Statistics and Applications {**ISSN 2452-7395 (online)**}
Volume 17, No. 1, 2019 (New Series), pp 85-103

Characterization and Optimal Designs for Discrete Choice Experiments

Feng-Shun Chai[1], **Ashish Das**[2] **and Soumen Manna**[2]
[1]*Academia Sinica, Taipei, Taiwan*
[2]*Indian Institute of Technology Bombay, Mumbai, India*

Received: November 29, 2017; Accepted: February 15, 2019

Abstract

In discrete choice experiments, a choice design involves n attributes (factors) with i-th attribute at l_i levels, and there are N choice sets each of size m. Demirkale, Donovan and Street (2013) considered the setup of symmetric factorials ($l_i = l$) and obtained D-optimal choice designs under main effects model in the absence of two or higher order interaction effects. They provide some sufficient conditions for a design to be D-optimal. In this paper, we first derive a modified Information matrix of a choice design for estimating the factorial effects of a $l_1 \times l_2 \times \cdots \times l_n$ choice experiment. For a 2^n choice experiment, following Singh, Chai and Das (2015), under the broader main effects model (both in the presence and in the absence of two-factor interactions) we give a simple necessary and sufficient condition for the Information matrix to be diagonal. Furthermore, we characterize the structure of the choice sets which gives maximum $trace$ of the Information matrix. Our characterization of such an Information matrix facilitates construction of universally optimal choice designs for estimating main effects, both in the presence and in the absence of two-factor interactions but, in the absence of three or higher order interaction effects.

Key words: Choice sets; Choice design; Factorial design; Resolution; Main effects; Hadamard matrix.

1 Introduction

Discrete choice experiments are widely used in various areas including marketing, transport, environmental resource economics and public welfare analysis. A choice experiment consists of a number of choice sets, each containing several options (alternatives, profiles or treatment combinations). Respondents are shown each choice set in turn and are asked which option they prefer, as per their perceived utility, in each of the choice sets presented. Each option in a choice set is described by a set of attributes (factors), each with some number of levels. We assume that there are no repeated options in a choice set. We describe the options which are being compared, by n attributes with i-th attribute at l_i levels ($l_i \geq 2$), and that all the choice sets in a $l_1 \times l_2 \times \cdots \times l_n$ choice experiment have m options. It is ensured that respondents choose one of the options in each choice

Corresponding Author: Ashish Das
E-mail: ashishdas.das@gmail.com

set (termed *forced choice experiment* in the literature). A choice design is a collection of choice sets employed in a choice experiment. A choice design comprises N such choice sets. Recently, Großmann and Schwabe (2015) present a review of designs for discrete choice experiments.

Earlier, Street and Burgess (2007) presented a comprehensive exposition of designs for choice experiments under multinomial logit (MNL) model. MNL model specifies the probability that an individual will choose one of the m alternatives, say s_i, from a choice set S (say). The probability is given as the exponential of the expected utility of that alternative s_i, divided by the sum of all the exponentiated expected utilities. Mathematically,

$$P(s_i|S) = \frac{e^{V_i}}{\sum_{j=1}^{m} e^{V_j}}, \tag{1.1}$$

where V_i is the utility measure represented by the treatment combination effect for a $l_1 \times l_2 \times \cdots \times l_n$ factorial. For more detailed discussion on MNL model and choice experiments, see Train (2009) and Street and Burgess (2007).

Demirkale, Donovan and Street (2013) considered the setup of symmetric factorials ($l_i = l$) and obtained D-optimal choice designs under main effects in the absence of two or higher order interaction effects. They provide some sufficient conditions for a designs to be D-optimal.

In this paper, we first derive a modified Information matrix of a choice design for estimating the factorial effects. Such a modification is fundamental to the study of optimal choice designs since the modification provides the desired additive property to the Information matrix. It overcomes the existing shortcoming of situations where with addition of a choice set the information content of the design decreases. For a 2^n choice experiment, under the broader main effects model (both in the presence and in the absence of two-factor interactions) we give a simple necessary and sufficient condition for the Information matrix to be diagonal. Furthermore, we characterize the structure of the choice sets which gives maximum $trace$ of the Information matrix. Our characterization of such an Information matrix facilitates construction of universally optimal choice designs giving more flexibility for choosing m. Finally, we provide universally optimal choice designs (optimal in the class of all designs with given N, n and m) for estimating main effects, both in the presence and in the absence of two-factor interactions but, in the absence of three or higher order interaction effects.

2 Information Matrix

In choice experiment we deal with multiple independent populations which have common parameters. In a choice experiment, each choice set represent a different population. We call this set of populations as *associated populations.* When sampling from such *associated populations*, Bradley and Gart (1962) have presented related assumptions and asymptotic properties of the ML estimators. Under these assumptions, EI-Helbawy and Bradley (1978) have derived large sample results for paired choice experiments when each choice item is coming from a factorial setup. Later Street and Burgess (2007)generalized the setup for choice set size m and obtained the Information matrix on similar lines. It is seen that their Information matrix is derived using the averaging principle leading to situations where adding more choice sets to a design leads to information matrix with less information content than the information matrix of the original design. In what follows, we adopt an approach different from Bradley and Gart (1962) and EI-Helbawy and Bradley (1978). We derive a slightly modified Information matrix of a choice design for estimating the factorial effects. Such a modification gives the Information matrix the desired additive property. Our approach addresses a possible lacuna in the current non-additive form of the Information matrix.

Let X_i be a random variable over the region R_i, independent of $\theta = (\theta_1, \theta_2, \ldots, \theta_k)'$, an unknown parameter vector lying on a k-dimensional interval Ω. Furthermore let $f_i(x_i; \theta)$, $i = 1, 2, \ldots, n^*$; be the pdf or pmf of X_i from n^* different *associated populations*. It is not necessary that each f_i depends on all $\theta_1, \theta_2, \ldots, \theta_k$. Let $\mathbf{X_i} = (X_{i_1}, X_{i_2}, \ldots, X_{i_{n_i}})$ be a random sample of size n_i, from f_i. Then the likelihood function corresponding to it is

$$\mathcal{L}_i = \prod_{j=1}^{n_i} f_i(X_{i_j}; \theta) = \mathbf{f_i}(\mathbf{X_i}; \theta). \tag{2.1}$$

According to Fisher (for more details see Rao (1973)), the Information contained in the sample $\mathbf{X_i}$ is denoted by the information matrix $\mathcal{I}_i = ((\mathcal{I}_{i(rs)}(\theta)))_{k \times k}$, where

$$\mathcal{I}_{i(rs)}(\theta) = \int_{R_i} \frac{\partial \ln \mathbf{f_i}}{\partial \theta_r} \frac{\partial \ln \mathbf{f_i}}{\partial \theta_s} f_i dx_i = E\left(\frac{\partial \ln \mathbf{f_i}}{\partial \theta_r} \frac{\partial \ln \mathbf{f_i}}{\partial \theta_s}\right) \tag{2.2}$$

is non-negative definite.

Now if we take random sample $\mathbf{X_i}$ of size n_i, from each of the n^* *associated populations* f_i, then the likelihood function of θ for all the samples $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_{n^*}}$; can be written as

$$\mathcal{L} = \prod_{i=1}^{n^*} \mathbf{f_i}(\mathbf{X_i}; \theta). \tag{2.3}$$

We define the information for θ contained in all the samples $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_{n^*}}$; from n^* *associated populations* by the Information matrix $\mathcal{I} = ((\mathcal{I}_{rs}(\theta)))_{k \times k}$ with

$$\mathcal{I}_{rs}(\theta) = \sum_{i=1}^{n^*} \mathcal{I}_{i(rs)}(\theta), \tag{2.4}$$

which is also non-negative definite.

We now derive the expression for the information matrix of a choice design with choice set size m. Consider a $l_1 \times l_2 \times \cdots \times l_n$ choice experiment with $L = \prod_{i=1}^{n} l_i$. Let the L treatments in the choice experiment be denoted by $T_1, T_2, \ldots, T_L$ where, $T_i = (i_1 i_2 \ldots i_h \ldots i_k \ldots i_n), i_r = 0, 1, \ldots, l_r - 1; r = 1, 2, \ldots, n$; is a typical treatment combination. In order to ensure that T_i's are arranged in a lexicographic order, let $i = i_1 \prod_{i=2}^{n} l_i + i_2 \prod_{i=3}^{n} l_i + \cdots + i_{n-1} l_n + i_n + 1$. In other words, i is the lexicographic order number of the treatment combination T_i.

