficient is equal to 0.6 by a numerical search. On the other hand, Oishi and Yamada (2018) has focused on the correlation between a main effect column and interaction column in Plackett Burman design \( N = 20 \). It is obtained that there are 48 pairs whose correlation coefficient is equal to 0.6 by a numerical search.

Designs are one of the class of fractional factorial designs, where the number of factors is larger than the number of runs. The proposed method improves the usage of design tables, so it can be substituted without burden. Table 2 summarizes the correlation coefficients of the main effects and interactions from \( N = 8 \) to 100 except 92 with \( 2^k \) type fractional factorial designs. We also consider the designs of \( N = 2^k \) because similar approach can be applied to improve the property.

| X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 |
|----|----|----|----|----|----|----|----|----|-----|-----|
| 1  | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1   | -1   |
| 2  | 1  | 1  | -1 | 1  | 1  | -1 | -1 | -1 | 1    | -1   |
| 3  | -1 | 1  | 1  | -1 | 1  | 1  | -1 | -1 | -1   | 1    |
| 4  | 1  | -1 | 1  | 1  | -1 | 1  | 1  | -1 | -1   | -1   |
| 5  | -1 | 1  | -1 | 1  | 1  | -1 | 1  | 1  | -1   | -1   |
| 6  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1  | 1    | -1   |
| 7  | -1 | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1    | 1    |
| 8  | 1  | -1 | -1 | 1  | -1 | 1  | 1  | -1 | 1    | 1    |
| 9  | 1  | 1  | -1 | 1  | -1 | 1  | 1  | -1 | 1    | -1   |
| 10 | 1  | 1  | 1  | -1 | -1 | 1  | 1  | -1 | 1    | 1    |
| 11 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | -1 | 1    | 1    |
| 12 | 1  | -1 | 1  | 1  | -1 | -1 | 1  | -1 | 1    | 1    |

2. Previous research

Wu (1993) uses Plackett Burman design \( N = 12 \) and 20 for constructing supersaturated designs. Supersaturated designs are one of the class of fractional factorial designs, where the number of factors is larger than the number of experiments. Wu (1993) focuses on the magnitude of correlations of paired columns from main effect columns and interaction columns and constructing supersaturated designs to avoid particularly large correlations. It is obtained that there are 48 pairs whose correlation coefficient is equal to 0.6 by a numerical search. On the other hand, Oishi and Yamada (2018) has focused on the correlation between a main effect column and interaction column in Plackett Burman design \( N = 20 \). It is obtained the all combinations of a main effect column and interaction column whose absolute value of correlation coefficient is equal to 0.6. The combinations are summarized in some graphs. It is found that the combination with a correlation of 0.6 had a relationship of the triplets. By removing columns that forms triplets from the design, the number of triplets included in the constructed design can be reduced than the case where triplets are not formed. The reduction of triplets means that the correlations are kept at small, so the possibility of the experimental result becoming unstable can be reduced. The number of triplets of the design constructed by this method is smaller than the random selecting approach by the following method. Therefore, Oishi et al., after finding large correlations, try to avoid the large correlation as much as possible using the same method. The verification method is as follows. First, \( m \) main effect columns are randomly selected from the design, therefore there are \( \binom{m}{2} \) interaction columns. The number of pairs whose correlation coefficient is equal to 0.6 in the correlation of \( m \) main effect columns and \( \binom{m}{2} \) interaction columns. This procedure is repeated 10000 times for every \( m \). Repeat this procedure for every \( m \). The proposed method provides lower or equal number of pairs whose absolute value of correlations coefficient is equal to the minimum number of the 10000 repetition.

