STUDIES ON OPTIMALITY OF SOME CLASSES OF DESIGNS

By

RITA SAHARAY

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements
for the award of the degree of
Doctor of Philosophy

CALCUTTA

1986

ACKNOWLEDGEMENT

My deepest debt of gratitude is due to Prof. B.K. Sinha for his valuable guidance, constant encouragement althrough the investigation and numerous instructive suggestions that have directed the course of this research.

I am also greatly indebted to Prof. A.C. Mukhopadhyay under whose guidance I started my research work first in 1980. I have benefited a lot from many fruitful discussions with him and his numerous helpful suggestions. I remain grateful to Prof. Sinha and Prof. Mukhopadhyay for their kind permission to include our joint work in this thesis. With the greatest pleasure I take this opportunity to record my sincerest gratitude to Prof. J.K. Ghosh for allowing me to work as a research fellow under Prof. B.K. Sinha in the Stat-Math. Division of Indian Statistical Institute since July, 1981.

I am also grateful to Dr. Rahul Mukerjee for many useful discussions and his kind permission to include in this thesis the work I did jointly with him.

My sincere thanks are due to all my Professors, friends and colleagues particularly Prof. G.M. Saha, Prof. P. Bhimsankaram, Mrs. Sunanda Bagchi, Miss Mousumi Sen and Sri D. Bhaumik for many a useful discussion with them.

Finally, I would like to thank Mr. D.K. Bardhan for his efficient and elegant typing with extreme courtesy at all stages and Mr. Mukta Lal Khanra for his efficient cyclostyling.

Rita SahaRay

CONTENTS

*****

CHAPTER 1

INTRODUCTION AND SUMMARY

1a. General Observations and Literature Review

Experimentation plays an essential role in most of the statistical investigations carried out for drawing inferences about certain unknown parameters of interest. If the situation allows for only one experiment to be executed out of a number of available alternative experiments, the experimenter should aim at performing the one which is "optimum" in some sense. This is how the problem of choosing the 'best' experiment comes up.

To judge the relative performances of various statistical experiments, Blackwell (1951, 1953) and Blackwell and Girshick (1954) introduced the concept of 'sufficient experiments'. However, in the context of design settings fitting into the usual Analysis of Variance (ANOVA) models, it is not straightforward to settle the question of existence or non-existence of sufficient experimental design in general terms, even though, in some simple settings (such as oneway ANOVA) the non-existence of a sufficient experiment is readily ascertained. In such situations, therefore, the choice of 'optimum' experiments is guided by specific 'optimality criterion' which evolves from different considerations depending on the particular problems of interest.

The early work of Smith (1918) introduced a formal definition of design optimality in the study of response surface function. The inaugural paper in the literature on optimality of block designs is due to Wald (1943). In this paper, he posed a very important optimality criterion and established optimality of a different kind of the Latin Square Designs (LSD's) (vide also Nandi (1950) in this context). But it was not until 1958, when the theory was given a precise and systematic formulation. Confining to the class of connected block designs Kiefer (1958) considered the general problem of estimating a <u>full set of orthonormal contrasts</u> $\eta = P\underline{\tau}$, where P is $n-1 \times n$ lower submatrix of an $n \times n$ orthogonal matrix $\bar{O}$ having the first row as $(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$. He gave precise definitions and investigated the interrelations of a number of standard optimality criteria, namely A-, D-, E-, L-, M- optimality criteria for judging the performances of the least square estimates of the orthonormal contrasts in η. So far as the non-randomized designs are concerned, his fundamental result establishes such optimality of Balanced Incomplete Block Designs (BIBD's) in one way elimination of heterogeneity set-up and that of the Latin Square Designs (LSD's) and Youden Square Designs (YSD's) in two way elimination of heterogeneity set-up, <u>whenever</u> such designs <u>exist</u>. At the same time, again for estimating a <u>full set</u> of <u>orthonormal treatment contrasts</u>, Mote (1958) independently proved the E-optimality of BIBD's among binary designs and Kshirsagar (1958) proved the A- and D-optimality of BIBD's and YSD's also among binary designs. Contemporarily, Roy (1958) came up with the following findings with respect to A-optimality criterion: For estimating <u>all elementary</u> treatment contrasts, in the class of proper, incomplete block designs, a <u>most efficient</u> design, whenever it exists, is necessarily a BIBD. All these results pertain to the optimality of symmetrical designs. In the absense of symmetrical designs, intuitively it appears that the design <u>closest</u> (in some appropriate sense) to the hypothetical symmetrical design is a reasonable design to use. With this idea, Shah (1960) introduced a new optimality criterion. Motivated by this, Eccleston and Hedayat (1974) developed an optimality criterion called (M.S) optimality criterion. Kiefer (1975) extended almost all of the previously defined optimality criteria to a very general class of criteria by bringing in the notion of " Universal optimality." His fundamental result establishes Balanced Block Designs (BBD's), a generalization of BIBD's for the case when the block size is larger than the number of test treatments, to be universally optimal. A useful subfamily of the general class of criteria, namely φ_p^* - criteria which includes the age old A-, D- and E - optimality criteria, was also discussed separately in this paper. Besides this, a major part of the paper was devoted to the study (optimality, characterization and construction (in brief)) of Generalized Youden Square Designs (GYD's) in the two way elimination of heterogeneity set-up. Generalizing GYD's to multiway settings Cheng (1978a) defined Youden Hyperrectangles and proved their optimality behavior in a way very similar to those of GYD's. Some of the constructions of such designs had been tackled later (Cheng 1979 b) . Cheng (1979a) prepared a list of all D-optimal designs for $v = 4$, and Gaffke (1982) extended the list to $v \leq 6$ with a few significantly general results. In another paper Cheng (1981a) defined a class of Pseudo Youden Designs (PYD's) in the two-way elimination of heterogeneity setting under the special case where number of levels in both the directions is the same and studied their optimality behavior. Extending these designs to analogous m-way setting, he obtained similar results. Cheng (1981b) deals with construction of Pseudo Youden designs with row size less than the number of symbols. In another paper, Cheng (1981c) discussed two related problems in graph theory and optimum design theory : maximizing the number of spanning trees in a graph and finding a D-optimum incomplete block design. Gaffke and Krafft(1979) and Gaffke (1981) considered general two-way models and characterized some classes of optimal designs.

Apart from the study of optimality of symmetrical designs, researches were also undertaken for exploring possible optimality properties of some classes of asymmetrical designs. Takeuchi (1961,1963) was the first to initiate such a study. With a new elegant technique, he showed E-optimality of Group Divisible Designs (GDDs) with $\lambda_2 = \lambda_1 + 1$ whenever they exist (in the absense of BIB designs). Sinha and Sinha (1969) studied relative performances of various PBIB designs with respect to A-optimality criterion. Conniffe and Stone (1974,1975) proved that in the same situation, any Most Balanced Group Divisible Partially Balanced Incomplete Block Design (MBGDPBIBD) of type 1 is A-optimal over a restricted class of competing designs. Exploring their technique, Cheng (1978b) derived some further results on optimality of asymmetrical designs, namely, the optimality of MBGDD of type 1 with respect to a large class of criterion, called " ψ_f - optimality " which includes, in particular, the commonly used A-, D-, E-, and the φ_p^* - criteria. φ_p^* - optimality of another important class of asymmetrical designs, called Linked Block Designs (LBD's), within the class of proper, connected equireplicate designs had been put forward by Sinha (1971) and also by Shah, Khatri and Raghavarao (1976). Cheng (1981d) established ψ_f - optimality of GD designs with $\lambda_2 = \lambda_1 + 1$ along with optimality of other designs within the class of two associate PBIB designs. In another paper by Cheng (1981e) it has been proved that quite often the problem of finding optimum designs can be reduced to considering designs of block sizes two. For block size two the problem can be further reduced to consideration of binary designs. Jacroux (1983) established GDD's with $m = 2$, $n = v/2$, $\lambda_2 = \lambda_1 + 2$ as being E-optimal. Jacroux (1984) also established D-optimality of GDD's with parameters $v = mn$, $m = v/2$, $n = 2$ and $\lambda_2 = \lambda_1 \pm 1$. Removing the condition of equireplicability, Cheng (1980a) proved E-optimality of LBD's and duals of some GDD's within the same class. In this asymmetrical situation, remarkable contributions in respect of E-optimal designs over the class of non-equireplicate designs are also due to Jacroux (1980a, 1980b) and Constantine (1981) among others. Hedayat (1984) characterized symmetrical BIBD as universally optimal design within the class of connected binary designs with arbitrary block sizes and arbitrary replications. In a recent paper, Jacroux (1985) derived a set of sufficient conditions for type 1 optimality of block designs.

All the above discussions pertain to the fixed effects model in which the effects of treatments and heterogeneity directions are assumed to be fixed. Under mixed effects models where the effects of heterogeneity levels are assumed to be random, extensive optimality study has been made by Bagchi(Mukhopadhyay)(1982). The optimality of BBD's, LBD's, GYD's and PYD's have been worked out under this model. In the multiway heterogeneity context, concept of " balanced multiway setting " was introduced in Mukhopadhyay and Mukhopadhyay (1984). The speciality of such a type of setting lies in the fact that the number of experimental units (which usually equals all possible level combinations of the factors involved) was considerably reduced. In the light of Cheng (1978a), Balanced Youden Hyperrectangles (BYHR) and Balanced Pseudo Youden Designs (BPYD) were defined and optimality studies carried out for these designs by Mukhopadhyay and Mukhopadhyay (1984). The asymmetrical case poses a real problem and optimality of MBGDD's has been proved confining to certain classes of designs with block size less than the number of treatments in Bagchi (1985). Optimality of MBGDD's under mixed effects model within the equireplicate designs and also within a wider class of designs including binary designs were also observed by Khatri and Shah (1984) and Bhattacharya and Shah (1984) respectively.

A study of optimality in the context of weighing designs involving the chemical and spring balance designs has received considerable attention of several researchers. Weighing problems originated in a casual illustration furnished by Yates (1935) and a precise formulation of the problem was given by Hotelling (1944). Hotelling (1944) and Mood (1946) also furnished basic optimality results in this area of study. Detailed references are to be found in Raghavarao (1971) and Banerjee (1975). Among other references mention may be made of Sinha (1971, 1972), Swamy (1980, 1982), Galil and Kiefer (1980a, 1980b, 1981, 1982a, 1982b), Cheng (1980b), Jacroux (1986), Jacroux and Notz (1983) and Jacroux, Wong and Masaro (1983).

Apart from the areas of block designs and weighing designs, there are a number of recent contributions to the study of optimality in the area of Repeated Measurements Designs (RMD). These designs have been mentioned in the literature under a variety of names such as cross-over, or change-over designs, time series designs or before-after designs. It was not until 1975, when a systematic study in this area was initiated by Hedayat and Afsarinejad. This paper gives construction of some sporadic families of balanced RMD's, along with an extensive list of references. So far as the optimality aspect is concerned, the first paper is also due to them (1978), where they proved uniform balanced RMD to be universally optimum within the class of non-circular uniform RMD's with number of periods (p) equal to the number of treatments (t). Apart from this result, the paper also contains some construction. Constantine and Hedayat (1982) also devoted to the construction of cyclic balanced and connected RMD's. The more important work of Cheng and Wu (1980) in this area removed the uniformity condition on the competing designs and established universal optimality of strongly balanced uniform RMD's (SBURMD) for direct as well as residual effects within a very general class of RMD's. Along with some other results this paper as also Cheng and Wu (1983) prove universal optimality of balanced RMD's over the class of designs which have no pairs of consecutive identical treatments. Magda (1980) undertook a study of RMD's under circular model. Hedayat (1981a) gives an overview of these results so far described. Dey, Gupta and Singh (1983) proved universal optimality of balanced designs within a subclass of all possible designs for $p < t$. Kunert (1983) dealt with the approach of orthogonality ensuring equality of two information matrices in the finer and simpler model, and using this technique in the special case of RMD's, he derived optimality results. His result shows that there are designs in this context which are optimal and orthogonal but neither strongly balanced nor balanced. Moreover, nearly strongly balanced designs were introduced and their universal optimality was proved. Recent papers by Kunert (1984a, 1984b) bring in optimality of balanced uniform RMD's under circular and non-circular models respectively over the class of all possible designs including those which may have identical pairs of treatments. Mukerjee and Sen (1984) studied the robustness of these optimality results introducing in the model interaction terms for direct and residual effects of treatments.

Another important direction of study and research is that of analyzing the data with reference to non-additive models. Even though we mainly work with additive models, there is no denying the fact that sometimes the simplicity of the underlying models are critically questioned and it is felt that appropriate models are to be developed, specifically for the interactions. Though it is not difficult to build up a model incorporating suitable and relevant interaction terms, often the data may not be adequate for a satisfactory analysis. Tukey (1949) was the first to suggest a non-additive model with the specific form of interaction term, as a constant multiplier θ of the product of corresponding block and treatment effects in the context of Randomized Block Designs (RBD) with b blocks and v varieties. He provided an estimate of θ and also a test for $H_0 : \theta = 0$ under the usual assumptions of the law of distribution of the errors. Since then several types of non-additive models (Scheffe (1959), Gollob (1968), Mandel (1961, 1969), Milliken and Graybill (1970)) have been developed and the work primarily consisted of testing the hypotheses regarding interaction terms and estimation of error variance (σ^2). In this respect, mention may be made also of later authors - Johnson and Graybill (1972a, 1972b), Hegemann and Johnson (1976a, 1976b), Yochmowitz and Cornell (1978), Marasinghe and Johnson (1981, 1982), Kshirsagar (1983). Surprisingly, nothing practically is said about efficient estimation of the non-additive parameter(s).

The above references in the area of study of optimality are not exhaustive. However, they give a strong view of how extensively optimality studies have created interest among research workers in the field of design of experiments.

In the present dissertation the author has made some further studies on relative performances of various designs to find out optimal designs, universal or specific, if any, in the following fields : (i) Two way and three way elimination of heterogeneity set ups with non-orthogonal base structure, (ii) Spring-balance weighing designs, (iii) Repeated Measurements designs and (iv) Tukey's non-additive model and its extended form.

In the sections below, we provide adequate motivations for undertaking the different problems as also brief summaries of the main results in the respective fields of study.

The following section gives an account of different optimality criteria discussed in the literature and cited in the thesis.

1b. Notions of Various Optimality Criteria

We start with the basic linear model

$$E(\underset{\sim}{Y}) = (X_{d1} : X_{d2}) \begin{bmatrix} \underset{\sim}{\theta}_1 \\ \underset{\sim}{\theta}_2 \end{bmatrix}$$

$$\mathrm{Cov}(\underset{\sim}{Y}) = \sigma^2 I_n$$

where $\underset{\sim}{Y}$ is the $n \times 1$ vector of observations, $X_d = (X_{d1} : X_{d2})$ is the $n \times p$ design matrix with known entries underlying the design d, $\underset{\sim}{\theta}' = (\underset{\sim}{\theta}_1' : \underset{\sim}{\theta}_2')$ is the $1 \times p$ vector of unknown parameters. In many cases, we are interested only in linear combinations of the form $\underset{\sim}{\lambda}'\underset{\sim}{\theta}$ of the components of $\underset{\sim}{\theta}_1$. Then the information matrix of $\underset{\sim}{\theta}_1$ under the design d is $X_{d1}'X_{d1} - X_{d1}'X_{d2}(X_{d2}'X_{d2})^- X_{d2}'X_{d1}$ where $(X_{d2}'X_{d2})^-$ stands for any generalized inverse of $X_{d2}'X_{d2}$.

In any design set-up, varietal contrasts are of major concern, and the relevant information matrix for varietal effects is denoted by the well known C-matrix. The reduced normal equations for the vector of varietal (treatment) effects assume the form (suffix d being used in the obvious sense)

$C_d \underset{\sim}{\tau} = \underset{\sim}{Q}_d$ where $\underset{\sim}{Q}_d = \left[X_{d1}' - X_{d1}'X_{d2}(X_{d2}'X_{d2})^- X_{d2}' \right] \underset{\sim}{Y}$.

Usually, row sums and column sums of C_d-matrix are zero. (Exception is observed in the case of residual effects in RMD, Cheng and Wu (1983), however, again a restriction may be imposed on the parametric space in order that this property of C_d is retained.) Since estimability of a linear function $\underset{\sim}{\ell}'\underset{\sim}{\tau}$ is ensured if and only if $\underset{\sim}{\ell}'$ belongs to the row space of C_d, usually the only estimable functions are essentially contrasts of treatments effects.

Following Kiefer (1958, 1959, 1975) we deal with the inferential problem involving a full set of orthonormal contrasts of $\underset{\sim}{\tau}$. Writing the vector $\underset{\sim}{\eta}'$ as $(\eta_1, \eta_2, \ldots, \eta_{v-1})$ and $\underset{\sim}{\tau}'$ as $(\tau_1, \ldots, \tau_v)$, the problem is thus to infer about $\underset{\sim}{\eta} = P\underset{\sim}{\tau}$, where P is a $v-1 \times v$ matrix of rank $v-1$, whose rows are orthonormal and orthogonal to the constant vector $\frac{1}{\sqrt{v}}\underset{\sim}{1}'$. To ensure estimability of all these $v-1$ linearly independent orthonormal contrasts of $\underset{\sim}{\tau}$ involved in $\underset{\sim}{\eta}$, one has to consider only those designs d for which rank of C_d is equal to $v-1$, that is to say one restricts only to what are known as <u>connected</u> designs. The usual least square estimator of $P\underset{\sim}{\tau}$ is $(PC_dP')^{-1}P\,\underset{\sim}{Q}_d$ with variance-covariance matrix $\sigma^2(PC_dP')^{-1} = \sigma^2 V_d$ (say). Thus it is natural to specify some optimality functional ψ on $(v-1)\times(v-1)$ matrices and to pose the problem as follows : Find d to minimize $\psi((PC_dP')^{-1})$. The resulting design may be said to be ψ-optimal with respect to the problem of inferring on $\underset{\sim}{\eta}$. The commonly used optimality criteria are defined below : (The word definition is abbreviated as DFN and $\mathcal{D}$ refers to the competing class of designs).

<u>DFN 1b.1.1</u> <u>(D-optimality)</u> : A design $d^* \varepsilon \mathcal{D}$ is D-optimal iff

$$\det(V_{d^*}) = \min_{d\,\varepsilon\,\mathcal{D}} \det(V_d).$$

Here 'det' stands for determinant. Under normality assumption, the statistical sense of this criterion is the following :

If d^* is D-optimal, d^* minimizes :

a) The volume of the smallest invariant confidence ellipsoid on $\eta_1, \dots, \eta_{v-1}$ for any confidence coefficient (Kiefer 1958).

b) The generalized variance (as defined by Wilks) of the best linear unbiased estimators (blue's) of the parameters.

<u>DFN 1b.1.2</u> <u>(A-optimality)</u> : A design $d^* \varepsilon \mathcal{D}$ is A-optimal iff

$$\operatorname{tr}(V_{d^*}) = \min_{d \varepsilon \mathcal{D}} \operatorname{tr}(V_d).$$

Here " tr " stands for the trace of a matrix and " A " - stands for average. Statistically, if d^* is A-optimal, it minimizes the average variance of the blue of $(\eta_1, \eta_2, \dots, \eta_{v-1})$.

<u>DFN 1b.1.3</u> <u>(E-optimality)</u> : A design $d^* \varepsilon \mathcal{D}$ is E-optimal iff

$$\lambda^{-1}_{d^*(v-1)} = \min_{d \varepsilon \mathcal{D}} \lambda^{-1}_{d(v-1)}.$$

Here " E " - stands for eigenvalue. In statistical sense it has the following interpretations.

a) <u>In hypothesis testing</u> the criterion states that under normality assumption an E-optimal design maximizes the minimum power of the associated F-test of size α on the contour $\underline{\eta}'\underline{\eta} = c$, for every α and every c.

b) <u>In point estimation</u>, an E-optimal design minimizes the maximum variance of the blue of $\underline{\ell}'\underline{\tau}$ over all $v \times 1$ vectors $\underline{\ell}$ with $\underline{\ell}'\underline{\ell} = 1$, $\underline{\ell}'\underline{1} = 0$.

Remark : We shall refer to two inferential problems η and η^* as nonsingular transforms of each other iff $\exists$ a nonsingular $M \ni \eta = M\eta^*$. If further M is orthogonal, η and η^* are said to be orthogonal transforms of each other. Then the following properties hold for the three criteria : If d^* is D-optimum for η, then it is D-optimum for all nonsingular transforms of η. If d^* is A-optimum (E-optimum) for η, then it is A-optimum (E-optimum) for all orthogonal transforms of η.

It is convenient (as pointed out by Kiefer (1975)) to define optimality criterion as a class of convex non-increasing functionals φ on the set of information matrices C_d rather than a class of convex non-decreasing functionals on the set of covariance matrices V_d (which depend on the nature of varietal contrasts considered), since the former is more general than the latter.

Notations used to illustrate some such optimality criteria are listed below :

$\mathcal{B}_{v,o}$ = the class of all $v \times v$ non-negative definite matrices with zero row and column sums.

$\mathcal{D}$ = the class of designs under consideration.

$\mathcal{C} = \{C_d : d \varepsilon \mathcal{D}\}$.

Also let $\lambda_{d1} \geq \lambda_{d2} \geq \cdots \geq \lambda_{d(v-1)} > 0$ be the non-zero eigenvalues of C_d. Note that $C_d \varepsilon \mathcal{B}_{v,o}$ and $\lambda_{dv} = 0$ for all $d \varepsilon \mathcal{D}$.

DFN 1b.1.4 (φ_p^*-Optimality) : Let $\varphi_p^*(C_d) = (\frac{1}{v-1} \sum_{i=1}^{v-1} \lambda_{di}^{-p})^{1/p}$, $0 < p < \infty$.

A design $d^* \varepsilon \mathcal{D}$ is φ_p^*-optimal iff $\varphi_p^*(C_{d^*}) = \min\limits_{d \varepsilon \mathcal{D}} \varphi_p^*(C_d)$.

Note that A-, D-, and E-optimality criteria with respect to η are connected with φ_p^* criterion as follows :

i) When $p = 1$, $\varphi_1^*(C_d) = \frac{1}{v-1} \sum\limits_{i=1}^{v-1} \lambda_{di}^{-1}$ is equivalent to A-optimality criterion.

ii) When p approaches 0, the limiting case of φ_p^* criterion i.e. $\varphi_0^*(C_d) = \lim\limits_{p \to 0} \varphi_p^*(C_d) = (\sum\limits_{i=1}^{v-1} \lambda_{di}^{-1})^{\frac{1}{v-1}}$ is equivalent to D-optimality criterion.

iii) When p approaches ∞, the limiting case of φ_p^* criterion, i.e. $\varphi_\infty^*(C_d) = \lim\limits_{p \to \infty} \varphi_p^*(C_d) = \lambda_{d(v-1)}^{-1}$ is equivalent to E-optimality criterion.

<u>DFN 1b.1.5</u> <u>Universal Optimality (Kiefer 1975).</u>

A design $d^* \varepsilon \mathcal{D}$ is a universally optimal design if d^* minimizes $\varphi(C_d)$, $d \varepsilon \mathcal{D}$, for any $\varphi : \mathcal{B}_{v,0} \to (-\infty, \infty)$ satisfying

(i) φ is convex

(ii) $\varphi(bC_d)$ is non-increasing in the scalar $b \geq 0$ (1b.1.1)

(iii) φ is permutation invariant.

It is not inappropriate to mention here Proposition 1, and Proposition 2 of Kiefer (1975) as these are important tools so far established in verifying universal optimality and/or φ_p^*-optimality of a large class of designs.

DFN 1b.1.6 A square matrix M of order v is called completely symmetric (c.s.) if M is of the form $aI_v + bJ_v$, where a and b are scalars.

Proposition 1

Suppose for a design $d^* \varepsilon \mathcal{D}$ the class $\mathcal{C}$ contains C_{d^*} for which

a) C_{d^*} is c.s.

b) $\text{tr } C_{d^*} = \max\limits_{d \varepsilon \mathcal{D}} \text{ tr } C_d$

Then d^* is universally optimal in $\mathcal{D}$ (since $-\text{ tr } C_d$ always satisfies (1b.1.1), it follows that condition (b) is always necessary for any universally optimal design to exist).

Clearly universal optimality criterion includes all φ_p^* criteria for $p \geq 0$.

Proposition 2

If $\varphi_1 \leq \varphi_2$, on $\mathcal{C}$, with equality for C_{d^*} and if C_{d^*} is φ_1-optimal, then C_{d^*} is φ_2-optimal as well.

As a consequence of Proposition 2, if C_{d^*} is c.s. and d^* is φ_p^*- optimal, then d^* is φ_q^*- optimal $\forall\, q > p$.

Apart from the problem of inferring on $\underline{\eta} = p\underline{\tau}$, the general linear inferential problem involving the parameters $(\tau_1, \tau_2, \ldots, \tau_v)$ can be looked upon as one of inferring on $\underline{\xi} = L\underline{\tau}$ where L is an $i \times v$ matrix, $i \geq 1$, $L\underline{1} = 0$.

With reference to this problem, we call a design d as acceptable iff all the components of $\underline{\xi}$ are estimable under d. We refer to this problem as nonsingularly estimable (full rank) problem iff rank $(L) = i$ (rank $(L) = i = v - 1$). For rank $(L) = v-1 < i$, we refer to this problem as a singularly estimable full rank problem. As in the case of $\underline{\eta} = P\underline{\tau}$, here also all the various relevant optimality criteria refer to the dispersion matrix $D(\hat{\underline{\xi}})$ of $\hat{\underline{\xi}}$, the blue of $\underline{\xi}$.

Consider again a general linear model $\underline{Y} = X\underline{\theta} + \underline{\varepsilon}$ and this time consider the problem of inferring on $\underline{\theta}$ as a whole, assuming that the components of $\underline{\theta}$ are also unrestricted (as in weighing designs, for example). All relevant design matrices are now necessarily of full column-rank and, as a result, the information matrices $X'X$ are positive definite (p.d.). Universal optimality criterion under such a full-rank set-up has been formulated by Kiefer (1975) and reformulated, among others, by Sinha and Mukerjee (1982). We will use the following result due to Sinha and Mukerjee (1982).

Proposition 1'

If a class of matrices $\mathcal{C} = \{C_d : d \in \mathcal{D}\}$ contains C_{d*} which is a multiple of the identity and which also maximizes $tr(C_d)$ for $d \in \mathcal{D}$, then d^* is universally optimum in $\mathcal{D}$ in the sense that d^* minimizes $\varphi(C_d)$ for any convex, permutation-invariant criterion φ satisfying $\varphi(aI + bJ) \geq \varphi(\alpha I)$ with $\alpha \geq a + b$.

We refer Hedayat (1978, 1981b) and Ash and Hedayat (1978) for further details on other optimality criteria.

1c. Chapterwise Summary of the Work of the Thesis.

In Chapter II, we take up the problem of characterization and construction of optimal designs underlying two-way and three-way elimination of heterogeneity set-ups with non-orthogonal framework. The term " non-orthogonal framework " connotes the following : For eliminating heterogeneity in two or more directions, " orthogonal framework " technically means that all possible level combinations of every pair of directions permit experimentation and are included in the experiment equally often. That is to say, the levels of the experimental units form an orthogonal array of strength two when the symbol in the i^{th} row and u^{th} column of the array is identified as the level of the i^{th} direction in u^{th} experimental unit, which means that the incidence pattern of every pair of directions is taken to be represented (except for a multiplier ≥ 1) by the matrix $J = ((1))$ of all 1's. However, in practice, situations may arise when the incidence structure of the basic frame no longer remains orthogonal. As for example, there are situations where some sporadic cells may remain empty because the corresponding level combinations are infeasible. Recently, Adhikary and Panda (1983) brought out and explained some concrete physical situations in two-way elimination of heterogeneity set-up where the above sort of peculiarity in row-column structure cannot be overlooked. This is that, some combinations of rows and columns may not be feasible when identified with levels of some organic and inorganic manures in the context of agricultural experiments. When this is the situation, the usual optimality results so far obtained for orthogonal framework break down and it necessitates a separate study of optimal designs with non-orthogonal framework. Motivated by this idea, a study has been carried out in this chapter underlying such two-way and three-way elimination of heterogeneity set-ups with non-orthogonal incidence structure for pairs of directions. Specifically, in two-way elimination of heterogeneity set-up (section 2a), we work with the non-orthogonality in the sense of empty cells along the principal diagonal of a $b \times b$ array, that is to say, the row-column incidence structure is assumed to be $(J-I)$. The cases of $b \equiv 1 \pmod v$ (subsection 2a.2) and $b \equiv 0 \pmod v$ (subsection 2a.3) are taken up separately. In case of $b \equiv 1 \pmod v$ intuitively it appears that a design d^*, which assigns each treatment equally frequently to each of the rows and columns should turn out to be optimal, and indeed this is universally optimal, as demonstrated by Theorem 2a.2.1. Construction of such designs is also undertaken in this situation. The case of $b \equiv 0 \pmod v$ poses a real problem. No universally optimal design could be identified, as Proposition 1 of Kiefer (1975) turns out to be inapplicable. Thus specific A-, D- and E-optimal designs have been characterized. Essentially Kiefer's technique (1975) of " Concave envelope " used to prove specific optimality of GYD's, has been applied with necessary modifications as required in this particular non-orthogonal set-up. Theorem 2a.3.1 gives the relevant result.

Theorem 2a.3.1. Suppose for given $b = mv$, $m \geq 2$, with $b \times b$ row-column incidence structure as $J-I$, there exists a design d^* for which

(i) treatment - row incidence pattern is a BBD (incidence is m or $m-1$)

(ii) treatment - column incidence pattern is a BBD (incidence is m or $m-1$)

(iii) and, moreover, the b pairs of treatments denoted by (h_i, h_i'), $i = 1,2,\ldots,b$ where h_i is the treatment which occurs $(m-1)$ times in the i^{th} row and h_i' is the treatment which occurs $(m-1)$ times in the i^{th} column, are such that they

(a) satisfy $h_i \neq h_i'$ $\forall$ $i = 1,2,\ldots,b$

and

(b) exhaust all possible $\binom{v}{2}$ pairs of treatments equally often.

Then d^* is D - optimal for all $m \geq 2$, $v \geq 5$

A - optimal for all $m \geq 2$, $v \geq 4$

E - optimal for all $m \geq 2$, $v \geq 3$.

It is to be noted that d^* does not exist for $m = 1$. Construction of optimal designs d^* has been undertaken for the cases of (i) v = even integer, (ii) v = odd prime or prime power (subsection 2a.4). These optimal designs are combinatorially quite involved and lead us to the study of relative efficiencies of designs having nice simple structure compared to the exact optimal designs d^* characterized by this study. A simple class of designs proposed by Aggarwal (1966b) is seen to possess high efficiency even for moderate values of m and v.

An analogous study for three-way elimination of heterogeniety has also been carried out (section 2b). Here non-orthogonality is understood and utilized in the following sense : Let E be an $OA(b^2,3,b,2)$ such that the three constraints have the same level combinations in each of the first b columns, i.e. the first b columns are of the type $(i\ i\ i)'$, $i = 1,2,\ldots,b$. Then the remaining $b(b-1)$ columns of this orthogonal array serve as experimental units for us in the three-way elimination of heterogeneity set-up, where the entry in the i^{th} row of the u^{th} column denotes the level of i^{th} direction in the u^{th} experimental unit $i = 1,2,3$; $u = 1,2,\ldots,b(b-1)$. That is to say, the incidence structure is taken to be $J - I$ for every pair of directions. The characterization of optimal designs for the case $b \equiv 1 \pmod v$ and $b \equiv 0 \pmod v$ are more or less similar to those under two-way elimination of heterogeneity set-up. In fact, for $b \equiv 0 \pmod v$ optimal design d^* is such that, for every direction, treatment - direction incidence matrix is a BBD, and for every pair of directions, condition (iii) of Theorem 2a.3.1 holds. We also consider the construction of such designs which are more involved than under two-way elimination of heterogeneity set-up. Detailed construction for (i) $v = 2^t$, $t > 2$ integer, and (ii) $v = tk + 1$, odd prime or prime power, k odd, $t > 2$ integer has been taken up (subsection 2b.4). As in the case of two-way elimination of heterogeneity set-up, here also we construct highly efficient simple designs appropriately generalizing Aggarwal's (1966b) designs to the three-way elimination of heterogeneity set-up (subsection 2b.5). The contents of sections 2a and 2b are primarily based on the paper SahaRay (1986) and the technical report Mukhopadhyay and SahaRay (1985) respectively.

In Chapter III we discuss further aspects of optimal weighing designs. We formulate the weighing design problem in the language of measurement of distances among a set of fixed objects along a line. To be specific, suppose there are $(n+1)$ objects, serially numbered $1,2,\ldots,n+1$, fixed along a line and we are interested in measuring distances between any two objects (or some function thereof) by taking N measuring operations. The mathematical and combinatorial aspects of the problem can be described as follows : A set of recorded observations follow the standard regression model

$$\underset{\sim}{Y}_{N\times 1} = X_{N\times n}\,\underset{\sim}{\theta}_{n\times 1} + \underset{\sim}{\varepsilon}_{N\times 1}$$

$$E(\underset{\sim}{\varepsilon}) = 0,$$

$$E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = \sigma^2 I_N$$

where X is a $(0,1)$ matrix and $\underset{\sim}{\theta}$ is the parameter vector of unknown consecutive distances between the objects. An interesting feature of this design matrix X is that each row of X contains only one run of 1's, as, in measuring distances between any two points, we automatically take account of the intermediate points, if any. This special property of X has been termed as " string property " in this thesis; this has also been referred to as " consecutive one's property " by others. The problems of interest here are inferences on $\underset{\sim}{\theta}$ (as a whole) and on a full set of orthonormal contrasts $\underset{\sim}{\xi} = P\underset{\sim}{\theta}$ where P is as usual $n-1\times n$ lower submatrix of the orthogonal matrix $\bar{O}$ with the first row vector as $(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$. Incidentally the following interesting combinatorial properties of the $n\times n$ $(0,1)$ matrices with string property have been observed and made use of in deducing the optimality results in the special case $N = n$.

Property 1 : $|X| = \pm 1$ or 0 according as X is non-singular or not. Whenever X is non-singular, the following proposition holds.

Property 2 : (i) The elements in X^{-1} are $(0, \pm 1)$ and, in each column, the non-zero elements occur with alternate signs.

(ii) The column totals (through all or any subset of consecutive rows) are $(0, \pm 1)$ - not all being 0's.

(iii) If a certain column total is $+1(-1)$ the first non-zero entry in that column is $+1\ (-1)$, if a certain column total is 0, the first non-zero entry in that column may be $+1$ or -1.

The main optimality results are stated below.

Let $\Omega(N,n)$ be the class of designs with string property for N observations with n objects.

Theorem 3a.2.1. Any X in $\Omega(n,n)$ is D-optimal for inferring on $\underline{\theta}$ and, moreover, $X^* = I_n$ is uniquely φ_p^*-optimal over the class $\Omega(n,n)$ for all $p > 0$ ($p = 0$ corresponds to D-optimality criterion).

Theorem 3a.2.2. $X^* = I_n$ is uniquely D-optimal over the class $\Omega(n,n)$ for inferring on $\underline{\xi} = P\underline{\theta}$.

Apart from this case of $N = n$, we have also suggested an optimum way, using D- and A-optimality criteria, of choosing one/two additional observations, given an optimal choice of the first n observations (section 3b). (The choice of E-optimal design for the general case of (N,n) has been recently solved by Jacroux (1986)). We also take up the general case of $N > n$ separately (section 3c). Here choice of the optimal " exact N- observation design " as such becomes intractable. Instead, following the search of " approximate " regression designs originally formulated by Kiefer and Wolfowitz (1959) and used effectively by Kiefer (1959), Fedorov (1972), Kiefer and Studden (1976), Studden (1978), Silvey (1980), we formulate here the design problem as that of assessing a " probability measure " η on the design space $\mathcal{X}$ of all possible $n \times 1$ (0,1) vectors with string property. With this approach, we find out A-, D-, and E- optimal probability measures and accordingly the " approximate " optimal designs are derived when N observations are taken. (The results pertaining to E-optimality have been independently found out by Mukerjee and SahaRay (1985)). For some particular combinations of N and n, these " approximate " designs reduce to " exact " optimal designs. As for example, in case of $N = \frac{tn(n+1)}{2}$, $t \geq 1$ integer, it has been demonstrated in Corollary 3c.2.1 that the exact D-optimal design puts equal mass to each of the $\frac{n(n+1)}{2}$ vectors with string property. On the other hand, for $N = tn$ (Corollary 3c.4.1) an E-optimal design puts equal mass only to the rows of identity matrix and no mass to other vectors with string property. The approximate A-optimal design measure is comparatively of more difficult nature and does not reduce to any exact result as in the case of D- or E-optimality even for particular combinations of N and n. Apart from these problems, we give a brief discussion on optimal choice for measuring the <u>total</u> length covering the set of objects (section 3d). The major part of sections 3a and 3c have been published in Sinha and Saha (1983) and Mukerjee and SahaRay (1985) respectively.

Chapter IV extends the optimality results already established in the context of Repeated Measurements Designs,(RMD (t,n,p)) assuming totally fixed effects model (Hedayat and Afsarinejad(1978)), Cheng and Wu (1980 , 1983), Magda (1980), Kunert (1983, 1984a, 1984b), to the case of mixed effects model. In practice, often the situation arises when the experimental units included in the experiment constitute a random sample from a population of all available experimental units. The present optimality study is carried out assuming an additive mixed effects model where effects due to periods, as well as direct and first order residual effects of treatments are retained as fixed effects while effects due to units are taken to be random. All the universal optimality results regarding strongly balanced or balanced designs both under non-circular and circular model of Hedayat and Afsarinejad(1978), Cheng and Wu (1980, 1983), Magda (1980) could be extended to mixed effects set-up (sub-sections 4b.1, 4b.2, 4b.3 and section 4d) except that Theorem 4.4 of Cheng and Wu (1980) was found to hold only under a suggested modification. The formal proof of Lemma 4b.3.1 is a new one, and it is not explicitly mentioned in Cheng and Wu (1980 or 1983) though in the proof of the theorem they assume this result. Apart from the so far established results under fixed effects model, we independently take up a study of optimality of balanced designs under non-circular model over the class of designs in which no treatment precedes itself and the number of periods is $\lambda_2 t+1$, (subsection 4b.4). The analogous results of Theorem 4b.4.1, Theorem 4b.4.3 under fixed effects model are not provided in Cheng and Wu (1980). Regarding the optimal designs of Kunert (1983) which are neither balanced nor strongly balanced but satisfy some orthogonality conditions, we have shown that the corresponding results for the case $(t/n,\ t/p)$ and $(t \not| n,\ t/p)$ hold good under mixed effects model (Theorem 4c.1.1, Theorem 4c.1.2) whereas the corresponding mixed effects analogue for the case of $t/n,\ t \not| p$ could not be established. The mixed effects version of Theorem 5.3 and Theorem 5.4 of Kunert (1983) have also been established. These results prove optimality of nearly strongly balanced designs over a subclass of all possible designs under non-circular model (i.e., without preperiod). But the techniques developed and used in the present chapter are found to be inadequate to establish the mixed effects analogue of Theorem 5.8 of the same paper of Kunert, which extends the above-mentioned results for nearly strongly balanced designs to the general class of <u>all</u> designs. Moreover, the results of Kunert (1984a, 1984b) which establish optimality of balanced uniform designs for non-circular and circular models respectively over the class of all possible designs including those which have pairs of consecutive identical treatments, could also not be established in the mixed effects situation. Apart from extending optimality results to the mixed effects model, we have also undertaken here construction of nearly strongly balanced RMD's (section 4e). The major part of the contents of subsections 4b.1, 4b.2, 4b.3, and section 4d of this chapter have been published in Mukhopadhyay and Saha (1983).

In Chapter V, we study some aspects of estimation of non-additive parameter with special reference to Tukey's non-additive model as applied to a general block design set-up. (subsection 5b.1 and subsection 5b.2). A general non-additive model as taken up by Milliken and Graybill (1970), as also discussed by Kshirsagar (1983), can be described as

$$\underline{Y} = X\underline{\gamma} + F\underline{\theta} + \underline{\varepsilon}$$

where $\underline{\gamma}$ is the vector of unknown additive parameters, $\underline{\theta}$ is the vector of interaction parameters and the elements of F are arbitrary function of estimable parametric functions of $\underline{\gamma}$ -estimable under the simple linear model $\underline{Y} = X\underline{\gamma} + \underline{\varepsilon}$. The existing literature gives emphasis on the tests for non-additivity based on residuals obtained under the model assuming $\underline{\theta} = \underline{0}$, and overlooks the problem of efficient estimation of interaction parameter $\underline{\theta}$. In the present work we make an attempt in the latter direction. Mainly with estimation of the non-additive parameter in view, formal definition of estimability of $\underline{\theta}$ in this non-linear model has been introduced in DFN 5b.1.1. Theorem 5b.1.1 and Theorem 5b.1.2 justify some intuitive feelings regarding estimability of $\underline{\theta}$ in the given sense. It turns out that under Tukey's non-additive model applied to the <u>general</u> block design set-up i.e.

$$Y_{ij} = \mu + \beta_i + \Sigma\delta_{ijh}\tau_h + \theta\beta_i(\Sigma\delta_{ijh}\tau_h) + \varepsilon_{ij}, \quad 1 \leq j \leq k_{di}, \quad 1 \leq i \leq b,$$

not all designs provide estimation of θ. Precisely, we establish the following :

<u>Theorem 5b.2.1</u> Under Tukey's model applied to the general block design set-up, a connected block design $d(v, b, k_{d1}, \ldots, k_{db})$ will provide estimation of θ if at least one pair of treatments (h, h') say, occur in two different blocks simultaneously.

Theorem 5b.2.2 and Corollary 5b.2.2 characterize a class of designs from which θ is <u>not</u> estimable.

Again, the problem of estimation of an interaction parameter vector $\underset{\sim}{\theta}$ under the above non-additive model generalized to involve up to second powers of τ_h's and β_i's has been focused (subsection 5b.3). As a matter of fact, it has been pointed out that whatever the block design adopted in a non-additive model, the interaction parameters corresponding to higher powers of β_i or τ_h <u>alone cannot</u> be estimated. For example, if the model involves terms like $\theta_1\beta_i^2$ or $\theta_2\Sigma\delta_{ijh}\tau_h^2$, θ_1 or θ_2 is not estimable and, as such, the multiparameter problem under consideration involves terms including both β_i and τ_h. As in the case of single non-additive parameter θ, here also sufficient condition on designs providing estimation of $\underset{\sim}{\theta}$ has been derived in Theorem 5b.3.3. Apart from judging estimability of θ and $\underset{\sim}{\theta}$, we focus our attention to the problem of efficient estimation of single interaction parameter θ under Tukey's non-additive model within the class of all connected block designs providing estimation of θ with fixed N (total number of experimental units) and block sizes $\leq v$. The relevant results are stated below :

Let $\Omega(N,v)$ denote the class of competing designs.

<u>Theorem 5c.2.1</u> Whenever v/N, an RBD is the uniformly best among all binary designs in $\Omega(N,v)$.

<u>Theorem 5c.3.2</u> An RBD is the minimax design within the class of designs $\Omega(N,v)$ if non-binary designs are also allowed to be judged.

Regarding multiparameter problems it has been observed that an RBD does not behave nicely in the given context and as such it is difficult to establish any optimality result for multiple interaction parameters (subsection 5c.4). The contents of this chapter have been taken from the paper Sinha, SahaRay and Mukhopadhyay (1985).

1d. Definitions and Notations.

In this section we give a brief description of the notations and definitions of some known designs widely used in the present thesis.

1d.1 Notations.

$\underline{x}$ denotes a vector of appropriate order. 1_n, J_n, I_n stand respectively for the $n \times 1$ vector of all 1's, $n \times n$ matrix of all 1's identity matrix of order n. Sometimes we omit the lower suffix 'n' to avoid complexity in notation.

$\mathrm{Diag}(r_1, r_2, \ldots, r_n)$ denotes the diagonal matrix of order n with the element r_i in the $(i,i)^{th}$ diagonal entry, $i = 1,2,\ldots,n$.

$(A)_{ij}$ stands for the $(i,j)^{th}$ element of the matrix A. Sometimes we write $A = ((a_{ij}))$ to denote the matrix A with elements a_{ij}. λ_{min} and λ_{max} stand respectively for the minimum and the maximum eigenvalue of the underlying matrix.

$\mathrm{tr}(A)$, rank (A), and $\det(A)$ or $|A|$ denote respectively trace, rank and determinant of the matrix A.

$\mathcal{M}(A)$ stands for the vector space spanned by the column vectors of the matrix A. A' and A^- denote respectively the transpose and a generalized inverse of the matrix A.

'$A \otimes B$' denotes the Kronecker-product between A and B (Rao 1965).

By the symbol " $A \geq B$ " we mean to say $A - B$ is non-negative definite (n.n.d).

$[x]$ denotes the greatest integer $\leq x$.

$(x_i)_{i=1}^{n}$ denotes the sequence $x_1, x_2, \ldots, x_n$.

A statement " f is $\uparrow (\downarrow)$ " means that the function f is increasing (decreasing) in its argument; " f is $\Uparrow (\Downarrow)$ " means that f is non-increasing (non-decreasing).

The notations 'a/b' and '$a \not| b$' mean respectively that " a divides b" and " a does not divide b" . " $(a,b) = k$" means that the greatest common divisor of a and b is k.

The standard symbols '$\forall$', 'ε', '$\exists$', '$\ni$', '$\Rightarrow$', '$\Leftrightarrow$', '$\cap$', '$\cup$', '$\supset$' '$\supseteq$', '$\propto$' etc. are used in their usual senses in this thesis.

'GF(v)' denotes the 'Galois Field' containing v elements (Chakrabarti 1962).

In general, for a design d, N_d stands for the treatment - block incidence matrix with elements n_{dij} and r_{di}, k_{dj}, n_{dij}, λ_{dij} stand respectively for replication of the i^{th} treatment, size of the j^{th} block, number of occurrences of the i^{th} treatment in the j^{th} block, and $(i,j)^{th}$ element of $N_d N_d'$.

1d.2 Definitions.

The terms and concepts used in one way elimination of heterogeneity setting i.e. when the experimental units are subject to heterogeneity along one direction are described as follows : For a one-way design, let v stand for the number of treatments and b for the number of blocks.

DFN 1d.2.1 Proper Design :

A block design d is said to be proper if the block sizes are equal i.e. $k_{d1} = k_{d2} = \ldots = k_{db} = k$ (say).

DFN 1d.2.2 Equireplicate Design :

A block design d is said to be equireplicate if all the treatments occur equally often i.e. $r_{d1} = r_{d2} = \ldots = r_{dv} = r$ (say).

DFN 1d.2.3 Binary Designs :

A block design d is said to be binary if n_{dij} can take only two values – 0 or 1 for all i,j. Naturally, then, $K_{di} < v$ for all i.

DFN 1d.2.4 Randomized Block Design (RBD) :

A proper connected one-way design with b blocks of k plots is said to be Randomized Block Design (RBD) if the following relations hold.

(i) $v = k$

(ii) $n_{dij} = 1 \quad \forall\ i = 1,2,\ldots,v; \quad j = 1,2,\ldots,b.$

DFN 1d.2.5 Balanced Block Design (BBD) :

A proper connected one-way design with b blocks of k plots each is said to be a Balanced Block Design (BBD) if the following relations are satisfied :

(i) v/bk and $r_{di} = \sum_{j=1}^{b} n_{dij} = \frac{bk}{v}$, $i = 1,2,\ldots,v$

(ii) $\left| n_{dij} - \frac{k}{v} \right| \leq 1, \quad j = 1,2,\ldots,b, \quad i = 1,2,\ldots,v.$

(iii) $\lambda_{dii'} = \sum_{j=1}^{b} n_{dij} n_{di'j} = \lambda \quad \forall\, i \neq i', \; 1 \leq i, i' \leq v.$

Balanced Incomplete Block Designs (BIBD's) form special classes of BBD's whenever $k < v$ and RBD's obtain whenever $k = v$.

The terms and concepts used in two-way elimination of heterogeneity designs are described below :

DFN 1d.2.6 Latin Square Design (LSD) :

An arrangement of v treatments in v^2 plots of v rows and v columns is said to be Latin Square Design of order v if each treatment occurs once and only once in each row and in each column.

DFN 1d.2.7 Mutually Orthogonal Latin Squares (MOLS)

If two Latin Squares of the same order and with the same letters be such that when the two squares are superimposed, each letter of one square pairs exactly once with each letter of the other square, then they are said to be mutually orthogonal.

DFN 1d.2.8 Graeco - Latin Squares

Graeco Latin Square is another name of a pair of orthogonal Latin Squares superimposed on one another, the treatments being represented by Greek letters in one square and Latin letters in the other. In this arrangement, every Greek letter (Latin letter) occurs once in each row, once in each column and once with each Latin letter (Greek letter).

DFN 1d.2.9 Generalized Youden Square Design (GYD)

An arrangement of v treatments into a rectangle of size $b_1 \times b_2$ such that rows constitute blocks of a BBD as also the columns constitute blocks of a BBD, has been termed as Generalized Youden Square Design (GYD) by Kiefer (1975).

In the special cases of rectangular setting, when v equals either b_1 or b_2, a GYD is simply called a Younden Square Design (YSD).

Specifically when v/b_1 and also v/b_2, a GYD has been termed as Generalized Latin Square (GLS) design by Kunert (1983).

DFN 1d.2.10 Weighing Designs :

Suppose n objects are to be weighed in a chemical balance (in which the objects can be placed on either of the two pans) or in a spring balance (in which the objects can be placed on a single pan), involving exactly N weighing operations. Then the design matrix will be a $N \times n$ matrix with each element equal to

a) $+1$, -1 or 0 in case of a chemical balance

b) 0 or $(+)1$ in case of a spring balance.

Any arrangement of objects results in a weighing design in general terms.

DFN 1d.2.11 Repeated Measurements Design (RMD).

An experiment in which each unit is exposed repeatedly to a sequence of identical or different treatments is called a Repeated Measurements Design (RMD).

DFN 1d.2.12 Orthogonal Array (OA).

A $k \times n$ matrix A with elements from a set of $s(\geq 2)$ elements is called an orthogonal array of size N, k constraints, s levels, strength t, and index λ if any $t \times n$ submatrix of A contains all possible $t \times 1$ column vectors with same frequency λ. Such an array is denoted by $OA(N,k,s,t)$; N is also called the number of assemblies. In view of this definition, trivially, we must have $N = \lambda s^t$.

Orthogonal arrays of strength two will be used in some chapters of this thesis.

Now we quote a well known lemma (without proof) which we will frequently use in the derivation of optimal designs in almost all chapters.