Let $\pi_i = e^{V_i}$ be the parameter associated to the treatment T_i. Our aim is to find the information matrix of certain parametric contrasts involving the parameters V_i, $i = 1, 2, \ldots, L$. A choice set of size m is denoted by $S_m = (T_{j_1}, T_{j_2}, \ldots, T_{j_m})$, where no two j_i's are equal. For a choice set S_m, we represent $(T_{j_i} > \{T_{j_1}, T_{j_2}, \ldots, T_{j_m}\})$ to mean T_{j_i} *is chosen over* $T_{j_1}, \ldots, T_{j_{i-1}}, T_{j_{i+1}}, \ldots, T_{j_m}$, by the respondent.

Consider an experiment in which there are N choice sets of size m. We define a set A_t as

$$A_t = \{(j_1, j_2, \ldots, j_m) : \text{if } (T_{j_1}, T_{j_2}, \ldots, T_{j_m}) \text{ is a choice set in the experiment}\}.$$

Let us consider an indicator function $N_{j_1 j_2 \cdots j_m}$ as

$$N_{j_1 j_2 \cdots j_m} = \begin{cases} 1 & \text{if } (j_1, j_2, \ldots, j_m) \in A_t \\ 0 & \text{if } (j_1, j_2, \ldots, j_m) \notin A_t. \end{cases}$$

Therefore,

$$N = \sum_{j_1 < j_2 < \cdots < j_m} N_{j_1 j_2 \cdots j_m}. \tag{2.5}$$

For any $(j_1, j_2, \ldots, j_m) \in A_t$, we can write from (1.1) that

$$P(T_{j_i} > \{T_{j_1}, T_{j_2}, \ldots, T_{j_m}\}) = \frac{\pi_{j_i}}{\sum_{i=1}^{m} \pi_{j_i}} \tag{2.6}$$

for $i = 1, 2, \ldots, m$. Let, $\pi = (\pi_1, \pi_2, \ldots, \pi_L)'$. Here each choice set $(T_{j_1}, T_{j_2}, \ldots, T_{j_m})$ represent an associate population with parameters $\pi_{j_1}, \pi_{j_2}, \ldots, \pi_{j_m}$. Therefore, the pmf $f_{j_1 j_2 \cdots j_m}$ of the multinomial random variable $(x_{j_1}, x_{j_2}, \ldots, x_{j_m})$ corresponding to the choice set $(T_{j_1}, T_{j_2}, \ldots, T_{j_m})$, is

$$f_{j_1 j_2 \cdots j_m}(x_{j_1}, x_{j_2}, \ldots, x_{j_m}; \pi) = \prod_{i=1}^{m} \left(\frac{\pi_{j_i}}{\sum_{i=1}^{m} \pi_{j_i}} \right)^{x_{j_i}}, \tag{2.7}$$

where for $i = 1, 2, \ldots, m$, we define $x_{j_i} = 1$ if $(T_{j_i} > \{T_{j_1}, T_{j_2}, \ldots, T_{j_m}\})$; and 0 otherwise. To be more precise x_{j_i} can be written as $x_{j_i}^{(j_1, j_2, \ldots, j_m)}$, but for notational convenience we retain the notation x_{j_i} corresponding to the choice set $(T_{j_1}, T_{j_2}, \ldots, T_{j_m})$. Note that $\sum_{i=1}^{m} x_{j_i} = 1$. Therefore from equation (2.3), the likelihood function can be written as

$$\mathcal{L} = \prod_{j_1 < j_2 < \cdots < j_m}^{L} \{f_{j_1 j_2 \cdots j_m}(x_{j_1}, x_{j_2}, \ldots, x_{j_m}; \pi)\}^{N_{j_1 j_2 \cdots j_m}}. \tag{2.8}$$

Let $V = (V_1, V_2, \ldots, V_L)'$ be the vector of treatment effects that the researcher can capture for a $l_1 \times l_2 \times \cdots \times l_n$ choice experiment. Furthermore, let $\Lambda = ((\lambda_{kl}))$ be a $L \times L$ matrix representing the information matrix of V. Then, since $V_i = \ln \pi_i$, it follows from (2.2) and (2.4) that

$$\begin{aligned} \lambda_{kl} &= \sum_{j_1 < j_2 < \cdots < j_m}^{L} N_{j_1 j_2 \cdots j_m} E\left[\frac{\partial \ln f_{j_1 j_2 \cdots j_m}}{\partial V_k} \frac{\partial \ln f_{j_1 j_2 \cdots j_m}}{\partial V_l} \right] \\ &= \sum_{j_1 < j_2 < \cdots < j_m}^{L} N_{j_1 j_2 \cdots j_m} E\left[\frac{\partial \ln f_{j_1 j_2 \cdots j_m}}{\partial \pi_k} \frac{\partial \ln f_{j_1 j_2 \cdots j_m}}{\partial \pi_l} \right] \pi_k \pi_l. \end{aligned} \tag{2.9}$$

It is clear from (2.9) that if (k,l) does not belong to any element of A_t, then

$$\lambda_{kl} = 0. \tag{2.10}$$

From (2.7) we note that $(x_{j_1}, x_{j_2}, \dots, x_{j_m})$ is a multinomial random variable with parameters $\frac{\pi_{j_i}}{\sum_{i=1}^m \pi_{j_i}}$, $i = 1,2,\dots,m$ and $E(x_{j_i}) = E(x_{j_i}^2) = \frac{\pi_{j_i}}{\sum_{i=1}^m \pi_{j_i}}$; $i = 1,2,\dots,m$. Also, from (2.7), we get

$$\ln(f_{j_1 j_2 \cdots j_m}(x_{j_1}, x_{j_2}, \dots, x_{j_m}; \pi)) = \sum_{i=1}^m x_{j_i} \ln(\pi_{j_i}) - \ln\left(\sum_{i=1}^m \pi_{j_i}\right),$$

and therefore,

$$\frac{\partial \ln f_{j_1 j_2 \cdots j_m}}{\partial \pi_{j_i}} = \frac{x_{j_i}}{\pi_{j_i}} - \frac{1}{\sum_{i=1}^m \pi_{j_i}}; i = 1,2,\dots,m.$$

If (k,l) belongs to an element $(j_1, j_2, \dots, j_m)$ of A_t, then from (2.9), both the partial derivatives are non-zero for the choice sets $(T_{j_1}, T_{j_2}, \dots, T_{j_m})$, which contains T_k and T_l as options. Thus, without loss of generality, when $(k,l) = (j_1, j_2)$ such that $(j_1, j_2, \dots, j_m) \in A_t$, we have $\lambda_{kl} = \lambda_{j_1 j_2}$ which is

$$\begin{aligned}
&= \sum_{j_3<j_4<\cdots<j_m}^{L} N_{j_1 j_2 \cdots j_m} E\left[\left(\frac{x_{j_1}}{\pi_{j_1}} - \frac{1}{\sum_{i=1}^m \pi_{j_i}}\right)\left(\frac{x_{j_2}}{\pi_{j_2}} - \frac{1}{\sum_{i=1}^m \pi_{j_i}}\right)\right] \pi_{j_1}\pi_{j_2} \\
&= \sum_{j_3<j_4<\cdots<j_m}^{L} N_{j_1 j_2 \cdots j_m} E\left[\frac{x_{j_1}x_{j_2}}{\pi_{j_1}\pi_{j_2}} - \frac{x_{j_1}}{\pi_{j_1}\sum_{i=1}^m \pi_{j_i}} - \frac{x_{j_2}}{\pi_{j_2}\sum_{i=1}^m \pi_{j_i}} + \frac{1}{(\sum_{i=1}^m \pi_{j_i})^2}\right] \pi_{j_1}\pi_{j_2} \\
&= \sum_{j_3<j_4<\cdots<j_m}^{L} N_{j_1 j_2 \cdots j_m} \left[0 - \frac{1}{(\sum_{i=1}^m \pi_{j_i})^2} - \frac{1}{(\sum_{i=1}^m \pi_{j_i})^2} + \frac{1}{(\sum_{i=1}^m \pi_{j_i})^2}\right] \pi_{j_1}\pi_{j_2} \\
&= -\sum_{j_3<j_4<\cdots<j_m}^{L} N_{j_1 j_2 \cdots j_m} \frac{\pi_{j_1}\pi_{j_2}}{(\sum_{i=1}^m \pi_{j_i})^2}.
\end{aligned} \tag{2.11}$$

