https://doi.org/10.56093/JISAS.V78I1.5
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 78(1) 2024 37–45

Efficient Block Designs for Mixed Level Factorial Microarray Experiments based on Baseline Parameterization

Sukanta Dash and Rajender Parsad

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 31 October 2023; Revised 09 February 2024; Accepted 15 April 2024

SUMMARY

A procedure of obtaining efficient block designs in block size 2 for *n*-factor mixed level factorial microarray experiments based on baseline parameterization has been given. A software module has been developed using C# programming language with ASP.NET platform for generation of efficient block designs in block size 2 for $s_1 \times s_2 \times \ldots \times s_n$ factorial experiments in $v-1$ arrays, where s_j denotes the number of levels of j^{th} factor and *n* denotes the number of factors and $v = \prod_{j=1}^{n} s_j$, the total number of treatment combinations. For *n* = 2, the software developed can also generate efficient block designs for $v-1 \le b \le (v-1)+(s_1-1)(s_2-1)$, where *b* is the number of arrays.

Keywords: 2-Colour Microarray Experiments; Factorial Experiments; Block Designs; Baseline parameterization.

1. INTRODUCTION

We consider the 2-colour microarray experiments in which same set of genes is spotted on each array and as a consequence gene (G), array-gene interaction (AG), dye-gene interaction (DG) and variety-gene interaction (VG)} known as genes / gene specific effects are orthogonal to array (A), dye (D) and variety (V), known as global effects. A design that is efficient under a model containing global effects remains efficient under a model containing both global and gene specific effects. Therefore, the search for efficient designs can be restricted to the models involving only global effects. In 2-colour cDNA microarray experiments, the treatments or varieties are different types of tissues, drug treatments or time points of a biological process, which may be unstructured or have a factorial structure. The design is called a single factor or varietal design if the treatments are unstructured and factorial if the treatments have factorial treatment structure. In single factor experiments, the interest is in estimation of all or some pair wise treatment comparisons. In designs for factorial experiments, the inference is drawn with respect to factorial effects {main effects of the factors and interactions among them}. In factorial experiments, main effects of factors and interactions among them are defined via orthogonal parameterization involving mutually orthogonal treatment contrasts.

In microarray experiments, natural baseline or null state may exist. For example, there may be tissues from two mutants, one of which proliferates a particular disease and other does not. Therefore, the mutant that does not proliferate into disease is baseline. In a toxicological study with binary factors, each representing the presence or absence of a particular toxin, the state of absence can be regarded as a natural baseline level of each factor. Null state or baseline of a factor need not strictly mean zero level on some scale, but may as well refer to a standard or control level like the one currently being used in practice. Such experimental situations involving the control or standard treatment (natural baseline) do occur even beyond the domain of microarray experiments.

Corresponding author: Sukanta Dash
E-mail address: sukanta.iasri@gmail.com

In multi-factor microarray experiments in which one of the levels of each factor is a natural baseline, the comparisons with the baseline are of importance rather than the usual main effects and interactions. To make the exposition clear, consider a 2-colour microarray experiment in which it is desired to compare two cell lines FIΔ and V449E at times zero hour and 24 hours {see e.g. Glonek and Solomon (2004)}. The cell line V449E proliferates into leukaemia while FIΔ is non-leukaemic. Therefore, there are two factors dictating the cell populations. The first factor (F_1), namely, mutant has two levels FIΔ and V449E of which FIΔ, being non-leukaemic, is baseline. The two levels of mutants may be coded as 0 and 1 respectively. The second factor (F_2) is time with two levels as 0 hours and 24 hours and first of these levels (0 hours) is baseline. These two levels are also coded as 0 and 1 respectively. Thus considering the two factors together, there are four treatment combinations, 00, 01, 10, 11 representing the cell population. Let $\tau_{00}, \tau_{01}, \tau_{10}, \tau_{11}$ denote the expected log intensities, that is, the effects of these treatment combinations.