Lemma 1d.2.1 For any positive integers s and t, the minimum of $\sum_{i=1}^{s} n_i^2$ subject to $\sum_{i=1}^{s} n_i = t$ where the n_i's are non-negative integers is obtained when $t - s\left[\frac{t}{s}\right]$ of the n_i's are each equal to $\left[\frac{t}{s}\right] + 1$ and the others are each equal to $\left[\frac{t}{s}\right]$, where $\left[\frac{t}{s}\right]$ is the largest integer contained in $\frac{t}{s}$.

CHAPTER 2

TWO-WAY AND THREE-WAY ELIMINATION OF HETEROGENEITY SETTINGS WITH NON-ORTHOGONAL FRAMEWORK

Introduction

There is a good deal of literature available on the combinatorial, constructional and analysis aspects of designs eliminating heterogeneity in two or more directions. In such set-ups, usually an orthogonal framework is assumed, that is to say, the incidence pattern of every pair of directions is taken to be represented (except for a multiplier ≥ 1) by the matrix $J = ((1))$ of all 1's. Technically, this means that all possible level combinations for every pair of directions permit experimentation and is included in the experiment equally often.

Available optimality results (studied by Kiefer (1958, 1975), Cheng (1978a), Mukhopadhyay and Mukhopadhyay (1984)) deal exclusively with this orthogonal framework. The multiway heterogeneity setting considered by Cheng (1978a) is a completely orthogonal setting in the sense that each one of all possible level combinations of the heterogeneity directions or factors appears a constant number of times in the experiment. Realising that the complete set of observations may not be available (as also may not be necessary, since the usual fixed effects additive model incorporates only the main effects for the different heterogeneity directions or factors assuming all interaction effects of the factors to be negligible), Cheng's (1978a) multiway setting has been extended to a balanced multiway setting by Mukhopadhyay and

Mukhopadhyay (1984). The latter setting does not necessarily involve all the levels of the heterogeneity directions or factors, but those present make possible orthogonal estimation of the main effects of each factor. Take, for example, a three-way orthogonal setting with b levels for each of the directions. Then a complete setting (Cheng(1978a)) involves all possible b^3 experimental units with incidence structure for every pair of directions as bJ whereas a balanced setting (Mukhopadhyay and Mukhopadhyay (1984)) can be easily constructed with only b^2 experimental units having incidence structure for pairs of directions as simply J. In other words, the level combinations of the experimental units in the earlier setting of Cheng (1978a) form an $OA(b^3,3,b,2)$ whereas in the latter setting involving only b^2 units, they form an $OA(b^2,3,b,2)$, identifying the i^{th} row in the u^{th} column of the array as the level of the i^{th} direction in the u^{th} experimental unit.

However, in practice, situations may arise where the incidence structure of various pairwise heterogeneity directions no longer remain orthogonal; as for example, there are situations where some sporadic cells remain empty because the corresponding level combinations are infeasible. In such cases, usual analysis for orthogonal set-up breaks down, and appropriate modifications are needed. In the context of two way elimination of heterogeneity (commonly known as row-column designs), Aggarwal (1966a) derived the distribution of adjusted row and column sum of squares and the conditions for orthogonality of estimable row, column and treatment contrasts for a general row-column incidence structure, allowing for such empty cells. In subsequent papers

(Aggarwal (1966b, 1966c), Aggarwal and Sharma (1976)) he presented a series of two-way designs covering the situations where the cells along the principal diagonal are empty. Recently Adhikary and Panda (1983) brought out and explained some concrete physical situations where the above sort of peculiarity in the row-column structure cannot be overlooked. This is that some combinations of rows and columns may not be feasible when identified with levels of some organic and inorganic manures in the context of agricultural experiments.

With this in mind, here we initiate a study on optimal designs underlying such two and three-way elimination of heterogeneity set ups with non-orthogonal incidence structure for pair of directions. Specifically, we work with non-othogonality in the sense of empty cells along a transversal[1] of a $b \times b$ array for two-way layout. Without loss of generality, the transversal can be taken along the principal diagonal and the incidence structure for the two directions is then assumed to be $J - I$. In case of three-way elimination of heterogeneity set-up, we consider a set-up where feasible experimental units are those for which levels along <u>any</u> two directions are not the same. That is to say, for

---

[1] <u>Footnote</u> : Here the term " transversal " is used in the following sense : (vide Mukhopadhyay and Mukhopadhyay (1981)).

In a $b \times b$ array, a transversal is a collection of cells (i_1, j_1), $(i_2, j_2), \ldots, (i_b, j_b)$ where $(i_1, i_2, \ldots, i_b)$ and $(j_1, j_2, \ldots, j_b)$ both represent permutations of the numbers $(1, 2, \ldots, b)$.

every two directions, level combinations along the principal diagonal are infeasible. Let each of the three directions assume b levels. Following Cheng (1978a), therefore, if we assume the incidence structure for every pair of directions as $(b-2)(J-I)$ the number of feasible experimental units for allocation of v treatments becomes $b(b-1)(b-2)$ which may be too large to be available to the experimenter in practice. On the other hand, following Mukhopadhyay and Mukhopadhyay (1984), the number of experimental units can be reduced to $b(b-1)$ as follows, thereby producing the incidence structure as $J-I$ for every pair of directions. Let E be an $OA(b^2,3,b,2)$ such that the three constraints have the same level combinations in each of the first b columns, i.e. the first b columns of the OA are of the type $\begin{pmatrix} i \\ i \\ i \end{pmatrix}$, $i = 1,2,\ldots,b$. Then the remaining $b(b-1)$ columns of this orthogonal array serve as experimental units for us in the three-way elimination of heterogeneity set-up, where the entry in the i^{th} row of the u^{th} column denotes the level of i^{th} direction in the u^{th} experimental unit. We will deal with this latter incidence structure for the three-way elimination of heterogeneity designs.

With the above sort of peculiarities in the framework for both two-way (section 2a) and three-way elimination of heterogeneity set-ups (section 2b) we initiate a study on some problems of characterization and construction of optimal designs for inference on varietal contrasts. We have only studied the cases of $b = mv$ and $b = mv + 1$.

2a. Two-way Elimination of Heterogeneity Settings with Row-Column Incidence Structure as $J - I$.

2a.1 Preliminaries

Let us consider an arrangement of v treatments in a square array of size $b \times b$ where the cells along the principal diagonal are supposed to be infeasible, all other cells being feasible. The usual fixed effects model is :

$$y_{jj'}(h) = \mu + \alpha_j + \beta_{j'} + \tau_h + e_{jj'} \qquad 1 \leq h \leq v, \quad 1 \leq j \neq j' \leq b$$

where $y_{jj'}(h)$ is the observation in $(j,j')^{th}$ cell receiving h^{th} treatment and $\mu, \alpha_j, \beta_{j'}, \tau_h$ stand respectively for general effect, j^{th} row effect, j'^{th} column effect, h^{th} treatment effect; $e_{jj'}$'s are i.i.d. $N(0,\sigma^2)$.

As usual, we are interested in linear inferential problem involving a full set of orthonormal varietal contrasts. The familiar concept of connectedness also applies here and one can construct a large class of connected designs for the cases of $b = mv$ or $mv + 1$.

For a specified design d, let $L_d = ((\lambda_{dhj}))$, $M_d = ((m_{dhj}))$, stand respectively for treatment - row, treatment - column incidence matrices. In the present non-orthogonal set-up, the row - column incidence matrix, denoted by $N = ((n_{jj'}))$, assumes the form $J - I$. Let $\underline{r}_d = (r_{d1}, r_{d2}, \ldots, r_{dv})'$ be the vector of treatment replications for the design d and $D_{r_d} = \text{Diag}(r_{d1}, \ldots, r_{dv})$. Let further

$$n_{j.} = \sum_{j'=1}^{b} n_{jj'} \; , \quad n_{.j'} = \sum_{j=1}^{b} n_{jj'} \, .$$

The following relations are fundamental :

$$n_{j.} = \sum_{h=1}^{v} \ell_{dhj}, \quad n_{.j'} = \sum_{h=1}^{v} m_{dhj'},$$

$$r_{dh} = \sum_{j=1}^{b} \ell_{dhj} = \sum_{j'=1}^{b} m_{dhj'}.$$

Following Aggarwal (1966a), the C-matrix of the design is given by

$$C_d = B_{22} - X_{12}' A_{11}^{-} X_{12}$$

where $B_{22} = \mathrm{Diag}(r_{d1}, r_{d2}, \ldots, r_{dv}) - L_d \,\mathrm{Diag}(\frac{1}{n_{1.}}, \frac{1}{n_{2.}}, \ldots, \frac{1}{n_{b.}})L_d'$

$$X_{12} = -M_d' + N' \,\mathrm{Diag}(\frac{1}{n_{1.}}, \frac{1}{n_{2.}}, \ldots, \frac{1}{n_{b.}})L_d'$$

$$A_{11} = \mathrm{Diag}(n_{.1}, n_{.2}, \ldots, n_{.b}) - N' \mathrm{Diag}(\frac{1}{n_{1.}}, \frac{1}{n_{2.}}, \ldots, \frac{1}{n_{b.}})N.$$

We will deduce explicit expression for C_d with $N = J - \varepsilon I$, for $\varepsilon = 0$ or 1. Let k denote the constant row and column sizes i.e. $k = n_{.j'} = n_{j.}$, $1 \leq j,j' \leq b$. Note that for $\varepsilon = 0$, $k = b$, and for $\varepsilon = 1$, $k = b-1$. Then we get

$$C_d = D_{r_d} - \frac{L_d L_d'}{k} - \left\{- M_d + \frac{L_d N}{k}\right\} \left\{kI - \frac{NN'}{k}\right\}^{-} \left\{- M_d' + \frac{N' L_d'}{k}\right\}$$

$$= D_{r_d} - \frac{L_d L_d'}{k} - \left\{- M_d + \frac{L_d J - \varepsilon L_d}{k}\right\} \frac{k}{k^2 - \varepsilon^2} \left[I - \frac{b - 2\varepsilon}{k^2 - \varepsilon^2} J\right]^{-}$$

$$\left\{- M_d' + \frac{J L_d' - \varepsilon L_d'}{k}\right\}.$$

Next we note that

$$\left[I - \frac{b-2\varepsilon}{k^2-\varepsilon^2} J\right]^{-} = \left[I - \frac{J}{b}\right]^{-} = I - \frac{J}{b} \quad \text{for both } \varepsilon = 0 \text{ and } \varepsilon = 1.$$

Hence,

$$C_d = D_{r_d} - \frac{L_d L_d'}{k+\varepsilon} - \frac{M_d M_d'}{k+\varepsilon} - \frac{\varepsilon}{k^2-\varepsilon^2}(L_d + M_d)(L_d + M_d)' + \frac{k+\varepsilon}{(k-\varepsilon)bk} r_d r_d'$$

For $\varepsilon = 0$, we get the usual form of C_d matrix under orthogonal set-up, and for $\varepsilon = 1$, we get

$$C_d = D_{r_d} - \frac{L_d L_d'}{b} - \frac{M_d M_d'}{b} - \frac{(L_d + M_d)(L_d + M_d)'}{b(b-2)} + \frac{r_d r_d'}{(b-1)(b-2)} \qquad \ldots(2a.1.1)$$

Let us define Ω to be the class of all connected designs with row-column incidence structure as $J - I$. In our later derivation we will use the above C_d matrix in (2a.1.1) and find optimal designs in Ω for the cases of $b = mv$ and $b = mv + 1$. For $b = mv + 1$, universally optimal designs are obtained (subsection 2a.2) whereas for $b = mv$, we come up with specific optimal designs (subsection 2a.3). We also construct such optimal designs (subsection 2a.4) and calculate efficiencies of Aggarwal's designs and their generalizations relative to the optimal designs (subsection 2a.5).

2a.2 Universally Optimal Designs for $b = mv + 1$

In this section let Ω stand for the class of connected designs for $b \times b$ arrays with $b = mv + 1$ and row-column incidence structure as $J - I$. We will see below that whenever $b = mv + 1$, universally optimal designs (vide DFN 1b.1.5, Chapter 1) exist. We use Proposition 1 in Kiefer (1975) (vide Chapter 1) for this purpose. In effect, we prove the following.

<u>Theorem 2a.2.1</u> Let $d^* \varepsilon \Omega$ be a design which assigns each treatment m times in each row and in each column. Then d^* is universally optimal.

The proof is straightforward. Moreover, since $b = mv+1$, such a d^* always exists.

First we show that d^* maximizes the trace of C_d over the class Ω. Referring to (2a.1.1)

$$\begin{aligned}
\operatorname{tr} C_d &= \sum_{h=1}^{v} r_{dh} - \sum_{h=1}^{v}\sum_{j=1}^{mv+1} \frac{\ell_{dhj}^2}{mv+1} - \sum_{h=1}^{v}\sum_{j=1}^{mv+1} \frac{m_{dhj}^2}{mv+1} \\
&\quad - \sum_{h=1}^{v}\sum_{j=1}^{mv+1} \frac{(\ell_{dhj} + m_{dhj})^2}{(mv+1)(mv-1)} + \sum_{h=1}^{v} \frac{r_{dh}^2}{mv(mv-1)} \\
&= \sum_{h=1}^{v} r_{dh} - \frac{1}{mv+1} \sum_{h=1}^{v} \left\{ \sum_{j=1}^{mv+1} (\ell_{dhj} - \frac{r_{dh}}{mv+1})^2 \right\} \\
&\quad - \frac{1}{mv+1} \sum_{h=1}^{v} \left\{ \sum_{j=1}^{mv+1} (m_{dhj} - \frac{r_{dh}}{mv+1})^2 \right\} \\
&\quad - \frac{1}{(mv-1)(mv+1)} \sum_{h=1}^{v} \left\{ \sum_{j=1}^{mv+1} (\ell_{dhj} + m_{dhj} - \frac{2r_{dh}}{mv+1})^2 \right\} \\
&\quad - \frac{1}{mv(mv+1)} \sum_{h=1}^{v} r_{dh}^2 \qquad \text{...(2a.2.1)}
\end{aligned}$$

since $\sum_{j=1}^{mv+1} \ell_{dhj} = \sum_{j=1}^{mv+1} m_{dhj} = r_{dh}, \quad \forall \ h = 1,2,\ldots,v.$

Clearly, $\sum_{h=1}^{v} r_{dh}^2$ assumes the least value for d^* as d^* is equireplicate with $r_{d^*1} = r_{d^*2} = \ldots = r_{d^*v} = m(mv+1)$. Moreover, each of the first three sum of squares in (2a.2.1) is " ≥ 0 " for any competing design and " $= 0$ " for d^* since in d^*, λ_{d^*hj}'s and m_{d^*hj}'s are each equal to $m = (\frac{r_{d^*i}}{mv+1})$. This settles the part on trace maximization. Next, it is also evident that C_{d^*} is completely symmetric. Hence an application of Proposition 1 of Kiefer (1975) asserts that d^* is universally optimal.

Such types of designs can be easily constructed. We construct a Latin Square with $b = mv+1$ symbols, say $0,1,2,\ldots,mv$ such that along the diagonal the symbol 0 occurs. (Such Latin Squares exist for all orders (vide Denes and Keedwell (1974))). Then we delete the diagonal and reduce the rest of the symbols mod v.

2a.3 Specific Optimality Results for $b = mv$

Here Ω stands for the class of connected designs for $b \times b$ arrays with $b = mv$ and row-column incidence structure as $J - I$. In case of $b = mv$, a completely symmetric C-matrix of a design does not necessarily produce maximum trace of C_d in Ω. Take, for example, $v = 3$, $m = 2$. Below are displayed two designs

$$d_1 = \begin{bmatrix} - & 1 & 2 & 0 & 2 & 0 \\ 1 & - & 0 & 2 & 2 & 1 \\ 2 & 0 & - & 0 & 1 & 1 \\ 0 & 0 & 1 & - & 1 & 2 \\ 0 & 2 & 2 & 1 & - & 0 \\ 1 & 2 & 1 & 2 & 0 & - \end{bmatrix} \qquad d_2 = \begin{bmatrix} - & 0 & 1 & 0 & 1 & 2 \\ 1 & - & 0 & 1 & 2 & 0 \\ 2 & 1 & - & 1 & 0 & 0 \\ 0 & 2 & 0 & - & 2 & 1 \\ 0 & 1 & 2 & 0 & - & 1 \\ 1 & 0 & 2 & 2 & 0 & - \end{bmatrix}$$

(treatments are represented by symbols 0,1,2) with $\text{tr}\, C_{d_1} = \frac{37}{2}$, $\text{tr}\, C_{d_2} = \frac{281}{15}$.

C_{d_1} is completely symmetric and will also be later demonstrated to be E-optimal. But it has smaller trace than that of C_{d_2} which is neither completely symmetric nor E-optimal. Thus it is evident that Proposition 1 of Kiefer (1975) (which rests exclusively on simultaneous realisation of complete symmetry and trace maximization of C_d-matrix) is not applicable here as regards universal optimality. So we look for specific optimality below : As a matter of fact, we are able to enunciate the following results.

Theorem 2a.3.1 Suppose for given $b = mv$, $m \geq 2$ with the $b \times b$ row-column incidence structure as $J - I$, there exists a design d^* for which

(i) treatment - row incidence pattern is a BBD, ($\ell_{d^*hj} = m$ or $m-1$)

(ii) treatment - column incidence pattern is a BBD, ($m_{d^*hj} = m$ or $m-1$)

(iii) and, moreover, the b pairsof treatments denoted by (h_i, h_i'), $i = 1,2,\ldots,b$ where h_i is the treatment which occurs $(m-1)$ times in the i^{th} row and h_i' is the treatment which occurs $(m-1)$ times in the i^{th} column, are such that, they

(a) satisfy $h_i \neq h_i'$ $\forall$ $i = 1,2,\ldots,b$

and (b) exhaust all possible $\binom{v}{2}$ pairs of treatments equally often.

Then d^* is D - optimal for all $m \geq 2$, $v \geq 5$

A - optimal for all $m \geq 2$, $v \geq 4$

E - optimal for all $m \geq 2$, $v \geq 3$.

Proof : Not to obscure the essential steps of reasoning, we will organize our proof as follows :

First we turn to the inequalities which can be used in the absence of universal optimality to obtain weaker optimality results. We treat only $\mathcal{B}_{v,o}$ context (vide Chapter 1) here.

Suppose φ^* is of the form

$$\varphi^*(\lambda_{d1}, \lambda_{d2}, \ldots, \lambda_{dv}) = \sum_{i=1}^{v-1} f(\lambda_{di}) \qquad \ldots(2a.3.1)$$

where f is convex on $[0, -\infty)$ and λ_{di}'s are the characteristic roots of C_d-matrix.

Following essentially the technique in Kiefer (1975) used to prove *specific optimality of non-regular* GYD's, we then develop arguments in general terms for φ^* optimality (Step I, II, III, IV) and establish D- and A- optimality of d^* assuming specific forms of φ^* (Step V). E-optimality of d^* (in case where D- and/or A-optimality do not hold) is also established in Step V starting from the criterion itself. However, some modifications in the arguments are called for in order to tackle the peculiarities arising in this context of non-orthogonal set-up. This will be clear later as we get to the technical details.

Step I Discussion on φ^*-optimality.

Using steps in Kiefer (1975) ((2.8), (2.9))

$$\varphi^*(\lambda_{d1}, \lambda_{d2}, \ldots, \lambda_{dv}) = \sum_{i=1}^{v-1} f(\lambda_{di}) \geq \frac{v-1}{v} f\left(\frac{v}{v-1} C_{dhh}\right) \qquad \ldots(2a.3.2)$$

with equality if all λ_{di}'s are equal, i.e. if C_d is c.s.

Next, observe, from Kiefer (1975), the following

Proposition 3

If φ^* is given by (2a.3.1), with f convex, and if there exists a design d^* in Ω, such that C_{d^*} is c.s. and d^* minimizes $\sum_{h=1}^{v-1} f(\frac{v}{v-1} C_{dhh})$ then d^* is φ^*- optimal.

We assume f to be non-increasing, (vide Kiefer (1975)) and let the function g be defined as (vide $c(r)$ as defined in Kiefer (1975))

$$g(r) = \max_{\{d \,:\, r_{dh}=r\}} C_{dhh} \qquad \ldots(2a.3.3)$$

where r is assumed to take non-negative integer values only.

Thus we get, from (2a.3.2) and (2a.3.3),

$$\varphi^*(\lambda_{d1},\lambda_{d2},\ldots,\lambda_{dv}) \geq \frac{v-1}{v} \sum_{h=1}^{v} f(\frac{v}{v-1} C_{dhh})$$

$$\geq \frac{v-1}{v} \sum_{h=1}^{v} f(\frac{v}{v-1} g(r_h)) \qquad \ldots(2a.3.4)$$

writing r_h for r_{dh}.

Now, to establish φ^*-optimality of the design d^* as described in the statement of the present theorem, we assume the following two properties of d^* for the moment. (respective proofs will be given in Step IV)

$$\left.\begin{array}{l} \text{Property 1 : } C_{d^*} \text{ is c.s.} \\ \text{Property 2 : } C_{d^*hh} = g(\bar{r}) \end{array}\right\} \qquad \ldots(2a.3.5)$$

where $\bar{r} = m(mv-1)$, the equireplication size.

Now making use of these two properties and (2a.3.4) the problem then reduces to checking that

$$\min_{H} \sum_{h=1}^{v} f\left(\frac{v}{v-1}\, g(r_h)\right) = vf\left(\frac{v}{v-1}\, g(\bar{r})\right) \qquad \ldots(2a.3.6)$$

where $H = \left\{ (r_1, r_2, \ldots, r_v) : r_j \text{ non-negative integer for } j = 1,2,\ldots,v,\ \Sigma r_j = v\bar{r} \right\}$.

In effect, we have to establish the validity of

$$q(\bar{r}) = q\left(\frac{\sum_{h=1}^{v} r_h}{v}\right) \geq \frac{1}{v} \sum_{h=1}^{v} q(r_h) \qquad \ldots(2a.3.7)$$

where $q(r)$ is defined as

$$q(r) = -f\left(\frac{v}{v-1}\, g(r)\right). \qquad \ldots(2a.3.8)$$

Then, (2a.3.7) would atonce follow if $q(r)$ were concave, i.e. if $g(r)$ were concave, but this is not the case always. This is what motivates further development of tools (vide Kiefer (1975)).

Suppose there exists a concave function $\bar{q}(r)$ i.e. $\bar{q}(r+1) - 2\bar{q}(r) + \bar{q}(r-1) \leq 0$ for all r, such that

$$\left.\begin{array}{ll} & \bar{q}(r) \geq q(r) \quad \text{for all } r \\ \text{and} & \bar{q}(\bar{r}) = q(\bar{r}). \end{array}\right\} \qquad \ldots(2a.3.9)$$

Then (2a.3.7) will still hold.

Kiefer took $\bar{q}$ as the concave envelope of q i.e. a concave function $\bar{q}$ which happens to be the minimum function $\geq q$ with $\bar{q}(\bar{r}) = q(\bar{r})$. But in fact, it can be easily seen that this proof goes through if there exists <u>any</u> concave function $\tilde{q}$ with

$$\tilde{q} \geq q \quad \text{and} \quad \tilde{q}(\bar{r}) = q(\bar{r}).$$

Now we study the behaviour of the q-function relevant to this particular non-orthogonal set-up (i.e. row-column incidence structure being of the form $J - I$) and find a set of sufficient conditions for existence of $\tilde{q}$ satisfying (2a.3.9). For this, we first investigate the properties of the function g relevant to this set-up.

<u>Step II</u> : <u>Peculiar properties of the function g under the present non-orthogonal set-up</u>.

$$\text{Let} \quad G = \{0,1,2,\ldots,mv(mv-1)\} \qquad \ldots(2a.3.10)$$

with $G_1 = \{n : n \in G, \text{ and } n \text{ is a multiple of } mv\}$.

Then for any design $d \in \Omega$, $r_{dh} \in G$, $\forall$ $h = 1,2,\ldots,v$, and the function g is defined over the range of values in G. In what follows, we write $[C,D]$ for an interval of (successive) integers with $C < D$, and C,D being two consecutive integers in G_1. Thus whenever we write $r \in [C,D]$, r is restricted to integer values only. We call $[C,D]$ <u>an elementary interval</u> (vide Kiefer (1975)). For $\bar{r} = m(mv-1) = (m-1)mv + mv - m$, we write $\bar{r} \in [C_o,D_o]$, (say) with $C_o = (m-1)mv$ and $D_o = m.mv$. The interval $[C_o,D_o]$ is termed as <u>basic interval</u> by Kiefer (1975).

Now referring to (2a.1.1) we get,

$$C_{dhh} = r_{dh} - \frac{1}{mv}\sum_{j=1}^{mv} \ell_{dhj}^2 - \frac{1}{mv}\sum_{j=1}^{mv} m_{dhj}^2$$

$$- \frac{1}{mv(mv-2)}\sum_{j=1}^{mv} (\ell_{dhj} + m_{dhj})^2 + \frac{r_{dh}^2}{(mv-1)(mv-2)} .$$

We now derive the expression for $g(r)$ i.e. maximum of C_{dhh} subject to the condition $\sum_{j=1}^{mv} \lambda_{dhj} = r$, and $\sum_{j=1}^{mv} m_{dhj} = r$.

Set $r = mv\left[\frac{r}{mv}\right] + t$

$= mv.u + t$ (say).

This means that we assume r to belong to the u^{th} elementary interval $[umv, (u+1)mv]$. It is clear that the minimum of $\sum_{j=1}^{mv} \lambda_{dhj}^2$ subject to $\sum_{j=1}^{mv} \lambda_{dhj} = r$ is attained when t of λ_{dhj}'s are each equal to $u+1$ and the rest are each equal to u. A similar result holds for the m_{dhj}'s. Now in order to attain minimum of $\sum_{j=1}^{mv} (\lambda_{dhj} + m_{dhj})^2$ such that $\sum_{j=1}^{mv} (\lambda_{dhj} + m_{dhj}) = 2r$, we argue as follows :

Case (i) $t < \frac{mv}{2}$

Clearly, $\left[\frac{2r}{mv}\right] = 2\left[\frac{r}{mv}\right] = 2u$

and hence, the minimum is attained when $2t$ of $(\lambda_{dhj} + m_{dhj})$'s are each equal to $2u+1$ and the rest are each equal to $2u$ i.e. whenever $\lambda_{dhj}(m_{dhj}) = u+1$ for some j, the corresponding $m_{dhj}(\lambda_{dhj}) = u$.

Case (ii) $t \geq \frac{mv}{2}$.

Here $\left[\frac{2r}{mv}\right] = 2\left[\frac{r}{mv}\right] + 1 = 2u + 1$.

Hence the minimum is attained when $2t - mv$ of $(\lambda_{dhj} + m_{dhj})$'s are each equal to $2u+2$ and the rest are each equal to $2u+1$ i.e. whenever $\lambda_{dhj}(m_{dhj}) = u$ for some j, the corresponding $m_{dhj}(\lambda_{dhj}) = u+1$.

It can be checked that

under Case (i) : $t < \frac{mv}{2}$

$$g(r) = r + \frac{2}{mv}\left\{mvu^2 - (2r - mv)u - r\right\} + \frac{r^2}{(mv-1)(mv-2)}$$
$$+ \frac{1}{mv(mv-2)}\left\{(4mvu^2 - 8ur) - (2r - 2mvu)\right\}$$

and

under Case (ii) : $t \geq \frac{mv}{2}$,

$$g(r) = r + \frac{2}{mv}\left\{mvu^2 - (2r - mv)u - r\right\} + \frac{r^2}{(mv-1)(mv-2)}$$
$$+ \frac{1}{mv(mv-2)}\left\{(4mvu^2 - 8ur) - (6r - 6mvu - 2mv)\right\}$$

so that

$$g(r) = \begin{cases} g_1(r) = A(r) + B_1(r) & \text{if } t < \frac{mv}{2} \text{ i.e. if } \frac{r}{mv} - \left[\frac{r}{mv}\right] < \frac{1}{2} \\ g_2(r) = A(r) + B_2(r) & \text{if } t \geq \frac{mv}{2} \text{ i.e. if } \frac{r}{mv} - \left[\frac{r}{mv}\right] \geq \frac{1}{2} \end{cases} \qquad \text{...(2a.3.11)}$$

where $A(r) = r + \frac{2}{mv}\left\{mvu^2 - (2r - mv)u - r\right\} + \frac{r^2}{(mv-1)(mv-2)}$

$$+ \frac{1}{mv(mv-2)}\left\{4mvu^2 - 8ur\right\}$$

$$= \frac{r^2}{(mv-1)(mv-2)} + r\left\{\frac{m^2v^2 - 4mv - 4umv + 4}{mv(mv-2)}\right\} + \frac{2mvu^2 + 2u(mv-2)}{mv-2}$$

$$B_1(r) = -\frac{2r - 2mvu}{mv(mv-2)}$$

$$B_2(r) = -\frac{6r - 6mvu - 2mv}{mv(mv-2)},$$

In other words,

$$g(r) = \begin{cases} g_1(r) & \text{if} \quad umv \leq r \leq umv + \left[\frac{mv}{2}\right] \\ g_2(r) & \text{if} \quad umv + \left[\frac{mv}{2}\right] + 1 \leq r < (u+1)mv \end{cases} \qquad \ldots(2a.3.12)$$

$$u = 0,1,2,\ldots,mv-2$$

Note at this stage, that, the behaviour of the function $g(r)$ is quite different from the one discussed by Kiefer under non-regular GYD set-up in the sense that within each elementary interval, this time $g(r)$ assumes two different functional forms unlike one single functional form assumed earlier. To get rid of the difficulty arising out of this, we may argue as follows. Since our problem is to find a concave function $\bar{q} \geq q$ satisfying (2a.3.9), it suffices to find a concave function $\bar{q} \geq q_1$ where q_1 is a function which is still larger than q. However, for (2a.3.9) to hold, it is necessary that the value of q_1 must remain unchanged at $\bar{r}$ i.e. q_1 should coincide with q at $\bar{r}$. Since $-f$ is non-decreasing, the above argument justifies that instead of working with two different expressions for $g(r)$ in the two halves of each elementary interval, we may work with the larger of the two at some or all points of the intervals except for the point $\bar{r}$ in the interval $[C_o, D_o]$. Projecting $g_1(r)$ beyond the first half of any such interval, using (2a.3.11), we observe that

$$g_1(r) - g_2(r) = \frac{4r - 4mvu - 2mv}{mv(mv-2)}$$

$$= \frac{4(r - mvu) - 2mv}{mv(mv-2)}$$

$$= \frac{4t - 2mv}{mv(mv-2)} \begin{cases} > 0 & \text{for} \quad t > \frac{mv}{2} \\ = 0 & \text{for} \quad t = \frac{mv}{2} \\ < 0 & \text{for} \quad t < \frac{mv}{2} \end{cases} \qquad \ldots(2a.3.13)$$

This suggests us to work with <u>modified $g(r)$</u> defined as follows :

$$g(r) = \begin{cases} g_1(r) & r \leq C_o + \frac{mv}{2}, \ r \geq D_o \\ g_2(r) & C_o + \frac{mv}{2} \leq r \leq D_o - 1 \end{cases} \qquad \ldots(2a.3.14)$$

as $\bar{r} = (m-1)mv + m(v-1) > C_o + \frac{mv}{2}$.

Accordingly, modified q would be defined as $q(r) = -f(\frac{v}{v-1}\, g(r))$ with g as in (2a.3.14).

(In other words, we decide to use the function $g_1(r)$ also in the second half of all the intervals except $[C_o, D_o]$ for which we retain the exact functional form $g_2(r)$ itself. One could as well use only one functional form $g_2(r)$ for the entire range of values of r. That way, one might achieve simplicity in the analysis but possibly with some weaker results, because of distortion in the exact functional form $g_1(r)$ in first half of the basic interval $[C_o, D_o]$ containing $\bar{r}$. On the other hand, same complexity in the analysis would prevail if in the interval $[C_o, D_o]$ exact expression for $g(r)$ was undertaken along with $g_2(r)$ elsewhere.)

Now we study some properties of the function $g(r)$ through those of $g_1(r)$ and $g_2(r)$ over the entire range of values of r.

We have for an elementary interval $\left[umv, (u+1)mv\right]$ and an intermediate point r, satisfying $\left[\frac{r}{mv}\right] = \left[\frac{r+1}{mv}\right] = u$,

$$\Delta_1(r) = g_1(r+1) - g_1(r)$$

$$= \frac{2mvr + (m^3v^3 - 5m^2v^2 + 7mv - 2) - 4umv(mv-1)}{mv(mv-1)(mv-2)} \qquad \ldots(2a.3.15)$$

For fixed u, as $r \uparrow$, this difference $\uparrow$. Further, at the least value of r viz. mvu, one has

$$\Delta_1(mvu) = \frac{(m^3v^3 - 5m^2v^2 + 7mv - 2) - 2umv(mv-2)}{mv(mv-1)(mv-2)} \qquad \ldots(2a.3.16)$$

and this is positive

$$\text{iff} \quad \frac{m^3v^3 - 5m^2v^2 + 7mv - 2}{2mv(mv-2)} > u > 0 \quad (\because mv > 2)$$

$$\text{i.e. iff} \quad u \leq \left[\frac{m^2v^2 - 3mv + 1}{2mv}\right] = u_0 \quad \text{(say)} \qquad \ldots(2a.3.17)$$

($[x]$ means the greatest integer $\leq x$, (vide Chapter 1))

Again at the penultimate value of r in this u^{th} elementary interval i.e. at $r = (u+1)mv - 1$, using $\left[\frac{r}{mv}\right] = u$, $\left[\frac{r+1}{mv}\right] = u+1$,

$$\Delta_1((u+1)mv-1) = \frac{m^3v^3 - 5m^2v^2 + 7mv - 2 - 2umv(mv-2)}{mv(mv-1)(mv-2)} \qquad \ldots(2a.3.18)$$

which is the same as (2a.3.16) and so this is positive iff $u \leq u_0$. This is precisely the condition (2a.3.17) stated above. Thus the above facts imply that $g_1(r) \uparrow$ right from the start (i.e. $r = 0$) to $r = (u_0 + 1)mv$, covering thereby <u>all</u> the intermediate points in the intervals corresponding to $u = 0$ to $u = u_0$.

A similar behaviour of the function $g_2(r)$ can also be observed.

Moreover, it is interesting to note that the interval $[C_0, D_0]$ is to the left of $[u_0 mv, (u_0+1)mv]$ for all combinations of (m,v) except for certain cases like $v = 3$, $m = 1$ and $v = 3$, $m = 2$, where the two intervals coincide. A study of the functions g_1 and g_2 inside the second half of the interval $[C_0, D_0]$ reveals the following : (The verification immediately follows using (2a.3.13) - (2a.3.17) and the above facts)

$$\text{(i)} \quad \frac{mv}{2} = \text{integer} : \quad g_1(C_0 + \tfrac{mv}{2}) = g_2(C_0 + \tfrac{mv}{2}) \leq g_2(C_0 + \tfrac{mv}{2} + 1) \leq \cdots \leq g_2(D_0 - 1) \leq g_1(D_0 - 1) \leq g_1(D_0)$$

$$\text{(ii)} \quad \frac{mv}{2} \neq \text{integer} : \quad g_1(C_0 + \tfrac{mv-1}{2}) \leq g_2(C_0 + \tfrac{mv-1}{2}) \leq g_2(C_0 + \tfrac{mv+1}{2}) \leq g_2(C_0 + \tfrac{mv+3}{2}) \leq \cdots \leq g_2(D_0 - 1) \leq g_1(D_0 - 1) \leq g_1(D_0)$$

From the assumed funtional form of $g(r)$, as given in (2a.3.14), it is now evident from the above study that $g(r) \uparrow$ right from the start i.e. $r = 0$ to $r = (u_0+1)mv$. It remains to study the pattern of $g(r)$ beyond $(u_0+1)mv$ in which case we actually work with $g_1(r)$.

Taking $r = (u+1)mv - 2$, we get using (2a.3.15)

$$\Delta_1((u+1)mv - 2) = \frac{(m^3v^3 - 3m^2v^2 + 3mv - 2) - 2umv(mv-2)}{mv(mv-1)(mv-2)}$$

and this is '< 0'

$$\text{iff} \quad u > \frac{m^2v^2 - 3mv + 1}{2mv} + 1 \qquad \ldots(2a.3.19)$$

Moreover, our earlier analysis with reference to the penultimate point $r = (u+1)mv - 1$ (vide (2a.3.18)) shows that whenever (2a.3.19) holds, $\Delta_1((u+1)mv - 1) < 0$. Thus it finally follows that $g(r) \downarrow$ in r for $u \geq (u_0+2)$ i.e. for $r \geq (u_0+2)mv$. In the interval $[(u_0+1)mv, (u_0+2)mv]$ we cannot infer about the specific behaviour of $g(r)$ as such. (This, however, does not pose any difficulty for our later analysis since we are not really concerned as such with signs of these differences in this interval, only the expressions of $g(r)$ will suffice). To finalise, we see that the function g exhibits the following pattern.

Graph of the function $g(r)$ as defined by (2a.3.14)

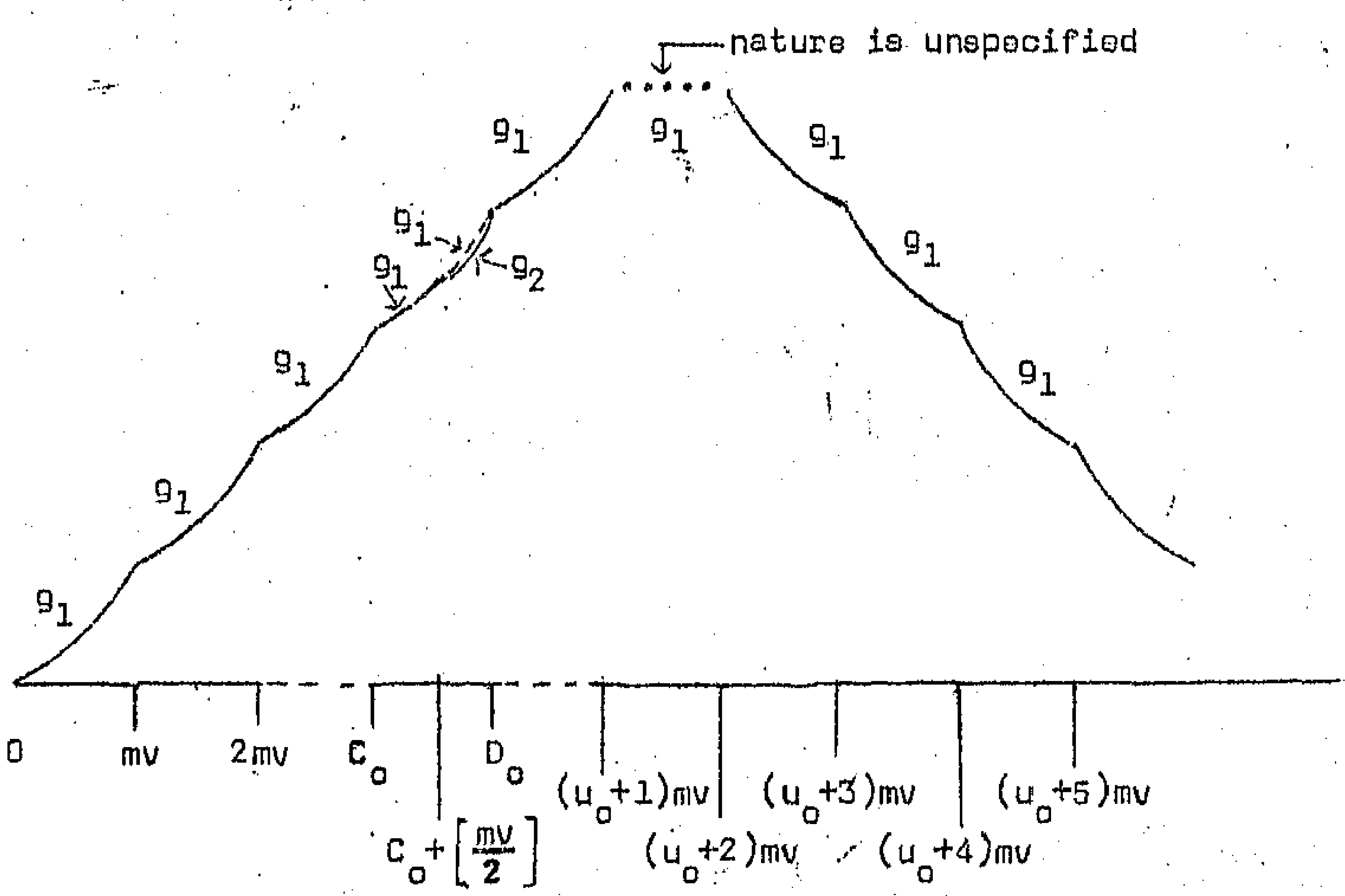

Figure 1

Now following Kiefer (1975), we derive a set of sufficient conditions for (2a.3.9) to hold using properties of g.

Step III : Derivation of a set of sufficient conditions for (2a.3.9) to hold.

Let $\tilde{D} \in G$ (vide(2a.3.10)) be the first integer where g attains its absolute maximum on G. By the above derivation, and from Figure 1, $D_0 \leq (u_0+1)mv \leq \tilde{D} \leq (u_0+2)mv$. Since $-f$ is monotone (vide page (45)) it is enough to prove existence of concave $\bar{q}$, $\bar{q} \geq q$ (q as defined by $-f(\frac{v}{v-1}\,g(r))$) satisfying (2a.3.9) with the domain G replaced by $G' = \left\{ r : r \in G,\ r \leq \tilde{D} \right\}$ (vide Kiefer (1975)). Again possibly in the elementary interval $[(u_0+1)mv, (u_0+2)mv]$ there may exist points r', r'', $r' < r''$ such that $g(r'') \leq g(r')$. Then because of the non-decreasing nature of $-f$, we have $q(r'') < q(r')$. Since we are concerned with a concave function $\bar{q} \geq q$, G' can be replaced by its subset G'' obtained by excluding such points as domain of q in defining $\bar{q}$, so that for any two points r_1 and r_2 in G'', we have $q(r_1) < q(r_2)$ whenever $r_1 < r_2$. Now a sufficient condition for existence of concave function $\bar{q} \geq q$, with $\bar{q}(\bar{r}) = q(\bar{r})$ is local concavity of q at $\bar{r}$ (vide Kiefer (1975)). That is

$$q(r_1+1) - q(r_1) \geq q(r_2) - q(r_2-1) \qquad \ldots(2a.3.20)$$

whatever $r_1+1 \leq \bar{r} \leq r_2-1, \qquad r_1, r_2 \in G''$.

We may note that r_2-1 need not belong to G''. As in Kiefer (1975), one can establish (2a.3.20) by proving (i) to (iii) stated below :

$$\left.\begin{aligned}
&\text{(i)} && \min_{0<r_1<\tilde{r}} \left\{q(r_1+1) - q(r_1)\right\} = q(\tilde{r}) - q(\tilde{r}-1) \\
&\text{(ii)} && \max_{\tilde{r}<r_2 \in G''} \left\{q(r_2) - q(r_2-1)\right\} = q(\tilde{r}+1) - q(\tilde{r}) \\
&\text{(iii)} && q(\tilde{r}) - q(\tilde{r}-1) \geq q(\tilde{r}+1) - q(\tilde{r})
\end{aligned}\right\} \quad \ldots(2a.3.21)$$

Now making use of the properties of g we will establish that in the present non-orthogonal set-up, a set of sufficient conditions for (2a.3.21) to hold is

$$\left.\begin{aligned}
&q(r+1) - q(r) \uparrow \text{ in } r \text{ for } C_0 \leq r \leq D_0 - 1 \\
&\text{i.e. } q \text{ is concave in } C_0 \leq r \leq D_0 - 1 \\
&\text{and moreover,} \\
&q(\tilde{r}+1) - q(\tilde{r}) \geq q(D_0) - q(D_0 - 1)
\end{aligned}\right\} \quad \ldots(2a.3.22)$$

Remark 1 : At this stage it may be noted that the above condition (2a.3.22) is slightly different from those in Kiefer (1975). While in Kiefer's formulation under non-regular GYD set-up, the function q may admit of concavity for $C_0 \leq r \leq D_0$, in the present framework, it turns out that q fails to be so at the end point D_0. In view of this, we need an extra condition to be satisfied. This is precisely the second condition in (2a.3.22) stated above.

Remark 2 : It has been pointed out later (in the pages to follow) that (2a.3.21) (iii) is not satisfied when $m = 1$.

We take up below a proof of sufficiency of (2a.3.22) assuming then $m \geq 2$.

Clearly the first condition in (2a.3.22) directly implies (2a.3.21) (iii) for $m \geq 2$ as in such cases $\bar{r}+1 \leq D_o - 1$.

To prove (2a.3.21) (i) take r_1 in $[C, D]$ where $[C, D] = [umv, (u+1)mv]$ is an elementary interval to the left of $[C_o, D_o]$ so that $D \leq C_o$.

By (2a.3.15)

$$\Delta_1(r_1) > \Delta_1(C) \quad \forall \quad C \leq r_1 \leq C + mv - 2.$$

Again for $r_1 = C + mv - 1$, by earlier study of penultimate point (vide (2a.3.18)),

$$\Delta_1(C + mv - 1) = \frac{m^3v^3 - 5m^2v^2 + 7mv - 2 - 2umv(mv-2)}{mv(mv-1)(mv-2)}$$

Also by (2a.3.16),

$$\Delta_1(C) = \frac{m^3v^3 - 5m^2v^2 + 7mv - 2 - 2umv(mv-2)}{mv(mv-1)(mv-2)}$$

so that $\Delta_1(C + mv - 1) = \Delta_1(C)$.

Hence, $\Delta_1(r_1) \geq \Delta_1(C) \quad \forall \quad r_1 \in [C, D], \; r_1 \leq D - 1.$

Let $D_1 = u_1 mv$ and $D_2 = u_2 mv$, $u_2 > u_1$

By (2a.3.16),

$$\Delta_1(D_1) - \Delta_1(D_2) = \frac{2(u_2 - u_1)}{(mv-1)} > 0 \qquad (\because \; mv > 2, \; u_2 > u_1)$$

This implies $\Delta_1(D_1) > \Delta_1(D_2)$.

Thus finally,

$$\Delta_1(r_1) \geq \Delta_1(C) > \Delta_1(C_o) > 0 \quad \text{for all} \quad r_1 \leq C_o - 1 \qquad \ldots(2a.3.23)$$

The last inequality follows from (2a.3.17) as $m-1 \leq \left[\frac{m^2v^2-3mv+1}{2mv}\right]$ for $v \geq 3$, $m \geq 2$.

Set now $g(r_1+1) = y_1$, $g(r_1) = y_2$, $g(C_o+1) = y_3$, $g(C_o) = y_4$, so that $y_2 < y_1 \leq y_4 < y_3$. Then $\Delta_1(r_1) > \Delta_1(C_o) > 0$ means

$$y_1 - y_2 > y_3 - y_4 > 0.$$

Since $-f$ is non-decreasing and concave, by Mean-Value Theorem,

$$\frac{-f(y_1)+f(y_2)}{y_1-y_2} \geq \frac{-f(y_3)+f(y_4)}{y_3-y_4}$$

i.e. $$-f(y_1)+f(y_2) \geq \frac{y_1-y_2}{y_3-y_4}\left(-f(y_3)+f(y_4)\right)$$

$$\geq \left(-f(y_3)+f(y_4)\right) \qquad \left[\because\ y_1-y_2 \geq y_3-y_4\right]$$

i.e. $$q(r_1+1)-q(r_1) \geq q(C_o+1)-q(C_o)$$

So $$\min_{r_1 \leq C_o} \left\{q(r_1+1)-q(r_1)\right\} = q(C_o+1)-q(C_o)$$

This, together with the first condition of (2a.3.22), i.e., concavity of q within $C_o \leq r \leq D_o - 1$ implies

$$\min_{0 \leq r_1 < \bar{r}} \left\{q(r_1+1)-q(r_1)\right\} = q(\bar{r})-q(\bar{r}-1)$$

so that (2a.3.21) (i) is settled.

We next prove that (2a.3.22) implies (2a.3.21) (ii). For this, take any $r_2 \in [C,D] = [\,umv,\ (u+1)mv\,]$ to the right of $[C_o, D_o]$ so that $C \geq D_o$, $D \leq (u_o+2)mv$.

By (2a.3.15), for any r_2 such that $\left[\frac{r_2}{mv}\right] = \left[\frac{r_2+1}{mv}\right] = \left[\frac{r_2+2}{mv}\right] = u$,

$$\Delta_1(r_2+1) - \Delta_1(r_2) = \frac{2}{(mv-1)(mv-2)} > 0 .$$

Again, when $r_2 = (u+1)mv - 2$, by (2a.3.15) and (2a.3.18)

$$\Delta_1(r_2+1) - \Delta_1(r_2) = \Delta_1((u+1)mv-1) - \Delta_1((u+1)mv-2)$$

$$= \frac{-2}{(mv-1)} < 0 .$$

Hence, $\Delta_1(r_2) \leq \Delta_1(D-2)$ for $C \leq r_2 \leq D-1$.

Also when $D_1 = u_1 mv$, $D_2 = u_2 mv$, $u_2 > u_1$, by (2a.3.15)

$$\Delta_1(D_2-2) - \Delta_1(D_1-2) = \frac{-2(u_2-u_1)}{(mv-1)} < 0$$

so that $\Delta_1(D_2-2) < \Delta_1(D_1-2)$.

Thus $0 < \Delta_1(r_2) \leq \Delta_1(D-2) < \Delta_1(D_0 + mv - 2)$

by repeated application of the preceding inequality for $D_0 \leq C \leq r_2 \leq D-1$.

Now, as before, we want to relate the point $D_0 + mv - 2$ to a point within the second half of $[C_0, D_0]$. We define $\Delta(r) = g(r+1) - g(r)$ and observe, using (2a.3.13) and (2a.3.14), that

$$\Delta(D_0-2) = g_2(D_0-1) - g_2(D_0-2) = \Delta_1(D_0-2) - \frac{4}{mv(mv-2)}$$

and $\Delta(D_0-1) = g_1(D_0) - g_2(D_0-1) = \Delta_1(D_0-1) + \frac{2}{mv}$.

Further, we may verify directly (using expressions for $\Delta_1((u+1)mv-2)$ $\Delta_1((u+1)mv-1)$ derived earlier) that

$$\Delta_1(D_0 + mv - 2) < \Delta(D_0 - 1).$$

so that we come up with

$$0 < \Delta_1(r_2) < \Delta(D_o-1) \quad \text{for all} \quad D_o \leq r_2 \leq (u_o+2)mv - 1.$$

(It may be noted that the subset G'' may not include the whole of $[(u_o+1)mv, (u_o+2)mv]$. However, the above arguments carry through to validate the specific inequality :

$$0 < \Delta_1(r_2) < \Delta(D_o-1) \quad \text{for all} \quad D_o \leq r_2 \leq \bar{D}-1$$

which is needed with reference to G''.)