Also, when $k = l = j_1$, $\lambda_{kk} = \lambda_{j_1 j_1}$ which is

$$\begin{aligned}
&= \sum_{j_2<j_3<\cdots<j_m} N_{j_1 j_2 \cdots j_m} E\left[\left(\frac{x_{j_1}}{\pi_{j_1}} - \frac{1}{\sum_{i=1}^m \pi_{j_i}}\right)^2\right] \pi_{j_1}\pi_{j_1} \\
&= \sum_{j_2<j_3<\cdots<j_m} N_{j_1 j_2 \cdots j_m} E\left[\frac{x_{j_1}^2}{\pi_{j_1}^2} - \frac{2x_{j_1}}{\pi_{j_1}\sum_{i=1}^m \pi_{j_i}} + \frac{1}{(\sum_{i=1}^m \pi_{j_i})^2}\right] \pi_{j_1}\pi_{j_1} \\
&= \sum_{j_2<j_3<\cdots<j_m} N_{j_1 j_2 \cdots j_m} \left[\frac{1}{\pi_{j_1}\sum_{i=1}^m \pi_{j_i}} - \frac{2}{(\sum_{i=1}^m \pi_{j_i})^2} + \frac{1}{(\sum_{i=1}^m \pi_{j_i})^2}\right] \pi_{j_1}\pi_{j_1} \\
&= \sum_{j_2<j_3<\cdots<j_m} N_{j_1 j_2 \cdots j_m} \frac{\pi_{j_1}\sum_{i=2}^m \pi_{j_i}}{(\sum_{i=1}^m \pi_{j_i})^2}.
\end{aligned} \tag{2.12}$$

Therefore, in terms of π_i's, Λ can be rewritten as

$$\lambda_{kl} = \begin{cases} \displaystyle\sum_{j_2<j_3<\cdots<j_m}^{L} N_{j_1 j_2 \cdots j_m} \frac{\pi_{j_1}(\sum_{i=2}^{m} \pi_{j_i})}{(\sum_{i=1}^{m} \pi_{j_i})^2} & \text{if} \quad k = l = j_1 \\ -\displaystyle\sum_{j_3<j_4<\cdots<j_m}^{L} N_{j_1 j_2 \cdots j_m} \frac{\pi_{j_1}\pi_{j_2}}{(\sum_{i=1}^{m} \pi_{j_i})^2} & \text{if} \quad k = j_1, l = j_2 \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

Since $P(T_{j_i} > \{T_{j_1}, T_{j_2}, \ldots, T_{j_m}\}) = \frac{\pi_{j_i}}{\sum_{i=1}^{m} \pi_{j_i}}$ is not dependent on parameter scale, we assume a convenient scale determining constraint

$$\sum_{i=1}^{L} V_i = 0. \tag{2.14}$$

Let

$$B_{(p+q)} = \begin{pmatrix} B_{(p)} \\ B_{(q)} \end{pmatrix} \tag{2.15}$$

be a partition of the orthonormal contrast matrix of order $(L-1) \times L$, with $p+q = L-1$. Here, our interest lies in finding the information matrix of $\Theta_1 = B_{(p)}V$, while $\Theta_0 = B_{(q)}V$ are the nuisance parameters. Now, with

$$B_{(q)} = \begin{pmatrix} B_{(q_1)} \\ B_{(q_2)} \end{pmatrix}, \tag{2.16}$$

the nuisance parameters Θ_0 can be partitioned as $\Theta_0 = (\Theta_{0_1}' \;\; \Theta_{0_2}')'$ where $\Theta_{0_1} = B_{(q_1)}V$, $\Theta_{0_2} = B_{(q_2)}V$ and $q_1 + q_2 = q$. Under the assumption

$$\Theta_{0_2} = B_{(q_2)}V = 0_{q_2}, \tag{2.17}$$

and with

$$B_{(p+q_1)} = \begin{pmatrix} B_{(p)} \\ B_{(q_1)} \end{pmatrix}, \tag{2.18}$$

we first find the information matrix of $\Theta = B_{(p+q_1)}V$, where $\Theta = (\Theta_1' \;\; \Theta_{0_1}')'$, $\Theta_1 = (\theta_1, \theta_2, \ldots, \theta_p)'$ and $\Theta_{0_1} = (\theta_{p+1}, \theta_{p+2}, \ldots, \theta_{p+q_1})'$.

Let I_p denote an identity matrix of order p. Also, let $G' = \left[L^{-\frac{1}{2}}1 \;\; B_{(p+q_1)}' \;\; B_{(q_2)}'\right]$, where 1 is a column vector of all ones. Then G is an orthogonal matrix of order $L \times L$, and $GG' = G'G = I_L$. Therefore,

$$B_{(p+q_1)}' B_{(p+q_1)} = I_L - \frac{11'}{L} - B_{(q_2)}' B_{(q_2)}. \tag{2.19}$$

Now, since $\Theta = B_{(p+q_1)}V$, using (2.14), (2.17) and (2.19), we have

$$\begin{aligned} & B'_{(p+q_1)}\Theta = B'_{(p+q_1)}B_{(p+q_1)}V \\ \Rightarrow\ & B'_{(p+q_1)}\Theta = [I_L - \frac{11'}{L} - B'_{(q_2)}B_{(q_2)}]V \\ \Rightarrow\ & B'_{(p+q_1)}\Theta = I_L V = V. \end{aligned} \tag{2.20}$$

Let $B_{(p+q_1)} = ((b_{r_1r_2}))$. Also, let the $(p+q_1)\times(p+q_1)$ information matrix of Θ be denoted by $C_{\{p+q_1\}} = ((c_{rs}))$. Then from (2.2) and (2.4), and using (2.20), we have

$$\begin{aligned} c_{rs} &= \sum_{j_1<j_2<\cdots<j_m}^{L} N_{j_1j_2\ldots j_m} E\left[\frac{\partial \ln f_{j_1j_2\ldots j_m}}{\partial\theta_r}\frac{\partial \ln f_{j_1j_2\ldots j_m}}{\partial\theta_s}\right] \\ &= \sum_k^L\sum_l^L\left\{\sum_{j_1<j_2<\cdots<j_m}^{L} N_{j_1j_2\ldots j_m} E\left[\frac{\partial \ln f_{j_1j_2\ldots j_m}}{\partial V_k}\frac{\partial \ln f_{j_1j_2\ldots j_m}}{\partial V_l}\right]\right\} b_{rk}b_{sl} \\ &= \sum_k^L\sum_l^L \lambda_{kl}b_{rk}b_{sl} \\ &= B_{(p+q_1)}\Lambda B'_{(p+q_1)}. \end{aligned} \tag{2.21}$$

Thus, the partitioned form of the information matrix of Θ is

$$C_{\{p+q_1\}} = \begin{bmatrix} B_{(p)}\Lambda B'_{(p)} & B_{(p)}\Lambda B'_{(q_1)} \\ B_{(q_1)}\Lambda B'_{(p)} & B_{(q_1)}\Lambda B'_{(q_1)} \end{bmatrix}, \tag{2.22}$$

and the information matrix of Θ_1 is

$$C_{\{p\}} = B_{(p)}\Lambda B'_{(p)} - B_{(p)}\Lambda B'_{(q_1)}[B_{(q_1)}\Lambda B'_{(q_1)}]^- B_{(q_1)}\Lambda B'_{(p)}, \tag{2.23}$$

where $B_{(p)}\Lambda B'_{(p)}$ and $B_{(p)}\Lambda B'_{(q_1)}[B_{(q_1)}\Lambda B'_{(q_1)}]^- B_{(q_1)}\Lambda B'_{(p)}$ are both non-negative definite matrices and Y^- represents a g-inverse of Y. Furthermore, the second term does not arise if $q_1 = 0$. For notational convenience we denote $C_{\{p\}}$ by C. A choice design for estimating Θ_1 is said to be *connected* if $rank(C) = p$. We restrict ourselves to the class of all *connected* designs. When a design is *connected*, it ensures the estimability of Θ_1. In general Θ_1 is estimable if and only if $rank(C) = p$.