Now the question "Are there any genes specific to V449E that result into leukaemic effects?" can be answered from the treatment contrast $\tau_{10} - \tau_{00}$. Further, the change in intensity of FIΔ (natural baseline of mutant) between zero and 24 hours can be estimated from the treatment contrast $\tau_{01} - \tau_{00}$. Further, the difference in FIΔ and V449E at time 24 hours can be estimated from the treatment contrast $\tau_{11} - \tau_{01}$. The difference between these two lines at 0 hours was estimated using $\tau_{10} - \tau_{00}$. The difference of the two $\{(\tau_{11} - \tau_{01}) - (\tau_{10} - \tau_{00})\}$ represents the differential expression between the two cell lines that exists at 24 hours beyond what was present at time zero. Similarly this difference can also be estimated as $(\tau_{11} - \tau_{10}) - (\tau_{01} - \tau_{00})\}$. Therefore, the inference is required on the three contrasts viz.

$\theta_{10} = \tau_{10} - \tau_{00}$;

$\theta_{01} = \tau_{01} - \tau_{00}$ and

$\theta_{11} = \tau_{11} - \tau_{01} - \tau_{10} + \tau_{00}$.

Here θ_{10}, θ_{01} and θ_{11} are baseline parameterization of main effect of F_1, F_2 and interaction F_1F_2 respectively. However, if at least one factor, like gender, lacks a natural baseline, then the baseline parameterization is inappropriate because this will arbitrarily single out one level of such a factor. In such situations, it is advisable to use the orthogonal parameterization.

The corresponding treatment contrasts of main effect F_1, main effect F_2 and interaction effect F_1F_2 are

$\theta^*_{10} = (\tau_{11} - \tau_{01} + \tau_{10} - \tau_{00})/2$,

$\theta^*_{01} = (\tau_{11} + \tau_{01} - \tau_{10} - \tau_{00})/2$ and

$\theta^*_{11} = (\tau_{11} - \tau_{01} - \tau_{10} + \tau_{00})/2$

From the above, it is clear that the definitions of main effects under the two parameterizations are entirely different. While θ_{11} is proportional to θ^*_{11}, this equivalence for two factor interaction also disappears in case of experiments involving more than two factors.

The main distinction between these two kinds of parameterization is that while the orthogonal parameterization defines the factorial effects via mutually orthogonal treatment contrasts, the baseline parameterization defines these effects with reference to natural baseline levels of the factors and, hence, entails non-orthogonality.

For efficient designs for two-level factorial microarray experiments for estimation of factorial effects of interest under orthogonal parameterization, a reference may be made to Yang and Speed (2002), Churchill (2002), Yang and Draper (2003), Wang (2004), Gupta (2006), Kerr (2006), Grossmann and Schwabe (2008).

Although the baseline parameterization looks simpler than the orthogonal parameterization, it renders the task of finding optimal or efficient designs somewhat more challenging due to lack of orthogonality. Glonek and Solomon (2004) were the first to study designs for multi-factor microarray experiments under baseline parameterization. They have introduced a criterion of statistical efficiency in terms of variances of the estimated parameters of interest. For given number of arrays, *b*, a design is said to be admissible if the variance of each of the estimated parameters of interest is less than or equal to the variance of the estimated parameters of interest through any other design in same number of arrays and strict inequality holds for at least one parameter. This criterion was illustrated in obtaining efficient designs for 2^2 factorial experiments for given number of arrays. They have also illustrated

the utility of admissible criterion for 2 × 3 factorial experiments.

The key reference for obtaining optimal/efficient designs for baseline parameterization is Banerjee and Mukerjee (2008) and also to know more about optimality aspect of factorial experiments with baseline parameterization one may refer to Mukerjee and Tang (2012). They studied *n*-factorial experiment with factors as $F_1, F_2, \ldots, F_j, \ldots, F_n$ with factor $F_j, 1 \le j \le n$ at $s_j \ge 2$ levels represented by 0, 1, …, $s_j - 1$. The total number of treatment combinations is $v = \prod_{j=-1}^{n} s_j$.

