Construction of Super Saturated Design Using Hadamard Matrix and Its $E(S^2)$ Optimality

Salawu, I.S.,

Department of Statistics, Air Force Institute of Technology, Kaduna.

Abstract

Supersaturated design is essentially a fractional factorial design in which the number of potential effects is greater than the number of runs. In this paper, a super-saturated design is constructed using half fraction of Hadamard matrix of order N. A Hadamard matrix of order N, can investigate up to N – 2 factors in N/2 runs. Result is shown in N = 16. The extension to larger N is adaptable.

Keywords: Super-saturated design, Hadamard Matrix, Optimality, Orthogonal, Fractional Factorial, Aberration.

ISSN 2229-5518

1.0 Introduction

A supersaturated design is a fractional factorial design in which the number of factors *m* exceeds the number of runs *n*. A design is called supersaturated when $m \geq n$-1. Satterthwaite (1959), Watson (1961), Booth and Cox (1962), Wu (1993), and Salawu (2015) reviewed work on the supersaturated design. Lin (1993) and Salawu (2014) uses half fractions of Hadamard matrices to construct supersaturated designs, taking a column as the branching column and then constructing two half fractions according to the sign of that column. Wu (1993) augments Hadamard matrices by adding interaction columns. Because of the special features of the supersaturated designs, there exist all kinds of construction methods. However, the designs may not be the best in terms of the criteria used for generation. Butler et.al (2001) proposed a construction method for improved supersaturated designs by the criterion of E(s^2).

Lin (1993), Li and Wu (1997) motivated towards the construction of super-saturated designs through column wise, pair wise exchanges. They differ from the k-exchange algorithms in two aspects, one is, they exchange columns instead of rows of the design matrix and another one is, they employ a pair wise adjustment in the search for a better column. Deng, Lin and Wang (1999) studied the properties of super-saturated designs and proposed a criterion based on the projection property called resolution rank. This was discussed in Ameen and Bhatra (2016)

Fang, Lin, and Ma (2000) proposed a construction procedure by embedding a saturated orthogonal design into a uniform design of the same row size. They adopt the collapsing method from Addelman. The basic idea of the construction method is, to collapse a multi-level factor into several low-level factors, where they collapse U-type uniform designs. They proposed five criteria for comparing multi-level super-saturated designs. As contained in Ameen and Bhatra (2016).

Fang Kaitai, Gennian G.E., and Liu Minqian (2004) proposed a combinatorial approach called the packing method. They studied the connection between orthogonal arrays and resolvable packing designs for constructing optimal super-saturated designs and properties of the resulting designs are also proposed. Aggarwal and Gupta (2004) proposed construction method for multi-level super-saturated designs based on Galois field theory. Xu and Wu (2005) proposed construction methods for multi-level super-saturated designs inspired by Addelman- Kempthorne of orthogonal arrays and proposed new lower bound for multi-level super-saturated designs.

Hung, C., Lin, D.K.J., and Liu M.Q. (2012) proposed a new criterion for supersaturated designs with quantitative factors. Mbegbu and Todo (2012) proposed E(s^2)-optimal supersaturated designs with an experimental run size n=20 and number of factors m=57 (multiple of 19). This construction is based on BIBD using a theorem proposed by Bulutoglu and Cheng. This was discussed in Ameen and Bhatra (2016).

Liu, Y., and Liu, M.Q. (2013) proposed complementary design method. The basic principle of this method is that for any existing E(fNOD) optimal supersaturated whose E(fNOD) value reaches its lower bound, its complementary design in the corresponding maximal balanced design is also E(fNOD) optimal. This method applies to both symmetrical and asymmetrical cases.

2.0 E(s^2)-Optimal Designs

We assume that each factor has two levels, coded +1 and −1, often written as "+" and "– ". As in almost all the literature, we assume that each factor is observed at each level an equal number of times, although Allen and Bernshteyn (2003) recently relaxed this assumption. Ameen and Bhatra (2013, 2014) etc. made attempts on the construction of super-saturated designs with their E(s^2) optimality.