Following the concept of Resolution (see e.g., Dey and Mukerjee 1999) in fractional factorial plans, we define Resolution (f,t) choice designs as ones which ensure estimability of the complete sets of contrasts belonging to factorial effects involving at most f factors under the absence of factorial effects involving $t+1$ or more factors $(1 \le f \le t \le n-1)$. Thus, the information matrix of $\Theta_1 = B_{(p)}V$, under a Resolution (f,t) model setup (henceforth called $\mathcal{R}(f,t)$ model) is

$$C = B_{(f)}\Lambda B'_{(f)} - B_{(f)}\Lambda B'_{(t)}[B_{(t)}\Lambda B'_{(t)}]^- B_{(t)}\Lambda B'_{(f)}, \tag{2.24}$$

where $B_{(f)}$ is the contrast matrix corresponding to the complete set of factorial effects involving at most f factors and $B_{(t)}$ is the contrast matrix corresponding to the complete set of factorial effects involving more than f factors but less than $t+1$ factors. Furthermore, the second term does not arise if $f = t$. Thus, under the usual nomenclature, in what follows we consider the model $\mathcal{R}(1,2)$ corresponding to Resolution $(1,2)$ choice designs, i.e, designs which ensure estimability of all the main effects under the absence of three or higher order interaction effects. We also consider the main effects model $\mathcal{R}(1,1)$.

3 C-matrix under $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$

For the purpose of optimal choice design, as in the literature, we assume that the options are equally attractive i.e., $\pi_1 = \pi_2 = \cdots = \pi_L$ ($= \pi_0$, say).

Then from (2.13), Λ turns out to be

$$\lambda_{kl} = \begin{cases} \dfrac{m-1}{m^2} \displaystyle\sum_{j_2<j_3<\cdots<j_m}^{L} N_{j_1 j_2 \ldots j_m} & \text{if} \quad k = l = j_1 \\ -\dfrac{1}{m^2} \displaystyle\sum_{j_3<j_4<\cdots<j_m}^{L} N_{j_1 j_2 \ldots j_m} & \text{if} \quad k = j_1, l = j_2 \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

Let $M^{(j_1 j_2 \ldots j_m)} = ((m_{st}))$ be a $L \times L$ matrix corresponding to a choice set $(T_{j_1}, T_{j_2}, \ldots, T_{j_m})$, where

$$m_{st} = \begin{cases} m-1 & \text{if} \quad s = t,\ t \in \{j_1, j_2, \ldots, j_m\} \\ -1 & \text{if} \quad s \neq t, (s,t) \in \{j_1, j_2, \ldots, j_m\} \\ 0 & \text{otherwise.} \end{cases}$$

Then for any choice experiment with N choice sets, we can write

$$\Lambda = \frac{1}{m^2} \sum_{j_1<j_2<\cdots<j_m}^{L} N_{j_1 j_2 \ldots j_m} M^{(j_1 j_2 \ldots j_m)}. \tag{3.2}$$

We can consider the matrix $M^{(j_1 j_2 \ldots j_m)}$ as the contribution of the choice set $(T_{j_1}, T_{j_2}, \ldots, T_{j_m})$ to Λ. The definition of $M^{(j_1 j_2 \ldots j_m)}$ suggests that we can write

$$M^{(j_1 j_2 \ldots j_m)} = \sum_{j_r < j_{r'}} M^{(j_r j_{r'})} \tag{3.3}$$

where, $j_r, j_{r'} \in \{j_1, j_2, \ldots, j_m\}$. This means, the contribution of the choice set $(T_{j_1}, T_{j_2}, \ldots, T_{j_m})$ to the Λ is equal to the sum of the individual contributions of the $\binom{m}{2}$ different *component pairs* that it contains. Therefore, Λ corresponding to choice sets of size m can be translated in terms of Λ corresponding to *component pairs*.

We now concentrate on 2^n choice experiments ($l_i = 2, i = 1, 2, \ldots, n$) under the models $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$. With $B_{(1)}$ (henceforth denoted as B) being the $n \times 2^n$ matrix of ± 1s with rows representing the orthogonal contrast vectors corresponding to the n main effects and $B_{(2)}$ being the $\binom{n}{2} \times 2^n$ matrix of ± 1s with rows representing the orthogonal contrast vectors corresponding to the $\binom{n}{2}$ two-factor interaction effects, from (2.24),(3.2) and (3.3), the C-matrix, $C = ((c_{hk}))$, for estimating the main effects under the models $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$ are, respectively,

$$2^n C = B\Lambda B' - B\Lambda B'_{(2)}[B_{(2)}\Lambda B'_{(2)}]^- B_{(2)}\Lambda B', \tag{3.4}$$

and

$$2^n C = B\Lambda B'. \tag{3.5}$$

Singh, Chai and Das (2015) obtained the information matrix under a broader main effects model, which is same as model $\mathcal{R}(1,2)$ described here. Also, from (3.2), (3.3) and (3.5), we can express $B\Lambda B'$ as

$$\begin{aligned} B\Lambda B' &= B\left(\frac{1}{m^2}\sum_{j_1<j_2<\cdots<j_m} N_{j_1j_2\cdots j_m} M^{(j_1j_2\cdots j_m)}\right)B' \\ &= \frac{1}{m^2}\sum_{j_1<j_2<\cdots<j_m} N_{j_1j_2\cdots j_m}\left\{BM^{(j_1j_2\cdots j_m)}B'\right\} \\ &= \frac{1}{m^2}\sum_{j_1<j_2<\cdots<j_m} N_{j_1j_2\cdots j_m}\left\{B\left(\sum_{j_r<j_{r'}} M^{(j_rj_{r'})}\right)B'\right\}. \end{aligned} \tag{3.6}$$

4 Characterization of C-matrix under $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$

In what follows, we find conditions under which the C-matrix has off-diagonal elements zero. First we have the following Lemma due to Manna and Das (2016).

Lemma 4.1. *Let $B_h = (x_{h1}, \ldots, x_{hj_r}, \ldots, x_{hj_{r'}}, \ldots, x_{h2^n})$ and $B_k = (x_{k1}, \ldots, x_{kj_r}, \ldots, x_{kj_{r'}}, \ldots, x_{k2^n})$ be row vectors with 2^n real elements. Then for a given component pair $(T_{j_r}, T_{j_{r'}})$, the value of $B_h M^{(j_rj_{r'})} B'_k = (x_{hj_r} - x_{hj_{r'}})(x_{kj_r} - x_{kj_{r'}})$.*

From Lemma 4.1, it follows that for a *component pair* $(T_{j_r}, T_{j_{r'}})$, the possible realized values of $B_h M^{(j_rj_{r'})} B'_k$ are:

(P1) If $x_{hj_r} = x_{hj_{r'}}$ or/and $x_{kj_r} = x_{kj_{r'}}$, then $B_h M^{(j_rj_{r'})} B'_k = 0$. (4.1)

(P2) If $x_{hj_r} = -x_{hj_{r'}} = \pm 1$ and $x_{kj_r} = -x_{kj_{r'}} = \pm 1$, then $B_h M^{(j_rj_{r'})} B'_k = 4$. (4.2)

(P3) If $x_{hj_r} = -x_{hj_{r'}} = \pm 1$ and $x_{kj_r} = -x_{kj_{r'}} = \mp 1$, then $B_h M^{(j_rj_{r'})} B'_k = -4$. (4.3)

Let $f_1, f_2, \dots, f_n$ be the factors corresponding to the 2^n choice experiment with treatment combination $T_i = (i_1 i_2 \dots i_r \dots i_n), i_r = 0, 1; r = 1, 2, \dots, n$. Let F_h represents the h-th factorial effect. Thus for $h = 1, \dots, n$, F_h represent the main effects and for $h = n+1, \dots, n + \binom{n}{2}$, F_h represent the two-factor interaction effects. For $h = 1, \dots, n$, we define the h-th positional value of F_h corresponding to the treatment T_i as i_h. Similarly, for $h = n+1, \dots, n+\binom{n}{2}$, we define the h-th positional value of F_h corresponding to the treatment T_i as $i_r + i_{r'}$ (mod 2) ($= i_h^*$, say) where F_h is the two factor interaction effect corresponding to the factors f_r and $f_{r'}$, $1 \le r < r' \le n$. Here one can use the *combinatorial number system* to have the correspondence between natural numbers $h = n+1, \dots, n + \binom{n}{2}$, and the 2-combinations (r, r'). For $h \neq k; (h, k) \in \{1, \dots, n\}$, the h-th and k-th positional value of the treatment T_i is denoted by $(i_h i_k)_{hk}$ and for the *component pair* (T_i, T_j), the h-th and k-th positional value is denoted by $(i_h i_k, j_h j_k)_{hk}$. Similarly, for $h \neq k; (h, k) \in \{n+1, \dots, n + \binom{n}{2}\}$, the h-th and k-th positional value of the treatment T_i is denoted by $(i_h^* i_k^*)_{hk}$ and for the *component pair* (T_i, T_j), the h-th and k-th positional value is denoted by $(i_h^* i_k^*, j_h^* j_k^*)_{hk}$. Finally, for $h \in \{1, \dots, n\}$ and $k \in \{n+1, \dots, n + \binom{n}{2}\}$, the h-th and k-th positional value of the treatment T_i is denoted by $(i_h i_k^*)_{hk}$ and for the *component pair* (T_i, T_j), the h-th and k-th positional value is denoted by $(i_h i_k^*, j_h j_k^*)_{hk}$.