The treatment combinations in lexicographic order are given by $\mathbf{a}_1 \times \mathbf{a}_2 \times \ldots \times \mathbf{a}_n$ where × denotes the symbolic direct product and $\mathbf{a}'_j = (0,1,\ldots,s_{j-1});\ j = 1,2,\ldots,n$. Corresponding to the treatment combination $i_1 \cdots i_n$, $0 \le i_j \le s_j - 1$, $1 \le j \le n$, Error! Digit expected. $\tau_{i_1 \ldots i_n}$ defines the expected log intensity (fluorescence intensity measured from application of treatment combination $i_1 \cdots i_n$ and transformed to log scale), *i.e.*, the effect, of the treatment combination $i_1 \ldots i_n$. As before, the baseline level of each factor is denoted by 0. Hence, $\theta_{00\cdots0} = \tau_{00\cdots0}$ stands for the baseline effect. Also, baseline parameterization for main effect, say, that of factor F_1, is represented by the $s_1 - 1$ parameters

$$\theta_{i_1 0\ldots0} = \tau_{i_1 0\ldots0} - \tau_{00\ldots0} \quad (1 \le i_1 \le s_1 - 1).$$

The baseline parameterization for a two-factor interaction, say, F_1F_2 is represented by the $(s_1 - 1)(s_2 - 1)$ parameters

$$\theta_{i_1 i_2 0\ldots0} = \tau_{i_1 i_2 0\ldots0} - \tau_{i_1 0\ldots0} - \tau_{0 i_2 0\ldots0} + \tau_{000\ldots0}$$
$$(1 \le i_1 \le s_1 - 1, 1 \le i_2 \le s_2 - 1).$$

Similarly, one can define $\theta_{i_1 \cdots i_n}$ for every $i_1 \cdots i_n \ne 0 \cdots 0 (0 \le i_j \le s_j - 1, 1 \le j \le n)$ so that any such $\theta_{i_1 \cdots i_n}$ represents a factorial effect as determined by its nonzero subscripts. The total number of parameters $\{\theta_{i_1 \cdots i_n} (i_1 \cdots i_n \ne 0 \cdots 0)\}$ to be estimated are $v - 1$ and these are collectively referred to as the θs for ease in presentation. Banerjee and Mukerjee (2008) have obtained lower bound to the variance of $\theta_{i_1 \cdots i_n}$ when number of arrays are equal to $v - 1$. A design which attains these lower bounds for each of θs, is called an optimal design. Banerjee and Mukerjee (2008) have also given a method of construction of optimal design in $v - 1$ arrays. If all main effects and interaction effects are of interest, then a design in $v - 1$ arrays is a saturated design and leaves no error degree of freedom for estimation of experimental error or testing of hypothesis regarding parameters of interest. Therefore, it is required to generate a design in number of arrays $b > v - 1$. For $b > v - 1$, a new criterion of optimality viz. ω-optimality was introduced for $s_1 \times s_2$ factorial. A design $d \in D(s_1 \times s_2, b, 2)$ for $s_1 \times s_2$ factorial in given number of arrays, *b*, is said to be ω-optimal, if it minimizes

$$T_1 = \sum_{i_1=1}^{s_1-1} var(\hat{\theta}_{i_1 0}) + \sum_{i_2=1}^{s_2-1} var(\hat{\theta}_{0 i_2}) + \omega \sum_{i_1=1}^{s_1-1} \sum_{i_2=1}^{s_2-1} var(\hat{\theta}_{i_1 i_2}) \quad (1)$$

One approach to get an ω-optimal design is to generate all possible $\binom{v(v-1)/2}{b}$ designs and select the design with minimum value of T_1. The optimal design may not be unique. Another approach suggested by Banerjee and Mukerjee (2008) is to (i) generate all optimal saturated design in $b = v - 1$ arrays, (ii) given *b*, augment each optimal design in (i) in all possible ways to generate design with *b* arrays and (iii) select one design as per chosen optimality criterion T_1 in (1). Using this approach, they have suggested the procedure of augmenting up to $b = (v-1) + (s_1 - 1)(s_2 - 1)$ arrays, *i.e.* adding any number of arrays from 1 to $(s_1 - 1)(s_2 - 1)$ and conjectured that these designs are ω-efficient. Here we have generalized the value of number of arrays (*b*) for *n* factors mixed level factorial experiments based on baseline parameterization i.e. number of arrays for *n* factors is

$$v - 1 \le b \le (v-1) + \sum_{j=2}^{n} \left\{ (j-1) \sum_{i_1 \ne i_2 \ne \cdots \ne i_j = 1}^{n} (s_{i_1} - 1)(s_{i_2} - 1) \cdots (s_{i_j} - 1) \right\}. \quad (2)$$