3.0 Criteria of Optimality

Consider the "main effects only" model,

$$Y = \beta_0 + \sum_{j=1}^{f} \beta_j X_j + \varepsilon, \quad (1)$$

where Y is a response variable, $\beta_0, \beta_1, \ldots, \beta_f$ are unknown parameters, $x_1, \ldots, x_f$ are the coded levels of the *f* factors, ε is an error term with E(ε) = 0 and V(ε) = σ^2, and error terms are independent. The model in matrix notation is

$$Y = X\beta + \varepsilon,$$

The E(s^2) optimality criterion has been the primary criterion used for comparing different supersaturated designs. A smaller value is preferred and a value of zero represented orthogonality. He constructed supersaturated designs based upon the E(s^2) criterion, which is used to measure orthogonality of design matrices. Let S_{ij} denote the dot product of columns i and j in the design matrix X, then the criterion is given by

$$E(s^2) = \frac{\sum s_{ij}^2}{\binom{m}{2}} \qquad (3)$$

where m is the number of columns of the design and $S_{ij} = \{C_i, C_j\}$ is the inner product of the i[th] column C_i and the j[th] column C_j. Then we propose a general method for constructing supersaturated designs, also using Hadamard matrices. The E(s²) criterion gives an intuitive measure of the degree of non-orthogonality, the smaller the better. Tang and Wu (1997) in Salawu (2014) derived a lower bound for E(s²), meaning optimal with respect to E(s²). The lower bound of E(s²), which is a function of n and m, is defined as

$$E(s^2) \geq \frac{n^2(m-n+1)}{(m-1)(n-1)} \qquad (4)$$

(m ≥ n), where m is the number of columns of the design and n is the number of runs. Similar results were independently obtained by Nguyen (1996), as contained in Salawu (2015). Among other things, Nguyen (1996) established the optimality of half fractions of Hadamard matrices.

Table 1: Designs for 14 factors in 8 runs.

Factors													
1	2	3	4	5	6	7	8	9	10	11	12	13	14
+	+	+	+	+	+	+	+	+	+	+	+	+	+
-	-	+	+	-	-	+	+	-	-	+	+	-	-
-	+	-	-	+	-	+	+	-	+	-	-	+	-
+	-	-	-	-	+	+	+	+	-	-	-	-	+
-	+	-	+	-	+	-	-	+	-	+	-	+	-
+	-	-	+	+	-	-	-	-	+	+	-	-	+
+	+	+	-	-	-	-	-	-	-	-	+	+	+
-	-	+	-	+	+	-	-	+	+	-	+	-	-

In a supersaturated design, even for this main effect only model, the matrix X'X is nonsingular, where 'denotes transpose, and so no unique least squares estimates of the parameters β can be obtained. Consider, for example, the small supersaturated design shown in Table 1.

$$X'X = \begin{bmatrix} 8&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&8&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&8&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&8&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&8&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&8&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&8&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&8&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&8&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&8&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&8&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&8&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&8&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&8 \end{bmatrix}$$

The diagonal elements of this matrix are fixed by the number of runs. Best of all is when the off-diagonal elements are all zero, in which case all main effects are estimated independently. It is, of course, impossible to get an X'X matrix for 14 factors in 8 runs with rank greater than 8. However, these considerations suggest that a good design will be one that makes the off-diagonal elements as small as possible (in absolute value). Letting the *(i, j)*[th] element of X'X be s_{ij}, Booth and Cox (1962) suggested two criteria based on the sizes of the s_{ij}. The first criterion they used was to choose a design with minimum $\max_{i \neq j} |S_{ij}|$, and among all such designs to choose one with the fewest s_{ij} that achieves this maximum.