The following Lemma on lines similar to Manna and Das (2016) provides a converse result of Lemma 4.1 in the sense that it establishes possible *component pairs* (T_i, T_j) that gives rise to specific values of $B_h M^{(ij)} B_k'$.

Lemma 4.2. *Given that the h-th row ($h = 1, 2, \dots, n$) of B is*

$$B_h = \otimes_{i=1}^{h-1} (1 \;\; 1) \otimes (-1 \;\; 1) \otimes_{i=h+1}^{n} (1 \;\; 1), \tag{4.4}$$

and the h-th row ($h = n+1, \dots, n + \binom{n}{2}$) of $B_{(2)}$ is

$$B_h = \otimes_{i=1}^{r-1} (1 \;\; 1) \otimes (-1 \;\; 1) \otimes_{i=r+1}^{r'-1} (1 \;\; 1) \otimes (-1 \;\; 1) \otimes_{i=r'+1}^{n} (1 \;\; 1), \tag{4.5}$$

the exhaustive cases leading to possible values of $B_h M^{(ij)} B_k'$ and its associated component pairs (T_i, T_j), are

- *Case 1: $h \neq k, (h, k) \in \{1, \dots, n\}$*
  a) $B_h M^{(ij)} B_k' = -4$ *when* $(i_h i_k, j_h j_k)_{hk} \equiv (01, 10)_{hk}$
  b) $B_h M^{(ij)} B_k' = 4$ *when* $(i_h i_k, j_h j_k)_{hk} \equiv (00, 11)_{hk}$
  c) $B_h M^{(ij)} B_k' = 0$ *for all other situations.*

- *Case 2: $h \in \{1, \dots, n\}, k \in \{n+1, \dots, n + \binom{n}{2}\}$*
  a) $B_h M^{(ij)} B_k' = 4$ *when* $(i_h i_k^*, j_h j_k^*)_{hk} \equiv (01, 10)_{hk}$
  b) $B_h M^{(ij)} B_k' = -4$ *when* $(i_h i_k^*, j_h j_k^*)_{hk} \equiv (00, 11)_{hk}$
  c) $B_h M^{(ij)} B_k' = 0$ *for all other situations.*

- *Case 3:* $h \neq k, (h,k) \in \{n+1, \ldots, n+\binom{n}{2}\}$

  a) $B_h M^{(ij)} B'_k = -4$ *when* $(i^*_h i^*_k, j^*_h j^*_k)_{hk} \equiv (01, 10)_{hk}$
  b) $B_h M^{(ij)} B'_k = 4$ *when* $(i^*_h i^*_k, j^*_h j^*_k)_{hk} \equiv (00, 11)_{hk}$
  c) $B_h M^{(ij)} B'_k = 0$ *for all other situations.*

Proof. Let $B_h = (x_{h1}, \ldots, x_{hj_r}, \ldots, x_{hj_{r'}}, \ldots, x_{h2^n})$ and $B_k = (x_{k1}, \ldots, x_{kj_r}, \ldots, x_{kj_{r'}}, \ldots, x_{k2^n})$. Note that $M^{(j_r j_{r'})}$ is a $2^n \times 2^n$ matrix with all elements 0 except $M^{(j_r j_{r'})}_{j_r j_r} = M^{(j_r j_{r'})}_{j_{r'} j_{r'}} = 1$ and $M^{(j_r j_{r'})}_{j_r j_{r'}} = M^{(j_r j_{r'})}_{j_{r'} j_r} = -1$. Then

$$\begin{aligned} B_h M^{(j_r j_{r'})} B'_k &= \big(0, \ldots, (x_{hj_r} - x_{hj_{r'}}), \ldots, -(x_{hj_r} - x_{hj_{r'}}), \ldots, 0\big) B'_k \\ &= (x_{hj_r} - x_{hj_{r'}}) x_{kj_r} - (x_{hj_r} - x_{hj_{r'}}) x_{kj_{r'}} \\ &= (x_{hj_r} - x_{hj_{r'}})(x_{kj_r} - x_{kj_{r'}}). \end{aligned} \tag{4.6}$$

From (4.4), (4.5) and the fact that T_i's are arranged in a lexicographic order, for any treatment combination T_j,
$x_{hj} = -1, 1$ if and only if $j_h = 0, 1$ respectively ($h = 1, \ldots, n$), and
$x_{hj} = -1, 1$ if and only if $j^*_h = 1, 0$ respectively ($h = n+1, \ldots, n+\binom{n}{2}$).
The proof then follows from (4.6). □

With F_h and F_k being any two effects, we now define two quantities N^+_{hk} and N^-_{hk} as follows:

- N^+_{hk} = Total number of *component pairs* of the type $(00, 11)_{hk}$ corresponding to h-th and k-th positional values across all $\binom{m}{2}$ possible pairs of a choice set of size m and among all such sets in the choice design.
- N^-_{hk} = Total number of *component pairs* of the type $(01, 10)_{hk}$ corresponding to h-th and k-th positional values across all $\binom{m}{2}$ possible pairs of a choice set of size m and among all such sets in the choice design.

Lemma 4.3. *For* $h \neq k, (h,k) \in \{1, \ldots, n\}$*, the* (h,k)*-th element of* $B\Lambda B'$ *will be zero if and only if* $N^+_{hk} = N^-_{hk}$.

Proof. The proof follows from (3.6) and Lemma 4.2 on noting the contribution towards the (h,k)-th element of $B\Lambda B'$ by N choice sets through its $\binom{m}{2}$ possible *component pairs*. The Case 1 of Lemma 4.2 leads to

- N^-_{hk} = Total number of *component pairs* falling under Case 1a.
- N^+_{hk} = Total number of *component pairs* falling under Case 1b.

- $N - (N^+_{hk} + N^-_{hk})$ = Total number of choice pairs falling under Case 1c.

Let c'_{hk} denote the (h,k)-th element of $m^2 B\Lambda B'$. Then, it follows from (3.6) and Case 1 of Lemma 4.2 that

$$\begin{aligned} c'_{hk} &= \sum_{j_1<j_2<\cdots<j_m} N_{j_1 j_2 \ldots j_m} \sum_{j_r<j_{r'}} \{B_h M^{(j_r j_{r'})} B'_k\} \\ &= \left\{(4N^+_{hk} - 4N^-_{hk}) + 0(N - (N^+_{hk} + N^-_{hk}))\right\}. \end{aligned}$$

Thus $c'_{hk} = 0$ if and only if $N^+_{hk} = N^-_{hk}$. □

Lemma 4.4. *For $h \in \{1, \ldots, n\}, k \in \{n+1, \ldots, n+\binom{n}{2}\}$, the (h,k)-th element of $B\Lambda B'_{(2)}$ will be zero if and only if $N^+_{hk} = N^-_{hk}$.*

Proof. The proof follows on lines similar to Lemma 4.3, since

$$B\Lambda B'_{(2)} = \frac{1}{m^2} \sum_{j_1<j_2<\cdots<j_m} N_{j_1 j_2 \ldots j_m} \left\{ B \left(\sum_{j_r<j_{r'}} M^{(j_r j_{r'})} \right) B'_{(2)} \right\}. \tag{4.7}$$

The Case 2 of Lemma 4.2 leads to

- N^-_{hk} = Total number of *component pairs* falling under Case 2a.
- N^+_{hk} = Total number of *component pairs* falling under Case 2b.
- $N - (N^+_{hk} + N^-_{hk})$ = Total number of choice pairs falling under Case 2c.