Thus for *n* factors mixed level factorial experiments we may add any number of array from 1 to $\sum_{j=2}^{n} \left\{ (j-1) \sum_{i_1 \ne i_2 \ne \cdots \ne i_j = 1}^{n} (s_{i_1} - 1)(s_{i_2} - 1) \cdots (s_{i_j} - 1) \right\}$ with (*v*-1) and conjectured that these designs are ω-efficient.

The generation of designs, however, is quite tedious. Therefore, in the present investigation, a software has been developed to generate optimal designs for any number of factors in $b = v - 1$ arrays and for two-factor factorial experiments in $b = (v-1)$ to $(v-1)+(s_1-1)(s_2-1)$ arrays. The details are given in Section 2. The method of generation of ω-efficient designs for two-factors in $b = (v-1)+(s_1-1)(s_2-1)$ has been extended to three factors in Section 3. A discussion on future scope is given in Section 4.

2. SATURATED AND UNSATURATED DESIGNS

In this section, we have generated optimal saturated block designs in block size 2 for estimation of factorial effects under baseline parameterization. As we have discussed earlier that in a saturated design, there is no error degree of freedom. If the interest is in only main effects and two factor interactions, then it leaves some degrees of freedom for error under the assumption of absence of higher order interactions except main effects and two factor interactions. Thus there is a need to obtain both saturated as well as unsaturated ω-efficient designs for mixed level factorial experiments based on baseline parameterization. In this section, we extend the procedure of obtaining block designs for mixed level factorial experiments of any number of factors in any number of blocks satisfying $v-1 \le b \le (v-1)+\sum_{j=2}^{n}\left\{(j-1)\sum_{i_1\neq i_2\neq\cdots\neq i_j=1}^{n}(s_{i_1}-1)(s_{i_2}-1)\cdots(s_{i_j}-1)\right\}$, where v, b, and n are number of treatment combinations, arrays and factors respectively. We begin with some preliminaries required for generation of designs with baseline parameterization.

In two-colour microarray experiments, only two treatment combinations can be accommodated on each array and one of them is labelled with red dye and the other with green dye. Let the treatment combinations $i_1\cdots i_n$ and $j_1\cdots j_n$ be labeled with red and green dyes respectively. Then an array is denoted by an ordered pair $(i_1\cdots i_n,\ j_1\cdots j_n)$.A design is represented by a collection of such pairs. Assuming the absence of dye-colour bias, the ordering within any pair is immaterial. Baseline of treatment combination $i_1\cdots i_n \neq 0...0$ $\rho(i_1...i_n)$ is obtained by replacing non-zero level of any factor by zero level and leaving the level of other factors unchanged. The procedure of obtaining the design is as follows:

Steps of Construction

Case I: When $b = v - 1$

1. Write all possible treatment combinations excluding the control treatment 00…0 in lexicographic order.
2. Obtain baseline of each treatment combination by replacing the first non-zero level by zero and keeping levels of other factors unchanged.
3. Keep all treatment combinations obtained in step 1 in lexicographic order with the upper dye (say red) and corresponding baseline treatment combination with the lower dye (say green) in an array

This yields an optimal saturated design in $b = v - 1$ arrays. The above procedure can give optimal saturated block designs for any $s_1 \times s_2 \times \ldots \times s_n$ factorial experiment.

Example 1: For a 2×3 factorial experiment, the optimal block design in 5 arrays is

Arrays	1	2	3	4	5
Dye 1	01	02	10	11	12
Dye 2	00	00	00	01	02

In this way we can construct a design with v–1 number of arrays.