The second suggestion of Booth and Cox was to choose a design that minimizes

$$E(S^2) = \frac{2}{f(f-1)} \sum_{i<j} S_{ij}^2 \qquad (5)$$

This has become the most used criterion in the literature on supersaturated designs. Sometimes these two criteria are combined, for example, by choosing a supersaturated design that minimizes E(s²) subject to some upper bound on $\max_{i \neq j} |S_{ij}|$. Cheng and Tang (2001) gave upper bounds on the number of factors that could be included in a supersaturated design subject to $\max_{i \neq j} |S_{ij}| \leq c$, where c is a constant.

4.0 Hadamard Matrix

The use of Hadamard matrices to construct supersaturated designs that can examine k = N – 2 factors in n = N /2 runs, where N is the order of the normalized Hadamard matrix used. The first column of all 1's is not taken into consideration since it is fully aliased with the mean. Then, we choose a branching column out of the remaining N – 1 column and we split the N runs into two groups. Group I contain all the runs with the sign +1 in the branching column and Group II contains the remaining runs. Then by deleting the branching column either from Group I or Group II causes the remaining N – 2 columns to form a super saturated design to examine k = N – 2 factors in N/2 runs.

ISSN 2229-5518

Table 2: Supersaturated Design Derived from the Hadamard Matrix of order 16

$$H_{16} = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & : & GI \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & : & GII \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & : & GII \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & : & GI \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & : & GII \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & : & GI \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & : & GI \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & : & GII \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & : & GII \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & : & GI \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & : & GI \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & : & GII \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & : & GI \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & : & GII \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & : & GII \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & : & GI
\end{bmatrix}$$

Deleting the first column and the branching columns from GI, results in a matrix with 8 rows and 14 columns. This resulting matrix is a supersaturated design that examines N-2=14 factors in $\frac{N}{2} = 8$ runs.

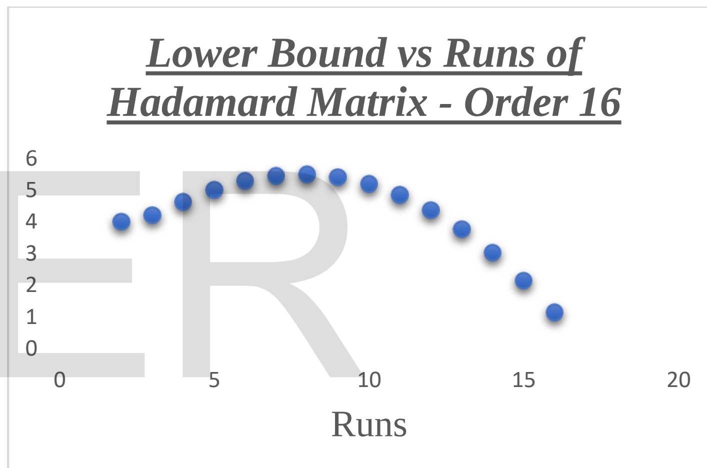
$$X = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
-1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 \\
-1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\
-1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1
\end{bmatrix}$$

In the above matrix, there are 8 rows and 14 candidate columns to construct a supersaturated design for 10 factors. Any combination consists of 10 factors out of the 14 candidate columns could be a candidate column could be a candidate solution. There will be 1001 candidate solutions in this case. As an example, one solution could be formed by selecting the first 10 columns of the matrix. The value of s_{ij} for the design matrix, for example the value of S_{23} is equal to and is calculated as follows:

$$S_{23} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = 0$$

The value of $E(s^2)$ for every design is calculated and the lowest value is picked to be the best design among the others.

Fig: Lower Bound against Runs of Hadamard Matrix of Order 16



5.0 Conclusion

The plot of lower bounds of Hadamard 16 at different run sizes indicated that optimum is reached in the Hadamard matrix of order 16 when the runs is at 8, and the designs is also at orthogonal. The result of the Hadamard Matrix of order 16 used in the construction of super-saturated design tallied with the findings of Nguyen (1996) and Salawu (2013).