3. Correlation coefficients list and classification of Plackett Burman design

3.1. Correlation coefficient list of Plackett Burman design

Table 2 summarizes the correlation coefficients of the main effects and interactions from \( N = 8 \) to 100 except 92 in the Plackett Burman designs. The type in Table 2 represents the method of constructing the Plackett Burman design. The one of the constructing methods is well known as a basic that constructs design by circulating one representative row. It is shown as C in the table 2. Plackett and Burman (1946) also provides a representative row for each \( N = 8, 16, 32 \) and 64, where the generated designs by circulating the representative row are equivalent with \( 2^k \) type fractional factorial designs. We also consider the designs of \( N = 2^k \) because similar approach can...
be applied to improve the property.
In this table, D is an abbreviation of Double, and indicates that the original design is fold over. For example, when N = 40 and the design table of N = 20 is A, it is constructed as follows.

\[
\begin{pmatrix}
-1 & A & A \\
1 & A & -A
\end{pmatrix}
\]

Block is different from the other method of construction and is constructed by circulating representative block units. In the table,” | Correlation |” indicates the absolute value of the correlation coefficient.
From this Table 2, we focus on those with particularly high correlation coefficients that are likely to affect the analysis results (in this study, the correlation coefficient is 0.6 or more, and those in bold). When the correlation coefficient is 0.6, the impact to sum of squares is equal to 36% that can be regarded as large influence. Therefore, we mainly focus on the case where the absolute value of correlation coefficient of 0.6 or higher. The design to be dealt with are designs with N = 8, 16, 20, 32, 40, 52, 56, 64, 88, 96 and 100. Because there are some characteristics depending on the size and method of design, N with large correlation coefficient is classified into four types in the following, and the discussion is advanced for each classification.

| N   | type | |Correlation| (between main effect and interaction) |
|-----|------|---|-------------|--------------------------------------|
| 8   | C    | 1 | 0           |
| 12  | C    | 0.33 | 0           |
| 16  | C    | 1 | 0           |
| 20  | C    | 0.6 | 0.2 | 0 |
| 24  | C    | 0.33 | 0           |
| 28  | Block | 0.429 | 0.143 | 0 |
| 32  | C    | 1 | 0           |
| 36  | C    | 0.333 | 0.111 | 0 |
| 40  | D20  | 1 | 0.6 | 0.2 | 0 |
| 44  | C    | 0.273 | 0.091 | 0 |
| 48  | C    | 0.333 | 0.167 | 0 |
| 52  | Block | 0.846 | 0.385 | 0.231 | 0.077 | 0 |
| 56  | D28  | 1 | 0.43 | 0.14 | 0 |
| 60  | C    | 0.2 | 0.067 | 0 |
| 64  | D32  | 1 | 0           |
| 68  | C    | 0.29 | 0.176 | 0.059 | 0 |
| 72  | C    | 0.22 | 0.11 | 0 |
| 76  | Block | 0.262 | 0.158 | 0.053 | 0 |
| 80  | C    | 0.2 | 0.1 | 0 |
| 84  | C    | 0.24 | 0.143 | 0.048 | 0 |
| 88  | D44  | 1 | 0.273 | 0.091 | 0 |
| 92  | NA   |     |            |
| 96  | D48  | 1 | 0.333 | 0.167 | 0 |
| 100 | Block | 0.92 | 0.28 | 0.2 | 0.12 | 0.04 | 0 |
3.2. Classification of Plackett Burman Design with large correlations

Table 3 shows classification category of the design to be improved where designs are extracted from Table 2 because of large correlation coefficient. In this table, $2^k$ is a two-level fractional factorial design constructed by power of 2. Double and Block are classifications according to how to make a design. N = 20 type has the property of N = 20. Here N = 40 is classified as N=20 type. Also, N = 40 and 64 are duplicates because they can be classified as Double, $2^k$, and N = 20 type. However, in this research, we consider N = 40 as N = 20 and N = 64 as $2^k$.

These classifications are performed because measures for large correlation coefficients are devised based on the method of constructing designs. Design in the category of Double include pairs whose correlation coefficient is equal to 1. The designs in Block category, there is a design without large correlation coefficient. In detail discussion is given in Section 4.3.

| classification | N   |
|----------------|-----|
| $2^k$          | 8, 16, 32, 64 |
| Double         | (40), 56, (64), 88, 96 |
| Block          | 52, 76 |
| N=20 type      | 20, 40 |

4. Suggestions for designs by classification

4.1. $2^k$ type

Table 4 summarizes the correlation coefficient values of main effects and interactions and the number of correlation coefficients in the $2^k$ type designs (N=8, 16, 32, 64). Table 4 shows that the design of $2^k$ has only -1 and 0 as the correlation coefficient between main effects and interactions. A correlation coefficient of -1 means that complete confounding exists. If a large interaction is completely confounded with the main effect, there is a good possibility that the analysis results may deviate from the real one. It is found that the category with the correlation coefficient equal to -1 is.