Case II: When

$$(v-1) < b \le (v-1)+\sum_{j=2}^{n}\left\{(j-1)\sum_{i_1\neq i_2\neq\cdots\neq i_j=1}^{n}(s_{i_1}-1)(s_{i_2}-1)\cdots(s_{i_j}-1)\right\}$$

1. Generate a design in $b = v - 1$ as given in Case I of Section 2.
2. Search for the first treatment combination having two non-zero levels of n factors in first row of the design obtained in Step 1.
3. Now identify its baseline treatment combination by replacing the non-zero level of factor 3 by zero and keeping the levels of factor 1 and 2 unchanged.
4. Add a column (array) containing the treatment combination in step 1 in upper dye and its baseline treatment combination in lower dye in the design with $b = v - 1$ obtained in Step 1.

This yields a design in $b = v$ arrays.

5. For getting a design in $b = v+1$ arrays, repeat steps 2 to 4 for second treatment combinations having all non-zero levels 2 factors in first row of the design obtained in Step 1.
6. For getting a design in any number of arrays $b = v+2$ to $(v-1)+(s_1-1)(s_2-1)+(s_1-1)(s_3-1)+(s_2-1)(s_3-1)$, repeat steps 2 to 4 for all other $(s_1-1)(s_2-1)+(s_1-1)(s_3-1)+(s_2-1)(s_3-1)-2$ treatment combinations having all non-zero levels of two factors in the lexicographic order.
7. Now identify first treatment combination having all non-zero levels of all 3 factors in first row of the design obtained in Step 1.
8. Now identify its baseline treatment combination by replacing the non-zero level of factor 2 by zero and keeping the levels of factor 1 and 3 unchanged and repeat step 4.
   This gives a design in $(v-1)+(s_1-1)(s_2-1)+(s_1-1)(s_3-1)+(s_2-1)(s_3-1)$ +1 blocks.
9. Now for the treatment combination identified in step 7, identify its baseline replacing the non-zero level of factor 3 by zero and keeping the levels of factor 1 and 2 unchanged and repeat step 4.
   This gives a design in $(v-1)+(s_1-1)(s_2-1)+(s_1-1)(s_3-1)+(s_2-1)(s_3-1)$ +2 arrays/blocks.
10. For generating all other $(s_1-1)(s_2-1)(s_3-1)-2$ arrays, repeat steps 7, 8, 9 for every other treatment combination having non-zero levels for all three factors.

A design $d \in D(s_1 \times s_2 \times s_3, b, 2)$ for $s_1 \times s_2 \times s_3$ factorial in given number of arrays, *b*, is said to be ω-optimal for main effects and two factor interactions, if it minimizes

$$T_2 = \sum_{i_1=1}^{s_1-1} var(\hat{\theta}_{i_1 00}) + \sum_{i_2=1}^{s_2-1} var(\hat{\theta}_{0 i_2 0}) + \sum_{i_3=1}^{s_3-1} var(\hat{\theta}_{00 i_3}) + \omega \sum_{i_1=1}^{s_1-1}\sum_{i_2=1}^{s_2-1} var(\hat{\theta}_{i_1 i_2 0}) + \sum_{i_1=1}^{s_1-1}\sum_{i_3=1}^{s_3-1} var(\hat{\theta}_{i_1 0 i_3}) + \sum_{i_2=1}^{s_2-1}\sum_{i_3=1}^{s_3-1} var(\hat{\theta}_{0 i_2 i_3}) \quad (3)$$

The efficiency of the block designs with $b > v-1$, can be obtained by the ratio of criterion T_2 of the design obtained to that of the design with minimum value of T_2 obtained through generating all possible $\binom{v(v-1)/2}{b}$ designs.

1. In the design obtained for *n* = 2 in Case I, search for the first treatment combinations having all non-zero levels in first row of the design obtained in Case I.
2. Now identify its baseline treatment combination by replacing the non-zero level of factor 2 by zero and keeping the levels of factor 1 unchanged.
3. Add a column (array) containing the treatment combination in step 1 in upper dye and its baseline treatment combination in lower dye in the design with $b = v-1$ obtained in Case I.

This yields a design in $b = v$ arrays.