Now as before, an application of Mean-Value Theorem yields

$$q(D_o) - q(D_o-1) \geq q(r_2+1) - q(r_2) \quad \forall \quad r_2 \geq D_o, \; r_2 \in G''.$$

Also, concavity of q for $C_o \leq r \leq D_o - 1$ ensures

$$q(\bar{r}+1) - q(\bar{r}) \geq q(\bar{r}+2) - q(\bar{r}+1) \geq \cdots\cdots \geq q(D_o-1) - q(D_o-2).$$

However, the above two inequalities cannot be linked together as it now turns out

$$\Delta(D_o-2) - \Delta(D_o-1) = \left\{\Delta_1(D_o-2) - \Delta_1(D_o-1)\right\} - \frac{4}{mv(mv-2)} - \frac{2}{mv}$$
$$= \frac{-2}{(mv-1)(mv-2)} < 0$$

which implies $q(D_o-1) - q(D_o-2) \leq q(D_o) - q(D_o-1)$.

This straightway explains Remark 1 made earlier. Remark 2 is also clear noting that $\bar{r}+1 = D_o$ for $m=1$. As such, we require an additional condition to be verified. This is that

$$q(\bar{r}+1) - q(\bar{r}) \geq q(D_o) - q(D_o-1).$$

which is precisely the second requirement under (2a.3.22). Thus, whenever condition (2a.3.22) obtains, we get

$$\max_{\tilde{r} < r_2 \varepsilon G''} \left\{ q(r_2) - q(r_2-1) \right\} = q(\tilde{r}+1) - q(\tilde{r})$$

which is (2a.3.21) (ii).

To summarise, in the present non-orthogonal set-up, φ^*-optimality of d^* (described in the statement of the Theorem) follows from a verification of (2a.3.22) along with its two properties displayed in (2a.3.5). The subsequent discussion does these verifications.

Step IV : Verification of Property 1 and Property 2 (vide (2a.3.5)) of d^*.

For the design d^*, for every $h = 1,2,\ldots,v$ let $\underline{\lambda_{d^*h}}$ and $\underline{m_{d^*h}}$ denote the two row-vectors with elements λ_{d^*hj} and m_{d^*hj} (vide subsection 2a.1) written in the order $j = 1,2,\ldots,mv$ respectively. Then, $\underline{\lambda_{d^*h}}$ and $\underline{m_{d^*h}}$ are represented as follows :

$$\underline{\lambda_{d^*h}} = (\underbrace{m,m,\ldots,m, \quad m,\ldots,m}_{m(v-1) \text{ times}} \quad \underbrace{m-1,m-1,\ldots,m-1}_{m \text{ times}})$$

$$\underline{m_{d^*h}} = (\underbrace{m-1,m-1,\ldots,m-1}_{m \text{ times}},m,\ldots,m \quad m,m,\ldots,m)$$

Clearly, this configuration follows from the fact that for d^*, with $mv-1$ feasible cells in each row (column), treatment-row incidence pattern is a BBD, and treatment-column incidence pattern is a BBD. Moreover, from condition (iii) (a) as stated in the theorem if the treatment h occurs $m-1$ times in j^{th} row (column) it must occur m times in j^{th} column (row). Now, noting that for $\bar{r} = (m-1)mv + m(v-1)$, $u = m-1$, $t = m(v-1) > \frac{mv}{2}$ for $m \geq 2$, $v \geq 3$, our arguments in subsection 2a.3, Step II, regarding minimization of $\sum_j \ell^2_{dhj}$, $\sum_j m^2_{dhj}$ and $\sum_j (\ell_{dhj} + m_{dhj})^2$ in the process of derivation of expression for $g(r)$ clearly reveals that

$$c_{d^*hh} = g(\bar{r}) = \max_{\{d \,:\, r_{dh} = \bar{r}\}} c_{dhh} .$$

To achieve complete symmetry of C_{d^*}, we first note that condition (i) and (ii) stated in Theorem 2a.3.1 namely BBD structure in treatment-row incidence as well as treatment-column incidence readily lead to complete symmetry of $L_{d^*}L'_{d^*}$ and $M_{d^*}M'_{d^*}$ (vide(2a.1.1)). Thus it remains to verify (vide(2a.1.1)) complete symmetry of $(L_{d^*} + M_{d^*})(L_{d^*} + M_{d^*})'$ as d^* is equireplicate. Consider any pair of treatments say (h,h'), $h \neq h'$ and look to the configuration of $(\underline{\ell_{d^*h}} + \underline{m_{d^*h}})$ and $(\underline{\ell_{d^*h'}} + \underline{m_{d^*h'}})$ which follows from the above discussion regarding separate configurations of $\underline{\ell_{d^*h}}$ and $\underline{m_{d^*h}}$ for any h.

$$\underline{\ell_{d^*h}} + \underline{m_{d^*h}} \;:\; (\underbrace{2m-1, 2m-1, \ldots, 2m-1, 2m-1, \ldots, 2m-1}_{2m \text{ times}},\ \underbrace{2m, \ldots, 2m, 2m, \ldots, 2m}_{m(v-2) \text{ times}})$$

$$\underline{\ell_{d^*h'}} + \underline{m_{d^*h'}} \;:\; (\underbrace{2m-1, 2m-1, \ldots, 2m-1}_{x \text{ times}},\ \underbrace{2m, \ldots, 2m}_{2m-x \text{ times}},\ \underbrace{2m-1, \ldots, 2m-1}_{2m-x \text{ times}},\ 2m, \ldots, 2m)$$

For any treatment h, occurrence of $2m-1$ in any position say j of $(\underline{\ell_{d^*h}} + \underline{m_{d^*h}})$ is due to the fact that h^{th} treatment occurs $m-1$ times in either j^{th} row or j^{th} column. Certainly the above two configurations represent (upto permutation) h^{th} and h'^{th} rows of the matrix $(L_{d*} + M_{d*})$. We note that the $(h,h')^{th}$ entry in $(L_{d*} + M_{d*})(L_{d*} + M_{d*})'$ will take the form $x(2m-1)^2 + 2(2m-x) \cdot 2m(2m-1) + (mv-4m+x)(2m)^2$ and so complete symmetry of $(L_{d*} + M_{d*})(L_{d*} + M_{d*})'$ requires constancy of x which is the frequency of $(2m-1, 2m-1)$ for every pair of rows in $(L_{d*} + M_{d*})$. Consider then a third row of $(L_{d*} + M_{d*})$ say h''^{th} row, $h'' \neq h \neq h'$, which represents nothing but the distribution of h''^{th} treatment over the rows and columns of d^* jointly. Since in any row or column of the design, only one treatment can occur $(m-1)$ times, while each of the rest occurs m times, it is evident that in the representation of $(\underline{\ell_{d*h''}} + \underline{m_{d*h''}})$ the first x positions cannot be $2m-1$. This clearly explains that in order to achieve c.s. property of C_{d*} the total frequency $2m$ of the element $(2m-1)$ in the representation $(\underline{\ell_{d*h}} + \underline{m_{d*h}})$ should get equally divided among the rest $v-1$ rows of $(L_{d*} + M_{d*})$. In other words, denoting by (j,j) the j^{th} diagonal cell in the design and by h_j (h_j') the treatment having replication $m-1$ in j^{th} row (j^{th} column respectively) we demand for c.s. property of C_{d*}, that in d^*, for the $b = mv$ cells of the type (j,j), the corresponding mv treatment pairs of the type (h_j, h_j') should get equally divided among all $\binom{v}{2}$ pairs of treatments. This is precisely the condition (iii) (b) of d^*, as mentioned in the statement of the theorem.

Now it remains to verify condition (2a.3.22) assuming specific functional forms of f in φ^*.

Step V : Specific optimality results for d^*

We now adopt specific optimality criteria assuming specific functional forms of f in φ^* (vide(2a.3.1)). For D-optimality (vide Chapter 1) f assumes the form $f(x) = -\log x$, and for A-optimality (vide Chapter 1) we take $f(x) = x^{-1}$. Thus d^* is A-, and D-optimal provided (2a.3.22) is satisfied for $q(r) = -f(g(r))$, the factor $\frac{v}{v-1}$ in the original definition of $q(r)$ (viz: $q(r) = -f(\frac{v}{v-1}\, g(r))$) can be conveniently dropped as with this form of q and the above two forms of f it makes no essential difference in the verification of (2a.3.22). So in our calculations from now onwards we adopt

$$q(r) = -f(g(r)). \qquad \ldots(2a.3.24)$$

D-optimality :

Here $f(\lambda_{di}) = -\log \lambda_{di}$

and $q(r) = -f(g(r)) = \log g(r)$ (vide (2a.3.24), and the discussion above)

Thus with this form of q, (2a.3.22) reads

$$\log g(r+1) - \log g(r) \leq \log g(r) - \log g(r-1) \qquad \ldots(2a.3.25)(i)$$

$$\text{for all } r : r-1 \geq C_0,\ r+1 \leq D_0 - 1$$

and

$$\log g(\bar{r}+1) - \log g(\bar{r}) \geq \log g(D_0) - \log g(D_0 - 1) \qquad \ldots(2a.3.25)(ii)$$

Recalling the expression for $g(r)$ (vide (2a.3.14)) we rewrite (2a.3.25)(ii) as

$$\frac{g_2(\bar{r}+1)}{g_2(\bar{r})} \geq \frac{g_1(D_o)}{g_2(D_o-1)} \qquad \ldots(2a.3.25)(ii)'$$

Now we define
$$\left.\begin{aligned} g_i^*(r) &= mv(mv-1)(mv-2)g_i(r) \\ \Delta_i^*(r) &= mv(mv-1)(mv-2)\Delta_i(r) \\ \Delta^*(r) &= mv(mv-1)(mv-2)\,\Delta\,(r) \end{aligned}\right\} \quad i = 1,2$$

Direct calculations yield, writing x for mv,

$$\begin{aligned} g_2^*(\bar{r}+1) &= mx^4 - x^3(m^2+4m-1) + x^2(4m^2+m-1) - x(3m^2-4m+1) - (2m-2) \\ &= m^5v^3(v-1) - 4m^4v^2(v-1) + m^3v(v^2+v-3) - m^2v(v-4) - m(v+2) + 2 \\ g_2^*(D_o-1) &= mx^4 - x^3(m^2+3m+1) + x^2(2m^2+4m+1) - x(4m-3) - 2 \\ &= m^5v^3(v-1) - m^4v^2(3v-2) - m^3v^2(v-4) + m^2v(v-4) + 3mv - 2 \\ g_2^*(\bar{r}) &= g_2^*(\bar{r}+1) - x^3 + (2m+1)x^2 - (2m-1)x - 2 \\ &= m^5v^3(v-1) - 4m^4v^2(v-1) + 3m^3v(v-1) + 2m^2v - 2m \\ g_1^*(D_o) &= mx^4 - x^3(3m+m^2) + mx^2(2m+2) \\ &= m^3v^2\left\{m^2v(v-1) - m(3v-2) + 2\right\} \end{aligned}$$

so that

$$\begin{aligned} &g_2^*(\bar{r}+1)g_2^*(D_o-1) - g_2^*(\bar{r})g_1^*(D_o) \\ &\quad = x^6(m-1) - x^5(6m^2-4m-2) + x^4(6m^3+11m^2-20m+3) \\ &\qquad - x^3(18m^3-15m^2-11m+8) + x^2(12m^3-33m^2+20m+1) \\ &\qquad + x(14m^2-22m+8) + (4m-4) \end{aligned}$$

$$= m^7v^4(v^2-6v+6) - m^6v^3(v^3-4v^2-11v+18)$$

$$+ m^5v^2(2v^3-20v^2+15v+12) + m^4v^2(3v^2+11v-33)$$

$$- 2m^3v(4v^2-10v-7) + m^2v(v-22) + 4m(2v+1)-4$$

and this is non-negative for all $v \geq 5$, $m \geq 2$. Also this is negative for $v = 3,4$, $m \geq 2$. This verifies (2a.3.25)(ii)' for $v \geq 5$, $m \geq 2$. Now to check (2a.3.25)(i) we have to show that

(i) $g_1^2(r) \geq g_1(r-1)g_1(r+1)$ for $C_o < r < C_o + \left[\frac{mv}{2}\right]$

(ii) $g_1^2(C_o + \left[\frac{mv}{2}\right]) \geq g_1(C_o + \left[\frac{mv}{2}\right]-1)g_2(C_o + \left[\frac{mv}{2}\right]+1)$

(iii) $g_2^2(C_o + \left[\frac{mv}{2}\right]+1) \geq g_1(C_o + \left[\frac{mv}{2}\right])g_2(C_o + \left[\frac{mv}{2}\right]+2)$

(iv) $g_2^2(r) \geq g_2(r+1)g_2(r-1)$ for $r > C_o + \left[\frac{mv}{2}\right]+1$.

Of course, when $\frac{mv}{2}$ is an integer, (iii) and (iv) above can be combined into a single statement since in such a case

$$g_1(C_o + \tfrac{mv}{2}) = g_2(C_o + \tfrac{mv}{2}).$$

Recall that

$$g_1^*(r) = r^2mv + r(mv-1)(m^2v^2-4m^2v+2) + 2(m-1)(mv-1)(m^3v^2-mv)$$

so that

$$g_1^{*2}(r) - g_1^*(r-1)g_1^*(r+1)$$

$$= r^2.2m^2v^2 + r.2mv(mv-1)(m^2v^2-4m^2v+2) - m^2v^2$$

$$+ (mv-1)^2(m^2v^2-4m^2v+2)^2$$

$$- 4(m-1)(mv-1)(m^3v^2-mv)mv \qquad \text{...(2a.3.26)}$$

Thus, in order to verify (i) above, equivalently one has to achieve non-negativity of (2a.3.26) for $C_o < r < C_o + \frac{mv}{2}$. Differentiating (2a.3.26) with respect to r, we get,

$$4m^2v^2r + 2mv(mv-1)(m^2v^2 - 4m^2v + 2)$$

which is positive for all $v \geq 5$. Therefore, replacing r by $C_o + 1 = mv(m-1) + 1$ in (2a.3.26), we get

$$\begin{aligned} g_1^{*2}(r) - g_1^*(r-1)g_1^*(r+1) \geq\ & 2m^4v^4(m-1)^2 + m^2v^2 + 4m^3v^3(m-1) \\ & + 2m^2v^2(mv-1)(m-1)(m^2v^2 - 4m^2v + 2) \\ & + 2mv(mv-1)(m^2v^2 - 4m^2v + 2) \\ & + (mv-1)^2(m^2v^2 - 4m^2v + 2)^2 \\ & - 4(m-1)(mv-1)mv(m^3v^2 - mv) \qquad \ldots(2a.3.27) \end{aligned}$$

In (2a.3.27), a part of the right hand side expression viz,

$$2m^2v^2(mv-1)(m-1)(m^2v^2 - 4m^2v + 2) - 4(m-1)(mv-1)mv(m^3v^2 - mv)$$

simplifies to $2mv(mv-1)(m-1)(m^3v^2(v-6) + 4mv)$ and this is non-negative for all $v \geq 6$ and $m \geq 2$. The remaining part of (2a.3.27) is always non-negative. Hence (2a.3.27) is non-negative for all $v \geq 6$, $m \geq 2$. Again, note that (2a.3.27) can be rewritten as

$$\begin{aligned} & g_1^{*2}(r) - g_1^*(r-1)g_1^*(r+1) \\ & \quad = 2m^4v^4(m-1)^2 + m^2v^2 + 4m^3v^3(m-1) + (mv-1)^2(m^2v^2 - 4mv + 2)^2 \\ & \qquad + 2mv(mv-1)(m^2v^2 - 4m^2v + 2) + 2mv(mv-1)(m-1)(m^3v^3 - 6m^3v^2 + 4mv) \\ & \qquad \qquad \ldots(2a.3.28) \end{aligned}$$

It can be checked that for $v = 5$, (2a.3.28) is always non-negative for any $m \geq 2$. Thus (i) holds for all $v \geq 5$, $m \geq 2$.

Remark 3 : It may be noted that in checking (i) we did not make use of the upper bound of r. So it still holds as an algebraic inequality for $C_o + \frac{mv}{2} \leq r \leq D_o - 1$.

Next, we proceed to check (ii). Recall that the inequality in (2a.3.13) ensures that

$$g_2(C_o + \left[\frac{mv}{2}\right] + 1) \leq g_1(C_o + \left[\frac{mv}{2}\right] + 1).$$

Hence, by (i), and the above remark, the claim in (ii) is verified.

As regards (iii), we have only to look to the case of $\frac{mv}{2} \neq$ integer. For this, using the relation (2a.3.13), we have,

$$\begin{aligned} & g_2^{*2}(C_o + \tfrac{mv+1}{2}) - g_1^{*}(C_o + \tfrac{mv-1}{2})g_2^{*}(C_o + \tfrac{mv+3}{2}) \\ &= g_1^{*2}(C_o + \tfrac{mv+1}{2}) - g_1^{*}(C_o + \tfrac{mv-1}{2})g_1^{*}(C_o + \tfrac{mv+3}{2}) \\ &\quad + 2(mv-1)(g_1^{*}(C_o + \tfrac{mv+1}{2}) - 3\Delta_1^{*}(C_o + \tfrac{mv-1}{2})) + 4(mv-1)^2 \end{aligned} \quad \ldots(2a.3.29)$$

In view of the above remark,

$$g_1^{*2}(C_o + \tfrac{mv+1}{2}) - g_1^{*}(C_o + \tfrac{mv-1}{2})g_1^{*}(C_o + \tfrac{mv+3}{2})$$

is non-negative for all $v \geq 5$, $m \geq 2$, and so (2a.3.29) would be non-negative whenever $g_1^{*}(C_o + \frac{mv+1}{2}) - 3\Delta_1^{*}(C_o + \frac{mv-1}{2})$ is so.

Now writing $C_o + \frac{mv+1}{2}$ as r,

$$g_1^*(r) - 3\Delta_1^*(r-1) = r^2 mv + r\{(mv-1)(m^2v^2 - 4m^2v + 2) - 6mv\}$$
$$+ 3(5m^2v^2 - 7mv + 2) + 2(m-1)(mv-1)(m^3v^2 + 5mv)$$
$$- 3m^3v^3 + 6mv$$

and this is clearly non-negative for all $v \geq 5$, $m \geq 2$.

Thus it remains to verify condition (iv).

Using the relations (2a.3.13), (2a.3.27) and expressions of $g_1(r)$ (vide (2a.3.11)) we get, on simplification,

$$g_2^{*2}(r) - g_2^*(r+1)g_2^*(r-1) = r^2 \cdot 2m^2v^2 + r \cdot 2mv\,(mv-1)(m^2v^2 - 4m^2v-2)$$
$$- m^2v^2 + (mv-1)^2(m^2v^2 - 4m^2v + 2)^2$$
$$+ 4mv(mv-1)(6m^2v + 3mv - 8m)$$
$$- 4m^2v^2(mv-1)(m^3v - m^2v - m + 2mv + 1)$$
$$\ldots(2a.3.30)$$

As in Case (i) (2a.3.30) is also an increasing function of r. So putting the least value of r, namely, $(m-1)mv + \frac{mv}{2} + 1$, one can see that (2a.3.30) is non-negative for $v \geq 5$, $m \geq 2$.

Hence, d^* is D-optimal for $v \geq 5$, $m \geq 2$.

<u>A-optimality</u> :

Here $f(\lambda_{di}) = \frac{1}{\lambda_{di}}$

and $q(r) = -f(g(r)) = -\frac{1}{g(r)}$ (vide (2a.3.24), and the discussion just above (2a.3.24).)

Thus, with this form of q, (2a.3.22) reads,

$$g(r)(g(r+1)+g(r-1)) \geq 2g(r-1)g(r+1) \quad \text{for all} \quad C_0 < r < D_0 - 1 \qquad \ldots(2a.3.31)(i)$$

and

$$\frac{1}{g(D_0-1)} - \frac{1}{g(D_0)} \leq \frac{1}{g(\bar{r})} - \frac{1}{g(\bar{r}+1)} \qquad \ldots(2a.3.31)(ii)$$

Recalling the expression for $g(r)$ we rewrite (2a.3.31) (ii) as

$$g_2(\bar{r}+1)g_2(\bar{r})\Delta(D_0-1) \leq g_1(D_0)g_2(D_0-1)\Delta(\bar{r}) \qquad \ldots(2a.3.31)(ii)'$$

Since d^* is completely symmetric, D-optimality of d^* for $v \geq 5$, $m \geq 2$, will automatically imply its A-optimality for $v \geq 5$, $m \geq 2$. For $v = 4$, $m \geq 2$, both the relations (2a.3.31)(i) and (2a.3.31)(ii)' and, consequently, A-optimality of d^* hold. For $v = 3$, $m \geq 2$ (2a.3.31)(i) fails to hold. (We omit the details). So we cannot infer about A-optimality of d^* for $v = 3$, $m \geq 2$.

<u>E-optimality</u> :

Since d^* is completely symmetric, E-optimality of d^* for $v \geq 4$, $m \geq 2$ will automatically follow from its A-optimality. So it remains to check E-optimality for $v = 3$, $m \geq 2$. We develop below arguments to establish E-optimality of d^*. The task of proving E-optimality is so fascinating that by now there have appeared in the literature a considerable number of articles dealing exclusively with E-optimality. In the same spirit we also provide a very general result on E-optimality of d^* for <u>all</u> $m \geq 2, v \geq 3$. The proof <u>does not</u> require knowledge of validity of D- and/or A-optimality for any combination of m and v.

First note that (vide Kiefer (1975))

$$\text{min eigenvalue of } (PC_dP') \leq \frac{v}{v-1} \min_{h \in \{1,2,\ldots,v\}} C_{dhh} \qquad \ldots(2a.3.32)$$

So it suffices to verify that d^* maximizes $\left\{\min_h C_{dhh}\right\}$. We do the verification below :

For any design, there exists a treatment say h_o, such that $r_{h_o} \leq \bar{r}$. Recall that $\bar{r} = (m-1)mv + m(v-1)$ belongs to second half of the elementary interval $[(m-1)mv, m.mv]$. We now distinguish between the following cases :

<u>Case (i)</u> : $umv \leq r_{h_o} < umv + \frac{mv}{2}$, $u = 0,1,2,\ldots,m-1$ (r_{h_o} covering all r values in the first halves of the elementary intervals upto and including the one containing $\bar{r}$).

<u>Case (ii)</u> a) $umv + \frac{mv}{2} \leq r_{h_o} < (u+1)mv$, $u = 0,1,\ldots,m-2$

(r_{h_o} covering all r values in the second halves of the elementary intervals upto but excluding the one containing $\bar{r}$).

b) $(m-1)mv + \frac{mv}{2} \leq r_{h_o} \leq \bar{r}$.

(r_{h_o} covering all the r values in the second half of the elementary interval $[C_o, D_o]$ containing $\bar{r}$).

<u>Under Case (i)</u> It is enough to establish that

$$C_{dh_oh_o} \leq g_1(r_{h_o}) \leq g_2(\bar{r}) = C_{d^*hh}$$

Also under Case (ii) (a) and (b) it is enough to verify

$$C_{dh_oh_o} \leq g_2(r_{h_o}) \leq g_2(\bar{r}) = C_{d^*hh}.$$

Clearly, these in their turn establish E-optimality of d^*. We proceed through the following steps.

Step I$'$: For r_{h_o} belonging to the first (second) half of any interval, upto and including $[C_o, D_o]$, we find upper bounds to $g_1(r_{h_o})$ (respectively $g_2(r_{h_o})$) involving g-values at points in the first (respectively second) half of the (basic) elementary interval $[C_o, D_o]$ containing $\bar{r}$.

Step II$'$: We establish that (i) $g_2(r)\uparrow$ in r in the second half of $[C_o, D_o]$ and further that

(ii) $g_2(C_o + \frac{mv+1}{2}) \geq g_1(C_o + \frac{mv-1}{2})$ in case mv is odd.

Step III$'$: Once we are through with the above two steps (verification given below), we argue as follows :

Case (i) : $g_1(r_{h_o}) \leq g_1(C_o + \left[\frac{mv}{2}\right])$ (by Step I$'$)

$\leq g_2(C_o + \left[\frac{mv}{2}\right] + 1)$ (by Step II$'$(ii))

$\leq g_2(\bar{r})$ (by Step II$'$(i))

Case (ii) (a) $g_2(r_{h_o}) \leq g_1(r_{h_o})$ (using (2a.3.13))

$\leq g_1(C_o + \left[\frac{mv}{2}\right])$ (vide Figure 1)

$\leq g_2(\bar{r})$ (by steps followed in Case (i))

Case (ii) (b) $g_2(r_{h_o}) \leq g_2(\bar{r})$ (by Step II$'$(i)).

Verifications : (Step I$'$) Figure 1 supports the claim that $g_1(r_{h_o}) \leq g_1(C_o + \left[\frac{mv}{2}\right])$ whenever r_{h_o} is in the first half of any interval upto and including $[C_o, D_o]$. As regards g_2, the property (2a.3.13) together with the increasing nature of g_1 at least upto $[C_o, D_o]$ justifies the claim that $g_2(r_{h_o}) \leq g_1(C_o + [\frac{mv}{2}])$ for any r_{h_o} in the second half of any interval upto but excluding $[C_o, D_o]$.

(Step II$'$(i)): Referring to (2a.3.13) and (2a.3.15) we get

$$\Delta_2^*(r) = \Delta_1^*(r) - 4(mv-1)$$

which is increasing in r with $\Delta_2^*(C_o + \frac{mv}{2}) > 0$ for all $v \geq 3$, $m \geq 2$. Hence the claim.

(Step II$'$(ii)): $g_2^*(C_o + \frac{mv+1}{2}) - g_1^*(C_o + \frac{mv-1}{2})$ simplifies to $m^2 v(mv^2 - 2mv - 2v + 4)$ which is non-negative for $v \geq 3$, $m \geq 2$.

Thus, finally E-optimality of d^* is settled for any $v \geq 3$, and all $m \geq 2$.

2a.4 Construction of Optimal Designs

In this section we develop methods of construction of some series of such optimal designs d^*.

The discussion regarding complete symmetry of d^* in Step IV of subsection 2a.3 - in other words, condition (iii) (b) in the statement of Theorem 2a.3.1 yields

$$b = mv = \lambda\binom{v}{2}, \text{ for some } \lambda \geq 1. \qquad \ldots(2a.4.1)$$

This means that for $m = 1$, d^* does not exist unless $v = 3$. In fact, it is clear that d^* does not exist even for $v = 3$, $m = 1$. We will present here construction of d^* for the least possible value of m, since for any other multiple of this value (= m (say)), denoted by $m^* = km$, the same design for the given m can be inserted as block diagonals k times, off - diagonals being filled up suitably by appropriate Latin Squares or F Squares.[2]

For $v = 3$, it is sufficient to consider d^* for $m = 2$, and $m = 3$ — the designs for all $m > 3$ following from them easily. These designs are shown below :

$v = 3,\ m = 2$

$$d^* = \begin{bmatrix} - & 1 & 2 & 0 & 2 & 0 \\ 1 & - & 0 & 2 & 2 & 1 \\ 2 & 0 & - & 0 & 1 & 1 \\ 0 & 0 & 1 & - & 1 & 2 \\ 0 & 2 & 2 & 1 & - & 0 \\ 1 & 2 & 1 & 2 & 0 & - \end{bmatrix}$$

$v = 3,\ m = 3$

$$d^* = \begin{bmatrix} - & 0 & 1 & 2 & 2 & 1 & 0 & 1 & 2 \\ 2 & - & 0 & 1 & 2 & 2 & 0 & 0 & 1 \\ 1 & 2 & - & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & - & 1 & 0 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 & - & 0 & 1 & 2 & 2 \\ 1 & 0 & 0 & 1 & 2 & - & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 1 & 2 & - & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 & 2 & - & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 1 & 2 & - \end{bmatrix}$$

We now proceed to the case of general v.

<u>Case (i)</u> : <u>v even integer</u>

For v even, a $v-1$ resolvable BIBD always exists with number of blocks $b^* = \binom{v}{2}$, and block size $k = 2$. The blocks of the BIBD

---

[2]<u>Foot-note</u> : Here we use special type of F-squares viz, $F(n,\lambda)$ which is an arrangement of t symbols (say), in a $n \times n$ array such that each symbol appears in each row and in each column $\lambda(\geq 1)$ times.

can be split into $v-1$ sets of $\frac{v}{2}$ blocks so that each set contains each of the v treatments exactly once. Let one of the representative sets say t^{th} set comprises of the blocks $(i_1,j_1), (i_2,j_2),\ldots,(i_{\frac{v}{2}}, j_{\frac{v}{2}})$ so that $i_1,i_2,\ldots,i_{\frac{v}{2}}, j_1,j_2,\ldots,j_{\frac{v}{2}}$ together exhaust the v treatments. Consider the square

$$A = \begin{bmatrix} - & v-1 & v-2 & v-3 & \ldots & & 1 \\ 0 & - & v-1 & v-2 & \ldots & & 2 \\ 1 & 0 & - & v-1 & \ldots & & 3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ v-3 & v-4 & v-5 & v-6 & \ldots & - & v-1 \\ v-2 & v-3 & v-4 & v-5 & \ldots & 0 & - \end{bmatrix}$$

In A, the pairs missing each twice in the same row and column numbers are precisely $(0, v-1),(1, v-2),\ldots,(\frac{v}{2}-1,\frac{v}{2})$. Corresponding to the t^{th} set of the given BIBD now construct the square A_t from A by permuting the symbols $(0,1,2,\ldots,v-1)$ to $(i_1,i_2,\ldots,i_{\frac{v}{2}}, j_{\frac{v}{2}}, j_{\frac{v}{2}-1},\ldots,j_1)$. Thus in A_t, the pairs missing (each twice) in the same row and column numbers are precisely given by the blocks of the t^{th} set of the BIBD). Each set of blocks thus gives rise to a similar square A_t, $t = 1,2,\ldots,v-1$. The resultant design with $m = v-1$ (which is its least value) is thus given by

$$d^* = \begin{bmatrix} A_1 & L & L & \ldots & L \\ L & A_2 & L & \ldots & L \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L & L & L & \ldots & A_{v-1} \end{bmatrix}$$

where L's are Latin Squares with v treatments.

Case (ii) : v = odd prime or prime power > 3 .

Let $\alpha_0 = 0$, $\alpha_1 = 1$, $\alpha_2, \ldots, \alpha_{\frac{v-1}{2}}, \alpha_{\frac{v-1}{2}+1} = -\alpha_1$, $\alpha_{\frac{v-1}{2}+2} = -\alpha_2$, $\ldots$, $\alpha_{v-1} = -\alpha_{\frac{v-1}{2}}$ be the v distinct elements of $GF(v)$.

Let us assume that it is possible to construct a $v \times v$ square A with missing diagonals such that the i^{th} row (column) contains all the v symbols of $GF(v)$ except α_i (respectively $(\alpha_i + \alpha)$) where $0 \neq \alpha \in GF(v)$, $i = 0,1,2,\ldots,v-1$. Let $A_i = \alpha_i A$ or $-\alpha_i A$, $i = 1,2,\ldots,\frac{v-1}{2}$. Then,

$$\begin{bmatrix} A_1 & L & L \ldots L \\ L & A_2 & L \ldots L \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ L & L & L \ldots A_{\frac{v-1}{2}} \end{bmatrix}$$

is the required design with $m = \frac{v-1}{2}$, where L is any Latin Square with v symbols of $GF(v)$.

Example : $v = 5$.

$\alpha_0 = 0$, $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = -\alpha_1 = 4$, $\alpha_4 = -\alpha_2 = 3$, are the elements of $GF(5)$.

A can be chosen as

$$A = \begin{bmatrix} - & 1 & 2 & 4 & 3 \\ 0 & - & 4 & 3 & 2 \\ 3 & 4 & - & 1 & 0 \\ 2 & 3 & 0 & - & 1 \\ 4 & 0 & 1 & 2 & - \end{bmatrix} \qquad \text{(with } \alpha = 1\text{)}.$$

Below we present the forms of the matrix A for $v = 7$ and $v = 9$ as well.

$$v = 7, A = \begin{bmatrix} - & 3 & 4 & 2 & 6 & 5 & 1 \\ 3 & - & 5 & 6 & 2 & 0 & 4 \\ 6 & 5 & - & 0 & 4 & 1 & 3 \\ 4 & 1 & 0 & - & 5 & 2 & 6 \\ 5 & 0 & 1 & 3 & - & 4 & 2 \\ 2 & 4 & 6 & 1 & 3 & - & 0 \\ 0 & 6 & 2 & 5 & 1 & 3 & - \end{bmatrix} \qquad (\text{with } \alpha = 1)$$

$v = 9$,

$$A = \begin{bmatrix} - & x+1 & x+2 & 1 & 2 & x & 2x & 2x+2 & 2x+1 \\ x+2 & - & 2 & x+1 & 2x & 2x+2 & 0 & 2x+1 & x \\ 2x+1 & 2x+2 & - & 2 & 1 & x+1 & x+2 & 0 & 2x \\ 0 & 2x & 2x+1 & - & 2x+2 & 2 & 1 & x & x+2 \\ 2x+2 & 0 & 1 & 2x+1 & - & 2x & x & x+1 & 2 \\ 2x & 2x+1 & 2x+2 & x & x+2 & - & x+1 & 1 & 0 \\ x & 1 & 0 & 2x+2 & 2x+1 & x+2 & - & 2 & x+1 \\ x+1 & x+2 & x & 2x & 0 & 2x+1 & 2 & - & 1 \\ 2 & x & 2x & 0 & x+1 & 1 & 2x+2 & x+2 & - \end{bmatrix} (\text{with } \alpha = 1)$$

2a.5 Efficiency of Aggarwal's Designs.

From the discussion in subsection 2a.4, it is clear that the construction of such optimal designs is combinatorially rather involved. This leads us to a study of the relative efficiencies of Aggarwal's (1966b) designs and their generalizations having certain nice simple structures compared to the actual optimal designs characterized herein.

Aggarwal (1966b) presented a series of two-way Latin Square designs with all distinct elements missing along the diagonal. These designs $\tilde{d}$ can be generalized to d_o, in the case of $mv \times mv$ arrays (for $m \geq 2$) by placing such $v \times v$ designs $\tilde{d}$ along the diagonal block matrices and ordinary Latin Squares L along the off-diagonal blocks. It is not difficult to verify that d_o does not maximize C_{dhh} and is not optimal, however, d_o as given by

$$d_o = \begin{bmatrix} \tilde{d} & L & L \dots L \\ L & \tilde{d} & L \dots L \\ \dots\dots & \dots\dots & \dots\dots \\ \dots\dots & \dots\dots & \dots\dots \\ L & L & L \dots \tilde{d} \end{bmatrix}$$

possesses a high degree of efficiency as is demonstrated below with respect to the A-optimality criterion.

Note that for the optimal design d^* as well as for the above mentioned design d_o, the C-matrices are completely symmetric. Hence, one gets for the efficiency of d_o, the expression

$$E_{d_o} = \frac{\sum\sum_{i \neq j} V(\hat{\tau}_i - \hat{\tau}_j) \text{ using } d^*}{\sum\sum_{i \neq j} V(\hat{\tau}_i - \hat{\tau}_j) \text{ using } d_o} \times 100\%$$

$$= \frac{\sum_i \frac{1}{\lambda_{d^*i}}}{\sum_i \frac{1}{\lambda_{d_o i}}} \times 100\%$$

$$= \frac{a_o}{a^*} \times 100\%.$$

Using the representations $C_{d_o} = a_o(I - \frac{J}{v})$ and $C_{d^*} = a^*(I - \frac{J}{v})$.

The difference in the diagonal elements of C_{d^*} and C_{d_o} arises out of the difference in the diagonal elements of $(L_{d^*} + M_{d^*})(L_{d^*} + M_{d^*})'$ and $(L_{d_o} + M_{d_o})(L_{d_o} + M_{d_o})'$ only. Recalling now the expressions for C_{dhh}, $A(\bar{r})$ and $g_2(\bar{r})$ (vide subsection 2a.3., Step II), we get

$$(1 - \frac{1}{v})a_o = \bar{r} - \frac{2}{mv} - mv(m-1)^2 + (2\bar{r} - mv)(m-1) + \bar{r}$$

$$+ \frac{\bar{r}^2}{(mv-1)(mv-2)} - \frac{4m^2.m(v-1) + 4m(m-1)^2}{mv(mv-2)}$$

$$= A(\bar{r}) - \frac{4(v-1)}{v(mv-2)} = E(\bar{r}) \quad \text{say,}$$

and $$(1 - \frac{1}{v})a^* = A(\bar{r}) - \frac{2(2v-3)}{v(mv-2)}.$$

Thus E_{d_o} now simplifies to $\dfrac{E(\bar{r})}{E(\bar{r}) + \frac{2}{v(mv-2)}} \times 100\%$.

Calculations indicate that even for small values of v and m (e.g. $v = 4$, $m = 2$) d_o has an efficiency as high as 99%.

2b. Three-way Elimination of Heterogeneity Settings with Incidence Structure as $J - I$ for Every Pair of Directions.

2b.1 Preliminaries

In this section, we take up a study of optimal designs in a set-up admitting of three-way elimination of heterogeneity where incidence structure for every pair of directions is nonorthogonal in the sense that level combinations along the principal diagonal are infeasible, all other level combinations being feasible. Let each of the directions assume b levels, and (i_1, i_2, i_3) denote the level combination for any experimental unit $1 \leq i_1, i_2, i_3 \leq b$, $i_1 \neq i_2 \neq i_3$. Thus, as it stands, the number of feasible experimental units for allocation of v treatments is $b(b-1)(b-2)$ which may be too large to be available for experimentation. So, following Mukhopadhyay and Mukhopadhyay (1984), we deal with a reduced set of experimental units, suitably arranged such that the incidence structure becomes $J-I$ for every pair of directions. Let E be an $OA(b^2,3,b,2)$ such that the three constraints have the same level combinations in each of the first b columns, i.e. the first b columns are of the type $\begin{pmatrix} i \\ i \\ i \end{pmatrix}$, $i = 1,2,\ldots,b$. Such an OA can always be constructed (Denes and Keedwell (1974)). Then the remaining $b(b-1)$ columns of this orthogonal array serve as experimental units for us in three-way elimination of heterogeneity set-up, where the entry in the i^{th} row of the u^{th} column denotes the level of i^{th} direction in the u^{th} experimental unit, $1 \le i \le 3$, $1 \le u \le b(b-1)$. By non-orthogonality of the set-up, we mean this sort of restriction in the incidence structure.

The usual fixed effects model is

$$y_{jj'k}(h) = \mu + \alpha_j + \beta_{j'} + \nu_k + \tau_h + e_{jj'k}$$

$$1 \le h \le v, \quad 1 \le j \ne j' \ne k \le b$$

where $y_{jj'k}(h)$ is the observation in the $(j,j',k)^{th}$ level combination receiving h^{th} treatment and $\mu, \alpha_j, \beta_{j'}, \nu_k, \tau_h$ stand respectively for general effect, effect of the j^{th} level along the 1^{st} direction, effect of j'^{th} level along the 2^{nd} direction, effect of k^{th} level along the 3^{rd} direction, and h^{th} treatment effect, $e_{jj'k}$'s are i.i.d $N(o,\sigma^2)$.

For a specified design d, let $N_{di} = ((n^{i}_{dh\ell}))$ $h = 1,2,\ldots,v$ $\ell = 1,\ldots,b$ stands for the treatment - i^{th} direction incidence matrix, $i = 1,2,3$.

As in the case of two way nonorthogonal set-up (section 2a), we are interested in linear inferential problems involving treatment contrasts only and as such we refer to the underlying C-matrix of the design d. Let $\underline{r}_d = (r_{d1},\ldots,r_{dv})'$ be the vector of treatment replications for d. Then it can be easily seen that with incidence structure as $J-I$ for every pair of directions, for any design d,

$$C_d = D_{r_d} - (N_{d1}, N_{d2}, N_{d3}) \begin{pmatrix} (b-1)I & J-I & J-I \\ J-I & (b-1)I & J-I \\ J-I & J-I & (b-1)I \end{pmatrix}^{-} \begin{pmatrix} N'_{d1} \\ N'_{d2} \\ N'_{d3} \end{pmatrix}$$

and with a particular choice of g-inverse, we get

$$C_d = D_{r_d} - \frac{1}{b(b-3)}(N_{d1}, N_{d2}, N_{d3}) \left(\begin{array}{c|c|c} (b-2)I - \frac{2}{(b-1)}\frac{J}{b} & I - \frac{J}{b} & I - \frac{J}{b} \\ \hline I - \frac{J}{b} & (b-2)(I - \frac{J}{b}) & I - \frac{J}{b} \\ \hline I - \frac{J}{b} & I - \frac{J}{b} & (b-2)(I - \frac{J}{b}) \end{array}\right) \begin{pmatrix} N'_{d1} \\ N'_{d2} \\ N'_{d3} \end{pmatrix}$$

After simplification, C_d becomes

$$C_d = D_{r_d} - \frac{b-4}{b(b-3)} \sum_{i=1}^{3} N_{di}N'_{di} - \frac{1}{b(b-3)} \sum_{\substack{i<j \\ i,j=1,2,3}}\sum (N_{di} + N_{dj})(N_{di} + N_{dj})'$$

$$+ \frac{2}{(b-1)(b-3)} \underline{r}_d \underline{r}'_d \qquad \ldots(2b.1.1)$$

As before, let Ω be the class of all connected designs with incidence structure as J-I for every pair of directions. In our later derivation, we make use of the above C_d matrix and find optimal designs for the cases of $b = mv$ and $b = mv+1$. For $b = mv+1$, universally optimal designs are again available (subsection 2b.2) whereas for $b = mv$, we come up with specific optimal designs (subsection 2b.3). We omit the details of the proofs of optimality results as they are essentially similar to those obtained in the two way elimination of heterogeneity set-up (section 2a) and point out only the major differences in the arguments. We mainly consider the construction of optimal designs for the case of $b = mv$ (subsection 2b.4) and calculate efficiencies of a class of designs by suitably extending Aggarwal's designs (subsection 2b.5).

2b.2 <u>Universal Optimality Results for $b = mv+1$</u>

Here Ω stands for the class of connected designs for the typical OA's with $b = mv+1$ and with the incidence structure as J-I for every pair of directions. As in the case of two way elimination of heterogeneity set-up, we have a similar result here :

<u>Theorem 2b.2.1</u> Let $d^* \varepsilon \Omega$ be a design which assigns each treatment m times to each level of every direction. Then d^* is universally optimal.

Following the same steps as in two way heterogeneity set-up, it can be easily verified that here also

(i) C_{d*} is c.s.

and (ii) $tr(C_d) \leq tr(C_{d*}) \ \forall \ d \ \varepsilon \ \Omega$.

Thus d^* is universally optimal.

Such types of designs can be easily constructed. We can always construct (vide Denes and Keedwell (1974)) an $OA((mv+1)^2, 4, mv+1, 2)$ such that the first constraint is arranged in the order $0,\ldots,0, 1,\ldots,1, 2,\ldots,2 \ldots, mv \ldots mv$, each symbol being repeated $(mv+1)$ times. Further each of the remaining three constraints has the same symbol in each of the first $(mv+1)$ columns. We identify first constraint as denoting the treatment, and the remaining three constraints as three directions. Then by construction the level combination (i,i,i), $i = 0,1,2,\ldots,mv$ receives treatment 0. We delete these first $(mv+1)$ columns from this OA and reduce the rest of the symbols mod v in row number 1. The resulting array produces the required d^*.

2b.3 <u>Specific Optimality in the Case $b = mv$: A Brief Discussion of Optimality Results : Major Differences with Two-way Heterogeneity Set-up.</u>

Here Ω stands for the collection of connected designs for the typical OA's with $b = mv$ and with the incidence structure as $J-I$ for every pair of directions. As in the case of two-way heterogeneity set-up, the C-matrix of an A-, D-, and E- optimal design here is completely symmetric but it does not necessarily produce maximum trace of C_d in Ω. Take for example $b = 9$, $m = 3$, $v = 3$. The design d_1 has larger trace than d_2, which will be shown to be E-optimal, where,

$d_1 =$

A_1 –	A_3 0	A_2 1	B_3 2	B_1 0	B_2 1	C_2 2	C_1 0	C_3 1
A_3 1	A_2 –	A_1 0	B_1 1	B_2 2	B_3 0	C_1 1	C_3 2	C_2 0
A_2 0	A_1 2	A_3 –	B_2 0	B_3 1	B_1 2	C_3 0	C_2 1	C_1 2
B_3 1	B_2 2	B_1 0	C_1 –	C_3 0	C_2 2	A_3 1	A_2 2	A_1 0
B_1 2	B_3 0	B_2 1	C_3 1	C_2 –	C_1 0	A_2 0	A_1 1	A_3 2
B_2 0	B_1 1	B_3 2	C_2 0	C_1 1	C_3 –	A_1 2	A_3 0	A_2 1
C_3 0	C_2 1	C_1 2	A_2 2	A_1 0	A_3 1	B_1 –	B_2 0	B_3 1
C_2 2	C_1 0	C_3 1	A_3 0	A_2 1	A_1 2	B_2 2	B_3 –	B_1 0
C_1 1	C_3 2	C_2 0	A_1 1	A_3 2	A_2 0	B_3 0	B_1 1	B_2 –

$d_2 =$

A_1 –	A_3 1	A_2 2	B_3 0	B_1 2	B_2 0	C_2 0	C_1 1	C_3 2
A_3 1	A_2 –	A_1 0	B_1 2	B_2 2	B_3 1	C_1 1	C_3 2	C_2 0
A_2 2	A_1 0	A_3 –	B_2 0	B_3 1	B_1 1	C_3 2	C_2 0	C_1 1
B_3 2	B_2 0	B_1 1	C_1 –	C_3 1	C_2 2	A_3 0	A_2 0	A_1 1
B_1 0	B_3 1	B_2 2	C_3 1	C_2 –	C_1 0	A_2 0	A_1 2	A_3 2
B_2 1	B_1 2	B_3 0	C_2 2	C_1 0	C_3 –	A_1 1	A_3 2	A_2 1
C_3 0	C_2 1	C_1 2	A_2 2	A_1 0	A_3 1	B_1 –	B_2 1	B_3 2
C_2 1	C_1 2	C_3 0	A_3 0	A_2 1	A_1 2	B_2 1	B_3 –	B_1 0
C_1 2	C_3 0	C_2 1	A_1 1	A_3 2	A_2 0	B_3 2	B_1 0	B_2 –

(In d_1 as well as in d_2, the three directions are given by rows, columns, and the Roman letters A_1, A_2, A_3, C_1, C_2, C_3, B_1, B_3, B_2, and the treatments are denoted by 0,1,2.)

Thus we fail to apply Proposition 1 of Kiefer (1975) as regards universal optimality. So we look for specific optimality results.

<u>Theorem 2b.3.1</u> Suppose for given $b = mv$, $m \geq 2$ with incidence structure for every pair of directions (in a three-way heterogeneity set-up) as $J - I$, there exists a design d^* for which

(i) treatment - 1st direction incidence pattern is a BBD
$(n^{(1)}_{d^*h\ell} = m$ or $m-1)$

(ii) treatment - 2^{nd} direction incidence pattern is a BBD
$(n^{(2)}_{d^*h\ell} = m$ or $m-1)$

(iii) treatment - 3^{rd} direction incidence pattern is a BBD
$(n^{(3)}_{d^*h\ell} = m$ or $m-1)$

(iv) and moreover, for every pair of directions (j,j'), $j<j'$, $1 \leq j, j' \leq 3$, the b pairs of treatments of the type $(h_{ji}, h'_{j'i})$ $i = 1,2,\ldots,b$ where h_{ji} is the treatment which occurs $m-1$ times in the i^{th} level of j^{th} direction and $h'_{j'i}$ is the treatment which occurs $(m-1)$ times in the i^{th} level of j'^{th} direction, are such that they

(a) satisfy $h_{ji} \neq h'_{j'i}$ $\forall$ $i = 1,2,\ldots,b$

(b) exhaust all possible $\binom{v}{2}$ pairs of treatments equally often.

Then d^* is D - optimal for all $v \geq 7$, $m \geq 2$

A - optimal for all $v \geq 6$, $m \geq 2$*

E - optimal for all $v \geq 3$, $m \geq 2$.

The proof goes through following essentially the technique in Kiefer (1975) and the modification of tools developed in the case of two way heterogeneity set-up (vide subsection 2a.3, Step II, Step III). We omit the details to avoid a lengthy repetition of analogous steps and only highlight below the functional form of $g(r)$ and the difference in the set of sufficient conditions required to ensure existence of a concave function $\bar{q}$ satisfying (2a.3.9). Referring to (2b.1.1) it can be easily seen that for

$$r = mv\left[\frac{r}{mv}\right] + t$$

$$= mvu + t \quad \text{say,}$$

i.e. for r in the u^{th} elementary interval $[\,umv, (u+1)mv\,]$

$$g(r) \underset{\text{(vide 2a.3.3)}}{=} \max_{\{d:\, r_{dh}=r\}} C_{dhh} = \begin{cases} g_1(r) = A(r) + B_1(r) & \text{if } t < \frac{mv}{2} \\ & \text{i.e. if } \frac{r}{mv} - \left[\frac{r}{mv}\right] < \frac{1}{2} \\ g_2(r) = A(r) + B_2(r) & \text{if } t \geq \frac{mv}{2} \\ & \text{i.e. if } \frac{r}{mv} - \left[\frac{r}{mv}\right] \geq \frac{1}{2} \end{cases} \qquad \text{...(2b.3.1)}$$

is defined over the range of values $\{0,1,2,\ldots,mv(mv-1)\}$

---

*For $v = 5$, d^* may be A-optimal for $m \geq m_0$ where $m_0 > 30$. So for all practical purposes, d^* is not A-optimal for $v = 5$.

where $A(r) = r + \frac{2r^2}{(mv-1)(mv-3)} - \frac{3(mv-4)}{mv(mv-3)} \left\{- u^2mv + r + 2ur - mvu\right\}$

$$- \frac{3}{mv(mv-3)} \left\{8ur - 4mvu^2\right\}$$

$$= \frac{2r^2}{(mv-1)(mv-3)} + r\left\{\frac{m^2v^2 - 6mv - 6umv + 12}{mv(mv-3)}\right\} + \frac{3mvu^2 + 3(mv-4)u}{mv-3}$$

$$B_1(r) = -\frac{3}{mv(mv-3)}(2r - 2mvu)$$

$$B_2(r) = -\frac{3}{mv(mv-3)}(6r - 6mvu - 2mv)$$

In other words,

$$g(r) = \begin{cases} g_1(r) & \text{if} \quad umv \leq r \leq umv + \left[\frac{mv}{2}\right] \\ \\ g_2(r) & \text{if} \quad umv + \left[\frac{mv}{2}\right] + 1 \leq r < (u+1)mv \end{cases} \qquad \ldots(2b.3.2)$$

$$u = 0,1,2,\ldots,mv-2$$

Moreover, as in the two way heterogeneity set-up, projecting $g_1(r)$ beyond the first half of any elementary interval we get, here,

$$g_1(r) - g_2(r) = \frac{3}{mv(mv-3)}(4(r - mvu) - 2mv)$$

$$= \frac{3}{mv(mv-3)}(4t - 2mv) \begin{cases} > 0 & \text{if} \quad t > \frac{mv}{2} \\ = 0 & \text{if} \quad t = \frac{mv}{2} \\ < 0 & \text{if} \quad t < \frac{mv}{2} \end{cases}$$

$$\ldots(2b.3.3)$$

Hence, following the same arguments given in subsection 2a.3, Step II, to find a concave function $\tilde{\tilde{q}} \geq q$ satisfying (2a.3.9) we work with the <u>modified g</u> as follows : (vide(2a.3.14))

$$g(r) = \begin{cases} g_1(r), & r \leq C_o + \frac{mv}{2}, \quad r \geq D_o \\ \\ g_2(r), & C_o + \frac{mv}{2} \leq r \leq D_o - 1 \end{cases} \qquad \ldots(2b.3.4)$$

where $[C_o, D_o]$ is the elementary interval containing $\bar{r} = m(mv-1)$ $= (m-1)mv + m(v-1) > C_o + \frac{mv}{2}$.