| N   | type | value correlation | # of correlation | correlation |
|-----|------|-------------------|------------------|------------|
| 8   | C    | Value correlation | -1               | 0          |
| 16  | C    | Value correlation | -1               | 0          |
| 32  | C    | Value correlation | -1               | 0          |
| 64  | D32  | Value correlation | -1               | 0          |

Next, we explain how to use these designs avoiding triplets. For example, let us consider the case of N=8. A color map representing a correlation coefficient of N=8 is summarized in Figure 1. There are seven triplets in all, (1,2,6), (1,3,4), (1,5,7), (2,3,7), (2,4,6), (3,5,6), (4,6,7). The combinations of main effect and interaction whose correlation is -1 are a relationship of triplets. Therefore, removing factors for each triplet will reduce the number of triplets in N=8, rather than assigning factors in order from the top. In fact, Table 5 compares the number of triplets by removing the factors that form a triplet of (1,2,6). The conventional type is to assign factors in order from the top. The proposed type is to remove the factors that form triplets. When m=4, a difference appears in the number of triplets. It can be seen that as N increases to 16, 32 and 64, the number of triplets is large in the conventional type and the proposed type.

[DOI : 10.17929/ Itq.6.43]  Copyright © 2020 Journal of the Japanese Society for Quality Control. All rights reserved.
Table 5. Comparison of the number of triplets between conventional and proposed type

| m   | base type (using N=8 in order from the top) | proposed type (removing triplet) |
|-----|------------------------------------------|----------------------------------|
|     | column to assign | # of triplets | column to assign | # of triplets |
| 7   | 1 2 3 4 5 6 7    | 7             | 1 2 3 4 5 6 7    | 7             |
| 6   | 1 2 3 4 5 6      | 4             | 2 3 4 5 6 7      | 4             |
| 5   | 1 2 3 4 5        | 2             | 3 4 5 6 7        | 2             |
| 4   | 1 2 3 4          | 1             | 3 4 5            | 0             |
| 3   | 1 2 3            | 0             | 3 4 5            | 0             |

Table 6. Double's correlation coefficient and its number

| N | type | value correlation | # of correlation |
|---|------|-------------------|------------------|
| 56 | D28  | -1                | 81               |
|    |      | -0.43             | 4212             |
|    |      | -0.14             | 14040            |
|    |      | 0                 | 46494            |
|    |      | 0.14              | 16848            |
| 88 | D44  | -1                | 129               |
|    |      | -0.273            | 25284            |
|    |      | -0.091            | 54180            |
|    |      | 0                 | 177246           |
|    |      | 0.091             | 57792            |
|    |      | 0.273             | 10836            |
| 96 | D48  | -1                | 141               |
|    |      | -0.333            | 12972            |
|    |      | -0.167            | 51888            |
|    |      | 0                 | 307286           |
|    |      | 0.167             | 51888            |

Table 7. Result of removing the first column from Table 6

| N | type | value correlation | # of correlation |
|---|------|-------------------|------------------|
| 56 | D28  | -0.43             | 4212             |
|    |      | -0.14             | 14040            |
|    |      | 0                 | 42174            |
|    |      | 0.14              | 16848            |
| 88 | D44  | -0.273            | 25284            |
|    |      | -0.091            | 54180            |
|    |      | 0                 | 166238           |
|    |      | 0.091             | 57792            |
|    |      | 0.273             | 10836            |
| 96 | D48  | -0.333            | 12972            |
|    |      | -0.167            | 51888            |
|    |      | 0                 | 294126           |
|    |      | -0.167            | 51888            |

4.2. Double type

Table 6 summarizes the correlation coefficient values of main effects and interactions and the number of correlation coefficients in Double category designs (N=56, 88, 96). In Double, the absolute value of correlation coefficient is the largest at -1. As the second largest one has correlation coefficient less than 0.6, the category with the correlation coefficient equal to -1 have to be considered.