Let c''_{hk} denote the (h,k)-th element of $m^2 B\Lambda B'_{(2)}$. Then, it follows from (4.7) and Case 2 of Lemma 4.2 that

$$\begin{aligned} c''_{hk} &= \sum_{j_1<j_2<\cdots<j_m} N_{j_1 j_2 \ldots j_m} \sum_{j_r<j_{r'}} \{B_h M^{(j_r j_{r'})} B'_k\} \\ &= \left\{(-4N^+_{hk} + 4N^-_{hk}) + 0(N - (N^+_{hk} + N^-_{hk}))\right\}. \end{aligned}$$

Thus $c''_{hk} = 0$ if and only if $N^+_{hk} = N^-_{hk}$. □

Theorem 4.5. *For $h \neq k, (h,k) \in \{1, \ldots, n\}$, under the model $\mathcal{R}(1,1)$, the (h,k)-th element of C-matrix will be zero if and only if $N^+_{hk} = N^-_{hk}$. Furthermore, for $h \neq k, (h,k) \in \{1, \ldots, n\}$, under the model $\mathcal{R}(1,2)$, the (h,k)-th element of C-matrix will be zero, if additionally, $B\Lambda B'_{(2)}$ is a null matrix, i.e., $N^+_{hk'} = N^-_{hk'}$, for $h \in \{1, \ldots, n\}, k' \in \{n+1, \ldots, n+\binom{n}{2}\}$.*

Proof. The proof follows from (3.4), (3.5), Lemma 4.3 and Lemma 4.4. □

We will now find the contribution of each choice set S_m of size m to the diagonal positions of $m^2 B\Lambda B'$.

Lemma 4.6. *Every component pair adds a value 4 in the (h,h)-th element of the $m^2 B\Lambda B'$, if and only if the pair has a change of level at the h-th position of its treatment combinations.*

Proof. Every *component pair* $(T_{j_r}, T_{j_{r'}})$ is adding a value $B_h M^{(j_r j_{r'})} B'_h$ at c'_{hh}. From (P2) in (4.2) it follows that this value will be 4 if and only if there is a change of level in the h-th position of the *component pair.* □

Let $n_h \in \{0, 1, 2, \ldots, m\}$ be the number of treatment combinations which have zero at the h-th position in the choice set S_m.

Lemma 4.7. *Every S_m adds a value $4n_h(m - n_h)$ to the (h,h)-th element of $m^2 B\Lambda B'$.*

Proof. Lemma 4.6 says that every *component pair* adds a value 4 to c'_{hh}, if and only if the pair has a change of level at the h-th position of its treatment combinations. There are a total of $\binom{m}{2}$ *component pairs* possible from S_m. The contribution of S_m to c'_{hh} is same as the sum of contributions of all the $\binom{m}{2}$ *component pairs* corresponding to S_m. Now there are n_h treatment combinations in S_m which have a 0 at the h-th position. We call this subset as A. Therefore the set $\bar{A}$ contains all treatment combinations which have a 1 at the h-th position. Every *component pair* which have one treatment from A and another treatment from $\bar{A}$, adds a value 4 to c'_{hh}. There are a total of $n_h(m - n_h)$ such pairs and they all together add a value $4n_h(m - n_h)$ to c'_{hh}. □

Corollary 4.8. *Every S_m adds a $4\sum_{h=1}^{n} n_h(m - n_h)$ value to the $trace(m^2 B\Lambda B')$.*

We will now find out the expression of $trace(m^2 B\Lambda B')$ when there are N choice sets. For this purpose we will use the following notations.

- S_{m_i} = the i-th choice set, $i = 1, 2, \ldots, N$.
- n_{h_i} = number of treatment combinations which have zero at the h-th position in the choice set S_{m_i}.

Lemma 4.9. *For N choice sets $S_{m_1}, \ldots, S_{m_N}$, $trace(m^2 B\Lambda B') = 4\sum_{i=1}^{N}\sum_{h=1}^{n} n_{h_i}(m - n_{h_i})$.*

Proof. From Corollary (4.8), every choice set S_{m_i} adds a value $4\sum_{h=1}^{n} n_{h_i}(m - n_{h_i})$ to $trace(m^2 B\Lambda B')$.

Therefore, for N choice sets $trace(m^2 B\Lambda B') = 4\sum_{i=1}^{N}\sum_{h=1}^{n} n_{h_i}(m - n_{h_i})$. □

Lemma 4.10. *Maximum of* $trace(m^2 B\Lambda B')$ *is attained when*

$$n_{h_i} = \begin{cases} \dfrac{m}{2} & \text{if } m \text{ even} \\ \dfrac{m-1}{2} \text{ or } \dfrac{m+1}{2} & \text{if } m \text{ odd} \end{cases}$$

for every position h *and for every choice set* S_{m_i}.

Proof. Maximum of $trace(m^2 B\Lambda B')$ is attained when every choice set S_{m_i} in the experiment contributes maximum value towards $trace(m^2 B\Lambda B')$. Each choice set S_{m_i} will contribute maximum value if and only if every h-th position of its treatments contribute maximum value to c'_{hh}. Lemma 4.7 says that if S_{m_i} has n_{h_i} zeros at the h_i-th position of its treatments then it will add a value $4n_{h_i}(m - n_{h_i})$ to $trace(m^2 B\Lambda B')$. We want to maximize $4n_{h_i}(m - n_{h_i})$ for n_{h_i}. Let $f(n_{h_i}) = 4n_{h_i}(m - n_{h_i})$ and let k_0 be the point at which the function attains its maximum. Then, $f(k_0 - 1) \leq f(k_0)$ implies $4(k_0 - 1)\{m - (k_0 - 1)\} \leq 4k_0(m - k_0)$, or $m(k_0 - 1) - k_0(k_0 - 1) + (k_0 - 1) \leq mk_0 - k_0^2$, or $2k_0 - m - 1 \leq 0$. Thus,

$$k_0 \leq \frac{m+1}{2}. \tag{4.8}$$

Also, $f(k_0) \geq f(k_0 + 1)$ implies

$$k_0 \geq \frac{m-1}{2} \tag{4.9}$$

Since k_0 only takes integer value, therefore from (4.8) and (4.9) we conclude that $f(n_{h_i})$ is maximum when (i) $n_{h_i} = \dfrac{m-1}{2}$ or $n_{h_i} = \dfrac{m+1}{2}$ (for m odd) and (ii) $n_{h_i} = \dfrac{m}{2}$ (for m even). Hence the proof.

□

Lemma 4.11. *For* N *choice sets of size* m*, the upper bound to* $trace(m^2 B\Lambda B')$ *is*

$$trace(m^2 B\Lambda B') \leq \begin{cases} Nnm^2 & \text{for } m \text{ even} \\ Nn(m^2 - 1) & \text{for } m \text{ odd} \end{cases}.$$

Proof. From Lemma 4.9 and Theorem 4.10 we can say that $trace(m^2 B\Lambda B')$ will be maximum if and only if each $n_{h_i} = k_0$ for every h and every i. Therefore, for m even, $k_0 = \dfrac{m}{2}$ and $trace(m^2 B\Lambda B') \leq 4\sum_{i=1}^{N}\sum_{h=1}^{n}\left(\frac{m}{2}\right)\left(m - \frac{m}{2}\right) = 4\sum_{i=1}^{N}\sum_{h=1}^{n}\frac{m^2}{4} = Nnm^2$. Also, for m odd, $k_0 = \dfrac{m-1}{2}$ or $\dfrac{m+1}{2}$ and $trace(m^2 B\Lambda B') \leq 4\sum_{i=1}^{N}\sum_{h=1}^{n}\left(\frac{m \pm 1}{2}\right)\left(m - \frac{m \pm 1}{2}\right) = 4\sum_{i=1}^{N}\sum_{h=1}^{n}\frac{m^2 - 1}{4}$ $= Nn(m^2 - 1)$. □

Thus an upper bound of $trace(C)$ is established under both $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$ models and is summarized as

Theorem 4.12. *Under model $\mathcal{R}(1,2)$, with N choice sets of size m, an upper bound to $trace(C)$ is*

$$trace(C) \leq \frac{1}{2^n} trace(B\Lambda B') \leq \begin{cases} \dfrac{Nn}{2^n} & \text{for } m \text{ even} \\ \dfrac{Nn(m^2-1)}{2^n m^2} & \text{for } m \text{ odd} \end{cases},$$

with equality attaining when the following two conditions are satisfied:

$$i) \qquad n_{h_i} = \begin{cases} \dfrac{m}{2} & \text{if } m \text{ even} \\ \dfrac{m-1}{2} \text{ or } \dfrac{m+1}{2} & \text{if } m \text{ odd} \end{cases}$$

for every position h and for every choice set S_{m_i} and

$ii)$ $B\Lambda B'_{(2)}$ is null matrix, i.e., for $h \in \{1,\dots,n\}, k \in \{n+1,\dots,n+\binom{n}{2}\}$, $N^+_{hk} = N^-_{hk}$.