Here also the behaviour of the function $g(r)$ has been similarly studied (vide subsection 2a.3, Step II). It turns out that for the case $\frac{mv}{2}$ = integer, $g(r)\uparrow$ right from the start $(r=0)$ to $r = u_o mv - 1$, then decreases at the point $u_o mv$, again increases upto $(u_o+1)mv-1$, covering thereby all the intermediate points in the intervals corresponding to $u = 0$ through $u = u_o$ where

$$u_o = \left[\frac{mv-4}{2}\right] \qquad \ldots(2b.3.5)$$

Then again $g(r)$ decreases starting from $(u_o+3)mv$. The exact behaviour of the function g cannot be specified for $(u_o+1)mv \leq r < (u_o+3)mv$. Also whenever $\frac{mv}{2} \neq$ integer, $mv > 6$, $g(r)\uparrow$ right from the start upto $r = (u_o+1)mv-1$, covering all the intermediate points in the intervals corresponding to $u = 0$ through $u = u_o$, and again starts decreasing from $(u_o+3)mv$. The exact functional behaviour cannot be specified for $(u_o+1)mv \leq r < (u_o+3)mv$. Carrying out similar calculations as in the two way heterogeneity set-up vide Step II and Step III, it turns out that with this form of g (vide(2b.3.4)) a set of sufficient conditions for existence of a concave function $\bar{q}$ satisfying $\bar{q} \geq q$ and $\bar{q}(\bar{r}) = q(\bar{r})$, in other words, a set of sufficient conditions for (2a.3.21) to hold is

$$\left.\begin{array}{ll} \text{(i)} & \text{concavity of } q \text{ within } C_0 \leq r \leq D_0-1 \\ \text{(ii)} & q(C_0) - q(C_0-1) \geq q(\bar{r}) - q(\bar{r}-1) \\ \text{(iii)} & q(\bar{r}+1) - q(\bar{r}) \geq q(D_0) - q(D_0-1) \end{array}\right\} \qquad \ldots(2b.3.6)$$

At this stage, it may be noted that the above condition is slightly different from those under the two way heterogeneity set-up. We need an extra condition namely condition (ii) here because of peculiarity of the function $\Delta_1(r) = g_1(r+1) - g_1(r)$ at the point $C_0 - 1$. It has been observed that in the present context,

$$\left.\begin{array}{l} \Delta_1(r) \geq \Delta_1(C_0) > 0 \quad \text{for all} \quad r \leq C_0-2 \\ \text{and } \Delta_1(C_0-1) < \Delta_1(C_0) \end{array}\right\} \qquad \ldots(2b.3.7)$$

Recall the analogous calculation in two way heterogeneity set-up (vide (2a.3.23)) where we obtained $\Delta_1(r) \geq \Delta_1(C_0) > 0$ for all $r \leq C_0 - 1$. Thus applying Mean-Value Theorem, (vide discussion of proof of sufficiecy of (2a.3.21) (i) in page (58)) we obtain from (2b.3.7)

$$\min_{\substack{r_1 \leq C_0 \\ r_1 \neq C_0-1}} \left\{ q(r_1+1) - q(r_1) \right\} = q(C_0+1) - q(C_0).$$

This together with concavity of q within $C_0 \leq r \leq D_0-1$ implies

$$\min_{\substack{0 \leq r_1 < \bar{r} \\ r_1 \neq C_0-1}} \left\{ q(r_1+1) - q(r_1) \right\} = q(\bar{r}) - q(\bar{r}-1).$$

So to achieve (2a.3.21) (i) we require the extra condition (2b.3.6) (ii) to be satisfied by q.

The other two conditions in (2b.3.6) follow as in the two way heterogeneity set-up.

Thus in the three-way heterogeneity set-up, φ^*-optimality of d^* will follow from a verification of (2b.3.6) along with two properties of d^* namely,

Property 1 : C_{d^*} is c.s.

Property 2 : $C_{d^*hh} = g(\bar{r})$.

As in the two way heterogeneity set-up, verifications are done here for D- and A- optimality of d^* (vide subsection 2a.3, Step IV, Step V). E-optimality follows along a direction very similar to that in the earlier set-up.

In the following, we concentrate on construction of such optimal designs which is rather involved and more complicated than in the case of two way heterogeneity set-up.

2b.4 Construction of Optimal Designs

Condition (iv) (b) as stated in Theorem 2b.3.1 yields, $mv = \lambda.\binom{v}{2}$, for some λ, a positive integer. We have already observed that for two way elimination of heterogeneity, in such cases, the required d^* does not exist for $m = 1$. We will present here construction of d^* for (i) $v = 3$, $m = 3$, (ii) $v = 2^t$, $t > 1$ integer, $m = v-1$, (iii) $v = tk+1$, odd prime or prime power, $t \geq 3$ integer, $k \geq 3$ odd integer, and $m = v-1$. We also exhibit how these d^* can be employed to get optimal designs for the multiples of m already chosen.

(i) $v = 3$, $m = 3$.

$d^* =$

A_1 –	A_3 1	A_2 2	B_3 0	B_1 2	B_2 0	C_2 0	C_1 1	C_3 2
A_3 1	A_2 –	A_1 0	B_1 2	B_2 2	B_3 1	C_1 1	C_3 2	C_2 0
A_2 2	A_1 0	A_3 –	B_2 0	B_3 1	B_1 1	C_3 2	C_2 0	C_1 1
B_3 2	B_2 0	B_1 1	C_1 –	C_3 1	C_2 2	A_3 0	A_2 0	A_1 1
B_1 0	B_3 1	B_2 2	C_3 1	C_2 –	C_1 0	A_2 0	A_1 2	A_3 2
B_2 1	B_1 2	B_3 0	C_2 2	C_1 0	C_3 –	A_1 1	A_3 2	A_2 1
C_3 0	C_2 1	C_1 2	A_2 2	A_1 0	A_3 1	B_1 –	B_2 1	B_3 2
C_2 1	C_1 2	C_3 0	A_3 0	A_2 1	A_1 2	B_2 1	B_3 –	B_1 0
C_1 2	C_3 0	C_2 1	A_1 1	A_3 2	A_2 0	B_3 2	B_1 0	B_2 –

where $A_1, A_2, A_3, C_1, C_2, C_3, B_1, B_3, B_2$ denote the 9 levels of the 3rd direction, the first two directions being the rows and the columns as usual, and 0, 1, 2 denote the three treatments.

In the following two cases $v = 2^t$ and $v = tk + 1$, we use some common notations. Let L be a Latin Square written with v symbols of $GF(v)$ such that the symbols appearing along the diagonal are all distinct. Let $(L + x_i)$, $x_i \in GF(v)$ denotes the Latin Square obtained from L by replacing each symbol y of L by the symbol $(y + x_i)$. L^* is obtained from L after deleting the diagonal (i.e. the diagonal positions of L^* are all empty) and $(L + x_i)^{**}$ is obtained from $(L + x_i)$ by only replacing its diagonal by the diagonal of L.

(ii) $v = 2^t$, $t > 1$ integer, $m = v-1$.

Define $x_i = 1 + \alpha + \alpha^2 + \ldots + \alpha^{i-1}$, $i = 1,2,\ldots,m$, where α is a primitive element of $GF(v)$. Obviously, x_i's are all distinct elements of $GF(v)$, and $x_m = 0$. We define a set of mv symbols divided into m sets, the i^{th} set written as $A_i(x)$, $x \in GF(v)$, $i = 1,2,\ldots,m$. Let A_i be a Latin Square written with the v symbols $A_i(x)$, such that along the diagonal of A_i, all the symbols are distinct. Each A_i, so constructed when superimposed on L is assumed to form a Graeco Latin Square. Let $\tilde{L}$ be any other Latin Square with elements of $GF(v)$, which need not have distinct elements along the diagonal and will form a Graeco Latin Square when superimposed on A_i. Now d^* can be constructed as follows.

First make a $m \times m$ Latin Square treating $\mathcal{A}_i$'s $i = 1,2,\ldots,m$ as m distinct symbols, such that along the principal diagonal and along the second direct diagonal (i.e., just above the principal diagonal) the symbols are all distinct. Since $m = v-1$ is odd, a cyclic Latin Square*

*(vide Denes and Keedwell (1974))

of order m with symbols $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m$ will possess these two properties. Now in place of symbol $\mathcal{A}_i$, insert the Latin Square A_i. The symbols $A_i(x)$, $i = 1,2,\ldots,m$, $x \in GF(v)$ in the order now they occur along the principal diagonal of this resulting $mv \times mv$ square will form the definite order of the levels of the third direction. Now arrange v treatments in this $mv \times mv$ square as follows :

Originally we started with an $m \times m$ Latin Square with m symbols $\mathcal{A}_i$, then $\mathcal{A}_i$'s are replaced by $v \times v$ squares A_i's as already indicated. Considering the original $m \times m$ square, its diagonal has m distinct elements $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m$, in some order. Superimpose the treatment square L^* on each of the A_i's (written in place of $\mathcal{A}_i$'s) in the principal diagonal. Then superimpose $(L + x_i)^{**}$, in the given order $i = 1,2,\ldots,m$ on A_i's along the second direct diagonal of the $m \times m$ original square. Superimpose $\tilde{L}$ on the remaining A_i's of this $m \times m$ square. Then d^* looks like this.

$d^* =$

A_1 L^*	A_2 $(L+x_1)^{**}$	A_3 $\tilde{L}$	A_4 $\tilde{L}$	... ...	A_{m-1} $\tilde{L}$	A_m $\tilde{L}$
A_2 $\tilde{L}$	A_3 L^*	A_4 $(L+x_2)^{**}$	A_5 $\tilde{L}$	... ...	A_m $\tilde{L}$	A_1 $\tilde{L}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_{m-1} $\tilde{L}$	A_m $\tilde{L}$	A_1 $\tilde{L}$	A_2 $\tilde{L}$	... ...	A_{m-3} L^*	A_{m-2} $(L+x_{m-1})^{**}$
A_m L	A_1 $\tilde{L}$	A_2 $\tilde{L}$	A_3 $\tilde{L}$	... ...	A_{m-2} $\tilde{L}$	A_{m-1} L^*

Clearly, from the construction procedure, d^* satisfies condition (i), (ii) and (iii) in the statement of Theorem 2b.3.1. As x_i's $i = 1, 2, \ldots, m-1$ are non-null distinct elements, d^* satisfies condition (iv(a)) also. To verify condition (iv(b)) we introduce vector notations $\underline{d(L)}$, and $\underline{d(L+x_i)}$, $i = 1,2,\ldots,m-1$, to denote the v distinct treatment symbols along the diagonals of L and $(L+x_i)$ respectively. The mv treatments (not all distinct) occurring with $(m-1)$ replications along the mv levels of row and column directions are given by the two vectors

$$(\underline{d(L+x_1)},\ \underline{d(L+x_2)},\ \underline{d(L+x_3)}, \ldots, \underline{d(L)})$$

and

$$(\underline{d(L)},\quad \underline{d(L+x_1)},\ \underline{d(L+x_2)}, \ldots, \underline{d(L+x_{m-1})})$$

respectively. As each of $\underline{d(L+x_i)}$ $i = 1,2,\ldots,m$ consists of v distinct treatment symbols and $(x_1, x_2-x_1, x_3-x_2, \ldots, x_{m-1}-x_{m-2}, -x_{m-1})$ consists of all non-null elements of $GF(v)$ each appearing exactly once, the mv treatment pairs occurring $(m-1)$ times along the (i,i) $i = 1,2,\ldots,mv$ level combinations of row and column directions respectively exhaust all possible $\binom{v}{2}$ pairs, each occurring equal number of times. Verification of this property for row vs third direction and column vs third direction is immediate.

(iii) <u>$v = tk+1$, odd prime or prime power, k odd, (> 1) $t > 2$ integer, $m = v-1 = tk$.</u>

Let x be a primitive element of $GF(v)$. Let A_i^j, $i = 1,\ldots,t$, $j = 1,\ldots,k$ be Latin Squares of size v with symbols $A_i^j(\alpha)$, $\alpha \in GF(v)$

such that along the principal diagonal, the symbols appearing are distinct and L and A_i^j's when superimposed, form Graeco Latin Squares. Let A_i's $i = 1,\ldots,t$ be Latin Squares of size k, formed with symbols $\mathcal{A}_i^j$, $j = 1,\ldots,k$, such that the symbols $\mathcal{A}_i^j$'s appearing along the principal diagonal and also the second direct diagonal are distinct. Since k is odd, this type of configuration is always possible. $\tilde{L}$ be any Latin Square of size v, with symbols from $GF(v)$, such that A_i^j superimposed on $\tilde{L}$, is a Graeco Latin Square. Now d^* can be constructed as follows.

First make a $t \times t$ Latin Square with symbols $\mathcal{A}_i$, $i = 1,\ldots,t$ such that along the principal diagonal the symbols are distinct. (As t is even, we can always construct a Latin Square having this property as follows*: First we construct a cyclic Latin Square L_0 (say) of size $t-1$ with symbols $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_{t-1}$. Then we add one more row and column to L_0 to form the resulting $t \times t$ Latin Square L_{00} (say) as follows. The $(t,t)^{th}$ cell in L_{00} is filled in by the symbol $\mathcal{A}_t$, the first $t-1$ cells of the t^{th} column as well as t^{th} row of L_{00} are filled in by the corresponding elements in the second direct diagonal of L_0, and all the elements in the second direct diagonal of L_0, are now replaced by the symbol $\mathcal{A}_t$.) Now in place of symbols $\mathcal{A}_i$'s, put the Latin Squares A_i's with symbols $\mathcal{A}_i^j$'s, $j = 1,\ldots,k$. Then in place of symbols $\mathcal{A}_i^j$'s put the corresponding Latin Squares A_i^j's. Now the order in which the symbols $A_i^j(\alpha)$, $i = 1,\ldots,t$, $j = 1,\ldots,k$, $\alpha \in GF(v)$, occur along the principal diagonal of this resulting $mv \times mv$ square represent the definite order of the levels of the third direction.

*(vide Franklin (1984))

We can now arrange v treatments on this square as follows. In the original $t \times t$ Latin Square, the symbols $\mathcal{A}_i$'s are replaced by the Latin Squares A_i's of order $k \times k$ with symbols $\mathcal{A}_i^j$'s. Superimpose L^* on each of the Latin Squares A_i^j's (written in place of $\mathcal{A}_i^j$'s) occurring in the principal diagonal of the resulting $tk \times tk$ square. Consider the second direct diagonal of each of the $k \times k$ squares, occurring in the principal diagonal positions of the original $t \times t$ square. Take a representative $k \times k$ square, say A_{i_1}. Then superimpose the treatment squares

$$\left(L+x^{i_1-1}\right)^{**},\ \left(L+x^{(i_1-1)+t}\right)^{**},\ldots,\ \left(L+x^{(i_1-1)+t(k-1)}\right)^{**}$$

in that order on the Latin Squares written corresponding to the elements occurring on the second direct diagonal of the square specified. The same is done for each such $k \times k$ square in the principal diagonal. Superimpose $\tilde{L}$ on each of the remaining $v \times v$ Latin Squares, written corresponding to the elements of the $tk \times tk$ square. The method is illustrated below :

$$\text{Let } C_i = \begin{array}{|c|c|c|c|}
\hline
\begin{matrix} A_i^1 \\ L^* \end{matrix} & \begin{matrix} A_i^2 \\ (L+x^{i-1})^{**} \end{matrix} & \cdots & \begin{matrix} A_i^k \\ \tilde{L} \end{matrix} \\
\hline
\begin{matrix} A_i^2 \\ \tilde{L} \end{matrix} & \begin{matrix} A_i^3 \\ L^* \end{matrix} & \cdots & \begin{matrix} A_i^1 \\ \tilde{L} \end{matrix} \\
\hline
\begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} \\
\hline
\begin{matrix} A^{k-1} \\ \tilde{L} \end{matrix} & \begin{matrix} A_2^k \\ \tilde{L} \end{matrix} & \cdots & \begin{matrix} A_i^{k-2} \\ (L+x^{i-1+t(k-2)})^{**} \end{matrix} \\
\hline
\begin{matrix} A_i^k \\ (L+x^{i-1+t(k-1)})^{**} \end{matrix} & \begin{matrix} A_i^1 \\ \tilde{L} \end{matrix} & \cdots & \begin{matrix} A_i^{k-1} \\ L^* \end{matrix} \\
\hline
\end{array}$$

$$B_i = \begin{array}{|c|c|c|c|}\hline A_i^1\,\tilde{L} & A_i^2\,\tilde{L} & \cdots & A_i^k\,\tilde{L} \\ \hline A_i^2\,\tilde{L} & A_i^3\,\tilde{L} & \cdots & A_i^1\,\tilde{L} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline A_i^k\,\tilde{L} & A_i^1\,\tilde{L} & \cdots & A_i^{k-1}\,\tilde{L} \\ \hline \end{array}\ , \quad i = 1,2,\ldots,t.$$

The design d^* is given by the matrix D, where

$$D = \begin{array}{|c|c|c|c|c|}\hline C_1 & B_t & B_3 & \cdots & B_2 \\ \hline B_2 & C_3 & B_t & \cdots & B_4 \\ \hline \cdots & \cdots & \cdots & \cdots & \cdots \\ \hline B_{t-1} & B_2 & B_4 & \cdots & C_t \\ \hline \end{array}$$

The sets $\{x^i, x^{i+t}, \ldots, x^{i+t(k-1)}\}$, $i = 0,1,\ldots,t-1$ have the following property : The sets

$$(x^i - x^{i+t(k-1)},\ x^{i+t} - x^i,\ x^{i+2t} - x^{i+t},\ \ldots,\ x^{i+t(k-1)} - x^{i+t(k-2)})\quad i = 0,1,\ldots,t-1$$

exhaust all the distinct non-null elements of $GF(v)$, each occurring exactly once. This ensures the property (iv(b)) of d^* along row vs column direction. Automatically the same property holds for row vs third and column vs third direction. The other properties (viz, condition (i) — iv(a) in the statement of Theorem 2b.3.1) are obviously satisfied.

In each of the above three cases, suppose the design is constructed for some m and we want to construct it for a multiple of m, say $m^* = pm$, $p > 2$. The $(mv \times mv)$ optimal d^* can then be used to construct $m^*v \times m^*v$ optimal designs. Let $d_1^*, d_2^*, \ldots, d_p^*$ be p optimal designs of size $mv \times mv$ constructed in the above manner with p different sets of mv symbols. d_i^{**} are designs obtained from d_i^*'s by replacing L^* and $(L+x)^*$, $x \varepsilon GF(v)$ of d_i^* by $\tilde{L}$ or L simply. Then d^* $(m^*v \times m^*v)$ can be constructed by making a $p \times p$ Latin Square with symbols d_i^* such that along the diagonal the symbols appearing are distinct. Then replace all the off diagonal d_i^*'s by d_i^{**}'s, and in place of d_i^* or d_i^{**} symbols, put the corresponding $(mv \times mv)$ squares.

2b.5 Efficiency of a Simple Class of Designs in the Line of Aggarwal (1966b).

The above discussion reveals that construction of these optimal designs d^* are highly involved. So in the following we give a simple design which is a trivial generalization of Aggarwal's designs considered in the case of two way elimination of heterogeneity and compute its efficiency with respect to the actual optimal designs.

Let us define a set of mv symbols divided into m sets, the i^{th} set written as α_j^i, $j = 1,\ldots,v$ $i = 1,\ldots,m$. Let GL_i be a $v \times v$ Graeco Latin Square formed with two sets of symbols $\{1,2,\ldots,v\}$ and $\{\alpha_1^i, \alpha_2^i, \ldots, \alpha_v^i\}$ such that along the diagonal $\{1\alpha_1^i, 2\alpha_2^i, \ldots, v\alpha_v^i\}$ occurs in the same order. $\widetilde{GL}_i$ be any Graeco Latin Square formed with the above two sets of symbols. GL_i^* is obtained from GL_i after deleting its diagonal. Then a new $mv \times mv$ design d_o can be formed as follows. First make a $m \times m$ Latin Square treating GL_i, $i = 1,\dots,m$, as m symbols such that along the diagonal the symbols appearing are distinct. Then replace the diagonal GL_i symbols by the Graeco Latin Square GL_i^*, and off diagonal GL_i's by the corresponding Graeco Latin Square $\widetilde{GL}_i$. For m odd, one representation of d_o can be

$$d_o = \begin{array}{|c|c|c|c|c|} \hline GL_1^* & \widetilde{GL}_2 & \widetilde{GL}_3 & \cdots & \widetilde{GL}_m \\ \hline \widetilde{GL}_2 & GL_3^* & \widetilde{GL}_4 & \cdots & \widetilde{GL}_1 \\ \hline \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hline \widetilde{GL}_m & \widetilde{GL}_1 & \widetilde{GL}_2 & \cdots & GL_{m-1}^* \\ \hline \end{array}$$

In d_o, the symbols α_j^i, $j = 1,\dots,v$, $i = 1,\dots,m$ denote the levels of third direction, the other two directions are represented by rows and columns as usual. Clearly d_o does not maximize C_{dhh}, and is not optimal. However, it possesses a high degree of efficiency as is demonstrated below with respect to A-optimality criterion.

Note that for the optimal design d^* as well as for the above mentioned extended Aggarwal's design d_o, the C-matrices are completely symmetric. Hence we get for the efficiency of d_o the expression,

$$E_{d_o} = \frac{\sum\sum_{i \neq j} V(\hat{\tau}_i - \hat{\tau}_j) \text{ using } d^*}{\sum\sum_{i \neq j} V(\hat{\tau}_i - \hat{\tau}_j) \text{ using } d_o} \times 100\%.$$

$$= \frac{\Sigma \frac{1}{\lambda_{d^*i}}}{\Sigma \frac{1}{\lambda_{d_o i}}} \times 100\%$$

$$= \frac{a_o}{a^*}$$

using the representations $C_{d^*} = a^* (I - \frac{J}{v})$ and $C_{d_o} = a_o (I - \frac{J}{v})$.

Recalling now expressions for C_{dhh}, $g(\bar{r})$, $A(\bar{r})$ (vide subsection 2b.3)

we get

$$(1-\frac{1}{v})\, a^* = A(\bar{r}) - \frac{6m(2v-3)}{mv(mv-3)}$$

and $$(1-\frac{1}{v})\, a^o = A(\bar{r}) - \frac{12m(v-1)}{mv(mv-3)}$$

$$= E(\bar{r}) \text{ say.}$$

E_{d_o} now simplifies to $\dfrac{E(\bar{r})}{E(\bar{r}) + \frac{6}{v(mv-3)}} \times 100\%$.

Calculations indicate that even for moderate values of v and m, E_{d_o} is close to unity.

Concluding Remarks :

In this chapter, our primary concern was to initiate a study of (analysis and) optimality of designs in situations where the " data-base " is non-orthogonal. The emphasis was on layouts for which each pairwise directional incidence pattern corresponds to that of a SBIBD. In general, however, the infeasible cells would correspond to any positions in the square, thereby rendering the optimality problem really intractable. We hope further researches in this direction will add interesting contributions in future years.

CHAPTER 3

OPTIMAL WEIGHING DESIGNS WITH A STRING PROPERTY

Introduction

Weighing problems were first posed and discussed by Hotelling (1944) and, subsequently, by Mood (1946). Over the past three decades, various aspects of such problems have been extensively studied. We refer to Raghavarao(1971) and Banerjee (1975) for all the basic results on this topic. An important point to be noted regarding these problems is that " the designs are applicable to a great variety of problems of measurement not only of weights, but of lengths, voltages and resistances, concentrations of chemicals in solutions, in fact any measurement such that the measure of a combination is a known linear function of the separate measures with numerically equal coefficients " (Mood 1946). Whereas the statisticians have discussed such problems exclusively in the framework of measurement of weights, the general technique seems to have received attention of researchers in other fields as well. Fulkerson and Gross (1965) and Ryser (1969) have studied some interesting combinatorial problems involving some classes of (0,1) matrices with biological applications. The application of weighing designs to optics seems to have been first pointed out by Marshall and Comisarow (1975) and independently by Sloane and Harwit (1976). Harwit and Sloane (1979) have nicely explained how the problem of designing Hadamard encoded optical instruments is related to the theory of weighing designs.

In this chapter, we intend to study one special type of weighing designs with possible scopes of application to optics and elsewhere as well. We formulate the problem below in the language of measurement of distances among a set of fixed objects along a line. To be specific, suppose there are $(n+1)$ objects, serially numbered $1,2,\ldots,n+1$, their positions are fixed along a line and we are interested in measuring consecutive distances between them (or some functions thereof) by undertaking N measuring operations. Clearly, in any such operation, we can measure the distance between any two objects along the line. However, the interesting point to be noted is that in doing so, we are automatically taking account of the intermediate objects, if any. This means that in every row of the resulting design matrix, there will appear exactly one run of 1's (the rest of the elements being 0's). We may call such designs spring balance designs with string property. Fulkerson and Gross (1965) have called them (0,1) matrices with the consecutive 1's property. The problem studied by them as well as by Ryser (1969) can be stated as follows : How much information about a (0,1) matrix is needed to decide whether it has consecutive 1's property or not ? (Do we need to know the whole of it or something less will suffice ?) We propose to study in this chapter some inferential aspects of the underlying design problem and discuss various results relating to optimal designs. Most of the findings of this chapter have been published in Sinha and Saha (1983) and Mukerjee and SahaRay (1985).

Assume that the recorded observations follow the standard regression model :

$$\underline{Y}_{N\times 1} = X_{N\times n}\,\underline{\theta}_{n\times 1} + \underline{\varepsilon}_{N\times 1},$$
$$E(\underline{\varepsilon}) = 0$$
$$E(\underline{\varepsilon}\,\underline{\varepsilon}') = \sigma^2 I_N .$$

where $X_{N\times n}$ is a (0,1) design matrix with the string property, and $\underline{\theta} = (\beta_1, \beta_2, \ldots, \beta_n)'$ is the parameter vector of unknown consecutive distances between the objects.

The particular case of $N = n$ has been extensively studied and φ_p^* - optimal designs obtained for the two familiar problems of inferring on $\underline{\theta}$ (as a whole) and $\underline{\xi} = P\underline{\theta}$ where P $(n-1\times n)$ is as usual the lower submatrix of an orthogonal matrix $\overline{O}$ with the first row vector as $(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$. Incidentally, some interesting features of the (0,1) matrices with the string property have been observed and made use of in deducing the optimality results. The case of $N > n$ seems to be complicated and only the problem of inferring on $\underline{\theta}$ has been studied. In particular, for the cases of $N = n+1$, and $N = n+2$, A- and D-optimal designs have been derived within the subclass of relevant designs formed by inclusion of the optimum design for n observations, viz, the identity matrix of order n as a submatrix of the whole design matrix (section 3b). For general $N > n+2$, to deal with the more intractable design problem, approximate design theory comes up, and the tool of Frechet derivative (vide Silvey 1980) has been employed to obtain approximate A- and D- optimal designs. E- optimal designs are also obtained using some other technique (section 3c). Incidentally, for some combinations of N and n, these approximate D- and E- optimal designs lead to exact optimal designs.

3a. Optimal Designs for $N = n$

3a.1 Preliminaries :

Throughout this section, our investigation relates to the case of $N = n$, i.e. to the square matrices $X_{n \times n}$ having string property. The optimality results here are based mainly on the very structural properties of the design matrices and their inverses. So before going into the discussion of optimality results, we first demonstrate some structural properties of $X_{n \times n}$. (We sometimes omit the lower suffix n of this matrix to avoid unnecessary complications in the notations).

Property 1. $|X| = \pm 1$, or 0, according as X is of full rank or not.

Proof : By elementary operations (of row permutation and row differencing) one can reduce any (0,1) square nonsingular $n \times n$ matrix with the string property to an identity matrix. Hence the result.

Property 2. For a square full rank (0,1) matrix X with the string property, the inverse matrix X^{-1} exhibits the following structure :

(i) The elements in X^{-1} are $(0, \pm 1)$ and, in each column, the non-zero elements occur with alternate signs.

(ii) The column totals (through all or any subset of consecutive rows) are $(0, \pm 1)$ — not all being 0's.

(iii) If a certain column - total is $+1(-1)$, the first non-zero entry in that column is $+1(-1)$, if a certain column - total is 0, the first non-zero entry in that column is either $+1$ or -1.

Proof : The proof is by induction on n, the order of the square matrix X. The result is easily verified for $n = 1,2$. Assume the properties to hold good for all such matrices of order n and let X_{n+1} be a full rank (0,1) square matrix of order $(n+1)$ with the stated string property. Since the properties (claimed) relate to the columns of X_{n+1}^{-1}, we may conveniently convert X_{n+1} to X_{n+1}^{o} by necessary row permutations such that the first (last) row of X_{n+1}^{o} has the smallest string - length among all rows starting (ending) with 1 in the first (last) position and these rows are arranged in increasing order of the string lengths from the top (bottom). Nonsingularity of X_{n+1} guarantees existence of these two rows in X_{n+1}^{o} as distinct from one another. For example,

$$X_6 = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad \text{is converted to} \quad X_6^{o} = \left[\begin{array}{c|cccc|c} 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

We will now deduce all the structural properties of X_{n+1}^{o} by induction argument.

Let $(1 \mid \underset{\sim}{\beta}')$ be the first row vector of X_{n+1}^{o} and suppose the first column of X_{n+1}^{o} contains $(t+1)$ 1's $(t \geq 0)$. Then we premultiply X_{n+1}^{o} by the matrix

$$M_{\overline{n+1} \times \overline{n+1}} = \left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & \dots & 0 & \\ -1 & 1 & 0 & 0 & \dots & 0 & \\ -1 & 0 & 1 & 0 & \dots & 0 & O \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot & \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot & \\ -1 & 0 & 0 & 0 & \dots & 1 & \\ \hline & & O & & & & I_{n-t} \end{array}\right].$$

and convert X^o_{n+1} to one with the first column vector as $(1,0,\dots,0)'$. This yields,

$$MX^o_{n+1} = \left[\begin{array}{c|c} 1 & \underline{\beta}' \\ \hline O & A_n \end{array}\right] \qquad \dots(3a.1.1)$$

Certainly, A_n is a (0,1) matrix of order n with the string property and hence, all the structural results apply to A_n^{-1}. Now, $X^{o\,-1}_{n+1}$ has the following representation,

$$X^{o\,-1}_{n+1} = \left[\begin{array}{c|c} 1 & -\underline{\beta}'A_n^{-1} \\ \hline O & A_n^{-1} \end{array}\right] M. \qquad \dots(3a.1.2)$$

The effect of post multiplication by M would be to change only the first column vector of the matrix on which it acts and thus the following arguments will demonstrate the structural results for all the columns of $X^{o\,-1}_{n+1}$ except for the first one, from the structural properties of the columns in the submatrix $\left[\dfrac{-\underline{\beta}'A_n^{-1}}{A_n^{-1}}\right]$. Suppose $\underline{\beta}'$ has unity in the first s positions $(s \geq 0)$. Then $\underline{\beta}'A_n^{-1}$ is a vector comprising of the first s row-sums of A_n^{-1}. Consider a specific column in A_n^{-1}. If the column total is $1(-1)$, then the partial sums from the top are 1 or 0 (-1 or 0) and hence it will contribute -1 or 0 (1 or 0) to the vector $-\underline{\beta}' A_n^{-1}$ and the result will be in accordance with the claimed structure. The same is true for other columns having the totals in A_n^{-1} as zero.

Finally, to deduce the structural results for the first column of $X_{n+1}^{o\,-1}$, this time we premultiply X_{n+1}^{o} by a matrix $N_{\overline{n+1} \times \overline{n+1}}$ of the form

$$N = \left[\begin{array}{c|c} I_{n-\ell} & O \\ \hline O & \underbrace{\begin{matrix} 1 & 0 & 0 & \dots & 0 & -1 \\ 0 & 1 & 0 & \dots & 0 & -1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{matrix}}_{\ell+1} \end{array}\right] \qquad \dots(3a.1.3)$$

(assuming that last column of X_{n+1}^{o} contains $(\ell+1)$ 1's, $\ell \geq 0$) so that X_{n+1}^{o} changes to

$$N X_{n+1}^{o} = \left[\begin{array}{c|c} B_n & O \\ \hline 0 \dots 0 \;\; 1 \dots 1 & 1 \end{array}\right] \qquad \dots(3a.1.4)$$

where B_n is again a (0,1) matrix of order n with the string property. Let $(\underline{\alpha}' \mid 1)$ be the last row-vector in the representation $N X_{n+1}^{o}$ in (3a.1.4), then

$$X_{n+1}^{o\ -1} = \left[\begin{array}{c|c} B_n^{-1} & 0 \\ \hline -\underline{\alpha}' B_n^{-1} & 1 \end{array}\right] N \qquad \ldots(3a.1.5)$$

and, hence, by induction argument, we are through with regard to all the elements in the first column of $X_{n+1}^{o\ -1}$ except for the last element (in the $(n+1,\ 1)^{th}$ position). If $\underline{\alpha}'$ contains 1's in the last r positions $(r \geq 0)$, then $\underline{\alpha}' B_n^{-1}$ is a vector consisting of the last r row-sums in B_n^{-1}. Now we argue as follows as regards the first column of B_n^{-1}. If this column total is $1(-1)$, then the partial totals from the bottom are 1 or 0 (-1 or 0) and hence the last element in the first column of $X_{n+1}^{o\ -1}$ is -1 or 0 (1 or 0) and this is in accordance with the structure indicated. Similar considerations apply if this column total is zero. Hence the claim.

Lemma 3a.1.1 Whatever be the design matrix X (of full rank) $\underline{1}'\ (X'X)^{-1}\ \underline{1}' \leq n$ with "=" for $X^o = I_n$ (and possibly, for other matrices as well).

Proof : Property 2 of the matrix X readily settles this.

3a.2 Optimal Designs for Inferring on $\underline{\theta}$ and $\underline{\xi} = P\underline{\theta}$.

We now proceed in a systematic manner to establish the optimality results using the above two properties of the matrix X.

Let $\Omega(n,n)$ be the class of full rank (0,1) matrices with string property of order n, $n \geq 3$. An appeal to universal optimality criterion for full rank models (Kiefer (1975), Sinha and Mukerjee (1982), vide Proposition 1' in Chapter 1) would settle the problem of inferring on $\underline{\theta}$ provided $X^* = I_n$ would maximize $\operatorname{tr}(X'X)$ among all design matrices X in $\Omega(n,n)$. However, as a matter of fact, $\operatorname{tr}(X'X)$ attains its minimum value for $X^* = I_n$. So we look for specific optimality for inferring on $\underline{\theta}$.

Theorem 3a.2.1 Any X in $\Omega(n,n)$ is D-optimal for inferring on $\underline{\theta}$ and moreover $X^* = I_n$ is uniquely φ_p^*-optimal over the class $\Omega(n,n)$ for all $p > 0$. ($p = 0$ corresponds to D-optimality criterion).

Proof : D-optimality follows from Property 1(vide subsection 3a.1). Unique φ_p^*-optimality for all $p > 0$ follows from an application of Proposition 2, Kiefer (1975) (vide Chapter 1).

As regards inference on $\underline{\xi}$ we first deduce the following lemma.

Lemma 3a.2.1 $\underline{\xi}$ is estimable if and only if $\operatorname{rank}(X) = n$.

Proof : Since $\underline{\xi}$ is a vector of $n-1$ independent orthogonal contrasts, estimability of $\underline{\xi}$ requires $\operatorname{rank}(X) \geq n-1$, and "=" leads to a contradiction, as then $\mathcal{M}(X') \subseteq \mathcal{M}(P')$ and consequently $\underline{1}'X' = 0$ which is impossible. Thus X has to be of full rank.

Regarding inference on $\underline{\xi}$, $D(\hat{\underline{\xi}}) = \sigma^2 P(X'X)^{-1}P'$, and as in the case of inference on $\underline{\theta}$, $X^* = I_n$, does not maximize the trace of $P(X'X)^{-1}P'$ over the class $\Omega(n,n)$. For example, for $n = 4$, $\operatorname{tr}(\frac{1}{\sigma^2} D(\hat{\underline{\xi}}))\Big|_{X^* = I_n} = 3$, whereas $\operatorname{tr}(\frac{1}{\sigma^2} D(\hat{\underline{\xi}}))\Big|_{X = X_0} = 3.25$, where

$$X_o = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Thus Proposition 1' regarding universal optimality for full rank models is ..inapplicable here, and we look for specific optimality result again.

<u>Theorem 3a.2.2</u> $X^* = I_n$ is uniquely D - optimal over the class $\Omega(n,n)$ for inferring on $\underline{\xi} = P\underline{\theta}$.

<u>Proof</u> : $1 = |\bar{O}(X'X)^{-1}\bar{O}'| = \left| \begin{bmatrix} \frac{1}{n}\underline{1}'(X'X)^{-1}\underline{1} & \frac{1}{\sqrt{n}}\underline{1}'(X'X)^{-1}P' \\ \frac{1}{\sqrt{n}}P(X'X)^{-1}\underline{1} & P(X'X)^{-1}P' \end{bmatrix} \right|$

$$\leq \left\{\frac{1}{n}\cdot\underline{1}'(X'X)^{-1}\underline{1}\right\} |P(X'X)^{-1}P'|$$

$$\leq |P(X'X)^{-1}P'| \quad \text{(using Lemma 3a.1.1)}$$

so that $|D(\hat{\underline{\xi}})| \Big|_{X^*} \leq |D(\hat{\underline{\xi}})| \Big|_{X}$

and "=" holds if and only if

(i) $P(X'X)^{-1}\underline{1} = 0$

and (ii) $\underline{1}'(X'X)^{-1}\underline{1} = n$ are simultaneously satisfied.

However (i) implies $(X'X)^{-1}\underline{1} \propto \underline{1}$.

i.e. $(X'X)\underline{1} \propto \underline{1}$

Now set $(X'X)\underline{1} = k\underline{1}$ for some k (k is necessarily a positive integer)
Then, $X\underline{1} = k(X')^{-1}\underline{1}$ and this is impossible unless $(X')^{-1}\underline{1} = \underline{1}$ (in

view of Property 2, vide subsection 3a.1). This means

$$X\underline{1} = k\underline{1} \quad \text{for some positive integer } k,$$

i.e. X has constant row sums.

Certainly a matrix can have at most $(n-k+1)$ linearly independent rows with each row sum equal to k. Hence, because of full rank of X, it turns out that $k=1$, and $X^* = I_n$ (up to row permutation). This settles the claim.

Using the notions of φ_p^*-optimality criteria, we readily have :

Corollary 3a.2.1 $X^* = I_n$ is uniquely φ_p^*-optimal over the class $\Omega(n,n)$ for inferring on $\underline{\xi} = P\underline{\theta}$ for all $p \geq 0$. (It is yet a stronger result than that regarding $\underline{\theta}$).

Proof : It is enough to note that $D(\hat{\underline{\xi}})\big|_{X^*}$ is a multiple of the identity matrix. Hence Theorem 3a.2.2 and Proposition 2 in Kiefer (1975) (vide Chapter 1) justify the claim.

3b. Choice of Additional Observations

3b.1 Preliminaries :

In the context of weighing designs the problem of selecting additional weighings so as to make the overall design optimum in some sense has received attention of researchers from time to time. Raghavarao (1971) gave a nice discussion on this topic. In this section we consider a similar problem in our present set-up, and try to find the optimum way, using D- and A- optimality, of choosing one/two additional observations, given the optimal choice for n observations. The problem of characterization of E-optimal designs for general (N,n) has recently been solved by Jacroux (1986).

From section 3a, it is known that for n observations, I_n is φ_p^*-optimal for estimating $\underset{\sim}{\theta}$ over $\Omega(n,n)$. So our problem here is to obtain optimal forms of $\underset{\sim}{\alpha}'$, $\underset{\sim}{\beta}'$ (both having string property) assuming the form of the design matrix as $X = \begin{bmatrix} I_n \\ \hline \underset{\sim}{\alpha}' \\ \hline \underset{\sim}{\beta}' \end{bmatrix}$. Let $\Omega(N,n)$ stand for the general class of $N \times n$ design matrices of rank n having string property.

3b.2 Optimum Choice of One Additional Observation

Taking the form of X as $X = \left(\frac{I_n}{\underset{\sim}{\alpha}'}\right)$, in the following we obtain D- and A- optimal design, where $\underset{\sim}{\alpha}'$ is a $n \times 1$ vector with string property.

D - optimal Choice

It can be checked that,

$$|X'X| = |I_n + \underset{\sim}{\alpha}\,\underset{\sim}{\alpha}'| = |I_n|(1 + \underset{\sim}{\alpha}'\underset{\sim}{\alpha}),$$

and this is a maximum for the choice $\underset{\sim}{\alpha}' = (1\ 1\ \ldots\ 1\ 1)$.
As a matter of fact, for an arbitrary X, as $X = \begin{bmatrix} X_1 \\ \hline \underset{\sim}{\alpha}' \end{bmatrix}$ where X_1 can be taken to be a full rank matrix with string property,

$$|X'X| = |X_1'X_1 + \underset{\sim}{\alpha}\,\underset{\sim}{\alpha}'| = |X_1'X_1|\,(1 + \underset{\sim}{\alpha}'(X_1'X_1)^{-1}\underset{\sim}{\alpha}).$$

Now referring to Property 1 (vide subsection 3a.1), we get

$$| X_1'X_1 | = 1$$

and, again by Property 2 (vide subsection 3a.2), we have

$$\underline{\alpha}'(X_1'X_1)^{-1}\underline{\alpha} \leq n$$

"=" holds for $X_1 = I_n$ and $\underline{\alpha}' = (1\ 1\ \ldots\ 1\ 1)$.

Thus $X^* = \left[\dfrac{I_n}{1\ 1\ \ldots\ 1\ 1}\right]$ is D - optimal for $N = n+1$ within the <u>general</u> class of designs $\Omega\,(n+1,\ n)$ having string property.

<u>A - optimal Choice</u>

Since $(X^{*'}X^*)$ (X^* as described in the discussion of D - optimal choice just above) is not a multiple of the identity, Proposition 2 of Kiefer (vide Chapter 1) cannot be applied here to obtain A - optimal design ; however, the D - optimal design also turns out to be A - optimal.

It can be easily verified that

$$\mathrm{tr}\,(X'X)^{-1} = \mathrm{tr}\,(I_n - \frac{\underline{\alpha}\,\underline{\alpha}'}{1 + \underline{\alpha}'\underline{\alpha}}) = n - \frac{\underline{\alpha}'\underline{\alpha}}{1 + \underline{\alpha}'\underline{\alpha}}$$

and this is a minimum for the choice of $\underline{\alpha}'$ as $\underline{\alpha}' = (1\ 1\ \ldots\ 1\ 1)$.

3b.3 <u>Optimum Choice of Two Extra Observations</u>

Here, we assume the form of X as

$$X = \begin{bmatrix} I_n \\ \underline{\alpha}' \\ \underline{\beta}' \end{bmatrix} \quad \text{(both } \underline{\alpha}',\ \underline{\beta}' \text{ having string structure)} \qquad \ldots(3b.3.1)$$

and further that the 2×1 column vectors of the type $\binom{0}{0}$, $\binom{1}{0}$, $\binom{1}{1}$ and $\binom{0}{1}$ occur as columns in the two-rowed submatrix $\binom{\underline{\alpha}'}{\underline{\beta}'}$ with frequencies f_1, f_2, f_3, f_4 respectively. Obviously, $f_1+f_2+f_3+f_4=n$. In the following we display D- and A- optimal choices of $\underline{\alpha}'$ and $\underline{\beta}'$.

D - optimal Choice

Taking the form of X as indicated above, it turns out that

$$|X'X| = (f_2+1)(f_4+1) + f_3(f_2+f_4+2).$$

Writing $f_2+1 = x$, $f_4+1 = y$, and $f_3 = z$ (say) one has to maximize

$$xy + z(x+y)$$

subject to $x+y+z \leq n+2$, $x \geq 1$, $y \geq 1$, $z \geq 0$ integers.

Now it can be verified easily that for fixed $x+y = t$ (say), $2 \leq t \leq n+2$ and arbitrary z, $0 \leq z \leq (n+2-t)$, the function $xy+z(x+y)$ attains its maximum when x, y differ by at most one and for given x, y the function is maximum when $z = n+2-x-y$. The following table summarises the final result obtained by applying the usual technique for handling maximization problem.

Table 3b.3.1

Values of n	Values of f_1	f_2	f_3	f_4
$3m$	0	m	m	m
$3m+1$	0	m	$m+1$	m
$3m+2$	0	$m+1$	$m+1$	m

From this, one can easily deduce the optimal structures of $\underline{\alpha}'$ and $\underline{\beta}'$ in any specific case.

A - optimal Choice

After some algebraic simplification, it turns out, for the form of X as in (3b.3.1), that

$$\operatorname{tr}(X'X)^{-1} = f_1 + f_3 + \frac{f_2^2}{f_2+1} + \frac{f_4^2}{f_4+1}$$

$$- \frac{f_3}{1+f_3\left(\frac{1}{f_2+1}+\frac{1}{f_4+1}\right)} \cdot \left(\frac{1}{(f_2+1)^2}+\frac{1}{(f_4+1)^2}\right)$$

Setting as before $f_2+1 = x$, $f_4+1 = y$, and $f_3 = z$ (say), this time one has to minimize

$$n-2+\frac{1}{x}+\frac{1}{y} - \frac{z}{1+z\left(\frac{1}{x}+\frac{1}{y}\right)}\left(\frac{1}{x^2}+\frac{1}{y^2}\right)$$

subject to $x+y+z \leq n+2$; $x \geq 1$, $y \geq 1$, $z \geq 0$ integers.

It can be checked that for fixed $x+y = t$, (say), $2 \leq t \leq n+2$, and arbitrary z, $o \leq z \leq (n+2-t)$, the above function is the least when x and y differ by atmost one, and, further, for fixed x and y, the function attains its minimum when $z = (n+2-x-y)$. Thus applying the usual technique of minimization it turns out that for minimization of the above function, one should try only with the values of t in the range $\left[\frac{4(n+2)}{3+\sqrt{3}}\right] - 2 \leq t \leq \left[\frac{4(n+2)}{3+\sqrt{3}}\right] + 2$. The following table gives the final result for some values of n viz., $7 \leq n \leq 15$.

Table 3b.3.2

Values of n	Values of f_1	f_2	f_3	f_4
7	0	3	1	3
8	0	3	2	3
9	0	3	2	4
10	0	4	2	4
11	0	4	2	5
12	0	5	2	5
13	0	5	2	6
14	0	6	2	6
15	0	6	3	6

Thus for $n = 9$, the A-optimum choice of $\binom{\underline{\alpha}'}{\underline{\beta}'}$ is $\left(\begin{smallmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{smallmatrix}\right)$. Unlike the case of one extra observation, this is different from the D-optimum choice of $\binom{\underline{\alpha}'}{\underline{\beta}'}$ viz., $\left(\begin{smallmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{smallmatrix}\right)$.

3c. Optimal Designs for General $N > n$, for Inferring on $\underline{\theta}$

In the previous sections, we have dealt with the N-observation design problem as one of choice of N vectors $\underline{x}'_{(1)}, \ldots, \underline{x}'_{(N)}$ from the relevant design space $\mathcal{X}$ — a collection of all possible $\frac{n(n+1)}{2}$ vectors with elements 0 and 1 and with the string property. For the case $N > n$, the search for optimal exact " N-observation design " as such becomes intractable (as is pointed out in case of $N = n+1$, $N = n+2$). Instead, we consider a broader class of designs where N is no longer explicitly present, and the design is given as a " probability measure " on the design space $\mathscr{X}$. Some justification for this formulation is given by Kiefer and Wolfowitz (1959), Kiefer (1959), Fedorov (1972), Kiefer and Studden (1976), Studden (1978), Silvey (1980) in the context of approximate regression designs.

3c.1 <u>Preliminaries</u> :

For $1 \leq u \leq v \leq n$, let $\underline{h}_{uv}$ be a (0,1) vector with the string property having the run of 1's starting at the u^{th} and ending at the v^{th} positions (both inclusive). Here, $\mathscr{X}$, the design space, is the set of all such $\frac{n(n+1)}{2}$ vectors $\underline{h}_{uv}$. Following Silvey (1980, p.15), let H be the class of probability distributions on the Borel sets of $\mathscr{X}$. Any $\eta \varepsilon H$ will be called a design measure. Since $\mathscr{X}$ is finite (compact) any such η defines a discrete distribution over $\mathscr{X}$ assigning a mass, say, π_{uv} at $\underline{h}_{uv}$ $(1 \leq u \leq v \leq n)$. For $\eta \varepsilon H$, under the standard Gauss Markov linear model with homoscedasticity and independence of errors, the $n \times n$ information matrix is given by

$$M(\eta) = E(\underline{x}\ \underline{x}')$$

$\underline{x}$ being a random vector with distribution η. Denoting the $(i,j)^{th}$ element of $M(\eta)$ by $m_{ij}(\eta)$, it can be checked that,

$$m_{ij}(\eta) = \Sigma \pi_{uv} \qquad \ldots(3c.1.1)$$

where the summation extends over $u \leq i$ and $v \geq j$, $(1 \leq i \leq j \leq n)$.