In this category, the first column (X1) is confounding with the other columns because of how to construct the designs. Select one arbitrary row a from original A. In the designs of Double, the correlation between an Hadamard product (element-wise product) of the first column and the a+1 column in left side of A and the a+1 column in the other side of A is always equal to -1. Therefore, all pairs whose correlation coefficient is equal to -1 have a relation of triplet including the first column. Table 7 summarizes the correlation coefficient when the first column is removed from the design table and the number of coefficients. By removing the first column, all the largest correlation coefficient can be eliminated.
4.3. Block type

This Block type includes N= 28, 76, 52 and 100, where the constructing method consists of circulation of matrices as block. For example, the N=28 design is generated by \((9 \times 9)\) design matrices. As shown in Table 3, we first focus on N= 52 and 100 designs because of existing high correlation between columns. Some comments on 28 and 76 are given after the improvements of N=52 and N=100. We think about dealing with these.

In this category, such as Double, the first column (X1) is confounding with the other columns. Except the first column (X1), other columns are circulated in block units. Also, the largest correlation is the relationship of triplets including all the first columns. In case of N=52, all triplets can be expressed as \((1, 2n, 2n+1)\) (where \(n=1,2, \ldots, 25\)). Therefore, the largest correlation coefficients can be eliminated by removing the first column. Table 9 summarizes the correlation coefficient values when the first column is removed from Table 8. It is confirmed that the largest absolute value of correlations are eliminated.

We do not discuss N=28, 76 designs in details. While these are the same construction method as N=52, 100, but there exists no pairs of large absolute value of correlation coefficient. The reason for this is that in N=28, there is no independent the first column, and it is constructed by circulating three types of design tables \((9 \times 9)\). It is considered that the correlation coefficient is distributed more uniformly than other designs in Block category. In N=76, because the circulating design table is \(2 \times 2\), the first column is alternately arranged 1 and -1. Therefore, it is considered that the first column in N=76 isn’t independent like other Block’s designs.

Table 8. Block's correlation coefficient and its number

| N  | type  | correlation          | # of correlation |
|----|-------|----------------------|------------------|
| 52 | Block | -0.385 \ -0.231 \ -0.077 | 1200 \ 5400 \ 33300 \ 2550 \ 15900 \ 5400 \ 1200 | 75                |
| 100| Block | -0.280 \ -0.200 \ -0.120 \ -0.040 | 7056 \ 28224 \ 28224 \ 234024 \ 9702 \ 109368 | 147               |

Table 9. Result of removing the first column from Table 8

| N  | type  | correlation          | # of correlation |
|----|-------|----------------------|------------------|
| 52 | Block | -0.385 \ -0.231 \ -0.077 | 1200 \ 5400 \ 30600 \ 2450 \ 15900 | 1200               |
| 100| Block | -0.280 \ -0.200 \ -0.120 \ -0.040 | 7056 \ 28224 \ 28224 \ 223440 \ 9506 | 105840 \ 28224 \ 28224 | 7056               |

4.4. N=20 type

Table 10 summarizes the correlation coefficients of a main effect column and an interaction column for designs in N = 20 type (N = 20, 40). As for N=20, it is omitted because it is all described in Oishi and Yamada (2018). In case of N=40, this design is constructed by double of N=20. Therefore, it has similar property of Plackett Burman design N=20. Therefore, the correlation coefficient -1 can be eliminated by removing the first column as described above. In addition, for 0.6-triplets, by removing the columns that forms triplet, it is possible to reduce the number of large correlations.
### Table 10. N=20 type's correlation coefficient and its number

| N type | value correlation | # of correlation |
|--------|-------------------|------------------|
| 20     | -0.600 0.200 0.200 | 171 1539 342 1197 |
| 40     | -1 -0.600 -0.200 0 0.200 | 57 684 6156 17214 4788 |

### 5. Conclusion

In this paper, the correlation coefficient between a main effect column and an interaction column in Plackett and Burman designs: N = 8, ..., 100 is discussed in a comprehensive approach. The proposed approach constructs designs by eliminating large correlation coefficients. By using this proposed method, it is expected to make the analysis result more stable than the random approach, sequential assignment approach such as from the first column.