1. For getting a design in $b = v+1$ arrays, repeat steps 1 to 3 for second treatment combinations having all non-zero levels in first row of the design obtained in Case I.
2. For getting a design in any number of arrays $b = v+2$ to $(v-1)+(s_1-1)(s_2-1)$, repeat steps 1 to 3 for all other $(s_1-1)(s_2-1)-2$ treatment combinations having all non-zero levels in the lexicographic order.

Example 2: For 2×3 factorial experiments, the optimal block design in 6 or 7 arrays is

Arrays	1	2	3	4	5	6	7
Dye 1	01	02	10	11	12	11	12
Dye 2	00	00	00	01	02	10	10

For computer aided generation of the efficient block designs under baseline parameterization obtained in Case I and Case II, a software module using C# programming language with ASP.NET platform has been developed and can be access in https://drs.icar.gov.in/dbp/. For running the software module, user is asked to select number of factors ($2 \le n \le 10$) and select the levels ($2 \le s_j \le 10, 1 \le j \le n$) of each of the *n* factors as shown in the following screen

On selecting the number of factors and the levels as, say 2 and 3, $(s_1-1)(s_2-1)+1$ block designs get generated in $b = v-1, v, ..., (v-1)+(s_1-1)(s_2-1)$ blocks/arrays as shown in Fig. 2.