Hence, it follows that

$$\pi_{uv} = m_{uv}(\eta) - m_{u,v+1}(\eta) - m_{u-1,v}(\eta) + m_{u-1,v+1}(\eta) \quad 1 \leq u \leq v \leq n \qquad \ldots(3c.1.2)$$

$m_{o,n+1}(\eta)$, $m_{ov}(\eta)$, $(1 \leq v \leq n)$ and $m_{u,n+1}(\eta)$ $(1 \leq u \leq n)$ being interpreted as zero.

Let $\mathcal{M} = \{M(\eta) : \eta \in H\}$; then the set $\mathcal{M}$ is convex. In fact, it is the convex hull of $\{\underline{x}\,\underline{x}' : x \in \mathcal{X}\}$. In particular, choice of η^0 as assigning probability 1 to the point $\underline{x}$ gives $\underline{x}\,\underline{x}'$ as a member of $\mathcal{M}$.

Suppose that φ is a functional defined on $k \times k$ symmetric matrices and bounded above on $\mathcal{M}$ such that

(i) φ is concave

(ii) $\varphi(M_1) \geq \varphi(M_2)$ if $M_1 - M_2$ is non-negative definite.

(iii) We allow φ to take the value $-\infty$ on all singular matrices in $\mathcal{M}$.

Then the problem is to determine η^* which maximizes $\varphi(M(\eta))$ over H. Any such η^* will be termed as φ - optimal.

In the sequel, we quote the definition of a directional derivative namely " Frechet-derivative " and two key theorems without proof (Silvey, 1980, Theorem 3.6, and Theorem 3.7) which play a basic role in the approximate design theory.

DFN 3c.1.1

Frechet derivative of φ at $M(\eta_1)$ in the direction $M(\eta_2)$ is defined as

$$F_\varphi(M(\eta_1), M(\eta_2)) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}\{\varphi((1-\varepsilon)M(\eta_1) + \varepsilon M(\eta_2)) - \varphi(M(\eta_1))\}$$

Theorem 3c.1.1 η^* is φ - optimal if and only if

$$F_\varphi(M(\eta^*), M(\eta)) \leq 0 \quad \text{for all} \quad \eta \in H.$$

Since $\mathcal{M}$ is convex, the essence of this theorem is in fact

Theorem 3c.1.2 If φ is differentiable at $M(\eta^*)$, then η^* is φ - optimal if and only if

$$F_\varphi(M(\eta^*), \underline{x}\,\underline{x}') \leq 0 \quad \text{for each} \quad \underline{x} \in \mathcal{X} .$$

In the following we use this theorem to obtain D- and A - optimal designs.

3c.2 D - optimal Designs

For D - optimality,

$$\varphi(M(\eta)) = \log |M(\eta)| \quad \text{if} \quad M(\eta) \quad \text{is nonsingular}$$

$$= -\infty \qquad \text{if} \quad M(\eta) \quad \text{is singular.}$$

With φ so defined, if M is nonsingular, it can be seen (Silvey, 1980, p.21) that

$$F_\varphi(M(\eta), \underline{x}\,\underline{x}') = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \log \left\{ \frac{|(1-\varepsilon)M(\eta) + \varepsilon\, \underline{x}\,\underline{x}'|}{|M(\eta)|} \right\}$$

$$= \underline{x}' \{M(\eta)\}^{-1} \underline{x} - n \qquad \ldots(3c.2.1)$$

after some simplification.

Let S be a matrix of order n with elements 2 along the principal diagonal, -1 along the diagonals just above and below the principal diagonal and 0 at each other position, e.g. with $n = 4$,

$$S = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

It can be readily seen that

$$\underline{x}' S \underline{x} = 2 \quad \text{for each } \underline{x} \in \mathfrak{X}$$

$$\text{i.e. } \frac{n}{2} \cdot \underline{x}' S \underline{x} - n = 0 \qquad \ldots(3c.2.2)$$

Thus it follows by Theorem 3c.1.2 and expression (3c.2.1) that if η^* be such that $M(\eta^*) = \frac{2}{n} \cdot S^{-1}$ then η^* is D-optimal. As

$$S^{-1} = \frac{1}{n+1} \begin{bmatrix} n & n-1 & n-2 & \ldots & 2 & 1 \\ n-1 & 2(n-1) & 2(n-2) & \ldots & 4 & 2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 2 & 4 & 6 & \ldots & 2(n-1) & n-1 \\ 1 & 2 & 3 & \ldots & n-1 & n \end{bmatrix},$$

by (3c.1.2) it becomes evident that $M(\eta^*) = \frac{2}{n} \cdot S^{-1}$ holds if and only if η^* assigns a probability mass $\frac{2}{n(n+1)}$ at each member of $\mathfrak{X}$. Hence the D-optimal design measure assigns equal mass to all members of $\mathfrak{X}$.

If N is a multiple of $\frac{n(n+1)}{2}$, we readily have the following exact optimality result.

<u>Corollary 3c.2.1</u> If $N = t \cdot \frac{n(n+1)}{2}$ for some integer $t \geq 1$, then the design matrix X^* in which each b_{uv} $(1 \leq u \leq v \leq n)$ occurs t times is D-optimal within the class $\Omega(N,n)$.

In particular, regarding exact D - optimality another result is anticipated. With N observations let n_{uv} be the number of times $\underline{h}_{uv}$ $(1 \leq u \leq v \leq n)$ occurs as a row of the design matrix. The findings above lead to the conjecture that when N is not a multiple of $\frac{n(n+1)}{2}$, in a D-optimum design every two n_{uv}'s should differ by at most 1. A complete enumeration of the possible situations confirms the conjecture for $n = 2$, or 3 and numerical examples suggest that this possibly holds for general n; however, an analytical proof is likely to involve complicated combinatorial techniques. At this stage we have been able to establish the following result on <u>restricted</u> D-optimality.

<u>Theorem 3c.2.2</u> Let $N = t.\frac{n(n+1)}{2} + \lambda$, $1 \leq \lambda \leq \left[\frac{n+1}{2}\right]$, and X^* be the exact D - optimal design for $N = \frac{n(n+1)}{2}$. Assume X is of the form

$$X = \begin{bmatrix} \left.\begin{matrix} X^* \\ \vdots \\ X^* \end{matrix}\right\} \text{ t copies} \\ \underline{\upsilon}_1' \\ \vdots \\ \underline{\upsilon}_\lambda' \end{bmatrix}$$

Then given this specified choice of first $t.\frac{n(n+1)}{2}$ vectors, a " D - optimum " choice of $\underline{\upsilon}_i$'s is as follows :

$\underline{\upsilon}_1'$ is any member in $\mathcal{X}_o$,

and $\underline{\upsilon}_i'(X^{*\prime}X^*)^{-1}\underline{\upsilon}_j = 0$ for $1 \leq i, j \leq \lambda$, $i \neq j$.

<u>Proof</u> : The proof follows by induction on ℓ.

Defining X_k as

$$X_k = \begin{bmatrix} \left.\begin{matrix} X^* \\ \vdots \\ X^* \end{matrix}\right\} \text{ t copies} \\ \underline{v}_1' \\ \vdots \\ \underline{v}_k' \end{bmatrix}$$

$$| X'X | = | X_\ell' X_\ell | \quad \text{and} \quad (X^{*'}X^*)^{-1} = \frac{1}{n+1} S.$$

Thus for $\ell = 1$,

$$| X_1'X_1 | = (1 + \underline{v}_1' \frac{(X^{*'}X^*)^{-1}}{t} \underline{v}_1) \, | \, t(X^{*'}X^*) \, |$$

$$= \left\{ 1 + \frac{2}{t(n+1)} \right\} | \, t(X^{*'}X^*) \, | ,$$

using (3c.2.2) whatever be the choice of $\underline{v}_1'$ in $\mathcal{X}$.

For $\ell = 2$,

$$| X_2'X_2 | = | X_1'X_1 | \Big\{ 1 + \underline{v}_2' \frac{(X^{*'}X^*)^{-1}}{t} \underline{v}_2$$

$$- \frac{1}{t} \frac{\underline{v}_2' \, (X^{*'}X^*)^{-1} \underline{v}_1 \underline{v}_1' (X^{*'}X^*)^{-1} \underline{v}_2}{t + \underline{v}_1'(X^{*'}X^*)^{-1} \underline{v}_1} \Big\}$$

and clearly this is a maximum when $\underline{v}_2'(X^{*'}X^*)^{-1}\underline{v}_1 = 0$ as $\underline{v}_2'(X^{*'}X^*)^{-1}\underline{v}_2$ is again a constant e.g. equal to $\frac{2}{n+1}$.

Now, we assume that the result holds for $\ell \leq s$. Then,

$$| X'_{s+1}X_{s+1} | = | X'_sX_s | \{ 1 + \underline{\upsilon}'_{s+1}(X'_sX_s)^{-1} \underline{\upsilon}_{s+1} \}$$

Now, $| X'_sX_s |$ is maximum for choices of $\underline{\upsilon}'_1, \ldots, \underline{\upsilon}'_s$ specified (by induction) and

$$(X'_sX_s)^{-1} = (X'_{s-1}X_{s-1})^{-1} - \frac{(X_{s-1}X_{s-1})^{-1}\, \underline{\upsilon}_s\underline{\upsilon}'_s(X'_{s-1}X_{s-1})^{-1}}{1+\underline{\upsilon}'_s(X'_{s-1}X_{s-1})^{-1}\underline{\upsilon}_s}$$

which implies

$$\underline{\upsilon}'_{s+1}(X'_sX_s)^{-1}\underline{\upsilon}_{s+1} = \underline{\upsilon}'_{s+1}(X'_{s-1}X_{s-1})^{-1}\upsilon_{s+1} \frac{\{\underline{\upsilon}'_{s+1}(X'_{s-1}X_{s-1})^{-1}\underline{\upsilon}_s\}^2}{1+\underline{\upsilon}'_s(X'_{s-1}X_{s-1})^{-1}\underline{\upsilon}_s}$$

Repeated arguments yield,

$$| X'_{s+1}X_{s+1} | \leq | X'_sX_s | \{ 1+\underline{\upsilon}'_{s+1} \frac{(X^{*'}X^{*})^{-1}}{t} \underline{\upsilon}_{s+1} \}$$

and " = " holds for the stated choice of $\underline{\upsilon}_{s+1}$

and hence the claim.

An application of this theorem yields a particular choice of $\underline{\upsilon}_i$'s as follows with the resulting design a D-optimum one.

$$\underline{\upsilon}'_1 = (1\ 1\ 1 \ldots 1\ 1\ 1)$$

$$\underline{\upsilon}'_2 = (0\ 1\ 1 \ldots 1\ 1\ 0)$$

$$\underline{\upsilon}'_3 \cdot (0\ 0\ 1 \ldots 1\ 0\ 0)$$

$$\vdots$$

$$\underline{\upsilon}'_{\left[\frac{n+1}{2}\right]} = (0\ 0 \ldots 0\ 1\ 1\ 0\ 0 \ldots 0)$$

Remark 1 Also for $N = t.\frac{n(n+1)}{2} - \ell$, $1 \leq \ell \leq \left[\frac{n+1}{2}\right]$, it can be easily checked that given the choice for $t.\frac{n(n+1)}{2}$ vectors as $X^0 = \left[\left.\begin{matrix} X^* \\ \vdots \\ X^* \end{matrix}\right\} \text{t copies}\right]$, the design matrix X, obtained from X^0 by deleting the same set of vectors $\underline{v}_1', \ldots, \underline{v}_\ell'$ (as specified in the above theorem) would be D - optimal in the restricted set-up.

3c.3 A-optimal Designs

For A - optimality,

$$\varphi(M(\eta)) = -\operatorname{tr}\{M(\eta)\}^{-1} \qquad \text{if } M(\eta) \text{ is nonsingular}$$

$$= -\infty \qquad \text{if } M(\eta) \text{ is singular.}$$

With φ so defined, if $M(\eta)$ is nonsingular, it can be shown that

$$F_\varphi(M(\eta), \underline{x}\,\underline{x}') = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}\{-\operatorname{tr}[(1-\varepsilon)M(\eta) + \varepsilon\,\underline{x}\,\underline{x}']^{-1} + \operatorname{tr}[M(\eta)]^{-1}\}$$

$$= \underline{x}'[M(\eta)M(\eta)]^{-1}\underline{x} - \operatorname{tr}[(M(\eta))^{-1}] \qquad \ldots(3c.3.1)$$

after some simplification.

Defining S as before, let λ_j's be the eigenvalues of S and $\underline{\xi}_j$'s be the corresponding eigenvectors $1 \leq j \leq n$. Since for each $\underline{x} \in \mathfrak{X}$, $\underline{x}' S \underline{x} = 2$ (vide(3c.2.2)) it follows by (3c.3.1) that if η^* be a design measure such that

$$M(\eta^*) = 2\left(\sum_{j=1}^{n} \lambda_j^{1/2}\right)^{-1} \sum_{j=1}^{n} \lambda_j^{-1/2}\, \underline{\xi}_j \underline{\xi}_j' \qquad \ldots(3c.3.2)$$

then $F_\varphi(M(\eta^*), \underline{x}\,\underline{x}') = 0$ for each $\underline{x} \in \mathfrak{X}$.

Consequently by Theorem 3c.1.2 such an η^* gives an A - optimal design measure.

It remains to obtain the π_{uv}'s corresponding to the above η^*. For this we need the expressions for λ_j's and $\underline{\xi}_j$'s as defined above.

Lemma 3c.3.1 $0 < \lambda_j < 4$ for all $j = 1, \ldots, n$.

Proof : It can be easily checked that S and $4I - S$ are both p.d. matrices and hence, the result.

Now, to determine the eigenvalues, let us denote $|S - \lambda I|$ by $V_n(\lambda)$, then we get

$$V_n(\lambda) - (2-\lambda)V_{n-1}(\lambda) + V_{n-2}(\lambda) = 0 \quad \text{for all} \quad n \geq 2$$

along with
$$V_0(\lambda) = 1$$
$$V_1(\lambda) = 2 - \lambda$$

Thus
$$V_n(\lambda) = \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}$$

where $\alpha = \dfrac{(2-\lambda) + \sqrt{\lambda(\lambda-4)}}{2}$ and $\beta = \dfrac{(2-\lambda) - \sqrt{\lambda(\lambda-4)}}{2}$

so that for any eigenvalue λ of S, $\dfrac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta} = 0$.

It can be checked directly that

$$\underline{\xi} = \left(1, \frac{\alpha^2 - \beta^2}{\alpha - \beta}, \ldots, \frac{\alpha^n - \beta^n}{\alpha - \beta}\right)' \qquad \ldots(3c.3.3)$$

is an (real) eigenvector of S corresponding to the eigenvalue λ.

Let γ be such that

$$\cos\gamma = \frac{2-\lambda}{2} \quad \text{so that} \quad \sin\gamma = \frac{\sqrt{\lambda(4-\lambda)}}{2}.$$

Then it turns out that

$$\alpha = \cos\gamma + i\ \sin\gamma\ , \quad \text{and} \quad \beta = \cos\gamma - i\ \sin\gamma\ , \qquad \ldots(3c.3.4)$$

so that $\dfrac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta} = 0$ implies $\sin(n+1)\gamma = 0$

and, hence $\gamma = \dfrac{p\pi}{n+1}$, $p = 0, \pm 1, \pm 2, \ldots\ldots$ $\qquad \ldots(3c.3.5)$

Thus, $\quad \lambda = 2(1-\cos\gamma) = 2(1-\cos(\frac{p\pi}{n+1}))$

where p is such that $0 < \lambda < 4$, i.e. $-1 < 1-\frac{\lambda}{2} < 1$ i.e. $-1 < \cos(\frac{p\pi}{n+1}) < 1$. So one should take $p = 1,2,3,\ldots,n,$ to determine the eigenvalues of S. The eigenvalues are

$$\begin{aligned}\lambda_j &= 2\left\{1-\cos(\tfrac{j\pi}{n+1})\right\} \qquad j = 1,2,\ldots,n \\ &= 4\sin^2(\tfrac{j\pi}{2(n+1)}) \\ &= 4\sin^2(\tfrac{j\theta}{2}) \quad \text{say,}\end{aligned} \qquad \ldots(3c.3.6)$$

where $\theta = \dfrac{\pi}{n+1}$,

It can be checked that, (recalling (3c.3.4), (3c.3.5))

$$\begin{aligned}\sum_{p=1}^{n}(\alpha_j^p-\beta_j^p)^2 &= -4\sum_{p=1}^{n}\sin^2(p\gamma_j) \\ &= -(2n+1) + \frac{\sin(2n+1)\gamma_j}{\sin\gamma_j} \\ &= -2(n+1)\end{aligned}$$

after simplification, so that the normalized eigenvector corresponding to the eigenvalue λ_j is (recall (3c.3.3))

$$\underset{\sim}{\xi}_j = \left(\frac{2}{n+1}\right)^{1/2} (\sin j\theta, \sin 2j\theta, \ldots, \sin nj\theta), \quad 1 \leq j \leq n. \quad \ldots(3c.3.7)$$

Now, from (3c.3.2),

$$m_{jk}(\eta^*) = \frac{2}{\sum_{p=1}^{n} \sqrt{\lambda_p}} \sum_{p=1}^{n} \frac{1}{\sqrt{\lambda_p}} \frac{(\alpha_p^j - \beta_p^j)(\alpha_p^k - \beta_p^k)}{\sum_{s=1}^{n} (\alpha_p^s - \beta_p^s)^2}$$

$$= \frac{1}{(n+1) \sum_{p=1}^{n} \sin(\frac{p\theta}{2})} \sum_{p=1}^{n} \frac{\sin(jp\theta)\sin(kp\theta)}{\sin(\frac{p\theta}{2})}$$

using the expressions for λ_j's ((3c.3.6)) and $\underset{\sim}{\xi}_j$'s ((3c.3.7)) above.

Hence, by (3c.1.2) and using the identity,

$$\sum_{p=1}^{n} \sin(\frac{pr\theta}{2}) = \frac{1}{2}\left[\cot \frac{r\theta}{4} - (-1)^{\frac{r-1}{2}}\right] \quad \text{where } r \text{ is an odd integer}$$

the π_{uv}'s corresponding to η^* are given by, say,

$$\pi_{uv}^* = \frac{2}{(n+1)(\cot\frac{\theta}{4} - 1)} \sum_{p=1}^{n} 2\left\{\sin(pv\theta) - \sin(p(v+1)\theta)\right\} \cos(p(2u-1)\frac{\theta}{2})$$

$$= \frac{2}{(n+1)(\cot\frac{\theta}{4} - 1)} \sum_{p=1}^{n} \left\{\sin(p(2v+2u-1)\frac{\theta}{2}) + \sin(p(2v-2u+1)\frac{\theta}{2})\right.$$

$$\left. - \sin(p(2u+2v+1)\frac{\theta}{2}) - \sin(p(2v-2u+3)\frac{\theta}{2})\right\}$$

$$= \frac{1}{(n+1)(\cot\frac{\theta}{4} - 1)} \left\{\cot((v+u-\tfrac{1}{2})\tfrac{\theta}{2}) + \cot((v-u+\tfrac{1}{2})\tfrac{\theta}{2}) - \cot((v+u+\tfrac{1}{2})\tfrac{\theta}{2})\right.$$

$$\left. - \cot((v-u+\tfrac{3}{2})\tfrac{\theta}{2})\right\} \qquad 1 \leq u \leq v \leq n.$$

It can be checked that $\pi^*_{uv} > 0 \quad (1 \le u \le v \le n)$ and $\sum_{u=1}^{n} \sum_{v=u}^{n} \pi^*_{uv} = 1$ which guarantees the existence of η^* satisfying (3c.3.2).

For example, for $p = 4$, the π^*_{uv}'s turn out to be as follows :

$$\pi_1 = \pi_4 = .2, \quad \pi_{12} = \pi_{34} = .0546915, \quad \pi_{13} = \pi_{24} = .0316769,$$

$$\pi_2 = \pi_3 = .1769854, \quad \pi_{23} = .0481535, \quad \pi_4 = .0251389 .$$

This approximate A-optimal design measure is of comparatively more involved nature and does not reduce to any exact result as in the case of D-optimality, even for particular combinations of N and n. For reasonably large N, an approximate design is closely approximated by an exact design with N observations where each n_{uv}, the number of times the row vector $\underline{h}_{uv}$ occurs as a row in the design matrix is taken as an integer close to $N\pi^*_{uv}$ (Fedorov (1972, Chapter 3), Silvey (1980, p.3)). This is of particular relevance when the available resources permit a fairly large number of observations and the problem is to take these observations in an efficient manner. It is known that such exact designs which are in a sense close to best approximate designs have only slightly reduced performance.

3c.4 E - optimal Designs

For E - optimality,

$$\varphi(M(\eta)) = \inf_{\underline{y}'\underline{y} \neq 0} \frac{\underline{y}'(M(\eta))\underline{y}}{\underline{y}'\underline{y}}$$

and let η^* be a design measure such that $M(\eta^*) = n^{-1} I_n$.

Now it is to be noted that $\varphi(M(\eta))$ is not differentiable at $M(\eta^*)$, and therefore, the technique of Frechet derivatives cannot be employed in obtaining the E-optimal designs. However, the following arguments establish η^* as the E-optimal design measure.

For any $\underline{y} = (y_1, \ldots, y_n)'$ and $\eta \in H$, by (3c.1.1),

$$\underline{y}'M(\eta)\underline{y} = \sum_{u=1}^{n} \sum_{v=u}^{n} (y_u + y_{u+1} + \ldots + y_v)^2 \pi_{uv}$$

In particular, with $\underline{y} = (1, -1, 1, -1, \ldots, (-1)^{n-1})' = \underline{y}_o$ say, we observe that in no set of consecutive elements of $\underline{y}_o$, the corresponding sum can exceed unity in magnitude and hence the above yields

$$\underline{y}_o'M(\eta)\underline{y}_o \leq \sum_{u=1}^{n} \sum_{v=1}^{n} \pi_{uv} = 1$$

so that for each $\eta \in H$,

$$\varphi(M(\eta)) \leq \frac{\underline{y}_o'M(\eta)\underline{y}_o}{\underline{y}_o'\underline{y}_o} = n^{-1}.$$

Since clearly $\varphi(M(\eta^*)) = n^{-1}$, it follows that η^* gives the E-optimal design measure. As $M(\eta^*) = n^{-1}I$, by (3c.1.2), the π_{uv}'s corresponding to η^* are,

$$\pi^*_{uu} = n^{-1} \quad (1 \leq u \leq n), \quad \pi^*_{uv} = 0 \quad (1 \leq u < v \leq n).$$

Thus we have arrived at the following exact E-optimality result.

<u>Corollary 3c.4.1</u> Let $N = tn$, $t \geq 1$, say, then the design matrix X^* in which $\underline{h}_{uu}$ $(1 \leq u \leq n)$ occurs as a row t times is E-optimal design

within the class $\Omega(N,n)$ of all relevant designs.

Very recently Jacroux (1986) independently studied the E-optimality problem and derived general results for all values of (N,n). Some results of Jacroux and Notz (1983) are also to be found here.

An overall review of D-, A-, E-optimal design measures developed here in the unrestricted set-up reveals that the E-optimal measure puts the entire mass on $\underline{h}_{uu}$ $(1 \leq u \leq n)$ and no mass on $\underline{h}_{uv}$ $(1 \leq u < v \leq n)$, while the A-optimal design measure puts greater mass on those $\underline{h}_{uv}$'s with $(v-u)$ small and less mass on those $\underline{h}_{uv}$'s with $(v-u)$ large. The D-optimal measure is " flattened " in this sense and puts equal mass at each $\underline{h}_{uv}$.

3d. Further Inferential Aspects

Suppose we are interested in estimating parameters of the form $\theta(i,j) = \sum_{r=i}^{j} \beta_r$, $1 \leq i \leq j \leq n$, retaining estimability of the individual lengths. In other words, this means that our interest is to estimate the length(s) between various pairs of points. Of particular interest is the problem of estimating the total length (the case $i = 1$, $j = n$) (vide, Banerjee (1975), Sinha (1971, 1972), Panda (1976) for various aspects of this problem in the framework of spring balance weighing designs). The results of subsection 3a.1 are useful for settling this problem completely when $N = n$.

To start with, let us fix (i,j), $1 \leq i \leq j \leq n$. We write

$$\theta(i,j) = \underline{v}_{ij}'\underline{\theta}$$

where $\underset{\sim}{\upsilon}_{ij} = (0,0,\ldots,0,1,1,\ldots,1,0,\ldots,0)'$ with 1's in the positions i through j. (Note that $\underset{\sim}{\upsilon}_{ij}$ is well-defined for every (i,j)).
We have

$$V(\theta(i,j)) = \sigma^2 \underset{\sim}{\upsilon}'_{ij}(X'X)^{-1}\underset{\sim}{\upsilon}_{ij} \geq \sigma^2$$

in view of (ii) of Property 2 of X^{-1} (vide subsection 3a.1).
Here "=" holds iff $\underset{\sim}{\upsilon}'_{ij} = \underset{\sim}{\upsilon}'_{hh}X$ for some combination of X and h. It is easy to verify that given any (i,j) we can have a choice of X for every h. (It is enough to choose the h^{th} row of X as $\underset{\sim}{\upsilon}'_{ij}$). Such a nonsingular X is then optimal. Clearly, this is true for any (i,j). Likewise, minimum possible variance (σ^2) is attainable for each component of a simultaneous inference problem on $\{\theta(i_1,j_1), \theta(i_2,j_2),\ldots,\theta(i_k,j_k)\}$ provided they are linearly independent (since we are dealing with the case where individual estimability of the β_i's is to be ensured using just n measuring operations). These results are highly interesting and are peculiar too to this set up ! (It may be recalled that the best unbiased spring balance weighing design for estimating the total weight of a set of n (≥ 3) objects in exactly n weighing operations is provided by $X_{oo} = \upsilon_{1n}\,\upsilon'_{1n} - I_n$ and the minimum variance is given by $\dfrac{n\sigma^2}{(n-1)^2}$ (Sinha, (1971, 1972)). This quantity is smaller than σ^2 and X_{oo} does not enjoy the string property. We refer to Sinha (1971, 1972), Panda (1976) and Swamy (1980), for further aspects of these problems under a restricted set-up. Analogous results are derivable in the present framework as well).

Concluding Remarks

We have investigated in this chapter a specific problem in weighing designs framed in the language of measurement of consecutive distances of a set of objects placed along a rectilinear line segment. The problem is supposed to have relevance in some concrete physical situations as explained in the introduction. Exact optimal designs for all values of $N > n$ could not be derived. However, it is strongly felt that the results discussed here might be useful in exploring further researches along this direction.

CHAPTER 4

REPEATED MEASUREMENTS DESIGNS

Introduction

So far, in the preceding chapters, we dealt with optimality studies for certain classes of designs assuming fixed effects models. In the present chapter we take up a study of optimality under mixed effects model in the context of Repeated Measurements Designs (RMD). These designs have been discussed in the literature under various names, viz. cross-over designs, change-over designs, time series designs, before-after designs.

Such RM experiments (vide DFN 1d.2.11, Chapter 1) are peculiar in that any treatment applied to a unit in a certain period influences the response of the unit not only in the current period but also leaves residual effects in the following periods. In practice, only the first order residual effect (carry-over effect) i.e., residual effect of any treatment upto just the next period is of importance. For a general review of such designs, including practical applications, reference is made to Hedayat and Afsarinejad (1975). The pioneering work in the area of optimal RMD's is due to Hedayat and Afsarinejad (1978) and further significant contributions have been made by Cheng and Wu (1980, 1983), Magda (1980) and Kunert (1983, 1984 a, 1984 b). All these authors considered the problem of characterization and construction of universally optimal designs under fixed effects additive linear models, incorporating effects due to units, periods, and direct and first order residual effects of treatments. But in practice, it is not uncommon to face situations where the experimental units included in the experiment constitute a random sample from a population of a large number of available experimental units. In such a case, a model incorporating random effects due to units has to be sought out. In the present chapter we assume an additive mixed effects model where period effects and direct and first order residual effects of treatments are retained as fixed while unit effects are taken to be random, and attempts are made to extend the optimality results already established in the context of completely fixed effects model to the above mixed effects model. We will constantly refer to Cheng and Wu (1980) and also sometimes to Hedayat and Afsarinejad (1978) - hereafter abbreviated as C & W (1980), and H & A (1978) respectively.

4a. Description of the Model, Definitions and Notations

An experiment based on t treatments, n experimental units and p periods, each unit being given one treatment during each period is abbreviated by RMD (t,n,p). Let $d(i,j)$ be the treatment assigned by a design d in the i^{th} period to the j^{th} experimental unit. The linear model assumed for the response obtained under the design d is

Model 1 :

$$y_{ij} = \mu + \alpha_i + \beta_j + \tau_{d(i,j)} + \rho_{d(i-1,j)} + e_{ij} \qquad \ldots(4a.1.1)$$

$$i = 1,2,\ldots,p$$

$$j = 1,2,\ldots,n.$$

$\underline{\beta} \sim N(0,\sigma_1^2 I)$, $\underline{e} \sim N(0,\sigma^2 I)$

$\underline{\beta}$ and $\underline{e}$ are uncorrelated.

Here μ, α_i, β_j, $\tau_{d(i,j)}$ and $\rho_{d(i-1,j)}$ represent respectively the general effect, the effect of the i^{th} period, the (random) effect of the j^{th} experimental unit, the (direct) effect of treatment $d(i,j)$ and the first order residual effect of treatment $d(i-1,j)$.

We distinguish between (a) designs with no residual effects on the first period i.e., with $\rho_{d(o,j)} = 0$ for $1 \leq j \leq n$ (vide H& A (1978), C & W (1980, 1983), Kunert (1983, 1984a)), and (b) designs with residual effects on the first period i.e., with a preperiod or conditioning treatment which means $\rho_{d(o,j)} \neq 0$ for $1 \leq j \leq n$ (vide Sampford (1957), Sinha (1975), Sonnemann (1982), Magda (1980), Kunert (1983, 1984b)). Usually the residual effects in the first period are derived using those in the last period as the preperiod or conditioning treatments (vide Sampford (1957), Magda (1980), Kunert (1984b) etc.). We call the model (4a.1.1) for designs with <u>no</u> residual effects in the first period i.e., the model with

$$\rho_{d(o,j)} = 0 \quad \text{for} \quad 1 \leq j \leq n \qquad \ldots(4a.1.2)$$

a <u>non-circular model</u>. Let $\Omega_{t,n,p}$ denote the collection of all such designs which are connected in the sense that all contrasts belonging to direct and residual treatment effects are estimable. Following Kunert (1983), the set of all connected RMD's <u>with preperiod</u> is denoted by $\tilde{\Omega}_{t,n,p}$. In particular, when

$$\rho_{d(o,j)} = \rho_{d(p,j)} \quad \forall\, j, \qquad \ldots(4a.1.3)$$

following Magda (1980) and Kunert (1984b), the model is termed as a <u>circular model</u>.

In vector notations, for a design $d \in \Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$, the model (4a.1.1) can be rewritten for np observations arranging serially unitwise as

$$\underline{Y} = \mu\underline{1} + P\underline{\alpha} + U\underline{\beta} + T_d\underline{\tau} + F_d\underline{\rho} + \underline{\varepsilon} \qquad \ldots(4a.1.4)$$

$E(\underline{\varepsilon}) = 0$, $D(\underline{\varepsilon}) = \sigma^2 I$, $\sigma^2 > 0$ (unknown)

$E(\underline{\beta}) = 0$, $D(\underline{\beta}) = \sigma_1^2 I$, $\sigma_1^2 > 0$ (unknown)

$\underline{\varepsilon}$ and $\underline{\beta}$ uncorrelated.

Let Σ denote the variance-covariance matrix of $\underline{Y}$. Then

$$\Sigma = \mathrm{Diag}(\sigma^2 I_p + \sigma_1^2 J_p, \ldots, \sigma^2 I_p + \sigma_1^2 J_p).$$

with $$\Sigma^{-1} = \omega I_{np} - \frac{\omega - \tilde{\omega}}{p}\,\mathrm{Diag}(J_p, J_p, \ldots, J_p) \qquad \ldots(4a.1.5)$$

where $\omega = \dfrac{1}{\sigma^2}$, $\tilde{\omega} = \dfrac{1}{\sigma^2 + \sigma_1^2 p}$

In the derivation of optimality results, besides the above model (4a.1.4), we sometimes make use of the following models as and when necessary.

<u>Model 2</u>

(i) $y_{ij} = \mu + \tau_{d(i,j)} + \beta_j + e_{ij}$

(assuming no differential period effects and no differential residual treatment effects)

(ii) $y_{ij} = \mu + \rho_{d(i-1,j)} + \beta_j + e_{ij}$ ..(4a.1.6)

(assuming no differential period effects and no differential direct treatment effects.)

contd...

<u>Model 3</u>

$$\text{(i)} \quad y_{ij} = \mu + \alpha_i + \tau_{d(i,j)} + \beta_j + e_{ij}$$

(assuming no differential residual treatment effects) ...(4a.1.6)

$$\text{(ii)} \quad y_{ij} = \mu + \alpha_i + \rho_{d(i-1,j)} + \beta_j + e_{ij}$$

(assuming no differential direct treatment effects)

<u>Model 4</u>

$$y_{ij} = \mu + \tau_{d(i,j)} + \rho_{d(i-1,j)} + \beta_j + e_{ij}$$

(assuming no differential period effects.)

It may be noted that in all the abovementioned models, unit effects are taken to be random. We use the symbol $C_d^{(i)}(\underline{\tau} \mid M)$ (respectively $C_d^{(i)}(\underline{\rho} \mid M)$) to denote the C_d-matrix of direct (respectively residual) treatment effects under model i, i = 1,2,3,4. The letter 'M' reflects the fact that we are working under a mixed effects model. Model 1 (vide (4a.1.1)) is most general and, for the most part, we deal with it. To avoid unnecessary complications in notation, we omit the upper suffix (1) for the corresponding C-matrices under Model 1.

Whatever the model adopted, the joint information matrix of direct and residual treatment effects has a general representation as

$$C_d(\underline{\tau}, \underline{\rho} \mid M) = \begin{bmatrix} C_{d11}(M) & C_{d12}(M) \\ C_{d21}(M) & C_{d22}(M) \end{bmatrix} \qquad \text{...(4a.1.7)}$$

and for direct effects, $C_d(\underset{\sim}{\tau} \mid M)$ assumes the form

$$C_d(\underset{\sim}{\tau} \mid M) = C_{d11}(M) - C_{d12}(M)C_{d22}^{-}(M)C_{d21}(M)$$

while for residual effects, we get

$$C_d(\underset{\sim}{\rho} \mid M) = C_{d22}(M) - C_{d21}(M)C_{d11}^{-}(M)C_{d12}(M)$$

with $C_{dij}(M)$'s appropriately derived under the model assumed.

Obviously, $C_d(\underset{\sim}{\tau} \mid M) = C_{d11}(M)$ whenever the model does not contain the term due to residual effects or whenever for a design d, $C_{d12}(M) = 0$ (a null matrix). A similar observation applies to the case $C_d(\underset{\sim}{\rho} \mid M) = C_{d22}(M)$.

For the design $d \varepsilon \Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$, we adopt the following notations from C & W (1980) and Kunert (1983).

n_{diu} = number of appearances of treatment i on unit u in the periods 1 to p,

$\tilde{n}_{diu}$ = number of appearances of treatment i on unit u in the periods 0 to $p-1$,

ℓ_{dik} = number of appearances of treatment i in period k over the units 1 to n,

$\tilde{\ell}_{dik}$ = number of appearances of treatment i in period $k-1$ over the units 1 to n,

m_{dij} = number of appearances of treatment i preceded by treatment j on the same unit and summed over all units,

r_{di} = number of appearances of treatment i in the periods 1 to p over the units 1 to n,

$\tilde{r}_{di}$ = number of appearances of treatment i in the periods 0 to $p-1$ over the units 1 to n,

s_{di} = number of appearances of treatment i in the periods 2 to p over the units 1 to n,

where $1 \leq u \leq n$, $1 \leq k \leq p$, $1 \leq i, j \leq t$.

Then the following relations are immediate.

$$r_{di} = \sum_{u=1}^{n} n_{diu} = \sum_{k=1}^{p} \ell_{dik}$$

$$\tilde{r}_{di} = \sum_{u=1}^{n} \tilde{n}_{diu} = \sum_{k=1}^{p} \tilde{\ell}_{dik} = \sum_{j=1}^{t} m_{dji} \qquad \ldots(4a.1.8)$$

$$s_{di} = \sum_{k=2}^{p} \ell_{dik} = \sum_{j=1}^{t} m_{dij} \text{ (for } d \in \Omega_{t,n,p}) \qquad \ldots(4a.1.9)$$

$$\sum_{i=1}^{t} n_{diu} = p, \quad \sum_{i=1}^{t} \tilde{n}_{diu} = p-1 \quad (\text{for } d \in \Omega_{t,n,p})$$

$$\sum_{i=1}^{t} \ell_{dik} = n, \quad \sum_{i=1}^{t} r_{di} = np, \quad \sum_{i=1}^{t} \tilde{r}_{di} = n(p-1) \quad (\text{for } d \in \Omega_{t,n,p})$$

Let $\underline{r}_d = (r_{d1},\ldots,r_{dt})$ and $\tilde{\underline{r}}_d = (\tilde{r}_{d1},\ldots,\tilde{r}_{dt})$.

Further, the following matrices are defined with the matrix elements suitably shown using the earlier notations. (Recall (4a.1.4) in this context).

$$D_d = T_d'T_d = \operatorname{Diag}(r_{d1},\ldots,r_{dt})$$

$$\tilde{D}_d = F_d'F_d = \operatorname{Diag}(\tilde{r}_{d1},\ldots,\tilde{r}_{dt})$$

$$M_d = T_d'F_d = ((m_{dij})), \quad 1 \leq i, j \leq t$$

$$N_{dp} = T_d'P = ((\ell_{dik})), \quad 1 \leq i \leq t, \quad 1 \leq k \leq p$$

$$\widetilde{N}_{dp} = F_d'P = ((\widetilde{\ell}_{dik})), \quad 1 \leq i \leq t, \quad 1 \leq k \leq p$$

$$N_{du} = T_d'U = ((n_{diu})), \quad 1 \leq i \leq t, \quad 1 \leq u \leq n$$

and $\widetilde{N}_{du} = F_d'U = ((\widetilde{n}_{diu})), \quad 1 \leq i \leq t, \quad 1 \leq u \leq n.$

The definitions given below are to be found in the references cited, but are presented here for the sake of completeness.

DFN 4a.1.1. A design $d \in \Omega_{t,n,p} \cup \widetilde{\Omega}_{t,n,p}$ is called uniform on periods if each treatment occurs the same number of times, say, λ_1 times in each period. A necessary condition for this to hold is $n = \lambda_1 t$.

DFN 4a.1.2. A design $d \in \Omega_{t,n,p} \cup \widetilde{\Omega}_{t,n,p}$ is called uniform on units if each treatment is assigned the same number of times, say, λ_2 times to each unit. This can occur only if $p = \lambda_2 t$.

DFN 4a.1.3. A design $d \in \Omega_{t,n,p} \cup \widetilde{\Omega}_{t,n,p}$ is called uniform if it is uniform on both periods and units. A uniform RMD $(t,\lambda_1 t,\lambda_2 t)$ is denoted by the symbol URMD $(t, \lambda_1 t, \lambda_2 t)$.

In the following, for $d \in \widetilde{\Omega}_{t,n,p}$ we take the last period as preperiod.

DFN 4a.1.4. A design $d \in \Omega_{t,n,p}$ $(\widetilde{\Omega}_{t,n,p})$ is called balanced if the ordered pairs $(d(i,j), d(i-1,j))$ for $i = 2,3,\ldots,p$ in a non-circular model (respectively for $i = 1,2,\ldots,p$ in a circular model) contain each ordered distinct pair of treatments equal number of times, say, λ times.

This means that $m_{dii} = 0 \quad 1 \leq i \leq t$, and $m_{dij} = \lambda$ for $i \neq j$, $1 \leq i, j \leq t$.

For a non-circular model a necessary condition for this to hold is $n(p-1) = \lambda t(t-1)$. For a circular model the corresponding condition is $np = \lambda t(t-1)$.

DFN 4a.1.5. A design $d \in \Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$ is called strongly balanced if the ordered pairs $(d(i,j), d(i-1,j))$ as given in DFN 4a.1.4 contain each ordered pair (including pairs of identical treatments) equal number of times, say, λ times. This means $m_{dij} = \lambda$ for $1 \leq i, j \leq t$.

For a non-circular model, a necessary condition for this to hold is $n(p-1) = \lambda t^2$. For a circular model, the corresponding condition is $np = \lambda t^2$.

DFN 4a.1.6. A design $d \in \Omega_{t,n,p}$ is called nearly strongly balanced if

(i) $M_d M_d'$ is completely symmetric (c.s.),

(ii) m_{dij} assumes one of the two values $\left[\frac{n(p-1)}{t^2}\right]$ and $\left[\frac{n(p-1)}{t^2}\right] + 1$, for all $1 \leq i, j \leq t$.

We recall here definitions of Generalised Youden Square Design (GYD) and Generalised Latin Square (GLS) design as given in DFN 1d.2.9 (vide Chapter 1.).

It can be checked from (4a.1.8) and (4a.1.9), that uniformity on units and equal m_{dij} $\forall$ i,j for any design $d \in \Omega_{t,n,p}$ imply uniformity in the first and last period. The same result is true even when $m_{dii} = 0$ $\forall$ i and all m_{dij}'s $i \neq j$ are equal.

In the subsequent sections we characterize and construct universally optimal designs under the non-circular model (section 4b), the circular model (section 4d) and also over the class $\Omega_{t,n,p} \cup \widetilde{\Omega}_{t,n,p}$ (section 4c). Except for a few new cases corresponding to $p = \lambda_2 t+1$, most of the optimality results established here are known to be valid under fixed effects model and the appropriate references are indicated within brackets in proper places. We refer to the theorems and lemmas of Hedayat and Afsarinejad (1978), Cheng and Wu (1980), Magda (1980), Kunert (1983, 1984 a , 1984 b) by appropriately prefixing the corresponding results in the respective papers by H & A, C & W, M, K(1), K(2), K(3) in this order. The technique for proving the results under mixed effects model does not differ substantially in many situations from those under fixed effects model and as such the detailed proofs are, for the most part, omitted and only the important modifications and derivations are highlighted.

4b. Non-Circular Model

4b.1 Preliminaries :

The model taken up in this section is as given in (4a.1.1) with $\rho_{d(o,j)} = 0 \ \forall\ j$. Under this model, the $C_{dij}(M)$'s as defined in (4a.1.7) turn out as follows (vide Bose 1975) (except for a multiplier ω which may be ignored allthrough the optimality considerations) :

$$\left.\begin{aligned}
C_{d11}(M) &= D_d - n^{-1}N_{dp}N'_{dp} - \frac{\omega-\tilde{\omega}}{\omega}\left\{p^{-1}N_{du}N'_{du} - n^{-1}p^{-1}N_{du}JN'_{du}\right\} \\
C_{d12}(M) &= C'_{d21}(M) = M_d - n^{-1}N_{dp}\tilde{N}'_{dp} - \frac{\omega-\tilde{\omega}}{\omega}\left\{p^{-1}N_{du}\tilde{N}'_{du} - n^{-1}p^{-1}N_{du}J\tilde{N}'_{du}\right\} \\
C_{d22}(M) &= \tilde{D}_d - n^{-1}\tilde{N}_{dp}\tilde{N}'_{dp} - \frac{\omega-\tilde{\omega}}{\omega}\left\{p^{-1}\tilde{N}_{du}\tilde{N}'_{du} - n^{-1}p^{-1}\tilde{N}_{du}J\tilde{N}'_{du}\right\}
\end{aligned}\right\} \quad \ldots(4b.1.1)$$

Thus the information matrices $C_d(\underset{\sim}{\tau}|M)$ and $C_d(\underset{\sim}{\rho}|M)$ for direct and residual effects respectively are to be computed using the formulae

$$C_d(\underset{\sim}{\tau}|M) = C_{d11}(M) - C_{d12}(M)C_{d22}^{-}(M)C_{d21}(M) \qquad \ldots(4b.1.2)$$

$$C_d(\underset{\sim}{\rho}|M) = C_{d22}(M) - C_{d21}(M)C_{d11}^{-}(M)C_{d12}(M) \qquad \ldots(4b.1.3)$$

Clearly, $C_d(\underset{\sim}{\tau}|M)$ and $C_d(\underset{\sim}{\rho}|M)$ are invariant under any choice of g-inverses of $C_{d11}(M)$ and $C_{d22}(M)$. Let us denote by $C_d(\underset{\sim}{\tau}|F)$, $C_d(\underset{\sim}{\rho}|F)$, $C_{dij}(F)$, $1 \leq i, j \leq 2$ the corresponding matrices under the fixed effects model. From Cheng and Wu (1980), we write

$$\left.\begin{aligned} C_{d11}(F) &= D_d - n^{-1}N_{dp}N_{dp}' - p^{-1}N_{du}N_{du}' + n^{-1}p^{-1}N_{du}JN_{du}' \\ C_{d12}(F) &= M_d - n^{-1}N_{dp}\widetilde{N}_{dp}' - p^{-1}N_{du}\widetilde{N}_{du}' + n^{-1}p^{-1}N_{du}J\widetilde{N}_{du}' = C_{d21}'(F) \\ C_{d22}(F) &= \widetilde{D}_d - n^{-1}\widetilde{N}_{dp}\widetilde{N}_{dp}' - p^{-1}\widetilde{N}_{du}\widetilde{N}_{du}' + n^{-1}p^{-1}\widetilde{N}_{du}J\widetilde{N}_{du}' \end{aligned}\right\} \ldots(4b.1.1)'$$

Note that the expression for $C_{dij}(F)$ follows from the expression for $C_{dij}(M)$ by replacing the scalar factor $\frac{\omega - \widehat{\omega}}{\omega}$ in (4b.1.1) by unity, $1 \leq i, j \leq 2$. Moreover, Lemmas 2.1, 2.2, 2.4 and 2.5 in C & W (1980) have the analogues valid for the above mixed effects model.

We now present results on universally optimal designs. The proofs of all the relevant theorems follow using the fundamental tool, viz., Proposition 1 of Kiefer (1975) (vide Chapter 1) and thus rest on verification of (i) complete symmetry and (ii) trace maximization of the relevant C-matrices over the competing class.

4b.2 Optimality of Strongly Balanced RMD's, $p = \lambda_2 t$

In this subsection we re-establish the results on universal optimality of strongly balanced RMD's when the number of periods (p) is a multiple of the number of treatments (t) and when the model is one of mixed effects. Recall the definition of strongly balanced RMD as given in DFN 4a.1.5. We omit the proofs altogether which follow essentially along the techniques employed in C & W (1980) or, otherwise, are to be found in Mukhopadhyay and Saha (1983).

Theorem 4b.2.1 (C & W Th.3.1) Let d^* be a strongly balanced uniform design in $\Omega_{t,\lambda_1 t,\lambda_2 t}$. Then d^* is universally optimal for the estimation of direct as well as residual effects over $\Omega_{t,\lambda_1 t,\lambda_2 t}$.

Theorem 4b.2.2. (C & W Th.3.3) Let $n = \lambda_1 t$, $p = \lambda_2 t+1$, $\lambda_1, \lambda_2 \geq 1$ and let d^* be a strongly balanced RMD (t,n,p) which is uniform on the units in the first $p-1$ ($= \lambda_2 t$) periods. Then d^* is universally optimal for the estimation of direct as well as residual effects over $\Omega_{t,n,p}$.

Following C & W (1980), we also have

Corollary 4b.2.2. (C & W Cor. 3.3.1) Let d^* be obtained by repeating the treatments in the last period of a balanced uniform RMD $(t,\lambda_1 t,t)$ in the following period. Then d^* is universally optimal for the direct as well as residual effects over $\Omega_{t,\lambda_1 t,t+1}$.

If we restrict the competing designs to a subclass, then the same stronger optimality results as in C & W (1980) can be re-established. Let

$\Omega^*_{t,n,p} = \{d \in \Omega_{t,n,p} : r_{d1} = r_{d2} = \ldots = r_{dt}\}$. Then we have

Theorem 4b.2.3. (C & W Th. 3.4) Let d^* be a strongly balanced uniform RMD (t,n,p). Then d^* minimizes the variance of the best linear unbiased estimator of any contrast among direct effects $\{\tau_i\}$, $i = 1,2,\ldots,t$, over the class $\Omega^*_{t,n,p}$.

Further, we also have

Theorem 4b.2.4. (C & W Th. 3.5) Let d^* be an RMD (t,n,p) satisfying the conditions in Theorem 4b.2.2. Then d^* minimizes the variance of the best linear unbiased estimator of any contrast among the residual effects $\{\rho_i\}$, $i = 1,2,\ldots,t$, over $\Omega^{**}_{t,n,p}$ where
$\Omega^{**}_{t,n,p} = \{d \in \Omega_{t,n,p} : \tilde{r}_{d1} = \tilde{r}_{d2} = \ldots = \tilde{r}_{dt}\}$.

4b.3. Optimality of Balanced RMD's, $p = \lambda_2 t$

We confine to the class $\Lambda_{t,n,p}$ of designs d in $\Omega_{t,n,p}$ for which $m_{dii} = 0$, $1 \leq i \leq t$ i.e., no treatment is allowed to be preceded by itself in d.

The first theorem here is a mixed effect analogue of the result given in H & A (1978) under the more restrictive assumption that the competing designs are uniform on both units and periods. It refers to the special case of $\lambda_2 = 1$ i.e., $p = t$.

Theorem 4b.3.1. (H & A Th. 3.1, Th. 3.2) Let d^* be a balanced uniform design in $\Lambda_{t,\lambda_1 t,t}$. Then d^* is universally optimal for direct as well as residual effects over the class of uniform designs in $\Lambda_{t,\lambda_1 t,t}$.

<u>Proof</u>. Since the competing designs are uniform, $C_d(\underset{\sim}{\tau}|M)$ and $C_d(\underset{\sim}{\rho}|M)$ assume the following forms (after simplification) :

$$C_d(\underset{\sim}{\tau}|M) = \lambda_1 t(I - \frac{J}{t}) - \left\{\lambda_1(t-1) - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{t}\right\}^{-1}\left\{M_d M_d' - \frac{\lambda_1^2(t-1)^2}{t} J\right\}$$

$$C_d(\underset{\sim}{\rho}|M) = \left\{\lambda_1(t-1) - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{t}\right\}(I - \frac{J}{t}) - \frac{1}{\lambda_1 t}\left\{M_d M_d' - \frac{\lambda_1^2(t-1)^2}{t} J\right\}.$$

The rest of the arguments is the same as in the proof of Theorem 3.1 in H & A (1978). We omit the details.