Proof. From (3.4), $trace(2^n C) = trace(B\Lambda B') - trace(B\Lambda B'_{(2)}[B_{(2)}\Lambda B'_{(2)}]^- B_{(2)}\Lambda B')$. Thus, noting that $B\Lambda B'_{(2)}[B_{(2)}\Lambda B'_{(2)}]^- B_{(2)}\Lambda B'$ is a non-negative definite matrix, the proof follows from Lemma 4.4, Lemma 4.10 and Lemma 4.11. □

Theorem 4.13. *Under model $\mathcal{R}(1,1)$, with N choice sets of size m, an upper bound to $trace(C)$ is*

$$trace(C) = \frac{1}{2^n} trace(B\Lambda B') \leq \begin{cases} \dfrac{Nn}{2^n} & \text{for } m \text{ even} \\ \dfrac{Nn(m^2-1)}{2^n m^2} & \text{for } m \text{ odd} \end{cases},$$

with equality attaining when

$$n_{h_i} = \begin{cases} \dfrac{m}{2} & \text{if } m \text{ even} \\ \dfrac{m-1}{2} \text{ or } \dfrac{m+1}{2} & \text{if } m \text{ odd} \end{cases}$$

for every position h and for every choice set S_{m_i}.

For $m = 2$, it follows from the following Theorem that the maximization of $trace(C)$ under model $\mathcal{R}(1,1)$ implies $B\Lambda B'_{(2)}$ is a null matrix.

Theorem 4.14. *For $m = 2$, let F_h be a main effect and F_k a two-factor interaction effect. Then, $B\Lambda B'_{(2)}$ is a null matrix if for every choice set $S_{2_i}, i = 1,\dots,N$, either (a) $n_{h_i} = 1$ for every position h or (b) $n_{h_i} \in \{0,2\}$ for every position h.*

Proof. Note that for any pair (T_i, T_j), since for every position h and for every choice set S_{2_i}, we have $n_{h_i} = 1$ or $n_{h_i} \in \{0, 2\}$, the h-th and k-th positional value, corresponding to main effect F_h and two-factor interaction effect F_k, is either $(i_h0, j_h0)_{hk}$ or $(i_h1, j_h1)_{hk}$. The result then follows from Case 2c of Lemma 4.2 and Lemma 4.4 since $N^+_{hk} = N^-_{hk} = 0$. □

Remark 4.1. *For given N and n, with respect to maximum of $trace(C)$, (i) all designs with m even are equivalent and (ii) a design with m odd is always inferior to a design with m even.*

5 Construction of Universally Optimal Designs

The criteria of *universal optimality* was introduced by Kiefer (1975) and is a strong family of optimality criteria which includes $A-$, $D-$, and $E-$ criteria as particular cases.

Let W_p denote the class of positive definite symmetric matrices of order p. A design $d^* \in \mathcal{D}$ is universally optimal over $\mathcal{D}$ if d^* minimizes $\phi(C_d), d \in \mathcal{D}$ for any $\phi : W_p \to (-\infty, \infty]$ satisfying

1. ϕ is matrix convex, i.e., $\phi\{aC_1 + (1-a)C_2\} \leq a\phi(C_1) + (1-a)\phi(C_2)$ for $C_i \in W_p, i = 1, 2$ and $0 \leq a \leq 1$,
2. $\phi(bC)$ is non increasing in the scalar $b \geq 0$ for each $C \in W_p$,
3. ϕ is invariant under each simultaneous permutation of rows and columns of C in W_p.

Kiefer (1975) obtained the following sufficient condition for universal optimality.

Suppose $d^* \in \mathcal{D}$ and C^*_d satisfies (a) C_{d^*} is scalar multiple of I_p i.e., $C_{d^*} = \alpha I_p$, and (b) $trace(C_{d^*}) = max_{d \in \mathcal{D}}\, trace(C_d)$, then d^* is universally optimal in $\mathcal{D}$.

We now provide few simple methods for constructing universally optimal designs for a 2^n choice experiment with choice set size m under models $\mathcal{R}(1, 2)$ and $\mathcal{R}(1, 1)$. Our characterization of the Information matrix facilitates construction of universally optimal choice designs giving more flexibility for choosing m. Let $\mathcal{D}_{N,n,m}$ be the class of all connected 2^n choice designs involving N choice sets of size m each. In view of Remark 4.1, for $m = 2$, we first provide a simple construction of universally optimal designs under the model $\mathcal{R}(1, 2)$.

Theorem 5.1. *Let $n = 4t - j$, where t is a positive integer and $j = 0, 1, 2, 3$. Also, given a Hadamard matrix H of order $4t$, let $Z_1 = H$ and $Z_2 = -H$. For $w = 1, 2$, let A_w be respective matrices obtained by replacing -1's by 0 and deleting rightmost j columns from Z_w, where $j = 4t - n, j \in \{0, 1, 2, 3\}$. Consider each row of A_w as treatment combination. Then under $\mathcal{R}(1, 2)$, the design $D_1 = (A_1, A_2)$ is universally optimal 2^n choice design in $\mathcal{D}_{4t,n,2}$.*

Proof. To prove that this construction gives universally optimal choice design, we will show that the C-matrix of the design is of the form αI_n, where α is a constant and $trace(C)$ is maximum. Therefore, first we show that every (h, k)-th element of the C-matrix is zero, where $h < k$ and $(h, k) \in \{1, 2, \ldots, n\}$. For design D_1, we first calculate N^+_{hk} and N^-_{hk}; $(h, k) \in \{1, 2, \ldots, n\}$.

Since H is a Hadamard matrix of order $4t$, for any two columns h and k of A_w, the combinations from the set $\{(00)_{hk}, (11)_{hk}\}$ and from the set $\{(10)_{hk}, (01)_{hk}\}$ occurs equally often. Therefore, it is easy to see that for the design D_1, $N^+_{hk} = N^-_{hk}$ for $(h, k) \in \{1, 2, \ldots, n\}$.

The construction of design D_1 also ensures that $n_{h_i} = 1$ for every position h and for every choice set. Therefore, using Theorem 4.5 and Theorem 4.14 it follows that the C-matrix has off-diagonal elements zero. Also, using Theorem 4.12 and Theorem 4.14 we can say that the diagonal elements of C-matrix are equal and $trace(C)$ is maximum for the design. Thus the designs D_1 is universally optimal 2^n choice experiment. □

Next, not restricting to $m = 2$, we give a construction of universally optimal designs under the model $\mathcal{R}(1, 1)$.

Theorem 5.2. *Let $n = 4t - j$, where t is a positive integer and $j = 0, 1, 2, 3$. Also, given a Hadamard matrix H of order $4t$, let for $u = 1, 2, \ldots, 4t$, H_u be the Hadamard matrix derived from H by multiplying the u-th column of H by -1. Let $Z_1 = H, Z_2 = -H, Z_{2u+1} = H_u, Z_{2u+2} = -H_u$. For $w = 1, 2, \ldots, 2n + 2$, let A_w be respective matrices obtained by replacing -1's by 0 and deleting rightmost j columns from Z_w, where $j = 4t - n, j \in \{0, 1, 2, 3\}$. Consider each row of A_w as treatment combination. Then under $\mathcal{R}(1, 1)$, $D_2 = (A_1, A_2), D_3 = (A_1, A_2, A_3), D_4 = (A_1, A_2, A_3, A_4), \ldots, D_{2n+2} = (A_1, A_2, A_3, A_4, \ldots, A_{2n+2})$ are universally optimal 2^n choice design in $\mathcal{D}_{4t,n,m}$ for $m = 2, 3, 4, \ldots, 2n + 2$, respectively.*

Proof. To prove that this construction gives universally optimal choice design, we will show that the C-matrix of the design is of the form αI_n, where α is a constant and $trace(C)$ is maximum. Therefore, first we show that every (h, k)-th element of the C-matrix is zero, where $h < k$ and $(h, k) \in \{1, 2, \ldots, n\}$. Note that the design D_w consists of the *component pair* designs $\{(A_\delta, A_{\delta'}), 1 \leq \delta < \delta' \leq w\}$. We denote the *component pair* designs of D_w by $D^{\delta\delta'}_w, 1 \leq \delta < \delta' \leq w$. We will now calculate N^+_{hk} and N^-_{hk} for the design $D_w, w = 2, \ldots, 2n + 2$.