References

[1]. T. T. Allen, and M. Bernshteyn, (2003). Supersaturated designs that maximize the probability of identifying active factors. Technometrics, 45, 90–97.

[2]. M. L. Aggarwal and S. Gupta (2004): A new method of construction of multi-level supersaturated designs. *Journal of Statistical Planning and Inference.* 2004. 121; 127-134.

[3]. S. K. Ameen Saheb and N. C. Bhatra (2016): Construction of Super-Saturated Designs, *International Journal of Advanced Mathematics and Statistics* 2016, Volume 2, Issue 1, pp. 58-63, Article ID Sci-500 ISSN 2348-5175

[4]. S. K. Ameen and N.C. Bhatra (2013) A New method for the construction of Super- Saturated design. *International Journal of Technology and Engineering Science*. 1 (6) 622-624.

[5]. S. K. Ameen Saheb and N. C. Bhatra (2016): Construction of Super-Saturated Designs. International Journal of Advanced Mathematics and Statistics, Volume 2, Issue 1, pp. 58-63, Article ID Sci-500 ISSN 2348-5175

[6]. K. H. V. Booth and D. R. Cox (1962) Some systematic supersaturated designs. *Technometrics.* 4; 489- 495.

[7]. N. A. Butler, R. Mead, K. M. Eskridge, and S. G. Gilmour (2001). A general method of constructing E(s^2)-optimal supersaturated designs. Journal of the Royal Statistical Society B, 63, 621–632.

[8]. C. S. Cheng and B. Tang (2001). Upper bounds on the number of columns in supersaturated designs. Biometrika, 88, 1169–1174.

[9]. L. Y. Deng, D. K. J. Lin, and J. N. Wang (1999): A resolution rank criterion for supersaturated designs. *Statistica Sinica.* 9; 605-610.

[10]. K. T. Fang, D. K. J. Lin, and C. X. Ma (2000): On the construction of multi-level supersaturated designs. *Journal of Statistical Planning and Inference*. 86; 239-252.

[11]. N. K. Nguyen (1996) An algorithm approach to constructing supersaturated designs. *Technometrics.* 38; 69-73.

[12]. D. K. J. Lin (1993): A new class of Supersaturated designs. *Technometrics.* 35; 28-31.

[13]. Y. Liu, and M. Q. Liu (2013): Construction of supersaturated design with large number of factors by the complementary design method. *Acta Mathematicae Applicatae Sinica*. English Series. 29; 253-262.

[14]. N. K. Nguyen (1996): An algorithmic approach to constructing supersaturated designs. Technometrics, 38, 69–73.

[15]. F. Satterthwaite (1959): Random Balance Experimentation. *Technometrics.* 1; 111-137.

[16]. I. S. Salawu. (2013): Construction of Near Balance and Orthogonal Mixed-Level Fractional Factorial Designs, Ph.d Dissertation

[17]. I. S. Salawu, B. L. Adeleke, and G. M. Oyeyemi, (2015): Review of Classical Methods in Supersaturated Designs (SSD) for factor Screening. *Journal of Mathematical Theory and Modeling*, www.iiste.org , ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online), Vol.5, No.5, 2015.

[18]. B. Tang and C. F. J. Wu (1997): A method for constructing supersaturated designs and its E(s2)-optimality. *Canadian Journal of Statistics.* 25; 191-201.

[19]. G. S. Watson (1961): A study of the group screening method. Technometrics, 4, 371-388.

[20]. C. F. J. Wu (1993): Construction of supersaturated designs through partially aliased interactions. Biometrika, 80, 661-669.

[21]. Xu, H., and Wu, C.F.J. (2005): Construction of optimal multi-level supersaturated designs. *Annals of Statistics.* 2005. 33; 2811-2836.

Salawu, I. Saheed has a Ph.D in Statistics and he is a lecturer in Statistics Department, Air Force Institute of Technology, Kaduna, Nigeria.

Email: salahus74@gmail.com