We now turn to the general class of RMD $(t,\lambda_1 t,\lambda_2 t)$'s with $\lambda_2 \geq 2$. For a balanced uniform d^* in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$, as C & W (1980) noted,

$$\left.\begin{aligned} & n_{d^*iu} = \lambda_2,\ \ell_{d^*ik} = \lambda_1,\ m_{d^*ii} = 0 \\ & m_{d^*ij} = \frac{\lambda_1(p-1)}{(t-1)} = \lambda \text{ (say) for all } i,u,k, \text{ and } i \neq j \end{aligned}\right\} \quad \ldots(4b.3.1)$$

Consequently, we get

$$\left.\begin{aligned} & C_{d^*11}(M) = \lambda_1 p(I - \frac{J}{t}) \\ & C_{d^*12}(M) = C_{d^*21}'(M) = -\lambda(I - \frac{J}{t}) \\ & C_{d^*22}(M) = \lambda_1(p-1-\frac{\omega-\tilde{\omega}}{\omega p})(I - \frac{J}{t}) \end{aligned}\right\} \quad \ldots(4b.3.2)$$

In the above, the expressions for $C_{d^*11}(M)$ and $C_{d^*12}(M)$ are the same as those of $C_{d^*11}(F)$ and $C_{d^*12}(F)$ respectively (vide C & W (1980) — expression (4.2)) as the matrix associated with $\frac{\omega-\tilde{\omega}}{\omega}$ in each of $C_{d^*11}(M)$ and $C_{d^*12}(M)$ (vide (4b.1.1.)) vanishes in the present set-up.

The following theorems re-establish the optimality results known for the fixed effects case.

Theorem 4b.3.2. (C & W Th.4.1) Let d^* be a balanced uniform design in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$, $t \geq 3$. Then d^* is universally optimal for the estimation of residual effects over the class of designs d in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$ with $\tilde{r}_{di} = \lambda_1(p-1)$ for all i i.e., among all designs which are equireplicate in the first (p-1) periods. (The fixed effects analogue also assumes this equireplicability condition as a technicality in the proof).

Theorem 4b.3.3. (C & W Th. 4.2) Let d^* be a balanced uniform design in $\Lambda_{t,\lambda_1 t,t}$. Then d^* is universally optimal for the estimation of residual effects over $\Lambda_{t,\lambda_1 t,t}$. (This refers to the case of $\lambda_2 = 1$ and the equireplicability condition stated above has been relaxed here. This also partially strengthens the result stated in Theorem 4b.3.1 as for residual effects the competing designs in $\Lambda_{t,\lambda_1 t,t}$ need not be uniform. This same phenomenon has already been observed in C & W (1980) with respect to the fixed effects model).

There is no substantial difference in the technique of proof of these theorems from those in Theorem 4.1 and Theorem 4.2 of C & W (1980) and as such we omit them.

Theorem 4b.3.4. (C & W Th. 4.3) Let d^* be a balanced uniform design in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$. Then d^* is universally optimal for the estimation of direct effects over the class of designs in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$ which are uniform on each unit and the last period.

[It may be mentioned that the original proof of the fixed effects analogue of this result (C & W Th. 4.3 (1980)) suffered from a mistake and it was subsequently rectified in C & W (1983). We propose to provide a sketch of the proof of our theorem].

Proof. For designs that are uniform on units and in the last period, it can be shown that (vide (4b.1.1))

$$C_{d11}(M) = \lambda_1 pI - \frac{1}{n} N_{dp}N_{dp}'$$

$$C_{d12}(M) = M_d - \frac{1}{n} N_{dp}\tilde{N}_{dp}'$$

(Note that in each of these matrices the coefficient of $\frac{\omega-\tilde{\omega}}{\omega}$ vanishes)

and $C_{d22}(M) = \left\{\lambda_1(p-1) - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p}\right\} I + \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{pt} J - \frac{1}{n}\tilde{N}_{dp}\tilde{N}_{dp}'$.

Next observe that $C_{d^*22}(M) = \left\{\lambda_1(p-1) - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p}\right\}(I - \frac{J}{t})$

so that $C_{d^*22}(M) - C_{d22}(M) = \frac{1}{n}\tilde{N}_{dp}\tilde{N}_{dp}' - \frac{\lambda_1(p-1)}{t} J$

$$= \frac{1}{n}(N_{dp}N_{dp}' - \lambda_1^2 pJ) = \frac{1}{n} N_{dp}(I - \frac{J}{p})N_{dp}' \geq 0$$

and hence $C_{d22}(M) \leq C_{d^*22}(M)$.

As we are concerned about the estimation of direct effects, the quantity to be maximized is $\operatorname{tr}\left\{C_{d11}(M) - C_{d12}(M)C_{d22}^{-}(M)C_{d21}(M)\right\}$ and this can be replaced by $\operatorname{tr}\left\{C_{d11}(M) - C_{d12}(M)C_{d^*22}^{-}(M)C_{d21}(M)\right\}$ which simplifies to

$$\lambda_1 pt - \frac{1}{n}\sum_{i=1}^{t}\sum_{k=1}^{p}\ell_{dik}^2 - \frac{1}{\lambda_1}\left(p-1-\frac{\omega-\tilde{\omega}}{\omega p}\right)^{-1}\sum_{i,j=1}^{t}\sum\left(m_{dij} - \frac{1}{n}\sum_{k=2}^{p}\ell_{dik}\ell_{dj,k-1}\right)^2.$$

The constraints are $\sum_{i=1}^{t}\ell_{dik} = n$, $m_{dii} = 0$, $\sum_{i=1}^{t} r_{di} = np$,

$\sum_{j=1}^{t} (m_{dij} - q_{dij}) = \sum_{i=1}^{t} (m_{dij} - q_{dij}) = 0$ (corresponding to the fact that the row sums and column sums of $C_{d12}(M)$ are zero) where

$$q_{dij} = \frac{1}{n} \sum_{k=2}^{p} \lambda_{dik} \lambda_{dj,k-1} + \frac{\omega - \tilde{\omega}}{\omega} \left\{ \frac{1}{p} \sum_{u=1}^{n} n_{diu} \tilde{n}_{dju} - \frac{1}{np} r_{di} \tilde{r}_{dj} \right\},$$

$$= \frac{1}{n} \sum_{k=2}^{p} \lambda_{dik} \lambda_{dj,k-1} \quad \text{as} \quad n_{diu} = \lambda_2,\ r_{di} = n\lambda_2,\ 1 \leq i \leq t,\ 1 \leq u \leq n$$

and, further, $\lambda_{dip} = \lambda_1$, $\sum_{k=1}^{p-1} \lambda_{dik} = \lambda_1(p-1)$ (as the competing designs are uniform on the last period and on each unit).

The maximization is now done following essentially the steps as in the proof of Theorem 4.1 in C & W (1980). At the final stage, it is enough to show that d^* minimizes $\left(\sum_{i=1}^{t} \sum_{k=2}^{p} \lambda_{dik} \lambda_{di,k-1}\right)$ as also $\sum_{i=1}^{t} \sum_{k=1}^{p} \lambda_{dik}^2$ subject to the constraints $\sum_{i=1}^{t} \lambda_{dik} = n = \lambda_1 t$, $\lambda_{dip} = \lambda_1$, $\sum_{k=1}^{p-1} \lambda_{dik} = \lambda_1(p-1)$, $1 \leq k \leq p$, $1 \leq i \leq t$. The second claim follows readily. As to the first claim, one needs the following result.

<u>Lemma 4b.3.1.</u> The minimum of $F(x) = \sum_{i=1}^{t} \sum_{k=2}^{p} x_{ik} x_{i,k-1}$ subject to $x_{ik} \geq 0$, $\sum_{i=1}^{t} x_{ik} = \lambda_1 t$, $x_{ip} = \lambda_1$, $\sum_{k=1}^{p-1} x_{ik} = \lambda_1(p-1)$, is achieved by taking $x_{ik} = \lambda_1$ for all i,k.

[It may be noted that this lemma is <u>not</u> explicitly mentioned in C & W (1980) though in the corrected proof of Theorem 4.3 in C & W (1983), this result has to be invoked]

<u>Proof of Lemma 4b.3.1.</u> It is obvious that the minimum must be attained at some bounded x_{ik} values and hence, it must be a local minimum.

We may rewrite $F(x)$ (using the constraints) as

$$F(x) = \sum_{i=1}^{t-1}\sum_{k=2}^{p-2} x_{ik}x_{i,k-1} + \sum_{k=2}^{p-2} (\lambda_1 t - \sum_{i=1}^{t-1} x_{ik})(\lambda_1 t - \sum_{i=1}^{t-1} x_{i,k-1})$$

$$+ \sum_{i=1}^{t-1} x_{i,p-1}(\lambda_1(p-1) - \sum_{k=1}^{p-2} x_{ik}) + \lambda_1 \sum_{i=1}^{t-1} (\lambda_1(p-1) - \sum_{k=1}^{p-2} x_{ik})$$

$$+ \left\{ \lambda_1 t - \sum_{i=1}^{t-1} (\lambda_1(p-1) - \sum_{k=1}^{p-2} x_{ik})\right\} (\lambda_1 t - \sum_{i=1}^{t-1} x_{i,p-2})$$

$$+ \lambda_1 \left\{ \lambda_1 t - \sum_{i=1}^{t-1} (\lambda_1(p-1) - \sum_{k=1}^{p-2} x_{ik})\right\}$$

A necessary condition for the existence of a local minimum is

$$\frac{\partial F(x)}{\partial x_{ik}} = 0 \quad \text{for} \quad 1 \leq i \leq t-1,\ 1 \leq k \leq p-2.$$

These lead to

$$x_{i2} - x_{t2} - x_{i,p-2} - \lambda_1 + x_{t,p-2} + \lambda_1 = 0$$

$$x_{i,k-1} - x_{t,k-1} + x_{i,k+1} - x_{t,k+1} + x_{t,p-2} - x_{i,p-2} = 0,\ 2 \leq k \leq p-3$$

$$x_{i,p-3} - x_{t,p-3} - \sum_{k=1}^{p-2} x_{ik} + \sum_{k=1}^{p-2} x_{tk} + x_{t,p-2} - x_{i,p-2} = 0,\ 1 \leq i \leq t-1$$

i.e., to $B\underline{Y}_i = 0$, $1 \leq i \leq t-1$ with $\underline{Y}_i = (x_{i1} - x_{t1}, \ldots, x_{ip-2} - x_{t,p-2})$
$1 \leq i \leq t-1$

and

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \ldots & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & 0 \ldots & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \ldots & 0 & 0 & 0 & 0 & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 \ldots & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \ldots & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & -1 & -1 \ldots & -1 & -1 & -1 & 0 & -2 \end{bmatrix}$$

i.e., $B = ((b_{ik}))$ with $b_{12} = 1$, $b_{1,p-2} = -1$, $b_{1k} = 0$ for $k \neq 2, p-2$

$b_{i,i-1} = 1$, $b_{i,i+1} = 1$, $b_{i,p-2} = -1$, $b_{i,k} = 0$, for $k \neq i-1, i+1, p-2$, $2 \leq i \leq p-4$

$b_{p-3,k} = 0$, $k \neq p-4$

$b_{p-3,p-4} = 1$, / $b_{p-2,1} = \ldots = b_{p-2,p-4} = -1$, $b_{p-2,p-3} = 0$, $b_{p-2,p-2} = -2$.

It can be checked that B is nonsingular. Hence $x_{ik} = z_k$, $1 \leq i \leq t$, $1 \leq k \leq p-2$. From the given constraints, it then follows that the minimum is attained when $x_{ik} = \lambda_1$ for all i and k.

To complete the proof of Theorem 4b.3.4., the final point to be noted is that $C_{d^*}(\tau | M)$ is c.s..

Theorem 4b.3.5 (C & W Th. 4.4 Second Part : Alternative Version).

Consider Model 4 (vide (4a.1.6)) i.e., assume there is no period effect. The design d^* which is balanced and uniform on units is universally optimal for the estimation of direct effects over the class of designs in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$ which are uniform on units and the last period.

Proof. We first note that uniformity of d^* in the last period will automatically follow from (4a.1.8) since d^* is balanced and uniform on units, and, hence, equireplicate. Under Model 4 (vide (4a.1.6)), we get, in general terms,

$$\left.\begin{aligned} C_{d11}(M) &= D_d - \frac{\omega - \tilde{\omega}}{\omega} \frac{N_{du} N'_{du}}{p} - \frac{\tilde{\omega}}{\omega} \frac{r_d\, r'_d}{np} \\ C_{d12}(M) &= C'_{d21}(M) = M_d - \frac{\omega - \tilde{\omega}}{\omega} \frac{N_{du} \tilde{N}'_{du}}{p} - \frac{\tilde{\omega}}{\omega} \frac{r_d\, \tilde{r}'_d}{np} \\ C_{d22}(M) &= \tilde{D}_d - \frac{\omega - \tilde{\omega}}{\omega} \frac{\tilde{N}_{du} \tilde{N}'_{du}}{p} - \frac{\tilde{\omega}}{\omega} \frac{\tilde{r}_d\, \tilde{r}'_d}{np} \end{aligned}\right\} \quad \ldots(4b.3.3)$$

For any design d which is uniform on units and the last period, it can be shown that

$$\left.\begin{aligned} C_{d11}(M) &= n\lambda_2\left(I - \frac{J}{t}\right) = C_{d*11}(M) \\ C_{d12}(M) &= M_d - \frac{\lambda_1(p-1)}{t} J \\ C_{d22}(M) &= \left\{\lambda_1(p-1) - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p}\right\} I - \left\{\frac{\omega-\tilde{\omega}}{\omega}(p-2)\lambda_1 + \frac{\tilde{\omega}}{\omega}\frac{\lambda_1(p-1)^2}{p}\right\}\frac{J}{t} \\ &= C_{d*22}(M) \end{aligned}\right\} \quad ..(4b.3.4)$$

It is <u>not</u> difficult to verify that d^* minimizes

$\operatorname{tr} C_{d12}(M)C^-_{d22}(M)C_{d21}(M) = \operatorname{tr} C_{d12}(M)C^-_{d*22}(M)C_{d21}(M)$ which simplifies to

$$c\sum_{i=1}^{t}\sum_{j=1}^{t}\left(m_{dij} - \frac{\lambda_1(p-1)}{t}\right)^2 + d\sum_{i=1}^{t}\left(s_{di} - \lambda_1(p-1)\right)^2 \quad ...(4b.3.5)$$

for some $c > 0$, $d > 0$,

over the class $\Lambda_{t,\lambda_1 t,\lambda_2 t}$ (in which for every d, $m_{dii} = 0$ for all i) as d^* is balanced and $s_{di} = \sum_{j=1}^{t} m_{dij}$ for all i. Moreover, $C_{d*}(\underline{\tau}|M)$ is c.s.

Hence the result.

The first part of Theorem 4.4 in C & W (1980) deals with estimation of residual effects. The original proof as in C & W (1980) suffers from an error. It was subsequently corrected in C & W (1983) wherein a peculiarity of the $C_d(\underline{\rho}|F)$ matrix under the assumption of no period effect has been pointed out. This is that the row sums of this matrix are, in general, not equal to zero. This makes Kiefer's Proposition 1 (as also Proposition 1' in fact) inapplicable. It has been suggested that the model may be reparametrized by imposing an additional constraint $\Sigma \rho_i = 0$ while estimating contrasts involving residual effects. So we consider the following

Model 4'

$$y_{ij} = \mu + \beta_j + \rho^*_{d(i-1,j)} + \tau_{d(i,j)} + \varepsilon_{ij} \qquad \ldots(4b.3.6)$$

with $\Sigma \rho^*_i = 0$ which is the same as considering a representation for $\underline{\rho}^*$ in the form $\underline{\rho}^* = (I - \frac{J}{t})\underline{\rho}$ with no restriction on $\underline{\rho}$. Working on this revised model, the C-matrix for the residual effects $\underline{\rho}^*$ assumes the form

$$C_d^{(4)}(\underline{\rho}^* | M) = (I - \frac{J}{t})C_d^{(4)}(\underline{\rho} | M)(I - \frac{J}{t})$$

where $C_d^{(4)}(\underline{\rho} | M)$ is to be derived using the formula

$$C_d^{(4)}(\underline{\rho} | M) = C_{d22}(M) - C_{d21}(M)C^-_{d11}(M)C_{d12}(M)$$

with $C_{dij}(M)$'s as given in (4b.3.3).

The optimality result is now stated as follows.

Theorem 4b.3.6 (C & W Th. 4.4 First Part (Weakened Version)). Consider Model 4'. The design d^* which is balanced and uniform over units is universally optimal for estimation of residual effects over the class of designs in $\Lambda_{t,\lambda_1 t,\lambda_2 t}$ with $\tilde{r}_{di} = \lambda_1(p-1)$ for all i.

Remark 1. In the above we need to assume equality of $\tilde{r}_{di}$'s for competing designs under the mixed effects model while the fixed effects analogue does away with this assumption.

Proof. Once more we refer to (4b.3.3). First note that for d^*,

$$C_{d^*11}(M) = n\lambda_2(I - \frac{J}{t}) \text{ so that } C^-_{d^*11}(M) = D^{-1}_{d^*}.$$

Further, $C_{d11}(M) \leq D_d$ for any other d. This time we have to establish that d^* maximizes

$$\mathrm{tr}\left\{(I - \frac{J}{t})(C_{d22}(M) - C_{d21}(M)C^-_{d11}(M)C_{d12}(M))(I - \frac{J}{t})\right\}.$$

In view of the above, it is enough to verify that

$$\mathrm{tr}\left\{(C_{d22}(M) - C_{d21}(M)D_d^{-1}C_{d12}(M))(I - \frac{J}{t})\right\}$$
$$\leq \mathrm{tr}\left\{(C_{d^*22}(M) - C_{d^*21}(M)D^{-1}_{d^*}C_{d^*12}(M))(I - \frac{J}{t})\right\}$$

Next observe that $\tilde{r}_{di}$'s are held all equal for any d and, further, $\sum_{i=1}^{t}\sum_{u=1}^{n}\tilde{n}^2_{diu}$ is minimized by d^*. Also $\mathrm{tr}(C_{d22}(M)J) = \mathrm{tr}(C_{d^*22}(M)J)$. Hence, it amounts to checking that d^* minimizes

$$\mathrm{tr}\left\{(C_{d21}(M)D_d^{-1}C_{d12}(M))(I - \frac{J}{t})\right\}$$

$$= \sum_{i=1}^{t} r^{-1}_{di}\sum_{j=1}^{t}(m_{dij} - \frac{\omega - \tilde{\omega}}{\omega\, p}\sum_{u=1}^{n} n_{diu}\tilde{n}_{dju} - \frac{\tilde{\omega}}{\omega np} r_{di}\tilde{r}_{dj} - \frac{s_{di}}{t} + \frac{p-1}{pt} r_{di})^2$$

$$= \sum_{i=1}^{t} r^{-1}_{di}\sum_{j=1}^{t}(m_{dij} - q^*_{dij})^2$$

where $q^*_{dij} = \frac{\omega - \tilde{\omega}}{\omega\, p}\sum_{u=1}^{n} n_{diu}\tilde{n}_{dju} + \frac{\tilde{\omega}}{\omega np} r_{di}\tilde{r}_{dj} + \frac{s_{di}}{t} - \frac{p-1}{pt} r_{di}$.

Now following essentially the same steps as in the proof of Theorem 4.1 in

C & W (1980), one can show that

$$\sum_{i=1}^{t} r_{di}^{-1} \sum_{j=1}^{t} (m_{dij} - q_{dij}^{*})^2 \geq \frac{t}{t-1} \sum_{i=1}^{t} r_{di}^{-1} q_{dii}^{*2} \quad \text{with " = " for } d^*.$$

Further, using Cauchy-Schwartz inequality,

$$\sum_{i=1}^{t} r_{di}^{-1} q_{dii}^{*2} \geq \frac{(\sum_{i=1}^{t} q_{dii}^{*})^2}{(\sum_{i=1}^{t} r_{di})} = \frac{(\frac{\omega - \widetilde{\omega}}{\omega\ p} \sum_{i=1}^{t} \sum_{u=1}^{n} n_{diu}\widetilde{n}_{diu} + \frac{\widetilde{\omega}}{\omega\, np} \sum_{i=1}^{t} r_{di}\widetilde{r}_{di})^2}{tn\lambda_2}$$

with " = " for d^*.

Since for every competing design, $\widetilde{r}_{di}$ is a constant, it now follows by an application of Lemma 5.3 in C & W (1980) and the fact (also stated and explained in the step (iii) of the proof of Theorem 4.1 in C & W (1980)) that $(n_{diu})_{u=1}^{n}$ and $(\widetilde{n}_{diu})_{u=1}^{n}$ are indeed similarly ordered for all i, that d^* minimizes the quantity

$$\frac{\omega - \widetilde{\omega}}{\omega\ p} \sum_{i} \sum_{u} n_{diu}\widetilde{n}_{diu} + \frac{\widetilde{\omega}}{\omega\, np} \sum_{i} r_{di}\widetilde{r}_{di}.$$

Finally, we observe that $C_{d^*}(\underline{\rho}^*|M)$ is c.s. Hence the result.

4b.4. <u>Optimality of Balanced Designs over the Class</u> $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$

In this subsection we propose to present some optimality results over the class $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$. These results are new even in fixed effects context and are <u>not</u> reported in C & W (1980) or elsewhere. As the number of periods (p) is taken to be $\lambda_2 t+1$, we will prove optimality of d^* which is balanced, uniform on periods and this time uniform on units in the first

$p-1$ periods. Here also we have to impose restrictions on competing designs.

Theorem 4b.4.1. Let d^* be balanced, uniform on periods, and uniform on each unit in the first $p-1$ periods in $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$, $t \geq 3$. Then d^* is universally optimal for the estimation of direct effects over the class of designs in $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$, which are equireplicate and uniform in the last period.

Proof : From (4b.1.1),

$$C_{d22}(M) \leq D_d \quad \text{and} \quad C_{d^*22}(M) = n\lambda_2(I - \frac{J}{t}).$$

Thus a choice of $C^-_{d^*22}(M)$ is $\frac{1}{n\lambda_2} I = \tilde{D}^{-1}_{d^*}$.

Hence, the maximization of $tr(C_{d11}(M) - C_{d12}(M)C^-_{d22}(M)C_{d21}(M))$ can be replaced by maximization of $tr(C_{d11}(M) - C_{d12}(M)\tilde{D}^{-1}_d C_{d21}(M))$. As we have assumed competing designs to be equireplicate and uniform on the last period, for any d, $\tilde{D}_d = n\lambda_2 I$. Thus we have to show,

$$tr(C_{d11}(M) - \frac{C_{d12}(M)C_{d21}(M)}{n\lambda_2}) \leq tr(C_{d^*11}(M) - \frac{C_{d^*12}(M)C_{d^*21}(M)}{n\lambda_2})$$

Now,

$$tr\left(C_{d11}(M) - \frac{C_{d12}(M)C_{d21}(M)}{n\lambda_2}\right) = \sum_{i=1}^{t} r_{di} - \frac{1}{n}\sum_{i=1}^{t}\sum_{k=1}^{p} \ell^2_{dik}$$

$$- \frac{\omega-\tilde{\omega}}{\omega}\frac{1}{p}\sum_{i=1}^{t}\sum_{u=1}^{n} n^2_{diu} + \frac{\omega-\tilde{\omega}}{\omega}\frac{1}{np}\sum_{i=1}^{t} r^2_{di}$$

$$- \frac{1}{n\lambda_2}\sum_{j=1}^{t}\sum_{i=1}^{t}\Big(m_{dij} - \frac{1}{n}\sum_{k=2}^{p}\ell_{dik}\ell_{dj,k-1}$$

$$- \frac{\omega-\tilde{\omega}}{\omega}\Big\{\frac{1}{p}\sum_{u=1}^{n} n_{diu}\tilde{n}_{dju} - \frac{1}{np} r_{di}\tilde{r}_{dj}\Big\}\Big)^2$$

Obviously, d^* minimizes $\sum_{i=1}^{t}\sum_{k=1}^{p}\lambda_{dik}^2$ and $\sum_{i=1}^{t}\sum_{u=1}^{n}n_{diu}^2$, and r_{di}'s and $\tilde{r}_{di}$'s are held constant for any competing design d.

Further, minimization of

$$\frac{1}{n\lambda_2}\sum_{j=1}^{t}\sum_{i=1}^{t}(m_{dij}-\frac{1}{n}\sum_{k=2}^{p}\lambda_{dik}\lambda_{dj,k-1}-\frac{\omega-\tilde{\omega}}{\omega}\{\frac{1}{p}\sum_{u=1}^{n}n_{diu}\tilde{n}_{dju}-\frac{1}{np}r_{di}\tilde{r}_{dj}.\})^2$$

can now be handled using the same technique as in the proof of Theorem 4.1 in C & W (1980). We omit the details.

Theorem 4b.4.2. Let d^* be a design which is balanced, uniform on each unit in the first p-1 periods, and uniform on each period. Then d^* is universally optimal for the estimation of residual effects over the class of designs in $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$ which are uniform on each unit in the first p-1 periods and also uniform in the last period.

Proof : The restrictions here on the competing designs are similar to those for establishing optimality results for the estimation of direct effects over the class $\Lambda_{t,\lambda_1 t,\lambda_2 t}$. Here $C^-_{d^*11}(M) \neq D^{-1}_{d^*}$ and hence the trace maximization of $(C_{d22}(M)-C_{d21}(M)C^-_{d11}(M)C_{d12}(M))$ cannot be replaced by the trace maximization of $(C_{d22}(M)-C_{d21}(M)D_d^{-1}C_{d12}(M))$.

On the other hand, the above restrictions on the competing designs yield,

$$C_{d11}(M) = D_d - \frac{1}{n}N_{dp}N'_{dp} - \frac{\omega-\tilde{\omega}}{\omega}\{\frac{1}{p}N_{du}N'_{du} - \frac{1}{np}N_{du}JN'_{du}\}$$

$$= (\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p})I + \frac{\lambda_1}{p}\ \frac{\omega-\tilde{\omega}}{\omega}\frac{J}{t} - \frac{1}{n}N_{dp}N'_{dp}$$

and $C_{d^*11}(M) = (\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p})(I-\frac{J}{t})$.

Thus, as $\frac{1}{n}(N_{dp}N'_{dp} - \lambda_1^2 p J) = N_{dp}(I - \frac{J}{p})N'_{dp}$ is n.n.d, we derive

$$C_{d11}(M) \leq C_{d^*11}(M).$$

Therefore, the problem reduces to that of establishing that d^* maximizes

$$\mathrm{tr}\{C_{d22}(M) - C_{d21}(M)C^-_{d^*11}(M)C_{d12}(M)\} = \mathrm{tr}\left(C_{d22}(M) - \frac{C_{d21}(M)C_{d12}(M)}{(\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p})}\right)$$

This can be achieved in a similar fashion as in the proof of Theorem 4.1 in C & W (1980). For an application of Lemma 5.6 in C & W (1980), we need only to check that

$$\frac{1}{n(t-1)(\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p})} \leq \frac{1}{2t(p-1)\lambda_1^2}$$

which is true since $2(p-1)p < (t-1)(p^2-1)$ for $t \geq 3$.

<u>Theorem 4b.4.3</u>. Assume there is no period effect, that is to say, consider Model 4 (vide (4a.1.6)). Then d^* which is balanced, uniform on units in the first $p-1$ periods and equireplicate is universally optimal for the direct effects in $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$ with $\tilde{r}_{di} = \lambda_1(p-1)$ for all i. The same design is universally optimal for the residual effects in the class $\Lambda_{t,\lambda_1 t,\lambda_2 t+1}$ of designs which are uniform on each unit in the first $p-1$ periods and also uniform in the last period.

<u>Proof</u> : First we prove the result for direct effects. From (4b.3.3), for the given d^*,

$$C_{d^*22}(M) = \lambda_1(p-1)I - \frac{\lambda_1\lambda_2^2 t}{p} J = aI + bJ \quad \text{(say)}$$

where $a = \lambda_1(p-1)$, $b = -\dfrac{\lambda_1\lambda_2^2 t}{p}$.

Further, imposing the restriction of equality of $\widetilde{r}_{di}$'s, we get, from (4b.3.3), for any competing design d,

$$C_{d22}(M) = \lambda_1(p-1)I - \frac{\omega-\widetilde{\omega}}{\omega}\,\frac{\widetilde{N}_{du}\widetilde{N}'_{du}}{p} - \frac{\widetilde{\omega}}{\omega}\,\frac{\lambda_1^2(p-1)^2}{np}\,J$$

and $\quad C_{d22}(M) \leq C_{d^*22}(M)$ as $\widetilde{N}_{du}\widetilde{N}'_{du} - n\lambda_2^2 J = \widetilde{N}_{du}(I - \frac{J}{n})\widetilde{N}'_{du}$ is n.n.d.

Thus maximization of $tr(C_{d11}(M) - C_{d12}(M)C^-_{d22}(M)C_{d21}(M))$ can be replaced by that of $tr(C_{d11}(M) - C_{d12}(M)C^-_{d^*22}(M)C_{d21}(M))$. Now using $C^-_{d^*22}(M) = cI + dJ$ where $c = \frac{1}{a} > 0$, $d = \frac{-b}{(a+bt)a} > 0$, we get

$$tr(C_{d11}(M) - C_{d12}(M)C^-_{d^*22}(M)C_{d21}(M)) = \sum_{i=1}^{t} r_{di} - \frac{\omega-\widetilde{\omega}}{\omega p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}^2$$

$$- \frac{\widetilde{\omega}}{\omega}\,\frac{\sum_{i=1}^{t} r_{di}^2}{np} - c\sum_{i=1}^{t}\sum_{j=1}^{t}\Big(m_{dij} - \frac{\omega-\widetilde{\omega}}{\omega p}\sum_{u} n_{diu}\widetilde{n}_{dju} - \frac{\widetilde{\omega}}{\omega}\,\frac{r_{di}\widetilde{r}_{dj}}{np}\Big)^2$$

$$- d\sum_{i=1}^{t}\Big(s_{di} - \frac{p-1}{p}\,r_{di}\Big)^2 .$$

Next observe that d^* minimizes $\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}^2$, $\sum_{i=1}^{t} r_{di}^2$ and also $\sum_{i=1}^{t}(s_{di} - \frac{p-1}{p} r_{di})^2$ as s_{d^*i}'s are all equal to the corresponding $\frac{p-1}{p} r_{d^*i}$'s. Further, the minimization of $\sum_{i=1}^{t}\sum_{j=1}^{t}(m_{dij} - \frac{\omega-\widetilde{\omega}}{\omega p}\sum_{u=1}^{n} n_{diu}\widetilde{n}_{dju} - \frac{\widetilde{\omega}}{\omega}\frac{r_{di}\widetilde{r}_{dj}}{np})^2$ can be handled using a technique similar to that adopted in the proof of optimality of residual effects in Theorem 4b.3.6.. Since

$C_{d^*11}(M) - C_{d^*12}(M)C^{-}_{d^*22}(M)C_{d^*21}(M)$ is c.s., the claim is established.

Now we discuss the case of residual effects. Following the same arguments given in Theorem 4b.3.6, we have to consider here Model 4' (vide (4b.3.6)). Here

$$C_{d^*11}(M) = (\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p})(I - \frac{J}{t})$$

and, further, because of the restrictions imposed on competing designs, we also have $C_{d11}(M) = C_{d^*11}(M)$ for any other design. Thus the problem reduces to checking that

$$\mathrm{tr}\left(\left(C_{d22}(M) - \frac{C_{d21}(M)C_{d12}(M)}{\left\{\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p}\right\}}\right)(I - \frac{J}{t})\right)$$

$$\leq \mathrm{tr}\left(\left(C_{d^*22}(M) - \frac{C_{d^*21}(M)C_{d^*12}(M)}{\left\{\lambda_1 p - \frac{\omega-\tilde{\omega}}{\omega}\frac{\lambda_1}{p}\right\}}\right)(I - \frac{J}{t})\right).$$

Again, following steps similar to those adopted in the proof of Theorem 4b.3.6, we can establish the claim.

4b.5. Optimality of Nearly Strongly Balanced RMD's

So far we have considered optimality of strongly balanced or balanced designs. It is known (C & W (1980)) that strongly balanced uniform RMD's can exist only if $t^2 | n$ and $p \geq 2t$. We now treat the situation where $p \geq 2t$ but $t^2 \nmid n$. We take $n = At^2 + Bt$, $A \geq 1$, $1 \leq B \leq t-1$. We will proceed along the line of work of Kunert (1983) and re-establish his results on optimality of 'nearly strongly balanced'

Generalized Latin Square (GLS) designs (vide DFN 4a.1.6 and DFN 1d.2.9) over a subset of all possible designs in the set-up of mixed effects model.

We first quote a lemma (without proof) from Magda (1980) (as also Proposition 2.3 of Kunert (1983)), which states that for the purpose of estimating linear functions of certain parameters, we only decrease the precision of our estimates by allowing more nuisance parameters in the model.

Under the general linear model,

$$E(\underset{\sim}{Y}) = X_1\theta_1 + X_2\theta_2 + \dots + X_r\theta_r$$

$$\mathrm{Cov}(\underset{\sim}{Y}) = \sigma^2 I,$$

we denote by $C_r(\theta_1)$ the information matrix of θ_1 under the above model.

<u>Lemma 4b.5.1</u> <u>(M Lemma 3.1, K(1) Proposition 2.3)</u>. For the two linear models,

$$E(\underset{\sim}{Y}) = \sum_{i=1}^{k} X_i\theta_i \quad \text{and} \quad E(\underset{\sim}{Y}) = \sum_{i=1}^{n} X_i\theta_i, \ \mathrm{Cov}(\underset{\sim}{Y}) = \sigma^2 I, \text{ with } k \leq n,$$

we have $C_n(\theta_1) \leq C_k(\theta_1)$.

It may be noted that Proposition 2.3 of Kunert (1983) states this result in terms of projection operator and gives the condition for equality in terms of orthogonality.

We now restate the optimality results for the mixed effects model.

<u>Theorem 4b.5.1.</u> <u>(K(1) Th. 5.3 & Th. 5.4)</u> Assume $n = At^2 + Bt$, $1 \leq B \leq t-1$, and $p = \lambda_2 t$. A nearly strongly balanced GLS $d^* \in \Omega_{t,n,p}$

is universally optimal for the estimation of direct effects over the class of designs in $\Omega_{t,n,p}$ which are uniform on the units and the last period. The same design is also universally optimal for the estimation of residual effects over the class of designs in $\Omega_{t,n,p}$ which are uniform on the units as also on the first and last periods.

Proof : Consider Model 4 assuming only unit effects and direct effects of treatments. Then, by an application of Lemma 4b.5.1,

$$C_d(\underset{\sim}{\tau} \mid M) \leq C_d^{(4)}(\underset{\sim}{\tau} \mid M)$$

where $C_d^{(4)}(\underset{\sim}{\tau} \mid M)$ corresponds to the $C(M)$-matrix for $\underset{\sim}{\tau}$ under Model 4 and can be derived from (4b.3.3). Thus, the claim made in this theorem will be established if we can prove

$$\text{(a)} \quad \text{tr } C_d^{(4)}(\underset{\sim}{\tau} \mid M) \leq \text{tr } C_{d^*}^{(4)}(\underset{\sim}{\tau} \mid M)$$

and $\quad$ (b) $\quad C_{d^*}^{(4)}(\underset{\sim}{\tau} \mid M) = C_{d^*}(\underset{\sim}{\tau} \mid M)$.

Now following the steps developed in the proof of Theorem 4b.3.5 for estimation of direct effects, we note that maximization of $\text{tr } C_d^{(4)}(\underset{\sim}{\tau} \mid M)$ boils down to minimization of $\text{tr } C_{d12}(M)C^-_{d^*22}(M)C_{d21}(M)$ with C_{dij}'s as defined in (4b.3.4). Consequently, $\text{tr } C_{d12}(M)C^-_{d^*22}(M)C_{d21}(M)$ has the same expression as (4b.3.5). Now, in the present case, we have assumed $n = At^2 + Bt$, and thus $\lambda_1 = At + B$. Therefore, minimization of

$$\Sigma\Sigma\left(m_{dij} - \frac{\lambda_1(p-1)}{t}\right)^2 = \Sigma\Sigma\left(m_{dij} - \frac{np}{t^2} + A + \frac{B}{t}\right)^2$$

subject to $\sum_{i=1}^{t} \left(m_{dij} - \frac{np}{t^2} + A + \frac{B}{t}\right) = 0$ for every j, will also be done by d^* over the class $\Omega_{t,\lambda_1 t,\lambda_2 t}$ as for d^* exactly B of the m_{dij}'s are equal to $\frac{np}{t^2} - A - 1$ and $(t-B)$ are equal to $\frac{np}{t^2} - A$ (for every j) (vide DFN 4a.1.6 in this context). The same quantity has been minimized in the proof of Theorem 4b.3.5 over the class $\Lambda_{t,\lambda_1 t,\lambda_2 t}$ assuming $m_{dii} = 0$ for all i, and for all competing design d. Obviously d^* also minimizes $\Sigma\,(s_{di} - \lambda_1(p-1))^2$ over $\Omega_{t,n,p}$. This settles the part (a) above.

Now to prove (b) we note that $C_{d^*11}(M)$ and $C_{d^*12}(M)$ remain unaltered for both Model 1 and Model 4, that is, irrespective of the period effects being present or absent. Moreover, under Model 4, (vide (4b.3.4))

$$C^-_{d^*22}(M) = cI + dJ \quad \text{where} \quad c = \left\{ \lambda_1(p-1) - \frac{\omega - \tilde{\omega}}{\omega} \; \frac{\lambda_1}{p} \right\}^{-1}$$

while under Model 1, (vide (4b.3.2) for a similar derivation here)

$$C^-_{d^*22}(M) = cI \quad \text{with the same expression for } c.$$

Since $C_{d^*12}(M)\,J = 0$, it now follows that

$$C_{d^*}(\underline{\tau} \mid M) = C^{(4)}_{d^*}(\underline{\tau} \mid M).$$

Also $C_{d^*}(\underline{\tau} \mid M)$ is c.s. as $M_{d^*}M'_{d^*}$ is c.s. Hence, the result.

Thus the first part of the theorem is established.

Now we discuss the case of residual effects. For this, we consider Model 4$'$ (vide (4b.3.6)). Then, by an application of Lemma 4b.5.1,

$$C_d(\underline{\rho} \mid M) = (I - \tfrac{J}{t})(C_d(\underline{\rho} \mid M))(I - \tfrac{J}{t}) \leq (I - \tfrac{J}{t})C^{(4)}_d(\underline{\rho} \mid M)(I - \tfrac{J}{t}) = C^{(4)}_d(\underline{\rho}^* \mid M)$$

(say)

(Recall that $C_d(\underset{\sim}{\rho} \mid M)J = 0$ (vide (4b.1.1)), whereas $C_d^{(4)}(\rho)J \neq 0$ in general, as discussed earlier).

Now, the claim in the theorem will be established if we can prove that

$$(a') \quad \text{tr} \;\; C_d^{(4)}(\underset{\sim}{\rho}^* \mid M) \;\leq\; \text{tr} \;\; C_{d^*}^{(4)}(\underset{\sim}{\rho}^* \mid M)$$

and $\quad (b') \quad C_{d^*}^{(4)}(\underset{\sim}{\rho}^* \mid M) = C_{d^*}(\underset{\sim}{\rho} \mid M).$

We have seen in the proof of Theorem 4b.3.5 that for any competing design d, being uniform on units and the last period,

$$C_{d11}(M) = \lambda_1 p(I - \tfrac{J}{t}) \quad \text{and} \quad C_{d22}(M) = aI + bJ \quad \text{(vide (4b.3.4))}$$

Thus, when there is no period effect, with C_{dij}'s as defined in (4b.3.3),

$$\text{tr } C_d^{(4)}(\underset{\sim}{\rho}^* \mid M) = \text{tr}\left\{ (C_{d22}(M) - \frac{C_{d21}(M)C_{d12}(M)}{\lambda_1 p}) \; (I - \tfrac{J}{t}) \right\}.$$

But $\text{tr}\{C_{d22}(M)(I - \frac{J}{t})\}$ is a constant as $C_{d22}(M) = aI + bJ$. Moreover,

$$C_{d12}(M)\,\underset{\sim}{1} = (M_d - \frac{\omega - \widetilde{\omega}}{\omega}\,\frac{N_{du}\widetilde{N}'_{du}}{p} - \frac{\widetilde{\omega}}{\omega}\,\frac{\underset{\sim}{r}_d\widetilde{\underset{\sim}{r}}_d}{np})\,\underset{\sim}{1}$$

$$= \underset{\sim}{s}_d - \lambda_1(p-1)\,\underset{\sim}{1} = \underset{\sim}{0}$$

since for any competing design which is uniform on the <u>first</u> period, $s_{di} = \lambda_1(p-1)$ for all i. Thus maximization of $\text{tr } C_d^{(4)}(\underset{\sim}{\rho}^* \mid M)$ boils down to minimization of $\text{tr } C_{d21}(M)C_{d12}(M) = \sum_{i=1}^{t} \sum_{j=1}^{t} \left(m_{dij} - \frac{\lambda_1(p-1)}{t}\right)^2$.

As in the case of direct effects (discussed in the first part above), it can be argued that d^* being nearly strongly balanced, minimizes $\text{tr } C_{d21}(M)C_{d12}(M)$. Thus (a') is through.

To prove (b'), we note that both $C_{d^*12}(M)$ and $C_{d^*11}(M)$ have the same expression under both the models Model 1 and Model 4.

Moreover, $\left(\frac{\tilde{N}_{d^*p}\tilde{N}'_{d^*p}}{n} - \frac{\tilde{r}_{d^*}\tilde{r}'_{d^*}}{np} \right) (I - \frac{J}{t}) = 0$.

Thus $(I - \frac{J}{t}) \left\{ C_{d^*22}(M) \text{ (without period effect)} + \frac{\tilde{N}_{d^*p}\tilde{N}'_{d^*p}}{n} - \frac{\tilde{r}_{d^*}\tilde{r}'_{d^*}}{np} \right\} (I - \frac{J}{t})$

$$= C_{d^*22}(M) \text{ (with period effect)}$$

and $C^{(4)}_{d^*}(\underset{\sim}{\rho}^* \mid M) = C_{d^*}(\underset{\sim}{\rho} \mid M)$ using the fact that $C_{d^*}(\underset{\sim}{\rho} \mid M) J = 0$.

As d^* is GLS and nearly strongly balanced, M_{d^*} is the incidence matrix of a BBD, with same number of treatments and blocks (equal to t). Thus, as noted by Kunert (1983), $M'_{d^*}M_{d^*}$ is also c.s. by application of Theorem 5.2.1 of Raghavarao(1971). Hence $C^{(4)}_{d^*}(\underset{\sim}{\rho}^* \mid M) = C_{d^*}(\underset{\sim}{\rho} \mid M)$ is c.s. Thus the result.

As already mentioned, Theorem 4b.5.1 provides the mixed effects analogue of the corresponding Theorems 5.3 and 5.4 of Kunert (1983). It may be noted that Kunert (1983) also extended the universal optimality result of d^* for the estimation of direct effects over the class of <u>all</u> designs $\Omega_{t,n,p}$ under the fixed effects model (his Theorem 5.8), imposing some restriction on parameters. However, the same result in the mixed effects situation is yet to be established.

4c. Optimality Results over the Class $\Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$

Optimality has so far been established exclusively for designs which are strongly balanced or nearly strongly balanced or simply balanced.

Following Kunert (1983) we give examples of optimal and orthogonal designs which are neither balanced nor strongly or nearly strongly balanced. For this we consider the class $\Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$ and use Lemma 4b.5.1.

<u>Theorem 4c.1.1</u> <u>(K(1) Th. 4.1)</u> Let $n = \lambda_1 t$ and $p = \lambda_2 t$, $\lambda_1, \lambda_2 \geq 1$. Assume there is a GLS $d^* \in \Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$, such that

$$m_{d^*ij} = t^{-1}\tilde{r}_{d^*j}, \quad 1 \leq i \leq t,\ 1 \leq j \leq t.$$

Then d^* is universally optimal for the estimation of direct effects over $\Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$.

<u>Proof</u> : We consider Model 2(i) (vide (4a.1.6)) assuming only general effect, unit effect (random) and direct effects of treatments. Then

$$C_d^{(2)}(\underset{\sim}{\tau}|M) = D_d - \frac{\omega-\tilde{\omega}}{\omega}\ \frac{N_{du}N_{du}'}{p} - \frac{\tilde{\omega}}{\omega}\ \frac{\underset{\sim}{r}_d\underset{\sim}{r}_d'}{np}$$

and by an application of Lemma 4b.5.1

$$C_d(\underset{\sim}{\tau}|M) \leq C_d^{(2)}(\tau|M).$$

Obviously, d^* being a GLS, maximizes $\mathrm{tr}\ C_d^{(2)}(\tau|M)$. Moreover, for d^*, $N_{d^*p}N_{d^*p}' = \frac{\underset{\sim}{r}_{d^*}\underset{\sim}{r}_{d^*}'}{p}$ and, hence, $C_{d^*11}(M)$ under Model 1 (vide (4b.1.1)) coincides with $C_{d^*}^{(2)}(\underset{\sim}{\tau}|M)$. Again, under Model 1 (vide (4b.1.1) once more),

$$C_{d^*12}(M) = M_{d^*} - \frac{N_{d^*p}\tilde{N}_{d^*p}'}{n} - \frac{\omega-\tilde{\omega}}{\omega}\left\{\frac{N_{d^*u}\tilde{N}_{d^*u}'}{p} - \frac{\underset{\sim}{r}_{d^*}\underset{\sim}{\tilde{r}}_{d^*}'}{np}\right\}$$

As d^* is a GLS, the matrix associated with $\frac{\omega-\tilde{\omega}}{\omega}$ above vanishes.

Also $M_{d^*} - \frac{N_{d^*p}\tilde{N}'_{d^*p}}{n} = M_{d^*} - \frac{J\,\tilde{N}'_{d^*p}}{t}$ = a null matrix since we have assumed for d^*, $m_{d^*ij} = t^{-1}\,\tilde{r}_{d^*j}$. Thus, under Model 1,

$$C_{d^*}(\underline{\tau}|M) = C_{d^*11}(M) - C_{d^*12}(M)C^{-}_{d^*22}(M)C_{d^*21}(M) = C_{d^*11}(M) = C^{(2)}_{d^*}(\underline{\tau}|M)$$

by our arguments above. Finally, $C^{(2)}_{d^*}(\underline{\tau}|M)$ is c.s. Hence the result. holds for GLS $\varepsilon\ \Omega_{t,n,p}$

Remark 2. If all $\tilde{r}_{d^*i}$'s are equal, which/the given condition $m_{d^*ij} = t^{-1}\,\tilde{r}_{d^*j}$ is just strong balance $(\lambda_2 \geq 2)$ i.e. m_{d^*ij}'s are all equal. But for designs with preperiod, i.e. for designs belonging to the class $\tilde{\Omega}_{t,n,p}$, this condition is more general.

Theorem 4c.1.2 (K(1) Th. 4.4) Let $t \nmid n$ and $p = \lambda_2 t$, $\lambda_2 \geq 1$. Assume there is a GYD $d^*\ \varepsilon\ \Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$ such that

$$m_{d^*ij} = n^{-1} \sum_{k=1}^{p} \ell_{d^*ik}\,\tilde{\ell}_{d^*jk} \qquad 1 \leq i \leq t,\ 1 \leq j \leq t.$$

Then d^* is universally optimal for the estimation of direct effects over $\Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$.

Proof : We consider Model 3 (vide (4a.1.6)) assuming no residual effects for the present. Then

$$C^{(3)}_{d}(\underline{\tau}|M) = D_d - \frac{N_{dp}N'_{dp}}{n} - \frac{\omega-\tilde{\omega}}{\omega}\left\{\frac{N_{du}N'_{du}}{p} - \frac{\underline{r}_d\,\underline{r}'_d}{np}\right\}.$$

Now, by an application of Lemma 4b.5.1,

$$C_d(\underline{\tau}|M) \leq C^{(3)}_{d}(\underline{\tau}|M)$$

and $\operatorname{tr} C_d^{(3)}(\underline{\tau}|M) = \sum_{i=1}^{t} r_{di} - \frac{1}{n}\sum_{i=1}^{t}\sum_{k=1}^{p}\lambda_{dik}^2$

$$- \frac{1}{p}\frac{\omega-\tilde{\omega}}{\omega}\left\{\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}^2 - \frac{1}{n}\Sigma\, r_{di}^2\right\}$$

$$= \sum_{i=1}^{t} r_{di} - \frac{1}{n}\sum_{i=1}^{t}\sum_{k=1}^{p}\lambda_{dik}^2 - \frac{1}{p}\frac{\omega-\tilde{\omega}}{\omega}\sum_{i=1}^{t}\left\{\sum_{u=1}^{n} n_{diu}^2 - \frac{r_{di}^2}{n}\right\}$$

is maximized by d^* over $\Omega_{t,n,p} \cup \tilde{\Omega}_{t,n,p}$ as $|\lambda_{d^*ik} - \lambda_{d^*i'k}| \leq 1\ \forall k$, $i \leq i \neq i' \leq t$ and n_{d^*iu}'s are same.

Again, under the general model,

$$C_{d^*12} = M_{d^*} - \frac{N_{d^*p}\tilde{N}'_{d^*p}}{n} - \frac{\omega-\tilde{\omega}}{\omega}\left\{\frac{N_{d^*u}\tilde{N}'_{d^*u}}{p} - \frac{\underline{r}_{d^*}\tilde{\underline{r}}'_{d^*}}{np}\right\}$$

vanishes because d^* is a GYD and satisfies the given condition

$$m_{d^*ij} = \frac{1}{n}\sum_{k=1}^{p}\lambda_{d^*ik}\tilde{\lambda}_{d^*jk}, \text{ for } 1 \leq i \leq t,\ 1 \leq j \leq t.$$

Remark 3. It can be observed that under the conditions of Theorem 4c.1.2, it may not be possible to prove that the same design is optimal for the estimation of residual effects. The design in Theorem 4c.1.1 also could not be proved to be optimal for the estimation of residual effects under the context of mixed effects model. Also we could not extend Kunert's Theorem 4.8 to the mixed effects model by following arguments similar to those used in this chapter throughout.

4d. Circular Model

The model taken up in this section is a model with special choice of preperiod, namely the last period is assumed here to be the preperiod.

Thus, here, under circular model,

$$\rho_{d(o,j)} = \rho_{d(p,j)} \quad \forall\, j=1,2,\ldots,n.$$

Following Magda's (1980) technique, we re-establish the results relating to optimality of circular RMD's in a unified way for several models invoking Lemma 4b.5.1 on right occasions. Recall that in our earlier discussion of optimality results we have adopted this technique.