Since H is a Hadamard matrix of order $4t$, for any two columns h and k of A_w, the combinations from the set $\{(00)_{hk}, (11)_{hk}\}$ and from the set $\{(10)_{hk}, (01)_{hk}\}$ occurs equally often. Therefore, in every *component pair* design $D^{\delta\delta'}_w$, it is easy to see that $N^+_{(\delta\delta')hk} = N^-_{(\delta\delta')hk}, 1 \leq \delta < \delta' \leq w$, where $N^+_{(\delta\delta')hk}$ is the total number of pairs of the type $(00, 11)_{hk}$ corresponding to h-th and k-th positional values in $D^{\delta\delta'}_w$, and $N^-_{(\delta\delta')hk}$ is the total number of pairs of the type $(01, 10)_{hk}$ corresponding to h-th and k-th positional values in $D^{\delta\delta'}_w$. In other words, for the design D_w, $N^+_{hk} = N^-_{hk}$.

Using the result of Theorem 4.5 it thus follows that the C-matrix has off-diagonal elements zero for the design $D_w, w = 2, \ldots, 2n + 2$.

The construction also ensures that $n_{h_i} = m/2$ for D_w's with m even, and $n_{h_i} = (m - 1)/2$ or $(m + 1)/2$ for D_w's with m odd, for every position h and for every choice set. Therefore using Theorem 4.13 we can say that the diagonal elements of C-matrix are equal and $trace(C)$ is maximum for the design. Thus the designs $D_w, w = 2, \ldots, 2n + 2$ are universally optimal design for $m = 2, 3, \ldots, 2n + 2$ respectively for a 2^n choice experiment. □

Remark 5.1. *The construction as provided in Theorem 5.2 can be extended to allow further increase in the choice set size by considering distinct Hadamard matrices H_u derived from H by* *multiplying any s columns of H by -1, $s = 1, 2, \ldots, 2t$. Though such a flexibly may allow having m large, it is desirable to select those H_u which minimizes repetitive sets of options within the constructed choice sets.*

Remark 5.2. *For $m = 4$, the construction as provided in Theorem 5.2 is also universally optimality under the model $\mathcal{R}(1,2)$. Starting with the Hadamard matrix H in normal form, corresponding to $(Z_1 = H, Z_2 = -H, Z_3 = H_1, Z_4 = -H_1)$, the choice design is $D_4^* = (A_1, A_2, A_3, A_4)$. Then, for $h \in \{1, \ldots, n\}, k \in \{n+1, \ldots, n+\binom{n}{2}\}$, because of Theorem 4.14 and the Hadamard property of H, it is seen that either $N^+_{(\delta\delta')hk} = N^-_{(\delta\delta')hk} = 0$ or $2t$ for $(\delta, \delta') = (1,2), (1,3), (2,4), (3,4)$, i.e., for each of the component pair designs $(A_1, A_2), (A_1, A_3), (A_2, A_4), (A_3, A_4)$. Furthermore, for $(\delta, \delta') = (1,4), (2,3)$, the respective component pair designs (A_1, A_4) and (A_2, A_3) have $N^+_{(\delta\delta')hk} = N^-_{(\delta\delta')hk} = 0$ or $2t$ for all $h \in \{1, \ldots, n\}, k \in \{n+1, \ldots, n+\binom{n}{2}\}$ except (h,k) corresponding to F_h, the h-th main effect, and F_k, the two factor interaction involving the first factor and the h-th main effect factor, $h = 2, 3, \ldots, n$. For such (h,k)'s, $N^-_{(14)hk} = N^+_{(23)hk} = 4t$, and $N^+_{(14)hk} = N^-_{(23)hk} = 0$. Therefore, from Lemma 4.4 it follows that $B\Lambda B'_{(2)}$ is a null matrix. The rest follows from Theorem 5.2 in establishing that the design D_4^* is universally optimal in $\mathcal{D}_{4t,n,4}$ under the model $\mathcal{R}(1,2)$.*

Remark 5.3. *In view of Remark 4.1, for given N and n, it follows that a universally optimal choice design in $\mathcal{D}_{N,n,2}$ is also universally optimal in a more broader class of all connected 2^n choice designs involving N choice sets and arbitrary m.*

Example 5.1. *Consider a 2^{8-j} choice experiment ($j = 0, 1, 2, 3$) conducted through 8 choice sets of size 4 each. The 2^8 ($j = 0$) choice design D_4^* (as below), under the model $\mathcal{R}(1,2)$ (as well as under the model $\mathcal{R}(1,1)$), is universally optimal in $\mathcal{D}_{8,8,4}$.*

$$
D_4 = \begin{array}{l}
(11111111, \;\; 00000000, \;\; 01111111, \;\; 10000000) \\
(10101010, \;\; 01010101, \;\; 00101010, \;\; 11010101) \\
(11001100, \;\; 00110011, \;\; 01001100, \;\; 10110011) \\
(10011001, \;\; 01100110, \;\; 00011001, \;\; 11100110) \\
(11110000, \;\; 00001111, \;\; 01110000, \;\; 10001111) \\
(10100101, \;\; 01011010, \;\; 00100101, \;\; 11011010) \\
(11000011, \;\; 00111100, \;\; 01000011, \;\; 10111100) \\
(10010110, \;\; 01101001, \;\; 00010110, \;\; 11101001)
\end{array}
$$

Deleting the last j factors we get the corresponding universally optimal design in $\mathcal{D}_{8,8-j,4}$, under both $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$. Also, taking the first 2 elements from each choice set we get the design D_2 which is universally optimal in $\mathcal{D}_{8,8-j,2}$ ($j = 0, 1, 2, 3$), under both $\mathcal{R}(1,2)$ and $\mathcal{R}(1,1)$. Finally, taking the first 3 elements from each choice set, under the model $\mathcal{R}(1,1)$ we get the design D_3 which is universally optimal in $\mathcal{D}_{8,8-j,3}$ ($j = 0, 1, 2, 3$).

Acknowledgment

Ashish Das's work is supported by the Science and Engineering Research Board, India (SR/S4/MS: 829/12).

References

Bradley, R. A. and Gart, J. J. (1962). The asymptotic properties of ML estimators when sampling from associated populations. *Biometrika*, **49**, 205–214.

Demirkale, F., Donovan, D. and Street, D. J. (2013). Constructing D-optimal symmetric stated preference discrete choice experiments. *Journal of Statistical Planning and Inference*, **143**, 1380–1391.

Dey, A. and Mukerjee, R. (1999). *Fractional Factorial Plans*. Wiley, New York.

El-Helbawy, A. T. and Bradley, R. A. (1978). Treatment contrasts in paired comparisons: large-sample results, applications, and some optimal designs. *Journal of American Statistical Association*, **73**, 831–839.

Großmann, H. and Schwabe, R. (2015). Design for discrete choice experiments. In *Handbook of Design and Analysis of Experiments*. Chapman and Hall/CRC, 787–831.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In: Srivasava, J.N. (Ed.), *A Survey of Statistical Design and Linear Models*. North-Holland, Amsterdam, 333–353.

Manna, S. and Das, A. (2016). Optimal two-level designs for partial profile choice experiments. *Statistics and Probability Letters*, **116**, 80–87.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Singh, R., Chai, F.-S. and Das, A. (2015). Optimal two-level choice designs for any number of choice sets. *Biometrika*, **102**, 967–973.

Street, D. J. and L. Burgess, L. (2007). *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. Wiley, New York.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.