Table 11 summarizes the values of the correlation coefficient and the number N of which the number of correlation coefficients has changed significantly. There is no significant change in $2^k$ type designs and the designs in N=20 type. However, the numbers of triplets included in the design are decreasing. Designs in Double and Block type have improved properties that the maximum values of correlation coefficients are all less than 0.6. Table 12 summarizes the results of reducing the number of those that removed particularly large correlation coefficients. Table 12 also shows that it is able to avoid the large absolute value of correlation such that larger than 0.6.
Table 11. Compare original with proposed design

| N  | type | original design | proposed design |
|----|------|-----------------|-----------------|
|    |      | Maximum of correlation coefficient | frequency | # of columns | Maximum of correlation coefficient | frequency | # of columns |
| 8  | C    | 1               | 1              | 4             | 1               | 0          | 4             |
| 20 | C    | 0.6             | 7              | 10            | 0.6             | 3          | 10            |
| 52 | Block | 0.846          | 75             | 51            | 0.358           | 1200       | 50            |
| 56 | D28  | 1               | 81             | 55            | 0.430           | 4212       | 54            |
| 88 | D44  | 1               | 129            | 87            | 0.273           | 25284      | 86            |
| 96 | D48  | 1               | 141            | 95            | 0.333           | 12972      | 94            |
| 100| Block | 0.920          | 147            | 99            | 0.280           | 7056       | 98            |

Table 12. Correlation coefficient list after improvement

| N  | type | [Correlation] |
|----|------|---------------|
| 8  | C    | 1             |
| 12 | C    | 0.330         |
| 16 | C    | 1             |
| 20 | C    | 0.600         |
| 24 | C    | 0.330         |
| 28 | Block | 0.429    |
| 32 | C    | 1             |
| 36 | C    | 0.333         |
| 40 | D20  | 0.600         |
| 44 | C    | 0.273         |
| 48 | C    | 0.333         |
| 52 | Block | 0.385    |
| 56 | D28  | 0.430         |
| 60 | C    | 0.200         |
| 64 | D32  | 1             |
| 68 | C    | 0.290         |
| 72 | C    | 0.220         |
| 76 | Block | 0.262    |
| 80 | C    | 0.200         |
| 84 | C    | 0.240         |
| 88 | D44  | 0.273         |
| 92 | NA   |               |
| 96 | D48  | 0.333         |
| 100| Block | 0.280    |

Acknowledgement

The authors sincerely thank the reviewers for their constructive comments.
References

[1] Oishi, Y. and Yamada, S., (2018), Evaluation on alias relation of interactions in Plackett Burman design and its application to guide assignment, proceedings of ANQ 2018, pp. 373-380.
[2] Wu, J., (1993), Construction of supersaturated designs through partially aliased interactions, Biometrika, 80, (3), pp. 661-669.
[3] Lin, D. K. L., (1993), A New Class of Supersaturated Designs, Technometrics, 35, (1), pp. 28-31.
[4] Hamada, M. and Wu, J., (1992), Analysis of designed Experiments with Complex Alisaing, Journal of Quality Technology, 24, (3), pp. 130-137.
[5] Plackett, R. L. and Burman, J. P., (1946), Some Generalizations in the Multifactorial Design, Biometrika, 33, (4), pp. 328-332.
[6] Yamada, S., (2004): Design of Experiments, Nikkagiren Publication, pp.108-119, (in Japanese).
[7] Iida, T., (1994), A construction Method of Two Level Supersaturated Designs Derived from L12, Japanese Journal of Applied Statistics, 23, (3), pp. 147-153.
[8] Montgomery, D. C., (2001): Design and Analysis of Experiments 5th edition, John Wiley and Sons, pp.155-164.
[9] John, J. A., (1986): Cyclic Designs –(Monographs on Statistics and Applied Probability), Chapman and Hall, pp. 40-60.

Author’s biographical notes:

Yoshiki Oishi is a graduate student in the Department of School of Science for Open and Environmental Systems, Keio University. His research interests are mainly statistical quality control.

Dr. Shu Yamada is a professor, Faculty of Science and Technology, Keio University. He teaches applied statistics and quality management at undergraduate and graduate schools. His research interests focus on total quality management, design of experiments for quality, statistical quality control and so forth.

[DOI : 10.17929/ tqs.6.43]

Received: May 6, 2019
Revised: March 29, 2020
Accepted: April 26, 2020