In what follows, d_1^* denotes a circular strongly balanced uniform RMD, d_2^* denotes a circular RMD uniform on units, d_4^* denotes a cicular strongly balanced RMD uniform on units.

<u>Theorem 4d.1.1</u> <u>(M Th. 3.1)</u> Under Model (i) $i = 1,2,4$ (vide (4a.1.1) and (4a.1.6)) whenever d_i^* exists, it is universally optimal for estimation of direct as well as residual effects over the collection of designs with the same parameters in $\tilde{\Omega}_{t,n,p}$ assuming the last period as the preperiod.

<u>Proof</u> : As in Theorem 4c.1.1 we first consider the simplest Model 2(i) for direct effects and by an application of Lemma 4b.5.1,

$$C_d^{(i)}(\underline{\tau}|M) \leq C_d^{(2)}(\underline{\tau}|M) \quad i = 1,4.$$

Following the same steps as developed in the proof of Theorem 4c.1.1, the result follows from orthogonality of the parameters with direct effects in d_i^*'s. The condition $m_{d^*ij} = t^{-1}\,\tilde{r}_{d^*j}$ of Theorem 4c.1.1 is nothing but that of strongly balanced design as under a circular model, for any design d, $r_{dj} = \tilde{r}_{dj}$ $\forall\, j = 1,2,\ldots,t.$

The same optimality result also holds simultaneously for residual effects as $C_d^{(2)}(\underline{\rho} \mid M) = C_d^{(2)}(\underline{\tau} \mid M)$ since $\tilde{N}_{du} = N_{du}$ and $\tilde{\underline{r}}_d = \underline{r}_d$ for any competing design d under a circular model.

Theorem 4d.1.2 (M Th. 3.2) Whenever a design d_i^* exists, $i = 1,2,4$, it minimizes the variance of the best linear unbiased estimate of any contrast of direct effects and any contrast of residual effects over the collection of equireplicate designs with same parameters.

We omit the proof which follows essentially along that of Theorem 4b.2.3.

In the following theorems, we consider the class of designs with $m_{dii} = 0$ and let δ_4^* denote a circular balanced RMD uniform on units and δ_1^* denote a balanced uniform circular RMD.

Theorem 4d.1.3 (M Th. 3.4) Whenever there exists a design $\delta_i^* \in \Lambda_{t,\lambda_1 t,\lambda_2 t}$, $i = 1,4$, it is universally optimal for the Model (i) over the class $\Lambda_{t,\lambda_1 t,\lambda_2 t}$.

Proof : We consider now Model 4 assuming no period effect. Then

$$C_d^{(4)}(\underline{\tau} \mid M) = C_{d11}(M) - C_{d12}(M)C_{d22}^{-}(M)C_{d21}(M)$$

where $C_{d11}(M) = C_{d22}(M) = D_d - \frac{\omega-\tilde{\omega}}{\omega p} N_{du}N_{du}' - \frac{\tilde{\omega}}{\omega np} \underline{r}_d \underline{r}_d'$.

and $C_{d12}(M) = C_{d21}'(M) = M_d - \frac{\omega-\tilde{\omega}}{\omega p} N_{du}N_{du}' - \frac{\tilde{\omega}}{\omega np} \underline{r}_d \underline{r}_d'$.

By an application of Lemma 4b.5.1,

$$C_d(\underline{\tau} \mid M) \leq C_d^{(4)}(\underline{\tau} \mid M).$$

As $C_{d11}(M) \leq D_d$ and $C_{\tilde{\delta}_4^* 11}(M) = D_{\delta_4^*}^{-1}$, as we have discussed earlier, maximization of $\operatorname{tr} C_d^{(4)}(\tau \mid M)$ is replaced by maximization of $\operatorname{tr}(C_{d11}(M) - C_{d12}(M)D_d^{-1}C_{d21}(M))$. δ_4^* trivially maximizes $\operatorname{tr} C_{d11}(M)$. Now,

$$\operatorname{tr} C_{d12}(M)D_d^{-1}C_{d21}(M) = \sum_{j=1}^{t} r_{dj}^{-1} \sum_{i=1}^{t} (m_{dij} - \frac{\omega - \tilde{\omega}}{\omega p} \sum_{u=1}^{n} n_{diu}n_{dju} - \frac{\tilde{\omega}}{\omega np} r_{di}r_{dj})^2$$

$$= \sum_{j=1}^{t} r_{dj}^{-1} \sum_{i=1}^{t} (m_{dij} - q_{dij})^2$$

where $q_{dij} = \frac{\omega - \tilde{\omega}}{\omega p} \sum_{u=1}^{n} n_{diu}n_{dju} + \frac{\tilde{\omega}}{\omega np} r_{di}r_{dj}$.

Now following the steps as developed in the proof of Theorem 4b.3.2, optimality of δ_4^* can be established as $m_{dii} = 0$ and

$$\sum_{j=1}^{t} q_{djj} = \frac{\omega - \tilde{\omega}}{\omega} \sum_{j=1}^{t} \sum_{u=1}^{n} \frac{n_{dju}^2}{p} + \frac{\tilde{\omega}}{\omega} \sum_{j=1}^{t} \frac{r_{dj}^2}{np}$$

$$\geq \frac{\omega - \tilde{\omega}}{\omega} \sum_{j=1}^{t} \frac{r_{dj}^2}{np} + \frac{\tilde{\omega}}{\omega} \Sigma \frac{r_{dj}^2}{np}$$

$$\geq \frac{(\sum_{j=1}^{t} r_{dj})^2}{tnp}$$

with "=" holding for δ_4^* in all the steps.

Moreover, $C_{\delta_4^*}(\underline{\tau} \mid M)$ is c.s. and $C_{\delta_4^*}(\underline{\tau} \mid M) = C_{\delta_1^*}(\underline{\tau} \mid M)$ as δ_1^* is uniform on periods and $\frac{N_{\delta_1^* p} N_{\delta_1^* p}'}{n} - \frac{\underline{r}_{\delta_1^*}\underline{r}_{\delta_1^*}'}{np} = 0$ (a null matrix).

Hence the result for direct effects is established and the same result for residual effects follows similarly.

4e. Construction of Nearly Strongly Balanced GLS d^*

So far we have dealt with the study of optimality of various RMD's assuming different models. Now, we focus our attention to the construction of such designs. Hedayat (1981a) gives a nice review of construction of balanced and strongly balanced RMD's. Further contributions to the construction of balanced and strongly balanced RMD's have been made by Constantine and Hedayat (1982), Afsarinejad (1983) and Mukerjee and Sen (1984). Regarding construction of strongly balanced GLS, Kunert (1983) gives a few sporadic examples.

We now provide construction of nearly strongly balanced GLS $d^* \varepsilon \Omega_{t,n,p}$ assuming $n = At^2 + Bt$, $p = \lambda_2 t$, where $B = 1$ or $t-1$. We first construct a strongly balanced URMD $(t, At^2, \lambda_2 t)$ and d^* is obtained from this either by appropriately adding Bt units or removing $t(t-B)$ units from a strongly balanced URMD $(t, (A+1)t^2, \lambda_2 t)$. Without loss of generality, we consider the case of $A = 1$, as a strongly balanced URMD $(t, At^2, \lambda_2 t)$, $A > 1$, is obtained from a strongly balanced URMD $(t, t^2, \lambda_2 t)$ by repeating the t^2 units A times.

Case (i) λ_2 an even integer.

The case of t odd and t even are considered separately here.

Let M denote the non-negative residues (mod t).

Sub-case : λ_2 even, t odd.

Let s be an element in M such that $(s,t) = 1$, $(s-1,t) = 1$. We first construct two periods of first t units of a strongly balanced

URMD $(t,t^2,\lambda_2 t)$ and then construct $2t$ periods for these t units. Finally the rest of the units for these $2t$ periods are obtained from the first t units.

Let G_o', formed as

$$G_o' = \begin{pmatrix} x_1, & x_2, \ldots, x_t \\ sx_1, & sx_2, \ldots, sx_t \end{pmatrix}$$

with $x_i \in M$, $x_i \neq x_j$ represent the first two periods of the first t units of strongly balanced URMD $(t,t^2,2t)$. Then G_h' is obtained from G_o' by adding $h \pmod t$ to each element of G_o'. Then B_o, defined as $B_o = (G_o, G_1, \ldots, G_{t-1})'$ form the $2t$ periods of these first t units, B_h, for $h = 1,2,\ldots,t-1$, is obtained from B_o by adding $h \pmod t$ to each element of B_o and finally $A = (B_o, B_1, \ldots, B_{t-1})$ form a strongly balanced URMD $(t,t^2,2t)$ with rows and columns identified with periods and units respectively.

Sub-case : λ_2 even, t even.

When t is even, strongly balanced URMD $(t,t^2,2t)$ is constructed in a similar fashion with some obvious modifications. The conditions this time naturally will be

$$(s,t) = 1, \quad (s-1,t) = 2.$$

Now, strongly balanced URMD $(t,t^2,\lambda_2 t)$ say d_o, is constructed from strongly balanced URMD $(t,t^2,2t)$ say d_1, by piecing $\frac{\lambda_2}{2}$ copies of d_1 together one after another. Finally, it can be verified easily that nearly strongly balanced GLS d^* for t^2+t units and $\lambda_2 t$ periods is obtained from d_o by adding extra t units chosen as the columns of any one of $\left(\begin{matrix} B_h \\ \vdots \\ B_h \end{matrix}\right\}$ repeated $\frac{\lambda_2}{2}$ times$\Bigg)$'s for $h = 0,1,2,\ldots,t-1$ to the units of d_o.

Similarly, nearly strongly balanced GLS d^* for $t^2+t(t-1)$ units is obtained by removing t units, chosen as the columns of any one of $\begin{pmatrix} B_h \\ \vdots \\ B_h \end{pmatrix}$'s, $h = 0,1,\ldots,t-1$ from the $2t^2$ units of strongly balanced URMD $(t,2t^2,\lambda_2 t)$.

Case (ii). λ_2 an odd integer.

We partially resolve the problem when λ_2 is an odd integer, considering some of the cases when t is an odd prime or prime power.

Let α be a primitive element of $GF(t)$ such that $2\alpha^{\lambda} - \alpha^{\lambda-1} - 1 \neq 0$ for $\lambda = 1,2,\ldots,t-2$. (We note here that the existence of such an α for all $GF(t)$, t odd prime or prime power, is not known. But the set is not vacuous in the sense that for some particular values of t say $t = 3$ or 5, the existence of such an α can be readily established.) Let $P = (x_1,x_2,\ldots,x_t)$ be an arrangement of the t elements of $GF(t)$, where $x_i \neq x_j$, $i,j = 1,2,\ldots,t$ and $P+y$ be defined as $(x_1+y,\ x_2+y,\ldots,x_t+y)$ for some $y \in GF(t)$. For $\lambda_2 = (2m+1)t$ for some $m \geq 1$, we first construct $3t$ periods of d_o — strongly balanced URMD $(t,t^2,\lambda_2 t)$ and then construct the rest $2(m-1)t$ periods of d_o following the steps of Case (i). Let $3\times t$ matrices $\bar{G}_o'$ and $\bar{G}_\lambda'$, $\lambda = 1,2,\ldots,t-1$ be represented as

$$\bar{G}_0' = \begin{pmatrix} p - (\alpha-1)^{-1} \\ p - \gamma(\alpha-1)^{-1} \\ p - \beta(\alpha-1)^{-1} \end{pmatrix}, \quad \bar{G}_\lambda' = \begin{pmatrix} p + (\alpha^\lambda - 1)(\alpha-1)^{-1} \\ p + \gamma(\alpha^\lambda-1)(\alpha-1)^{-1} \\ p + \beta(\alpha^\lambda - 1)(\alpha-1)^{-1} \end{pmatrix}$$

where $\gamma, \beta \in GF(t)$, $\beta = 1-\alpha$, and $(\gamma, \beta) = 1$. Then $\bar{B}_0$ is defined as $\bar{B}_0 = (\bar{G}_0, \bar{G}_1, \ldots, \bar{G}_{t-1})'$. (It may be checked that as a consequence of the assumption $2\alpha^\lambda - \alpha^{\lambda-1} - 1 \neq 0$ for $\lambda = 1,2,\ldots,t-2$ the pairs of elements appearing in the last row of $\bar{G}_k'$ and the first row of $\bar{G}_{k+1}'$ for $k = 0,1,2,\ldots,t-2$ are all distinct). $\bar{B}_h$, for $h = 1,2,\ldots,t-1$ is obtained from $\bar{B}_0$ by adding $h(\text{mod } t)$ to each element of B_0. Thus $A = (\bar{B}_0, \bar{B}_1, \ldots, \bar{B}_{t-1})$ form the first $3t$ periods of strongly balanced URMD $(t,t^2,\lambda_2 t)$. Then the rest $2(m-1)t$ periods of this design are constructed following the same steps as in the case of λ_2 even, with the special choice of G_0' as defined in Case (i) as $G_0' = \binom{p+1}{s(p+1)}$. Once strongly balanced URMD $(t,t^2,\lambda_2 t)$, λ_2 odd is constructed, it can be easily checked that nearly strongly balanced GLS d^* for either t^2+t units or $t^2+t(t-1)$ units are constructed following the similar method as in the case of λ_2 even. Here the additional or redundant t units are chosen as the columns of any one of

$$\begin{pmatrix} \bar{B}_h \\ \left.\begin{matrix} B_h \\ \vdots \\ B_h \end{matrix}\right\} \text{repeated } (m-1) \text{ times} \end{pmatrix}'s$$

for $h = 0,1,2,\ldots,t-1$.

Concluding Remarks.

The main purpose of this chapter has been to extend the known optimality results in the context of RMD's under the fixed effects model to the situation of mixed effects model, the unit effects being assumed to be random. It is pleasing to observe that most of the optimality results already proved under fixed effects model remain valid under the mixed effects model. We have incidentally added a few new results on optimality under mixed effects model. A few optimality results recently established by Kunert for balanced and nearly strongly balanced RMD's are found to be quite intricate and the technique used and developed in the present chapter are found to be inadequate to establish the mixed effects analogues of these results, although in case of nearly strongly balanced designs some results have been partially established for the mixed effects models.

CHAPTER 5

NON-ADDITIVE LINEAR MODELS : EFFICIENT ESTIMATION OF NON-ADDITIVE PARAMETERS

5a. Introduction and Literature Review

In practical situations often questioned is the validity of the *simplest linear model showing* additivity of block and varietal effects in two-way classification. However, as is well known, with at most one observation on each block $\times$ variety combination, it is not possible to analyze the data under the usual models involving fixed additive effects with interaction. Thus it appears that if at all interaction has to be checked, one has to evolve a different model in such situations. Tukey (1949) was the first to make an attempt in this direction and the model suggested by him is known in recent literature as Tukey's non-additive model. This model relates to an RBD with b blocks and v varieties with the specific form of interaction term as a constant multiple of the product of corresponding block and treatment effects. More specifically, the model proposed by Tukey is as follows (assuming $\sum_{i=1}^{b} \beta_i = 0,\ \sum_{j=1}^{v} \tau_j = 0)$:

$$y_{ij} = \mu + \beta_i + \tau_j + \theta\beta_i\tau_j + \varepsilon_{ij} \qquad \text{...(5a.1.1)}$$

where θ is the interaction parameter, the other notations/parameters having their usual significance. Tukey provided an estimator of θ as also a test for $H_0 : \theta = 0$ under the usual assumptions on the law of distribution of ε_{ij}'s. Basically, this part of the analysis rests on a study of the ordinary residuals $e_{ij} = y_{ij} - \hat{\mu} - \hat{\beta}_i - \hat{\tau}_j$ where $\hat{\mu}, \hat{\beta}_j, \hat{\tau}_j$ are ordinary LS estimates of the parameters under the usual additive model (i.e. without the interaction term). Tukey demonstrated that it is indeed possible to provide a valid F-test for the above hypothesis. Ward and Dick (1952) also came up with the above finding. Ghosh and Sharma (1963) studied the properties of the power function of Tukey's test. The generalization to LSD's was carried through by Mandel (1959).

Tukey's non-additive model (5a.1.1) has been generalized in two directions by later authors.

Scheffe (1959) hinted on one type of generalization which was, of late, taken up by Milliken and Graybill (1970). These latter authors considered an extension of the general linear hypothesis model in its most general form. The implication of such a study with reference to an RBD is the following. Instead of just one term $\theta\beta_i\tau_j$ in (5a.1.1) to describe non-additivity, one could introduce more terms. As a matter of fact, one could consider a model of the type :

$$y_{ij} = \mu + \beta_i + \tau_j + \theta_1 f_1 + \theta_2 f_2 + \dots + \theta_t f_t + \varepsilon_{ij} \qquad \dots(5a.1.2)$$

where $f_1, f_2, \dots, f_t$ are <u>any</u> known functions of β_i and/or τ_j. Again, an F-test for $H_o : \underline{\theta} = \underline{0}$ has been derived under the usual assumption on the law of distribution of ε_{ij}'s. (It is somewhat uncommon to note that testing problem has been formulated and solved without any reference to the estimation of $\underline{\theta}$ as such).

The non-additive models of the type (5a.1.1) or (5a.1.2) are not always satisfactory and may fail to reveal the true nature of interaction in some cases. Consider, for example, a hypothetical situation where the block effects are the same, as also the varietal effects are same. Still there may exist a certain degree of block x variety interaction. In such a case, the above models become inappropriate. It thus appears that in order that the non-additive terms may properly reflect the extent of interaction, the functions $f_1, f_2, \ldots, f_t$ in (5a.1.2) should not just involve β_i and τ_j. Instead they should possibly be functions of some other quantities like u_i and/or v_j (seemingly unrelated with (β_i, τ_j)). This motivated Mandel (1969) to make another generalization of Tukey's model. He replaced (5a.1.1) or (5a.1.2) by a model of the form :

$$y_{ij} = \mu + \beta_i + \tau_j + \theta_1 u_i^{(1)} v_j^{(1)} + \theta_2 u_i^{(2)} v_j^{(2)} + \ldots + \theta_t u_i^{(t)} v_j^{(t)} + \varepsilon_{ij} \quad \ldots(5a.1.3)$$

where u_i's and v_j's are assumed to be standardized but otherwise unknown and, moreover, not functionally related to β_i's and/or τ_j's. Almost at the same time Gollob (1968) also independently developed a similar model which combines features of factor analysis with ANOVA technique. The problem of estimation of θ_i's, u_i's and v_j's in (5a.1.3) has been tackled satisfactorily by applying the LS technique again on the ordinary residuals e_{ij}'s. It turns out that analysis of the residuals is very much similar to the principal component analysis. Mandel (1961, 1969) also discussed some variants of the model in (5a.1.3) by incorporating special choices of u's and v's. Recently, Kettenring (1983) discussed some applications of such models with real data.

Apart from these, assuming such types of non-additive models, a series of papers have been devoted to the problem of testing for interaction and estimation of error variance σ^2. (Johnson and Graybill (1972a, 1972b) Hegemann and Johnson (1976), Yochmowitz and Cornell (1978), Marasinghe and Johnson (1981, 1982)).

Going back to the models of the type (5a.1.1) or (5a.1.2), surprisingly, nothing noteworthy is available in the literature concerning efficient estimation of non-additive parameter(s). Consider Tukey's non-additive model applied to the general block design set-up, so that

$$y_{ij} = \mu + \beta_i + \sum_{h=1}^{v} \delta_{ij.h}\tau_h + \theta\beta_i(\Sigma\delta_{ij.h}\tau_h) + \varepsilon_{ij}$$

$$1 \leq j \leq k_{di},\ 1 \leq i \leq b \qquad \ldots(5a.1.4)$$

where k_{di} is i^{th} block size for a design d, and $\delta_{ij.h} = 1$, if h^{th} treatment occurs in j^{th} plot of i^{th} block; $= 0$, otherwise.

We assume further that $\sum_i \beta_i = 0, \sum_h \tau_h = 0$.

In this paper, we particularly take up this model and point out that the non-additive parameter is not necessarily estimable (through the analysis of residuals) for any arbitrary choice of the design (section 5b). With this observation, we next provide a simple characterization of designs allowing estimation of θ (subsection 5b.2). Regarding the model of the type (5a.1.2) assuming special forms of f_i's, we make an investigation on estimability of the corresponding θ_i's (subsection 5b.3). We also try to develop reasonable optimality criteria and make a relative comparison of designs for efficient estimation of θ (section 5c).

5b. Estimability of Non-additive Parameter

5b.1 Concept of Estimability of $\underset{\sim}{\theta}$ under General Linear Model

Following Milliken and Graybill (1970) (as also Kshirsagar (1983)) we may consider a general linear model

$$\underset{\sim}{Y}_{N\times 1} = X_{N\times p}\,\underset{\sim}{\gamma}_{p\times 1} + F_{N\times k}\underset{\sim}{\theta}_{k\times 1} + \underset{\sim}{\varepsilon} \qquad \ldots(5b.1.1)$$
$$E(\underset{\sim}{\varepsilon}) = 0, \quad D(\underset{\sim}{\varepsilon}) = \sigma^2 I_N .$$

where $\underset{\sim}{\gamma}$ = vector of unknown (additive) parameters.

$\underset{\sim}{\theta}$ = vector of interaction parameters.

F = matrix associated with the interaction parameters.

The functional forms of elements of F, $f_{ij}(\cdot)$ say, are known and in general, they are arbitrary functions of estimable parametric functions of $\underset{\sim}{\gamma}$ under the simple linear model

$$\underset{\sim}{Y} = X\underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} \qquad \ldots(5b.1.2)$$

Estimation of $\underset{\sim}{\theta}$ from the model (5b.1.1) as such is formidable as basically it is non-linear in $\underset{\sim}{\gamma}$. One may adopt the ad hoc procedure (explained, for example, in Kshirsagar (1983)), of getting rid of $\underset{\sim}{\gamma}$ and converting the model (5b.1.1) to one, which is a linear model involving $\underset{\sim}{\theta}$. Thus, premultiplying (5b.1.1) by $(I-P_X)$, the orthogonal projection operator of X, we get

$$(I-P_X)\underset{\sim}{Y} = (I-P_X)F\underset{\sim}{\theta} + (I-P_X)\underset{\sim}{\varepsilon} \qquad \ldots(5b.1.3)$$

using the fact that $(I-P_X)X$ is null. Denoting $(I-P_X)\underset{\sim}{Y}$, $(I-P_X)F$ and $(I-P_X)\underset{\sim}{\varepsilon}$ by $\underset{\sim}{Z}$, M and $\underset{\sim}{\varepsilon}^*$ respectively, the model (5b.1.3) can be rewritten as

$$\underset{\sim}{Z} = M\underset{\sim}{\theta} + \underset{\sim}{\varepsilon}^{*} \qquad \ldots(5b.1.4)$$

where $E(\underset{\sim}{\varepsilon}^{*}) = \underset{\sim}{0}$, $D(\underset{\sim}{\varepsilon}^{*}) = \sigma^2(I - P_X) = \Sigma$, say. Assume momentarily that the matrix M (even though its elements involve the unknown parameter $\underset{\sim}{\gamma}$ through F) is completely known. Following Rao (1965, Chapter 4, section 4i.4), we get formally,

$$\begin{aligned} \hat{\underset{\sim}{\theta}} &= (M'\Sigma^{-} M)^{-} M'\Sigma^{-}\underset{\sim}{Z} \quad \text{as} \quad \mathcal{M}(\Sigma) \supseteq \mathcal{M}(M) \\ &= [\, F'(I - P_X)F \,]^{-} F'(I - P_X)\underset{\sim}{Y} \qquad \ldots(5b.1.5) \end{aligned}$$

In practice, for estimation and testing purpose, we replace the functions of $x\underset{\sim}{\gamma}$ in F by their blue $x\hat{\underset{\sim}{\gamma}}$, obtained under the model (5b.1.2). In other words, we compute

$$\hat{\hat{\underset{\sim}{\theta}}} = [\, F'_{(x\hat{\gamma})}(I - P_X)F_{(x\hat{\gamma})} \,]^{-} F'_{(x\hat{\gamma})}(I - P_X)\underset{\sim}{Y} \qquad \ldots(5b.1.6)$$

(here $F_{(x\hat{\gamma})}$ denotes F with $x\underset{\sim}{\gamma}$ replaced by $x\hat{\underset{\sim}{\gamma}}$). At this stage, one might wonder as to whether $\underset{\sim}{\theta}$ is as such estimable or not irrespective of the choice of $\underset{\sim}{\gamma}$ in the relevant parameter space. To settle this, we first incorporate the following formal definition of estimability of $\underset{\sim}{\theta}$.

<u>DFN 5b.1.1</u> Under model (5b.1.1), the interaction parameter vector $\underset{\sim}{\theta}$ is estimable iff

$$\text{rank}\,\{[\, F'(I - P_X)F \,]\} = k \quad \text{i.e.,} \quad F'(I - P_X)F \text{ is nonsingular for}$$

all choices of $\underset{\sim}{\gamma}$ in the relevant parameter space.

The next theorems justify some intuitive feelings regarding estimability of $\underline{\theta}$.

Theorem 5b.1.1 Suppose $\underline{\theta}$ is estimable (in the above sense) under the model

$$\underline{Y} = X_{N\times p}\,\underline{\gamma}_{p\times 1} + F_{N\times k}\underline{\theta}_{k\times 1} + \underline{\varepsilon}_1$$

Then $\underline{\theta}$ is also estimable under an extended set-up with additional observations (involving the same $\underline{\gamma}$-parameter as in the first set of model expectation).

Proof : Clearly, it suffices to prove the result with one additional observation. Let the extended set-up be

$$\underline{Y} = X\underline{\gamma} + F\underline{\theta} + \underline{\varepsilon}_1$$

$$y = \underline{\lambda}'\underline{\gamma} + \underline{g}'\underline{\theta} + \varepsilon_2$$

$$\underline{\varepsilon} = \begin{pmatrix} \underline{\varepsilon}_1 \\ \varepsilon_2 \end{pmatrix} \sim N(o, \sigma^2 I)$$

Let $\textcircled{H}$ be the relevant parameter space of $\underline{\gamma}$.

Now under the set-up,

$$\underline{Y} = X\underline{\gamma} + F\underline{\theta} + \underline{\varepsilon}_1$$

estimability of $\underline{\theta}$ is ensured by the fact that

$$\text{rank}\left\{\left[F'(I - P_X)F\right]\right\} = k \quad \text{for all} \quad \underline{\gamma} \in \textcircled{H}$$

Thus, whenever this happens, we have to show

$$\text{rank}\left\{\left[I - P_{\binom{X}{\underline{\lambda}'}}\right]\begin{bmatrix} F \\ \underline{g}' \end{bmatrix}\right\} = k \quad \text{for all} \quad \underline{\gamma} \in \textcircled{H}.$$

<u>Case (i)</u> $\underset{\sim}{\ell} \notin \mathcal{M}(x')$

Then,

(a) $(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-}$ is a g-inverse of $x'x$

(b) $\underset{\sim}{\ell}'(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} x' = 0$

(c) $\underset{\sim}{\ell}'(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} \underset{\sim}{\ell} = 1$

Though these three results seem to be fairly simple, for the sake of completeness we sketch below a brief outline of the proofs.

$$(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} x'x = x'x \quad [\text{ vide Rao-Mitra (2.2)}]$$

i.e. $\underset{\sim}{\ell}\underset{\sim}{\ell}'(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} x'x = x'x[\, I - (x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} x'x \,]$.

As $\mathcal{M}(\underset{\sim}{\ell}) \cap \mathcal{M}(x') = \{\underset{\sim}{0}\}$, the above identity implies

$$x'x[\, I - (x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} x'x \,] = \underset{\sim}{\ell}\underset{\sim}{\ell}'(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} x'x = 0$$

Writing $\underset{\sim}{\ell} = \underset{\sim}{\ell}\underset{\sim}{\ell}'D_1$, and $x' = x'xD_2$ for some D_1 and D_2,

(a) and (b) follow immediately.

Further, (c) follows as a consequence of

$(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-} \underset{\sim}{\ell}\underset{\sim}{\ell}' = \underset{\sim}{\ell}\underset{\sim}{\ell}'$ and through an application of the result stated in (b).

Thus, whenever $\underset{\sim}{\ell} \notin \mathcal{M}(x')$

$$\left[\, I - P_{\binom{x}{\underset{\sim}{\ell}'}} \right] \begin{bmatrix} F \\ \underset{\sim}{g}' \end{bmatrix} = \begin{bmatrix} (I - P_x)F \\ 0 \end{bmatrix} \qquad \ldots(5b.1.7)$$

and hence,

$$\begin{aligned} \text{rank} \left\{ \left[\, I - P_{\binom{x}{\underset{\sim}{\ell}'}} \right] \begin{bmatrix} F \\ \underset{\sim}{g}' \end{bmatrix} \right\} &= \text{rank} \{ (I - P_x)F \} \\ &= k \quad \text{for all } \underset{\sim}{\gamma} \in \text{Ⓗ}. \end{aligned}$$

Case (ii) $\underset{\sim}{\ell} \in \mathcal{M}(x')$

Then, $\underset{\sim}{\ell} = x'x\underset{\sim}{\eta}$ for some $\underset{\sim}{\eta}$.

Now $\left[I - P_{\binom{x}{\underset{\sim}{\ell}'}}\right] = I - \binom{x}{\underset{\sim}{\ell}'}(x'x + \underset{\sim}{\ell}\underset{\sim}{\ell}')^{-}(x' : \underset{\sim}{\ell})$

$$= I - \binom{x}{\underset{\sim}{\ell}'}\left[(x'x)^{-} - \frac{(x'x)^{-}\underset{\sim}{\ell}\underset{\sim}{\ell}'(x'x)^{-}}{1+\underset{\sim}{\ell}'(x'x)^{-}\underset{\sim}{\ell}}\right](x' : \underset{\sim}{\ell})$$

$$= \begin{bmatrix} I - P_x + \dfrac{x\underset{\sim}{\eta}\underset{\sim}{\eta}'x'}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}} & \dfrac{-x\underset{\sim}{\eta}}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}} \\ \dfrac{-\underset{\sim}{\eta}'x'}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}} & \dfrac{1}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}} \end{bmatrix}$$

$$= \begin{bmatrix} I & -x\underset{\sim}{\eta} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I - P_x & 0 \\ \dfrac{-\underset{\sim}{\eta}x'}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}} & \dfrac{1}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}} \end{bmatrix} \quad \ldots(5b.1.8)$$

Since $\begin{bmatrix} I & -x\underset{\sim}{\eta} \\ 0 & 1 \end{bmatrix}$ is nonsingular, from (5b.1.8) we get,

$$\text{rank}\left\{\left[I - P_{\binom{x}{\underset{\sim}{\ell}'}}\right]\begin{bmatrix} F \\ \underset{\sim}{g}' \end{bmatrix}\right\} = \text{rank}\left\{\begin{bmatrix} (I - P_x)F \\ \underset{\sim}{\delta} \end{bmatrix}\right\} \quad \ldots(5b.1.9)$$

where $\underset{\sim}{\delta} = \dfrac{1}{1+\underset{\sim}{\eta}'x'x\underset{\sim}{\eta}}(-\underset{\sim}{\eta}'x'F + \underset{\sim}{g}')$

Now it is readily seen that

$$\text{rank}\left\{\begin{bmatrix} (I - P_x)F \\ \underset{\sim}{\delta} \end{bmatrix}\right\} = k \quad \text{for all } \underset{\sim}{\gamma} \in (H)$$

as rank $\{(I - P_x)F\} = k$ for all $\underset{\sim}{\gamma} \in (H)$.

Hence, the result.

<u>Theorem 5b.1.2</u> Suppose $\underset{\sim}{\theta}$ is <u>not</u> estimable under the model

$$\underset{\sim}{Y} = x\underset{\sim}{\gamma} + F\underset{\sim}{\theta} + \underset{\sim}{\varepsilon}_1$$

then, $\underline{\theta}$ is also not estimable under the extended set-up

$$\underline{Y} = X\underline{\gamma} + F\underline{\theta} + \underline{\varepsilon}_1$$

$$y = \underline{e}_i'X\underline{\gamma} + \underline{e}_i'F\underline{\theta} + \varepsilon_2$$

where $\underline{e}_i$ is a $N \times 1$ vector with 1 in i^{th} position and 0 elsewhere.

Proof : Since $\underline{\theta}$ is not estimable under the model

$$\underline{Y} = X\underline{\gamma} + F\underline{\theta} + \underline{\varepsilon}_1$$

$$\text{rank}\left\{\left[(I-P_X)F\right]\right\} < k \quad \text{for some} \quad \underline{\gamma}_0 \in (H)$$

Now from (5b.1.9),

$$\text{rank}\left\{\left[I - P_{\begin{pmatrix}X\\ \underline{e}_i'X\end{pmatrix}}\right]\begin{bmatrix}F\\ \underline{e}_i'F\end{bmatrix}\right\} = \text{rank}\left\{\begin{bmatrix}(I-P_X)F\\ \cdots\cdots\cdots\\ c\cdot\underline{e}_i'((I-P_X)F)\end{bmatrix}\right\} \text{ where } c = \frac{1}{\underline{e}_i'X(X'X)^{-}X'\underline{e}_i+1}$$

$$= \text{rank}\left\{\left[(I-P_X)F\right]\right\} < k \quad \text{for} \quad \underline{\gamma}_0 \in (H)$$

Hence, the result.

In the context of block designs, an application of these theorems simplifies verification of estimability of the interaction parameter (as a scalar) in Tukey's non-additive model and also of the vector parameter $\underline{\theta}$ in the generalized set-up.

5b.2 Estimability of θ under Tukey's Model in General Block Design Set-up

Suppose under the model (5a.1.4), $\beta_i = 0$ ($\tau_h = 0$) for some $i(h)$. This means that for the design d, for each of the $k_{di}(r_{dh})$ observations in the i^{th} block (involving h^{th} treatment) the multiplicative interaction term would vanish and hence, no information on the non-additive parameter θ

would be available from this i^{th} block (the set of observations under h^{th} treatment). Again, if $\beta_i = \beta_{i'}(\tau_h = \tau_{h'})$ this would mean that for estimation of θ one has to effectively work with a model with the treatment - block incidence matrix suitably modified in the sense that the corresponding columns (rows) of the original incidence matrix are merged to form a new column (row). Therefore, the relevant parameter space can be and will be taken as

$$\begin{rcases} \text{H} = \Big\{(\mu, \underline{\tau}, \underline{\beta}) : -\infty < \mu < \infty, \underline{\beta}'\underline{1} = 0, \underline{\tau}'\underline{1} = 0, \\ \quad \beta_i \neq \beta_{i'} \neq 0 \ \text{ for all } \ i \neq i', 1 \leq i, i' \leq b, \\ \quad \tau_h \neq \tau_{h'} \neq 0 \ \text{ for all } \ h \neq h', 1 \leq h, h' \leq v \end{rcases} \quad \ldots(5b.2.1)$$

Without any loss of generality, we will be working with connected designs only.

Arranging the observations serially blockwise, the model (5a.1.4) can be rewritten as

$$\underline{Y} = [\, \underline{1} : x_\beta : x_\tau \,] \begin{bmatrix} \mu \\ \underline{\beta} \\ \underline{\tau} \end{bmatrix} + \theta \underline{f} + \underline{\varepsilon}$$

$$= x\underline{\gamma} + \theta \underline{f} + \underline{\varepsilon} \quad \text{say},$$

where $\underline{1}$ is $N \times 1$ vector of all 1's

x_β is $N \times b$ coefficient matrix corresponding to the block effects

x_τ is $N \times v$ coefficient matrix corresponding to the treatment effects

x is $N \times (b+v+1)$ coefficient matrix $[\, \underline{1} : x_\beta : x_\tau \,]$

$\underline{\gamma}$ is $\overline{b+v+1} \times 1$ parameter vector $\begin{bmatrix} \mu \\ \underline{\beta} \\ \underline{\tau} \end{bmatrix}$ and $\underline{\gamma} \in$ H

$\underline{f} = ((\beta_i \sum_h \delta_{ij.h}\tau_h))$ is the $N \times 1$ column vector associated with the interaction parameter θ.

Now, the question of estimability of θ through an analysis of the residuals $\underline{Z}$ (vide 5b.1.4) is equivalent to the following : Is the column vector $(I - P_x)\underline{f}$ non-null for all γ in Ⓗ for <u>any</u> choice of the design matrix x ? The answer is hopelessly in the negative as we will demonstrate below. For this we use the following property of $(I - P_x)$, in the context of block-designs.

<u>Property 1</u> : The form of $x = [\underline{1} : x_\beta : x_\tau]$ above indicates $(I - P_x)x_\beta = 0$, $(I - P_x)x_\tau = 0$. That is, in each row of $(I - P_x)$ the sum of entries corresponding to each block and each treatment is zero.

The following lemma demonstrates that it suffices to settle estimability of θ in a binary design.

<u>Lemma 5b.2.1</u> Let $d(v,b,k_{d1}, \ldots ,k_{db})$ be any connected block design reduced to $\bar{d}(v,b,k_{\bar{d}1}, \ldots ,k_{\bar{d}b})$ as follows :

$$\left.\begin{array}{ll} n_{\bar{d}hi} = 1 & \text{if } n_{dhi} \geq 1 \\ \quad\;\; = 0 & \text{otherwise} \end{array}\right\} \quad 1 \leq h \leq v, \; 1 \leq i \leq b.$$

Then θ is estimable under d iff θ is estimable under $\bar{d}$. (In our latter discussions we will refer to such $\bar{d}$ as binary reduction of d).

<u>Proof</u> : We first note that connectedness of d retains connectedness of $\bar{d}$.

" If part " :

Follows as an immediate application of Theorem 5b.1.1. (It may be noted that the condition for applicability of Theorem 5b.1.1 written within parenthesis is readily verified here).

" Only if part " :

We can reconstruct d starting from $\bar{d}$ by adding the observations one by one for all treatments appearing more than once in a block. For every such observation, the row vector in the coefficient matrix coincides with one row of the coefficient matrix X_o in $\bar{d}$ and hence the result follows from Theorem 5b.1.2.

The essence of this theorem is that if θ is not estimable in a design, then θ is also not estimable in the extended design with additional observations corresponding to the same block x treatment combinations which are already present in the original design.

The following theorem provides a simple characterization of a large class of connected block designs ensuring estimability of θ.

Theorem 5b.2.1 Under Tukey's model, applied to the general block-design set-up (5a.1.4), a connected block design $d(v,b,k_{d1},k_{d2},\ldots,k_{db})$ will provide estimation of θ whenever at least one pair of treatments (h,h') say, occur in two different blocks.

Proof : The proof is by contradiction. Suppose h and h' occur together in i^{th} and i'^{th} blocks. Let

$d_o =$

Blocks	Treatments	
i	h	h'
i'	h	h'

be the subdesign of d formed by only i^{th} and i'^{th} blocks, and h and h'^{th} treatments with the incidence structure

$$n_{d_o hi} = n_{d_o h'i} = n_{d_o hi'} = n_{d_o h'i'} = 1$$

Let the partitions in x and $\underline{f}$ corresponding to the subdesign d_o be respectively x_o and $\underline{f}_o$.

Now, suppose θ is not estimable in the original design d. Then there exists a $\underline{\gamma}_o$ in the parameter space Ⓗ such that for

$$x = \begin{bmatrix} x_o \\ x_1 \end{bmatrix} \quad \text{and} \quad f = \begin{bmatrix} \underline{f}_o \\ \underline{f}_1 \end{bmatrix},$$

$$\left[I - P_{\begin{pmatrix} x_o \\ \cdots \\ x_1 \end{pmatrix}} \right] . \begin{bmatrix} \underline{f}_o \\ \cdots \\ \underline{f}_1 \end{bmatrix} = \underline{0} \quad \text{for } \underline{\gamma}_o \varepsilon \text{ Ⓗ}$$

$$\Rightarrow \begin{pmatrix} \underline{f}_o \\ \cdots \\ \underline{f}_1 \end{pmatrix} \varepsilon \mathcal{M} \begin{pmatrix} x_o \\ \cdots \\ x_1 \end{pmatrix} \quad \text{for } \underline{\gamma}_o \varepsilon \text{ Ⓗ}$$

$$\Rightarrow \underline{f}_o \varepsilon \mathcal{M}(x_o) \quad \text{for } \underline{\gamma}_o \varepsilon \text{ Ⓗ}$$

$$\Rightarrow (I - P_{x_o})\underline{f}_o = \underline{0} \quad \text{for } \underline{\gamma}_o \varepsilon \text{ Ⓗ}.$$

However, $(I - P_{x_o})$ assumes a very simple form as rank $\{(I - P_{x_o})\} = 1 =$ error d.f. under d_o. Using Property 1 (vide subsection 5b.2) one can develop any 1×4 row vector of $(I - P_{X_o})$, say i^{th} row vector as

$$a_i(1, -1 : -1, 1)$$

where a_i's are real numbers, not all zeros. The four positions here correspond to observations serially placed in the order in which the plots appear in the blocks of d_o. Note that in $(I - P_{X_o})$, sum of the entries corresponding to each block and each treatment is zero. As the elements of $\underline{f}_o$ are products of corresponding treatment and block effects, a typical element of $(I - P_{X_o})\underline{f}_o$ is

$$a_i \{(\beta_i - \beta_{i'})(\tau_h - \tau_{h'})\} \qquad \ldots(5b.2.2)$$

and thus $(I - P_{X_o})\underline{f}_o = \underline{0}$ for $\underline{\gamma}_o \in \text{Ⓗ}$ implies

$$(\beta_i - \beta_{i'})(\tau_h - \tau_{h'}) = \underline{0} \quad \text{for} \quad \underline{\gamma}_o$$

which can only happen if for $\underline{\gamma}_o$, $\beta_i = \beta_{i'}$ and/or $\tau_h = \tau_{h'}$ holds. This leads to a contradiction in the description and coverage of the parameter space Ⓗ (vide(5b.2.1)).

Hence the result.

Remark 1 Though θ is estimable under the subdesign d_o, it does not straightway guarantee estimability of θ under the original design d via Theorem 5b.1.1 as in the subdesign d_o and original design d, the two parameter spaces are different.

Theorem 5b.2.2 : Let $d(v, b, k_{d1}, k_{d2}, \ldots, k_{db})$ be a connected block design for which each elementary treatment contrast has a unique (unbiased) estimator under the simple linear model (5b.1.2). Then θ is not estimable from d.

Proof : We first establish that the implication of the condition of the theorem is that $(I - P_X)$ is a null matrix so that the claim is trivially justified.

It is evident that under the given hypothesis, from i^{th} block of d, we are getting exactly $k_{di} - 1$ distinct independent elementary treatment contrasts. Again, any estimable treatment contrast is a linear function of the within block elementary treatment contrasts. Since d is a connected design, we conclude that under the given hypothesis,

$$(k_{d1} - 1) + (k_{d2} - 1) + \dots + (k_{db} - 1) = v - 1.$$

$$\text{i.e. } \sum_{i=1}^{b} k_{di} = b + v - 1.$$

$$\text{i.e. rank } \{(I - P_X)\} = (\Sigma k_{di} - 1) - (b - 1) - (v - 1)$$

$$= 0$$

Hence, the result.

Corollary 5b.2.2 Let the binary reduction $\bar{d}(v,b,k_{\bar{d}1},\dots,k_{\bar{d}b})$ of a design $d(v,b,k_{d1},\dots,k_{db})$ as defined in Lemma 5b.2.1 be such that for $\bar{d}$, each of the elementary treatment contrasts has unique (unbiased) estimator under model (5b.1.2). Then θ is not estimable under d.

The proof immediately follows from an application of Lemma 5b.2.1 and Theorem 5b.2.2.

Remark 2 : There are still many designs which neither satisfy the condition of Theorem 5b.2.1 (ensuring estimability of θ) nor do they satisfy the condition of Corollary 5b.2.2 (leading to non-estimability of θ). For example, for a BIBD with $\lambda \geq 2$, the condition of Theorem 5b.2.1 is satisfied, hence θ is estimable; on the other hand, for a BIBD with $\lambda = 1$, estimability of θ could not be settled. However, we have a strong feeling (as indicated by several examples) that the " sufficient " condition stated in Theorem 5b.2.1 will turn out to be " necessary " also. Below we demonstrate two examples in justification of our conjecture.

Example 1. BIBD ($b = v = 3$, $\lambda = 1$).

Blocks	Varieties	
1	1	2
2	1	3
3	2	3

For this design, rank $\{(I - P_x)\} = 1$.

Hence, using Property 1 (vide subsection 5b.2) of $(I - P_x)$,

$$(I - P_x)\underline{f} = \underline{0} \Leftrightarrow \beta_1(\tau_1 - \tau_2) + \beta_2(\tau_3 - \tau_1) + \beta_3(\tau_2 - \tau_3) = 0.$$

Clearly, $\underline{\beta} = (1, -2/3, -1/3)'$ and $\underline{\tau} = (1, 2, -3)'$ is a solution to this equation.

Hence θ is <u>not</u> estimable under such a design.

Example 2 :

d :

Blocks	Varieties		
1	1	2	3
2	2	4	6
3	3	4	5
4	1	5	6

$\{1, 2, \ldots, 6\}$ denotes the set of treatments.

d is a design consisting of 4 blocks from the BIBD($b = v = 7$, $r = k = 3$, $\lambda = 1$).

For this design, rank $\{(I - P_X)\} = 3$.

Using Property 1 (vide subsection 5b.2), any typical row of $(I - P_X)$ is of the form,

$$(a_i, b_i, -(a_i+b_i) \vdots -b_i, c_i, b_i - c_i \vdots a_i + b_i, -c_i, c_i - a_i - b_i \vdots$$

$$-a_i, a_i + b_i - c_i, c_i - b_i)$$

where all a_i, b_i and c_i are real numbers and not equal to zero simultaneously.

Now, taking a choice of $\underline{\tau}$ and $\underline{\beta}$ as

$$\underline{\tau} = (1, -1, 2, -4, 4, -2)'$$

and $$\underline{\beta} = (-11, 13, 1, -3)'$$

one can easily verify that $(I - P_X)\underline{f} = \underline{0}$. Thus, θ is <u>not</u> estimable under this design.

Moreover, in view of Lemma 5b.2.1 any design d, for which binary reduction is one of the above two designs, does not provide any estimator of θ in the sense indicated here.

5b.3. Estimation of Interaction-parameter Vector $\underline{\theta}$ in a General Block Design

In this section we take up a model of the type (5a.1.2) in the context of general block design set-up and focus our attention on the problem of estimation of interaction-parameter vector $\underline{\theta}$.

Take for example, a non-additive model of the type

$$y_{ij} = \mu + \beta_i + \Sigma \delta_{ij.h}\tau_h + \theta_1\beta_i \sum_h \delta_{ij.h}\tau_h + \theta_2\beta_i^2 + \theta_3 \sum_h \delta_{ij.h}\tau_h^2 + \varepsilon_{ij} \quad \text{...(5b.3.1)}$$

Writing the observations blockwise in order, the model can be rewritten as

$$\underline{Y} = [1 : x_\beta : x_\tau] \begin{bmatrix} \mu \\ \underline{\beta} \\ \underline{\tau} \end{bmatrix} + [F_1 \; F_2 \; F_3] \begin{bmatrix} \underline{\theta}_1 \\ \underline{\theta}_2 \\ \underline{\theta}_3 \end{bmatrix} + \underline{\varepsilon}$$

$$= x\underline{\gamma} + F\underline{\theta} + \underline{\varepsilon}$$

where x_β and x_τ are the same as defined in subsection 5b.2. It can be easily seen that under this model, $F_2 = x_\beta \underline{\beta}^{(2)}$ and $F_3 = x_\tau \underline{\tau}^{(2)}$ where the elements of $\underline{\beta}^{(2)}$ and $\underline{\tau}^{(2)}$ are β_i^2 and τ_h^2 respectively. Thus Property 1 of $(I - P_x)$ (vide subsection 5b.2) yields that both the second and third columns of $(I - P_x)F$ are null vectors. Hence the following lemma.

Lemma 5b.3.1. Under the model (5b.3.1) for no block design, θ_2 and θ_3 are estimable.

This shows that it is not generally true that all the components of $\underline{\theta}$ in the model of the type (5a.1.2) are estimable through the analysis of residuals. Scheffe (1959) developed a different argument to ascertain a similar result in the context of RBD set-up. As a matter of fact in view of Property 1 of $(I - P_X)$, it turns out that whatever the block-design set-up adopted in a non-additive model of the type (5b.1.1) the interaction parameters corresponding to higher powers of β_i or τ_h alone cannot be estimated. These limitations in studying general models of the type (5b.1.1) cannot be overlooked.

What if instead we introduce more terms involving both β_i and τ_h in (5b.3.1) ? Take, for example, the model involving up to second powers of τ_h's and β_i's i.e.

$$y_{ij} = \mu + \beta_i + \Sigma\delta_{ij.h}\tau_h + \theta_1\beta_i\,\Sigma\delta_{ij.h}\tau_h + \theta_2\beta_i^2\,\Sigma\delta_{ij.h}\tau_h$$

$$+ \theta_3\beta_i\,\Sigma\delta_{ij.h}\tau_h^2 + \theta_4\beta_i^2\,\Sigma\delta_{ij.h}\tau_h^2 + \varepsilon_{ij} \qquad \ldots(5b.3.2)$$

Clearly, not all designs will provide estimation of $\underline{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)'$ as we have seen in case of single interaction parameter.

Next we characterize a class of designs which will provide estimation of $\underline{\theta}$ under model (5b.3.2). Following the same arguments given in subsection 5b.2, under (5b.3.2) also, the relevant parameter space can be taken to be the same as (H). Following essentially the same arguments as those given in the proofs of Lemma 5b.2.1, Theorem 5b.2.2

and Corollary 5b.2.2, one gets the following results concerning estimation of interaction parameter vector $\underset{\sim}{\theta}$.

Lemma 5b.3.2 Let $\bar{d}(v, b, k_{\bar{d}1}, \ldots, k_{\bar{d}b})$ be a binary reduction of $d(v, b, k_{d1}, \ldots, k_{db})$ as defined in Lemma 5b.2.1. Then, $\underset{\sim}{\theta}$ is estimable under d iff $\underset{\sim}{\theta}$ is estimable under $\bar{d}$.

Theorem 5b.3.1 Let $d(v, b, k_{d1}, \ldots, k_{db})$ be a connected block design for which each elementary treatment contrast has unique (unbiased) estimator under simple linear model (5b.1.2). Then $\underset{\sim}{\theta}$ is not estimable under d.

Corollary 5b.3.2 Let the binary reduction $\bar{d}(v, b, k_{\bar{d}1}, k_{\bar{d}2}, \ldots, k_{\bar{d}b})$ of a design $d(v, b, k_{d1}, k_{d2}, \ldots, k_{db})$ as defined in Lemma 5b.2.1 be such that for $\bar{d}$, each of the elementary treatment contrasts has unique (unbiased) estimator under the model (5b.1.2). Then $\underset{\sim}{\theta}$ is not estimable under d.

The following theorem establishes estimability of $\underset{\sim}{\theta}$ under RBD, and using this result the next theorem gives a sufficient condition under which $\underset{\sim}{\theta}$ is estimable.

Theorem 5b.3.2 An RBD with $b \geq 3$ and $v \geq 3$ provides estimation of $\underset{\sim}{\theta}$ under the model (5b.3.2).

Proof . Since $F'(I - P_X)F$ is invariant under any rearrangement of treatments among the plots within a block, without loss of generality, we assume that in the RBD considered, the j^{th} plot within any block receives the j^{th} treatment $(j = 1, 2, \ldots, v)$. Thus, writing the observations blockwise and in order, $(I - P_X) = (I - \frac{J}{b}) \otimes (I - \frac{J}{v})$. Under model (5b.3.2), $[(I - P_X)F]$

is a $N \times 4$ matrix, $(N = bv)$ and let its p^{th} row $(p = (i-1)b+j)$ corresponding to the observation in j^{th} plot of i^{th} block, be denoted by $\underline{\alpha}'_{ij}$. It is not difficult to verify that this typical row takes the form,

$$\underline{\alpha}'_{ij} = (\beta_i\tau_j \, , \tau_j(\beta_i^2 - \sum_{t=1}^{b} \frac{\beta_t^2}{b}) \, , \beta_i(\tau_j^2 - \sum_{\ell=1}^{v} \frac{\tau_\ell^2}{b}) \, , (\beta_i^2 - \sum_{t=1}^{b} \frac{\beta_t^2}{b})(\tau_j^2 - \sum_{\ell=1}^{v} \frac{\tau_\ell}{v}))$$

Now, to ensure estimability of $\underline{\theta}$ under RBD, we have to establish that

$$\text{rank}\,\{(I - P_X)F\} = 4 \quad \text{for all } \underline{\gamma} \,\varepsilon\, (H).$$

For this, we consider only one 9×4 submatrix A of $[(I - P_X)F]$ formed by the rows corresponding to $i, j = 1,2,3$.

contd..

$$A = \begin{bmatrix}
\beta_1\tau_1 & \tau_1\left(\beta_1^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_1\left(\tau_1^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_1^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_1^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_1\tau_2 & \tau_2\left(\beta_1^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_1\left(\tau_2^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_1^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_2^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_1\tau_3 & \tau_3\left(\beta_1^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_1\left(\tau_3^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_1^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_3^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_2\tau_1 & \tau_1\left(\beta_2^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_2\left(\tau_1^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_2^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_1^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_2\tau_2 & \tau_2\left(\beta_2^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_2\left(\tau_2^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_2^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_2^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_2\tau_3 & \tau_3\left(\beta_2^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_2\left(\tau_3^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_2^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_3^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_3\tau_1 & \tau_1\left(\beta_3^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_3\left(\tau_1^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_3^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_1^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_3\tau_2 & \tau_2\left(\beta_3^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_3\left(\tau_2^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_3^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_2^2 - \frac{\Sigma\tau_\ell^2}{v}\right) \\
\beta_3\tau_3 & \tau_3\left(\beta_3^2 - \frac{\Sigma\beta_t^2}{b}\right) & \beta_3\left(\tau_3^2 - \frac{\Sigma\tau_\ell^2}{v}\right) & \left(\beta_3^2 - \frac{\Sigma\beta_t^2}{b}\right)\left(\tau_3^2 - \frac{\Sigma\tau_\ell^2}{v}\right)
\end{bmatrix}$$

Side by side, let us define another 9 vectors $\underline{x}_1$, $\underline{x}_2$, $\underline{x}_3$, $\underline{x}_4$, $\underline{x}_5$, $\underline{x}_6$, $\underline{x}_7$, $\underline{x}_8$, $\underline{x}_9$ by

$$\underline{x}_1 = \underline{1}_3 \otimes \underline{1}_3$$
$$\underline{x}_2 = \underline{1}_3 \otimes \underline{\tau}_o$$
$$\underline{x}_3 = \underline{1}_3 \otimes \underline{\tau}_o^{(2)}$$
$$\underline{x}_4 = \underline{\beta}_o \otimes \underline{1}_3$$
$$\underline{x}_5 = \underline{\beta}_o \otimes \underline{\tau}_o$$
$$\underline{x}_6 = \underline{\beta}_o \otimes \underline{\tau}_o^{(2)}$$
$$\underline{x}_7 = \underline{\beta}_o^{(2)} \otimes \underline{1}_3$$
$$\underline{x}_8 = \underline{\beta}_o^{(2)} \otimes \underline{\tau}_o$$
$$\underline{x}_9 = \underline{\beta}_o^{(2)} \otimes \underline{\tau}_o^{(2)}$$

where $\underline{\tau}_o = (\tau_1, \tau_2, \tau_3)'$, $\underline{\beta}_o = (\beta_1, \beta_2, \beta_3)'$, $\underline{1}_3 = (1,1,1)'$, and $\underline{\tau}_o^{(2)}$ and $\underline{\beta}_o^{(2)}$ consist of elements τ_h^2, β_i^2 respectively, for $h,i = 1,2,3$.

Denote the i^{th} column of A by $\underline{c}_i$. Then,

$$\underline{c}_1 = \underline{x}_5 ,$$

$$\underline{c}_2 = \underline{x}_8 - \frac{\sum_{t=1}^{b} \beta_t^2}{b} \cdot \underline{x}_2$$

$$\underline{c}_3 = \underline{x}_6 - \frac{\sum_{\ell=1}^{v} \tau_\ell^2}{v} \cdot \underline{x}_4$$

and

$$\underline{c}_4 = \underline{x}_9 - \frac{\sum_{t=1}^{b} \beta_t^2}{b} \cdot \underline{x}_3 - \frac{\sum_{\ell=1}^{v} \tau_\ell^2}{v} \cdot \underline{x}_7 + \frac{\sum_{t=1}^{b} \beta_t^2}{b} \cdot \frac{\sum_{\ell=1}^{v} \tau_\ell^2}{v} \cdot \underline{x}_1$$

Thus, it follows at once that whenever the 9 vectors $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_9$ are independent, $\underline{c}_1, \underline{c}_2, \underline{c}_3, \underline{c}_4$ are independent.

$$(\text{as } \Sigma \delta_i \underline{c}_i = \underline{0} \Rightarrow \Sigma \delta_i' \underline{x}_i = 0 \Rightarrow \delta_i' = 0 \Rightarrow \delta_i = 0)$$

However, it is trivially known that the rank of the 9×9 matrix

$$B = \begin{bmatrix} 1 & \beta_1 & \beta_1^2 \\ 1 & \beta_2 & \beta_2^2 \\ 1 & \beta_3 & \beta_3^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & \tau_1 & \tau_1^2 \\ 1 & \tau_2 & \tau_2^2 \\ 1 & \tau_3 & \tau_3^2 \end{bmatrix}$$

formed by these 9 vectors $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_9$ written as columns, is exactly equal to 9 whenever $\tau_1 \neq \tau_2 \neq \tau_3$, and $\beta_1 \neq \beta_2 \neq \beta_3$.

Since for all $\underline{\gamma}$ in (H) (vide (5b.2.1)) β_i's and τ_h's are distinct, the vectors $\underline{x}_1, \cdots, \underline{x}_9$ and hence $\underline{c}_1, \underline{c}_2, \underline{c}_3, \underline{c}_4$ are independent and, consequently,

$$\text{rank}\,\{[(I - P_x)F]\} = 4 \quad \text{for all } \underline{\gamma} \in (H).$$

Hence the result.

<u>Theorem 5b.3.3</u> A connected block design $d(v, b, k_{d1}, k_{d2}, \cdots, k_{db})$ with $b \geq 3$, $v \geq 3$, and $k_{di} \geq 3$ for at least three blocks, will provide estimation of $\underline{\theta}$ under the model (5b.3.2) whenever at least one triplet of treatments (h, h', h'') say, occur in three different blocks.

We omit the proof which follows (through contradiction) essentially along similar line of arguments as that of Theorem 5b.2.1.

Remark 3 : It is difficult to establish that the sufficient condition under Theorem 5b.3.3 is also necessary for estimability of $\underline{\theta}$. Thus, for the designs which are not covered by Theorem 5b.3.3 and Corollary 5b.3.2, estimability of $\underline{\theta}$ could not be settled.

5c. Optimal Estimation of $\underline{\theta}$ in a Block-Design under Tukey's Model.

5c.1 Preliminaries

Now we focus our attention to the problem of efficient estimation of single interaction parameter θ under Tukey's non-additive model (5a.1.4). For fixed N (total number of experimental units) and v (number of treatments), let $\Omega(N,v)$ denote the class of all connected block designs (with block sizes $\leq v$) providing estimation of θ. For a design $d \in \Omega(N,v)$ let x_d denote the underlying x matrix, $v_d(\hat{\hat{\theta}})$ denote the variance of $\hat{\hat{\theta}}$ (vide (5b.1.6)).

Assume at this stage that $\hat{\underline{\tau}}$ and $\hat{\underline{\beta}}$ are consistent for $\underline{\tau}$ and $\underline{\beta}$. This could be achieved, for example, using the existing data set in combination with otherwise available consistent estimators for $\underline{\tau}$ and $\underline{\beta}$ from independent sources. Note that the passage from $\hat{\underline{\theta}}$ in (5b.1.5) to $\hat{\hat{\underline{\theta}}}$ in (5b.1.6) is not affected by use of such independent auxiliary information. Since we are primarily interested in making a relative comparison of various designs $\in \Omega(N,v)$, the above assumption is not unrealistic. On the other hand, this justifies the approximation

$$V_d(\hat{\hat{\theta}}) \approx V_d(\hat{\theta}) = [\underline{f}'(I - P_{X_d})\underline{f}]^{-1}\sigma^2 \quad \text{(vide (5b.1.5))}$$

and it depends on $\underline{\tau}$ and $\underline{\beta}$ through $\underline{f}$. We will make use of this approximate expression for $V_d(\hat{\theta})$ while comparing the relative efficiencies of different designs. For this sort of comparison, one may certainly restrict to the effective parameter space :

$$\textcircled{H}_1 = \Big\{ (\underline{\tau},\underline{\beta}) : \underline{\tau}'\underline{\tau} = \underline{1}, \underline{\tau}'\underline{1} = 0, \underline{\beta}'\underline{\beta} = 1, \underline{\beta}'\underline{1} = 0$$
$$\tau_h \neq \tau_{h'} \neq 0 \quad h \neq h',$$
$$\beta_i \neq \beta_i' \neq 0 \quad i \neq i' \Big\} \qquad \ldots(5c.1.1)$$

Note that in the above we have not unnecessarily bothered to include μ in the description of *effective* parameter space.

Now we recall the following two definitions.

DFN. 5c.1.1 A design $d^* \varepsilon \Omega(N,v)$ is uniformly best if for every other $d \varepsilon \Omega(N,v)$, $V_{d^*}(\hat{\hat{\theta}}) \leq V_d(\hat{\hat{\theta}})$ for all $(\underline{\tau},\underline{\beta}) \varepsilon \textcircled{H}_1$ with strict inequality at some point.

DFN. 5c.1.2 A design d^* in $\Omega(N,v)$ is a minimax design if

$$\max_{\textcircled{H}_1} V_{d^*}(\hat{\hat{\theta}}) = \min_{d \varepsilon \Omega(N,v)} \max_{\textcircled{H}_1} V_d(\hat{\hat{\theta}})$$

5c.2 Relative Efficiencies of Designs within the Binary Sub-class of $\Omega(N,v)$

If we confine only to binary designs in $\Omega(N,v)$, the following results may be derived.

Lemma 5c.2.1 For an RBD, $\sigma^{-2}V_d(\hat{\hat{\theta}}) = 1$ for all $(\underline{\tau},\underline{\beta}) \varepsilon \textcircled{H}_1$.

Proof : For an RBD with parameters b and v, $(I - P_{X_d}) = (I - \frac{J}{b}) \otimes (I - \frac{J}{v})$ and $\sigma^{-2} V_d(\hat{\hat{\theta}}) = \dfrac{1}{\left(\sum_{t=1}^{b} \beta_t^2\right)\left(\sum_{\lambda=1}^{v} \tau_\lambda^2\right)} = 1$ in $Ⓗ_1$.

This result is reported in Kshirsagar (1983).

Lemma 5c.2.2 For any binary design d, $\sigma^{-2} V_d(\hat{\hat{\theta}})$ varies between 1 and ∞.

Proof : We have $\sigma^{-2} V_d(\hat{\hat{\theta}}) = (\underline{f}'(I - P_{X_d})\underline{f})^{-1}$.

Since for an incomplete binary block design all combinations of $\beta_i\tau_h$, $i = 1,2,\dots,b$ and $h = 1,\dots,v$ do not appear in $\underline{f}$, we augment $(I - P_{X_d})$ to a $bv \times bv$ matrix by inserting suitable null columns and null rows corresponding to the combinations $\beta_i\tau_h$ which are missing in $\underline{f}$. With this in view, we may write

$\sigma^{-2} V_d(\hat{\hat{\theta}}) = (\underline{f}^o(I - P_{X_d})^o\underline{f}^o)^{-1}$, where $\underline{f}^o = (\underline{\beta} \otimes \underline{\tau})$ and $(I - P_{X_d})^o$ is the augmented $(I - P_{X_d})$.

Now, $\underline{f}^{o\prime}\underline{f}^o = 1$ since $\underline{\tau}'\underline{\tau} = 1$ and $\underline{\beta}'\underline{\beta} = 1$ in $Ⓗ_1$.

Also $\underline{f}^o\underline{1} = 0$ since $\underline{\tau}'\underline{1} = 0$ and $\underline{\beta}'\underline{1} = 0$ both hold in $Ⓗ_1$ even though one of them would have been enough.

$$Ⓗ_2 = \left\{(\underline{\tau},\underline{\beta}) : \underline{f}^{o\prime}\underline{f}^o = (\underline{\beta} \times \underline{\tau})'(\underline{\beta} \times \underline{\tau}) = 1, \quad \underline{f}^{o\prime}\underline{1} = (\underline{\beta} \times \underline{\tau})'\underline{1} = 0\right\}.$$

Let $Ⓗ_2$ be as above.

Obviously $Ⓗ_2 \supset Ⓗ_1$ and, consequently,

$$\min_{Ⓗ_2} \underline{f}^{o\prime}(I - P_{X_d})^o\underline{f}^o \le \min_{Ⓗ_1} \underline{f}^{o\prime}(I - P_{X_d})^o\underline{f}^o$$

and

$$\max_{\textcircled{H}_2} \underline{f}^{o'}(I - P_{X_d})^o \underline{f}^o \geq \max_{\textcircled{H}_1} \underline{f}^{o'}(I - P_{X_d})^o \underline{f}^o .$$

As $(I - P_{X_d})$ is idempotent, so also is $(I - P_{X_d})^o$ and $(I - P_{X_d})^o$ has the eigenvalues 0 (multiplicity > 1) and 1 (multiplicity ≥ 1) as

$$\begin{aligned} \text{trace of} \quad (I - P_{X_d})^o &= \text{rank } (I - P_{X_d})^o \\ &= \text{rank } (I - P_{X_d}) \\ &= \Sigma k_{di} - b - v + 1 < bv - b - v + 1. \end{aligned}$$

So $\min\limits_{\textcircled{H}_2} \underline{f}^{o'}(I - P_{X_d})^o \underline{f}^o = 0$ and $\max\limits_{\textcircled{H}_2} \underline{f}^{o'}(I - P_{X_d})^o \underline{f}^o = 1$ and, hence, $\sigma^{-2} V_d(\hat{\hat{\theta}}) = \left\{ \underline{f}^{o'}(I - P_{X_d})^o \underline{f}^o \right\}^{-1}$ lies between 1 and ∞.

Thus, summing up the above two lemmas, we immediately obtain the following result.

<u>Theorem 5c.2.1</u> Whenever v/N, an RBD is the <u>uniformly best</u> among all <u>binary</u> designs in $\Omega(N,v)$.

5c.3 <u>Relative Efficiencies of Designs within the General Class $\Omega(N,v)$.</u>

If non-binary designs are allowed to be considered, the RBD no longer remains uniformly best, but it turns out to be minimax. To prove this, we first establish one structural property of non-binary designs with block sizes $\leq v$.

<u>Lemma 5c.3.1</u> For any non-binary design d with $b > 2$ and $k_{di} \leq v$ for all i, there exist at least one pair of blocks say (i, i') such that

considering these two blocks only, the h^{th} treatment replication $n_{dhi} + n_{dhi'}$ is strictly less than 2 for some $h \in \{1,2,\ldots,v\}$. Moreover,

(i) if for some h, $n_{dhi} + n_{dhi'} = 0$, then there exists some $h'(\neq h) \in \{1,2,\ldots,v\}$ such that $n_{dh'i} + n_{dh'i'} \leq 3$;

(ii) if for some h, $n_{dhi} + n_{dhi'} = 1$, then there exists some $h'(\neq h) \in \{1,2,\ldots,v\}$ such that $n_{dh'i} + n_{dh'i'} \leq 2$.

<u>Proof</u> : To observe the essential steps in the argument, we first establish the following fact : For some treatment h and for some pair of blocks (i,i'), $n_{dhi} + n_{dhi'} < 2$.

Suppose in the design d, there exists at least one block say i, such that $k_{di} < v$. Then pairing this block with any other block, say i', we get

$$\bar{r}(i,i') = \frac{\sum_{h=1}^{v}(n_{dhi} + n_{dhi'})}{v} = \frac{k_{di} + k_{di'}}{v} < 2 .$$

Hence, for at least one treatment $h \in \{1,2,\ldots,v\}$, $n_{dhi} + n_{dhi'} < 2$.

Now suppose the design has constant block size $k = v$. Even then for some treatment h and for a pair of blocks (i,i'), one would get $n_{dhi} + n_{dhi'} < 2$. For, otherwise, one has $n_{dhi} + n_{dhi'} = 2 \ \forall h$, $\forall i, i'(i \neq i')$. We argue below that the design being <u>strictly</u> non-binary, the above equality cannot hold. First note that $n_{dhi} + n_{dhi'} = 2$ for all h, and (i,i'), $(i \neq i')$ implies $n_{dhi} = 0,1,2$ and so one must have $n_{dhi} = 2$ for some combination (i,h) (as, otherwise, the design would be binary) implying thereby that $n_{dhi'} = 0 \;\forall\; i' \;(\neq i)$ which is a contradiction to $n_{dhi'} + n_{dhi''} = 2$ for that h and $(i', i''), (i' \neq i'' \neq i)$.

Clearly the argument would break down for $b = 2$.

Now we are ready to prove (i) and (ii).

(i) If for any pair of blocks (i,i'), $n_{dhi} + n_{dhi'} = 0$ for some h, there exists some other h' with $n_{dh'i} + n_{dh'i'} \leq 3$ as otherwise $k_{di} + k_{di'} = \sum_{h=1}^{v} (n_{dhi} + n_{dhi'}) \geq 4(v-1) > 2v$ whenever $v \geq 3$ and this is a contradiction to $k_{di} \leq v$ for all i.

(ii) If $n_{dhi} + n_{dhi'} \geq 1$ for all h, and $n_{dh_oi} + n_{dh_oi'} = 1$ for some h_o, then there exists some other h' with $n_{dh'i} + n_{dh'i'} \leq 2$, as otherwise, by a similar argument $k_{di} + k_{di'} \geq 1 + 3(v-1) > 2v$ whenever $v \geq 3$. So again a contradiction is achieved. Hence the lemma.

The following theorem furnishes a relative comparison of efficiencies in the class $\Omega(N,v)$ <u>including</u> non-binary designs.

<u>Theorem 5c.3.1</u> For any non-binary design $d \in \Omega(N,v)$, $\sigma^{-2} V_d(\hat{\theta}) \nleq 1$ uniformly in $(\underline{\tau}\, \underline{\beta}) \in \text{Ⓗ}_1$

<u>Proof</u> : We prove this theorem by contradiction. Suppose, if possible,

$$\sigma^{-2} V_d(\hat{\theta}) = (f'(I - P_{X_d})\underline{f})^{-1} \leq 1 \quad \text{for all } (\underline{\tau},\underline{\beta}) \in \text{Ⓗ}_1$$

Then it follows that

$$\underline{f}'\underline{f} \geq \underline{f}'(I - P_{X_d})\underline{f} \geq 1 \quad \text{for all } (\underline{\tau},\underline{\beta}) \in \text{Ⓗ}_1 .$$

Now $\underline{f}'\underline{f} = \sum_{i=1}^{b} \sum_{h=1}^{v} n_{dhi}\beta_i^2\tau_h^2$.

We will show that,

$$\Sigma\Sigma n_{dhi}\beta_i^2\tau_h^2 \nless 1 \quad \text{for all} \quad (\underline{\tau},\underline{\beta}) \in (H)_1$$

The Lemma 5c.3.1 above guarantees that we can always get hold of a pair of blocks (i,i') say, in non-binary design d with $b > 2$ such that considering these two blocks only, there exist at least a pair of treatments (h,h') say, for which the sum of four terms corresponding to replications in these two blocks $(n_{dhi} + n_{dh'i}) + (n_{dhi'} + n_{dh'i'})$ is strictly less than 4. We set these two block effects viz. β_i and $\beta_{i'}$ as $\pm \frac{1}{\sqrt{2}}$ and other block effects as zero. Similarly, setting these two treatment effects viz., τ_h and $\tau_{h'}$ as $\pm \frac{1}{\sqrt{2}}$ and other treatment effects as zero, we get for this particular choice of $\underline{\tau}$ and $\underline{\beta}$, $\underline{f}'\underline{f} < 1$. But this point obviously does not belong to the relevant parameter space $(H)_1$. However, as $V_d(\hat{\hat{\theta}})$ is a continuous function of $\underline{\tau}$ and $\underline{\beta}$, this particular choice indicates that in the neighbourhood of this point, there exists a point in $(H)_1$ for which $\underline{f}'\underline{f} < 1$.

Hence, for any non-binary design d with $b > 2$, $V_d(\hat{\hat{\theta}}) \nless 1$ for all $(\underline{\tau},\underline{\beta}) \in (H)_1$.

It remains to prove the theorem for $b = 2$. For $b = 2$, the only non-trivial case corresponds to $r_{dh} = 2$ for all h. Since the design is connected, the structure of the design in this case is as follows :

$$
\begin{aligned}
n_{dh1} &= 2 \qquad h = 1,2,\ldots,v_1 \\
&= 1 \qquad h = v_2 + 1,\ldots,v \\
&= 0 \qquad h = v_1 + 1,\ldots,v_2 \\
n_{dh2} &= 0 \qquad h = 1,2,\ldots,v_1 \\
&= 1 \qquad h = v_2 + 1,\ldots,v \\
&= 2 \qquad h = v_1 + 1,\ldots,v_2
\end{aligned}
$$

$$(2v_1 \leq v_2 < v, \text{ all intergers})$$

From this very structure of the design it follows that for all points in $\textcircled{H}_1$ (clearly with $(\beta_1, \beta_2) = \pm 1/\sqrt{2}$ up to a permutation)

$$\underline{f}'\underline{f} = \sum_{h=1}^{v} \sum_{i=1}^{b} n_{dhi}\beta_i^2\tau_h^2 = \frac{1}{2} \sum_{h=1}^{v} \tau_h^2 r_{dh} = \sum_{h=1}^{v} \tau_h^2 = 1 .$$

Moreover, $P_x\underline{f}$ can be shown to be $\neq \underline{0}$, yielding thereby $\underline{f}'(I - P_{x_d})\underline{f}' < \underline{f}'\underline{f} = 1$ for some points in $\textcircled{H}_1$.

Thus, summing up all the lemmas and theorems in subsections 5c.2 and 5c.3, we get the final result :

<u>Theorem 5c.3.2</u> RBD is the minimax design within the class of designs Ω (N,v).

<u>Remark</u> 4 : It is interesting to note that, there are non-binary designs, for which $V_d(\hat{\hat{\theta}})$ is strictly less than 1 for some parameter points. As for example, with b blocks, having constant block size v, take the design d_o, (1,2,...,v denoting treatments)

$$d_o = \begin{bmatrix} 1 & 1 & 3 & 4 & \cdots & v \\ 1 & 2 & 3 & 4 & \cdots & v \\ \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots & \cdots \\ \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots & \cdots \\ 1 & 2 & 3 & 4 & \cdots & v \end{bmatrix}$$

which differs from an RBD only in the first block, i.e. $n_{d_o 11} = 2$, $n_{d_o 21} = 0$, $n_{d_o h1} = 1$, $\forall\ h = 3,4,\ldots,v$. Now, making the choice,

$$\tau_1 = \frac{1}{\sqrt{2}}\ ,\ \tau_v = -\frac{1}{\sqrt{2}}\ ,\ \tau_h = 0\ \forall\ h \neq 1, \text{ and } v$$

$$\beta_1 = \frac{1}{\sqrt{2}}\ ,\ \beta_2 = \frac{-1}{\sqrt{2}}\ ,\ \beta_i = 0\quad \forall\ i \neq 1,2,$$

$V_{d_o}(\hat{\hat{\theta}})$ turns out to be strictly less than 1. Hence, by continuity of the variance function, there exists a parameter point in the neighbourhood of this point for which $V_{d_o}(\hat{\hat{\theta}})$ is strictly less than 1.

5c.4 Discussion of Optimality Results Concerning Interaction-Parameter Vector $\underline{\theta}$.

In the above section we have established that RBD is the " best " (in some sense) for estimation of single interaction parameter. With this in view, as regards multiple interaction parameters, we first study the dispersion matrix of $\hat{\hat{\underline{\theta}}}$ under an RBD. Let the dispersion matrix be denoted by $D(\hat{\hat{\underline{\theta}}})$. Then $D^{-1}(\hat{\hat{\underline{\theta}}})$ is given by

$$\left[\begin{array}{c|c|c|c} \underline{\beta}'\underline{\beta} \times \underline{\tau}'\underline{\tau} & \underline{\beta}'\underline{\beta}^{(2)} \times \underline{\tau}'\underline{\tau} & \underline{\beta}'\underline{\beta} \times \underline{\tau}'\underline{\tau}^{(2)} & \underline{\beta}'\underline{\beta}^{(2)} \times \underline{\tau}'\underline{\tau}^{(2)} \\ & \underline{\beta}^{(2)'}(I-\frac{J}{b})\underline{\beta}^{(2)} \times \underline{\tau}'\underline{\tau} & \underline{\beta}^{(2)'}\underline{\beta} \times \underline{\tau}'\underline{\tau}^{(2)} & \underline{\beta}^{(2)'}(I-\frac{J}{b})\underline{\beta}^{(2)} \times \underline{\tau}'\underline{\tau}^{(2)} \\ & & \underline{\beta}'\underline{\beta} \times \underline{\tau}^{(2)'}(I-\frac{J}{v})\underline{\tau}^{(2)} & \underline{\beta}'\underline{\beta}^{(2)} \times \underline{\tau}^{(2)}(I-\frac{J}{v})\underline{\tau}^{(2)} \\ & & & \underline{\beta}^{(2)'}(I-\frac{J}{b})\underline{\beta}^{(2)} \times \underline{\tau}^{(2)'}(I-\frac{J}{v})\underline{\tau}^{(2)} \end{array}\right]$$

($\underline{\beta}^{(2)}$ and $\underline{\tau}^{(2)}$ as defined earlier, are vectors with elements β_i^2 and τ_h^2 respectively)

Over the restricted parameter space $\textcircled{H}_3 = \left\{(\underline{\tau},\underline{\beta}) : (\underline{\tau},\underline{\beta}) \in \textcircled{H}_1, \sum_{i=1}^{b} \beta_i^3 = 0, \sum_{h=1}^{v} \tau_h^3 = 0 \right\}$ the dispersion matrix takes a simple form, viz.,

$$D^{-1}(\hat{\hat{\theta}}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \Sigma\beta_i^4 - \frac{1}{b} & 0 & 0 \\ 0 & 0 & \Sigma\tau_h^4 - \frac{1}{v} & 0 \\ 0 & 0 & 0 & (\Sigma\beta_i^4 - \frac{1}{b})(\Sigma\tau_h^4 - \frac{1}{v}) \end{bmatrix}$$

Now $\lambda_{min}\left[D^{-1}(\hat{\hat{\theta}})\right] \leq (\Sigma\beta_i^4 - \frac{1}{b})(\Sigma\tau_h^4 - \frac{1}{v})$ and this RHS expression can be made arbitrarily small for some choice of $\underline{\tau}$ and $\underline{\beta}$ in the parameter space $\textcircled{H}_3$. Thus, if we bring in E-optimality criterion which may be relevant to this multiparameter problem, RBD behaves very badly. Since $F'(I-P_X)F$ becomes intractable for any arbitrary design, it is very difficult to establish any such optimality result for multiple interaction parameter. May be, we have to re-define the parameter space appropriately so that the RBD achieves a better standing.

Concluding Remarks.

Our efforts in this work have been concentrated around a small query regarding " estimability " and " efficient estimation " of the non-additive parameter in Tukey's non-additive model. We do admit that this is a humble introduction to the subject matter but at the same time, we strongly feel that this sort of problem will attract attention of researchers in the field of optimal designs and more fruitful work will be reported in the future. We see that there could be a large variety of problems related to this non-additive parameter situation. Thus, for example, a suitable mixed-effects model could admit the non-additive parameter(s) (involving the fixed effects) or one could also look to Mandel's Model (5a.1.3) and investigate the optimality aspect of estimation of $\underline{\theta}$.

REFERENCES

1. Adhikary, B. and Panda, R. (1983). On Group Divisible Response Surface (GDRS) designs, Jour. Statist. Plan. Infer., 7, 387-405.

2. Afsarinejad, K. (1983). Balanced Repeated Measurements Designs, 70, no.1, 199-204.

3. Aggarwal, H.L. (1966a). Two way elimination of heterogeneity, Cal. Statist. Assoc. Bull., 15, 32-38.

4. Aggarwal, H.L. (1966b). Some systematic methods of construction of designs for two-way elimination of heterogeneity, Cal. Statist. Assoc. Bull., 15, 93-108.

5. Aggarwal, H.L. (1966c). Some methods of construction of designs for two way elimination of heterogeneity, Jour. Amer. Statist. Assoc., 61, 1153-1171.

6. Aggarwal, H.L. and Sharma, K.L. (1976). On construction of a series of two-way designs, Cal. Statist. Assoc. Bull., 23, 139-150.

7. Ash, A. and Hedayat, A. (1978). An introduction to design optimality with an overview of the literature, Comm. Statist., Theory and Methods, A7, 14, 1295-1325.

8. Bagchi (Mukhopadhyay), S. (1982). Optimal Block Designs in One and Multiway Settings. Ph.D. Thesis, unpublished, Indian Statistical Institute.

9. Bagchi, S. (1985). On the optimality of MBGDD's under mixed effects model, Tech. report ASC85/11, Indian Statistical Institute.

10. Banerjee, K.S. (1975). Weighing Designs, Mercel-Dekker, New York.

11. Bhattacharya, C.G. and Shah, K.R. (1984). On the optimality of block designs under mixed effects model,Utilitas, Math. 26, November, 339-345.

12. Blackwell, D. (1951). Comparison of experiments. Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability, Univ. California Press 93-102.

13. Blackwell, D. (1953). Equivalent comparison of experiments, Ann. Math. Statist., 24, 265-272.

14. Blackwell, D. and Girshick, M.A. (1954). Theory of Games and Statistical Decisions., Wiley, New-York.

15. Bose, R.C. (1975). Combined intra and interblock estimation of treatment effects in incomplete block designs. A Survey of Statistical Designs and Linear Models, J.N. Srivastava Edited, 31-51.

16. Chakrabarti, M.C. (1962). Mathematics of Designs and Analysis of Experiments, Asia Publishing House.

17. Cheng, C.S. (1978a). Optimal designs for the elimination of multiway heterogeneity, Ann. Statist., 6, 1262-1272.

18. Cheng, C.S. (1978b). Optimality of certain asymmetrical designs, Ann. Statist., 6, 1239-1261.

19. Cheng, C.S. (1979a). Optimal incomplete block designs with four varieties, Sankhyā, ser. B, 41, 1-14.

20. Cheng, C.S. (1979b). Construction of Youden Hyperrectangles, Jour. Statist. Plan. Infer., 3, 109-118.

21. Cheng, C.S. (1980a). On the E-optimality of some block designs, Jour. Roy. Statist. Soc., ser. B, 42, 199-204.

22. Cheng, C.S. (1980b). Optimality of some weighing designs and 2^n factorial designs, Ann. Statist., 8, 436-446.

23. Cheng, C.S. (1981a). Optimality and construction of Pseudo Youden Designs, Ann. Statist., 9, 201-205.

24. Cheng, C.S. (1981b). A family of Pseudo Youden Designs with row size less than the number of symbols, Jour. Comb. Theory, ser. A, 31, no.2, 219-221.

25. Cheng, C.S. (1981c). Maximising the total number of spanning trees in a graph : Two related problems in Graph Theory and Optimum Design Theory, Jour. Comb. Theory., ser. B, 31, no.2, 240-248.

26. Cheng, C.S. (1981d). On the comparison of PBIB designs with two associate classes, Ann. Inst. Statist. Math., 33, 155-164.

27. Cheng, C.S. (1981e). Graph and Optimum Design Theories — some connections and examples, Bull. Intern. Statist. Inst., Proceedings of the 43rd session, Volume XLIX, Book 1, Buenos Aires, 1981, 580-590.

28. Cheng, C.S. and Wu, C.F. (1980). Balanced Repeated Measurements Designs, Ann. Statist. 8, 1272-1283.

29. Cheng, C.S. and Wu, C.F. (1983). Corrections to balanced Repeated Measurements Designs, Ann. Statist., 11, 349.

30. Conniffe, D. and Stone, J. (1974). The efficiency factor of a class of incomplete block designs. Biometrika, 61, 633-636.

31. Conniffe, D. and Stone, J. (1975). Some incomplete block designs of maximum efficiency. Biometrika, 62, 685-686.

32. Constantine, G. and Hedayat, A. (1979). Repeated Measurements Designs, III. Tech. report, Univ. Illinois at Chicago Circle.

33. Constantine, G. and Hedayat, A. (1982). A construction of Repeated Measurements Designs with balance for residual effects, Jour. Statist. Plan. Inf. 6, 153-164.

34. Constantine, G.M. (1981). Some E-optimal block designs. Ann. Statist. 9, 886-892.

35. Constantine, G.M. (1982). On the E-optimality of PBIB designs with a small number of blocks. Ann. Statist., 10, no.3, 1027-1031.

36. Denes, J. and Keedwell, A.D. (1974). Latin Squares and their Applications, Academic Press, New York and London.

37. Dey, A., Gupta, V.K. and Singh, M. (1983). Optimal Change - over designs, Sankhyā, ser. B, 45, pt.2, 233-239.

38. Eccleston, J.A. and Hedayat, A. (1974). On the theory of connected designs : Characterization and Optimality, Ann. Statist., 2, 1238-1255.

39. Ehrenfled, S. (1955). Complete Class Theorems in experimental design. Proceedings of Third Berkley Symposium, Univ. California Press, 1, 57-67.

40. Fedorov, V.V. (1972). Theory of Optimal Experiments, Academic Press, New York.

41. Franklin, M.F. (1984). Cyclic generation of orthogonal Latin Squares Ars. Combinatoria, 17, 129-140.

42. Fulkerson, D.R. and Gross, O.A. (1965). Incidence matrices and interval graphs, Pacific Jour. Math., 15, 835-855.

43. Gaffke, N. and Krafft, O. (1979). Optimum designs in complete two-way layouts, Jour. Statist. Plan. Infer. 3, 119-126.

44. Gaffke, N. and Krafft, O. (1980). D-optimum designs for the inter-block model : Mathematical Statistics and Probability Theory. (Proceedings of Sixth International Conference, 1978), 134-143, Lecture notes in Statistics, Springer, New York.

45. Gaffke, N. (1981). Some classes of optimality criteria and optimal designs for complete two-way layout, Ann. Statist., 9, 893-898.

46. Gaffke, N. (1982). D-optimal design with at most 6 varieties, Jour. Statist. Plan. Infer. 6, 183-200

47. Galil, Z. and Kiefer, J. (1980a). D-optimum weighing designs. Ann. Statist., 8, 1293-1306.

48. Galil, Z. and Kiefer, J. (1980b). Optimum weighing designs, Recent Developments in Statistical Inference and Data-Analysis 183-189.

49. Galil, Z. and Kiefer, J. (1981). Comparison of D-optimum design, Technometrics, 16, 211-220.

50. Galil, Z. and Kiefer, J. (1982a). On the characterization of D-optimal weighing designs for $n \equiv 3 \pmod 4$. Statistical Decision Theory and Related Topics. 1-35.

51. Galil, Z. and Kiefer, J. (1982b). Construction methods for D-optimum weighing designs when $n \equiv 3 \pmod 4$. Ann. Statist., 10, 502-510.

52. Ghosh, M.N. and Sharma, D. (1963). Power of Tukey's test for non-additivity. Jour. Roy. Statist., Soc. ser.B, 25, no.1, 213-219.

53. Gollob, H.F. (1968). A statistical model which combines features of Factor Analytic and Analysis of Variance techniques. Psychometrica, 33, 73-116.

54. Harwit, M. and Sloane, N.J.A. (1979). Hadamard Transform Optics. Academic Press, New York, London.

55. Hedayat, A. and Afsarinejad, K. (1975). Repeated Measurements Designs I, A Survey of Statistical Designs and Linear Models, J.N. Srivastava edited, 229-242.

56. Hedayat, A. and Afsarinejad, K. (1978). Repeated Measurements Designs, II, Ann. Statist., 6, 619-628.

57. Hedayat, A. (1978). Optimal Design Theory, Comm. Stat., Theory and Methods, A7, no.14.

58. Hedayat, A. (1981a). Repeated Measurements Designs, IV : Recent Advances, Bull. Inter. Statist., Inst., Proceedings of the 43rd session, Volume XLIX, Book 1, 591-605.

59. Hedayat, A. (1981b). Study of optimality criteria in Design of Experiments, Statistics and Related Topics, North-Holland, 39-56.

60. Hedayat, A. (1981c). A characterization of family of incomplete block designs with completely symmetric information matrices. Inst. Math. Stat. Bull. 10(5), 202

61. Hedayat, A. (1984). A characterization of a Universally Optimal designs within a class of block designs, Jour. Statist. Plan. Infer., 9, 143-145.

62. Hegemann, V.J. and Johnson, D.E. (1976a). On analysing two-way AOV with interaction, Technometrics, 18, 273-281.

63. Hegemann, V.J. and Johnson, D.E. (1976b). The power of two tests for non-additivity, Jour. Amer. Statist. Assoc., 71, 945-948.

64. Hotelling, H. (1944). Some improvements in weighing and other experimental techniques, Ann. Math. Statist., 15, 297-306.

65. Jacroux, M. (1980a). On the determination and construction of E-optimal block designs with unequal number of replications, Biometrika 67, 661-667.

66. Jacroux, M. (1980b). On the E-optimality of Regular Graph Designs, Jour. Roy. Statist., Soc., ser.B, 42, 205-209.

67. Jacroux, M. (1982). On the E-optimality of designs for the one-way and two-way elimination of heterogeneity, Jour. Roy. Statist., Soc. ser.B, 44, 253-261.

68. Jacroux, M. (1983). On the E-optimality of block designs, Sankhyā, ser.B, 45, pt.3, 351-361.

69. Jacroux, M., Wong, C.S. and Masaro, J.C. (1983). On the optimality of chemical balance weighing designs, Jour. Statist. Plan.Infer. 8, 231-240.

70. Jacroux, M. and Notz, W. (1983). On the optimality of spring balance weighing designs, Ann. Statist., 11, 970-978.

71. Jacroux, M. (1984). On the D-optimality of Group Divisible Designs. Jour. Statist. Plan. Infer. 9, no.1, 119-129.

72. Jacroux, M. (1985). Some sufficient conditions for the type 1 optimality of block designs Jour. Statist. Plan. Infer., 11, no.3, 385-398.

73. Jacroux, M. (1986). On the E and MV-optimality of spring balance designs with string property. Jour. Statist. Plan. Infer. 13, no.1, 89-96.

74. John, J.A. and Williams, E.K. (1982). Conjectures for optimal block designs. Jour. Roy. Statist. Soc. ser.B, 44, 221-225.

75. Johnson, D.E. and Graybill, F.A. (1972a). An analysis of two-way model with interaction and no replication. Jour. Amer. Statist. Assoc. 67, 862-868.

76. Johnson, D.E. and Graybill, F.A. (1972b). Estimation of σ^2 in a two-way classification model with interaction. Jour. Amer. Statist., Assoc., 67, 388-394.

77. Kettenring, J.R. (1983). Components of interaction in Analysis of Variance Models with no replications, Contributions to Statistics (Essays in Honour of N.L. Johnson, Edited by P.K. Sen), 283-298, North-Holland.

78. Khatri, C.G. and Shah, K.R. (1984). Optimality of block designs under mixed models, Proceedings of the Indian Statistical Institute, Golden Jubilee Conference : Statistics, Applications and New directions, 326-332.

79. Kiefer, J. (1958). On the non-randomized optimality and randomized non-optimality of symmetrical designs. Ann. Math. Statist., 29, 675-699.

80. Kiefer, J. (1959). Optimum experimental designs, Jour. Roy. Statist. Soc., B21, 272-319.

81. Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. Ann. Math. Stat., 30, 271-294.

82. Kiefer, J. (1974). Lectures on Design Theory. Tech. report 397, Dept. of Statistics, Purdue Univ., West Lafayette, IN.

83. Kiefer, J. (1975). Construction and optimality of Generalised Youden Designs. A Survey of Statistical Design and Linear Models, edited by J.N. Srivastava, North Holland, Amsterdam, 333-353.

84. Kiefer, J. and Studden, W.J. (1976). Optimal designs for large degree polynomial regression, Ann. Statist., 4, 1113-1123.

85. Kshirsagar, A.M. (1958). A note on incomplete block designs, Ann. Math. Statist. 29, 907-910.

86. Kshirsagar, A.M. (1983). A Course in Linear Models, Mercel-Dekker.

87. Kunert, J. (1983). Optimal design and refinement of the linear model with application to Repeated Measurements Designs, Ann. Statist., 11, no.1, 247-257.

88. Kunert, J. (1984a). Optimality of balanced uniform Repeated Measurements Designs. Ann. Stat., 12, no.3, 1006-1017.

89. Kunert, J. (1984b). Designs balanced for circular residual effects. Comm. Statist., Theory and Methods, B (21), 2665-2671.

90. Magda, G.C. (1980). Circular balanced Repeated Measurements Designs. Comm. Statist., Theory and Methods, A9, no.18, 1901-1918.

91. Mandel, J. (1959). The analysis of Latin Squares with a certain type of row-column interaction, Technometrics, 1, 379-387.

92. Mandel, J. (1961). Non-additivity in two-way Analysis of Variance, Jour. Amer. Statist., Assoc., 17, 878-888.

93. Mandel, J. (1969). The partitioning of interaction in Analysis of Variance, Jour. of Research of the National Bureau of Standards B37 8, no.4, 309-328.

94. Mandel, J. (1971). A new Analysis of Variance Model for non-additive data.Technometrics, 13, 1-18.

95. Marasinghe, M.G. and Johnson, D.E. (1981). Testing subhypothesis in the multiplicative interaction model, Technometrics, 23, no.4, 385-393.

96. Marasinghe, M.G. and Johnson, D.E. (1982). Estimation of σ^2 in the multiplicative interaction model, Comm. Statist., Theory and Methods, 11, 315-324.

97. Marshall, A.G. and Comisarow, M.B. (1975). Fourier and Hadamard transform methods in Spectroscopy. Analyt. Chem. 47, 491A-504A.

98. Milliken, G.A. and Graybill, F.A. (1970). Extensions of the General Linear Hypothesis Model. Jour. Amer. Statist., Assoc., 65, 797-807.

99. Milliken, G.A. and Graybill, F.A. (1971). Tests for interaction in the two-way model with missing data. Biometrics, 27, 1079-1083.

100. Mood, A.M. (1946). On Hotelling's weighing problem. Ann. Math. Statist., 17, 432-446.

101. Mote, V.L. (1958). On a minimax property of a balanced incomplete block design. Ann. Math. Statist., 29, 910-914.

102. Mukerjee, R. and SahaRay, R. (1985). Asymptotically optimal weighing designs with string property, Jour. Statist., Plan. Infer., 12, 87-91.

103. Mukerjee, R. and Sen, M. (1984). Optimal Repeated Measurements Designs under interaction, Tech. report, Stat-Math. 27/84, Indian Statistical Institute.

104. Mukhopadhyay, A.C. and Mukhopadhyay, S. (1981). Construction of balanced Youden Hyperrectangles. Tech. report no.ASC 181/14, Indian Statistical Institute, Calcutta.

105. Mukhopadhyay, A.C. and Mukhopadhyay, S. (1984). Optimality in a balanced multiway heterogeneity set up, Proceedings of the Indian Statistical Institute, Golden Jubilee International Conference : Statistics Applications and New directions, 466-477.

*106. Mukhopadhyay, A.C. and Saha, R. (1983). Repeated Measurements Designs. Cal. Statist., Assoc. Bull. 32, nos.127-128, 153-168.

107. Mukhopadhyay, A.C. and SahaRay, R. (1985). Three-way elimination of heterogeneity with non-orthogonal incidence structure for every two directions,Tech. report no.ASC.85/6, Indian Statistical Institute.

108. Mukhopadhyay, S. (1984). ψ_r - optimality of MBGDD of type 1 under mixed effects model within the restricted class of binary designs, Sankhyā, ser.B, 46, 113-117.

109. Nandi, H.K. (1950). On the efficiency of experimental designs, Cal. Statist., Assoc. Bull., 3, 167-171.

110. Panda, R. (1976). Asymptotically highly efficient spring balance weighing designs under restricted set-up, Cal. Statist., Assoc. Bull., 25, 179-184.

111. Preece, D.A. (1972). Non-additivity in two-way classification with missing values. Biometrics 28, 574-577.

112. Raghavarao,D. (1971). Construction and combinational problems in Design of Experiments. John-Wiley, New York.

113. Rao, C.R. (1965). Linear Statistical Inference and Its Applications John. Wiley, New York.

114. Rao, C.R. and Mitra, S.K. (1971). Generalized Inverse of Matrices and its Applications. New York, John Wiley and Sons.

115. Roy, J. (1958). On the efficiency factor of block designs, Sankhyā 19, 181-188.

116. Ryser, H.J. (1969). Combinatorial configurations, SIAM Jour. Appl. Math., 17, 593-602.

117. SahaRay, R. (1986). Optimal designs under a certain class of non-orthogonal row-column structure, Sankhyā, series B Vol.48, 44-72.

118. Sampford, M.R. (1957). Methods of construction and analysis of Serially Balanced Sequences, Jour. Roy. Statist. Soc., ser.B, 19, 286-304.

119. Scheffe, H. (1959). The Analysis of Variance, John Wiley.

120. Shah, K.R. (1960). Optimality criteria for incomplete block designs. Ann. Math. Statist., 31, 791-794.

121. Shah, K.R., Khatri, C.G. and Raghavarao, D. (1976). Optimality of two and three factor designs. Ann. Statist., 4, 419-422.

122. Silvey, S.D. (1980). Optimal Design. Chapman-Hall London.

123. Sinha, B.K. (1971). Contribution to Comparison of Experiments : Optimum Experiments for Linear Inference. Ph.D. Thesis, unpublished. Univ. of Calcutta.

124. Sinha, B.K. (1972). Optimum spring balance weighing designs, Proceedings of All India Convention on Quality and Reliability. Indian Institute of Technology, Kharagpur.

125. Sinha, B.K. (1975). On some optimum properties of Serially Balanced Designs. Sankhyā, ser.B, 37, 173-192.

126. Sinha, B.K. and Mukerjee, R. (1982). A note on the Universal Optimality criterion for full rank models. Jour. Statist. Plan. Infer. 7, 97-100.

*127. Sinha, B.K. and Saha, R. (1983). Optimal weighing designs with a string property, Jour. Statist. Plan. Infer., 8, 365-374.

128. Sinha, B.K., SahaRay, R. and Mukhopadhyay, A.C. (1985). Non-additive Linear Models : Estimability and efficient estimation of interaction parameters. A preliminary version of this paper was presented in 72nd Session of Indian Science Congress, January, 1985.

129. Sinha, B.K. and Sinha, B.K. (1969). Comparisons of relative efficiencies of some classes of designs. Cal. Stat., Assoc. Bull. 18, 99-122.

130. Sloane, N.J.A. and Harwit, M. (1976). Masks for Hadamard Transform Optics and Weighing Designs, Appl. Opt., 15, 107-114.

131. Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. Biometrika 12, 1-85.

132. Sonnemann, E. (1982). D-optimality of complete Latin Squares. Math. operations forschung und Statistik, Ser. Stat., 13, 387-394.

133. Studden, W.J. (1978). Designs for large degree polynomial regression. Comm. Statist, Theory and Methods, A7, 14, 1391-1397.

134. Swamy, M.N. (1980). Optimum spring balance weighing designs for estimating the total weight. Comm. Statist., Theory and Methods, A9, 1185-1190.

135. Swamy, M.N. (1982). Use of balanced bipartite weighing designs as chemical balance designs. Comm. Statist., Theory and Methods, 11, no.7, 769-785.

136. Takeuchi, K. (1961). On the optimality of certain type of PBIB designs, Rep. Statist., Appl. Res., Un. Japan. Sci. Engrs, 8, 140-145.

137. Takeuchi, K. (1963). Remark added to " On the optimality of certain types of PBIB designs ". Rep., Statist., Appl. Res. Un. Japan. Sci. Engrs, 10, 47.

138. Tukey, J.W. (1949). Onedegree of freedom for non-additivity, Biometrika, 5, 232-242.

139. Tukey, J.W. (1955). Answer to query 43, Biometrika 11, 111-113.

140. Tukey, J.W. (1962). The future of data analysis. Ann. Math. Statist. 33, 1-67.

141. Wald, A. (1943). On the efficient design of statistical investigations. Ann. Math. Statist. 14, 134-140.

142. Ward, G.C. and Dick, I.D. (1952). Non-additivity in Randomized Block Designs and Balanced Incomplete Block Designs, Newzeland Jour. of Science and Technology, 33, 430-436.

143. Wilks, S.S. (1932). Certain generalisations in the Analysis of Variance, Biometrika 24, 471-494.

144. Yates, F. (1935). Complex experiments, Supplementary Jour. of the Roy. Statist. Soc., 2, 181-247.

145. Yochmowitz, M.G. and Cornell, R.G. (1978). Stepwise tests for multiplicative components of interaction, Technometrics, 20, 79-84.

---

* Saha was the maiden tittle of R. SahaRay.