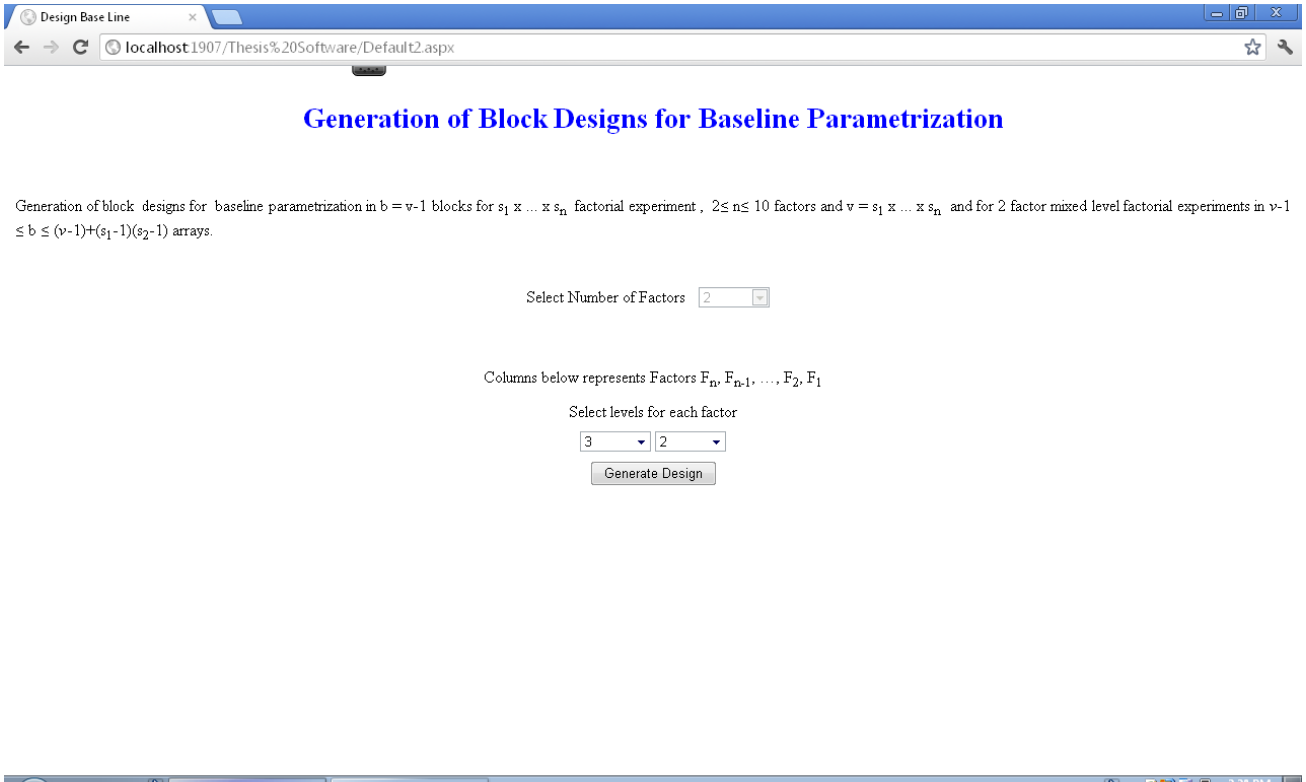

Fig. 1. Selection of number of factors and their levels

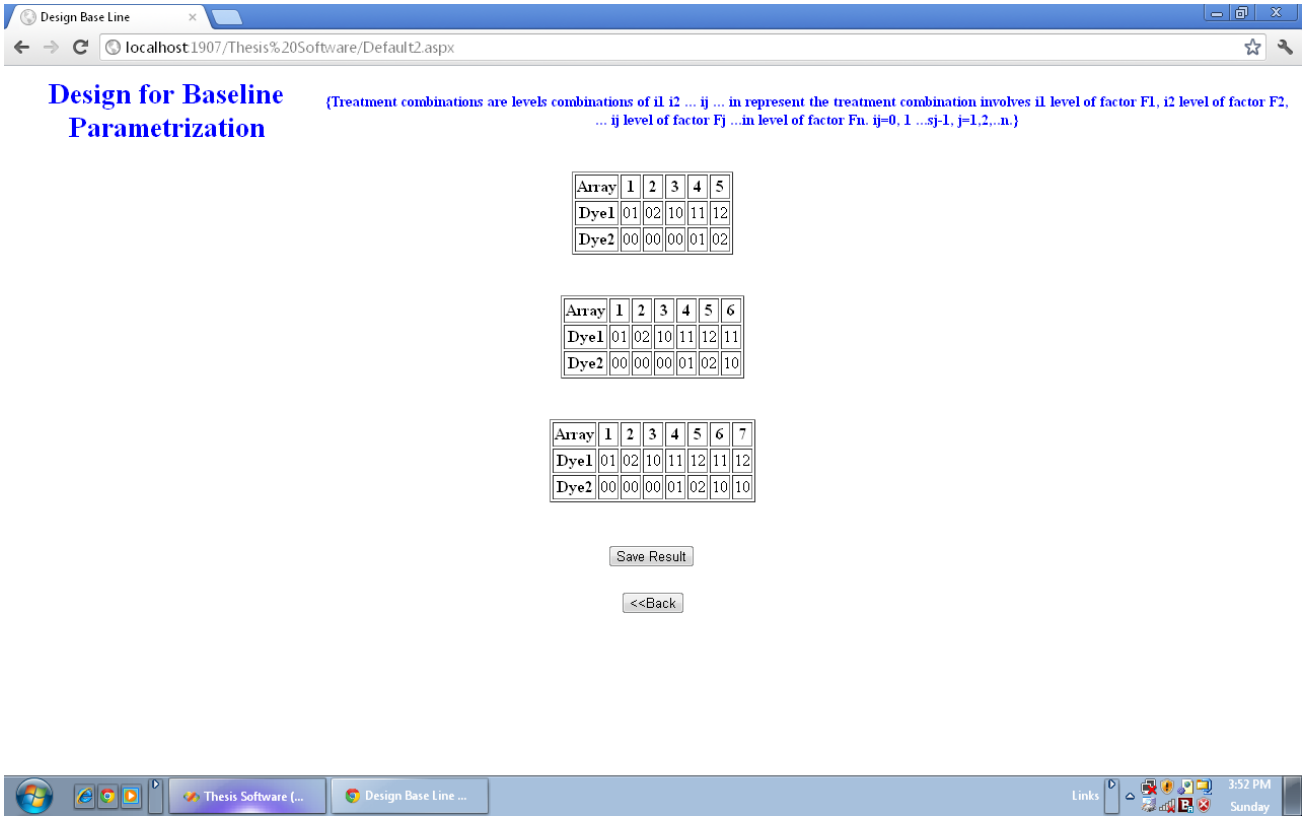

Fig. 2. Block designs in block size 2 in 5, 6 and 7 blocks for 2 ×3 factorial experiment for estimation of factorial effects based on baseline parameterization

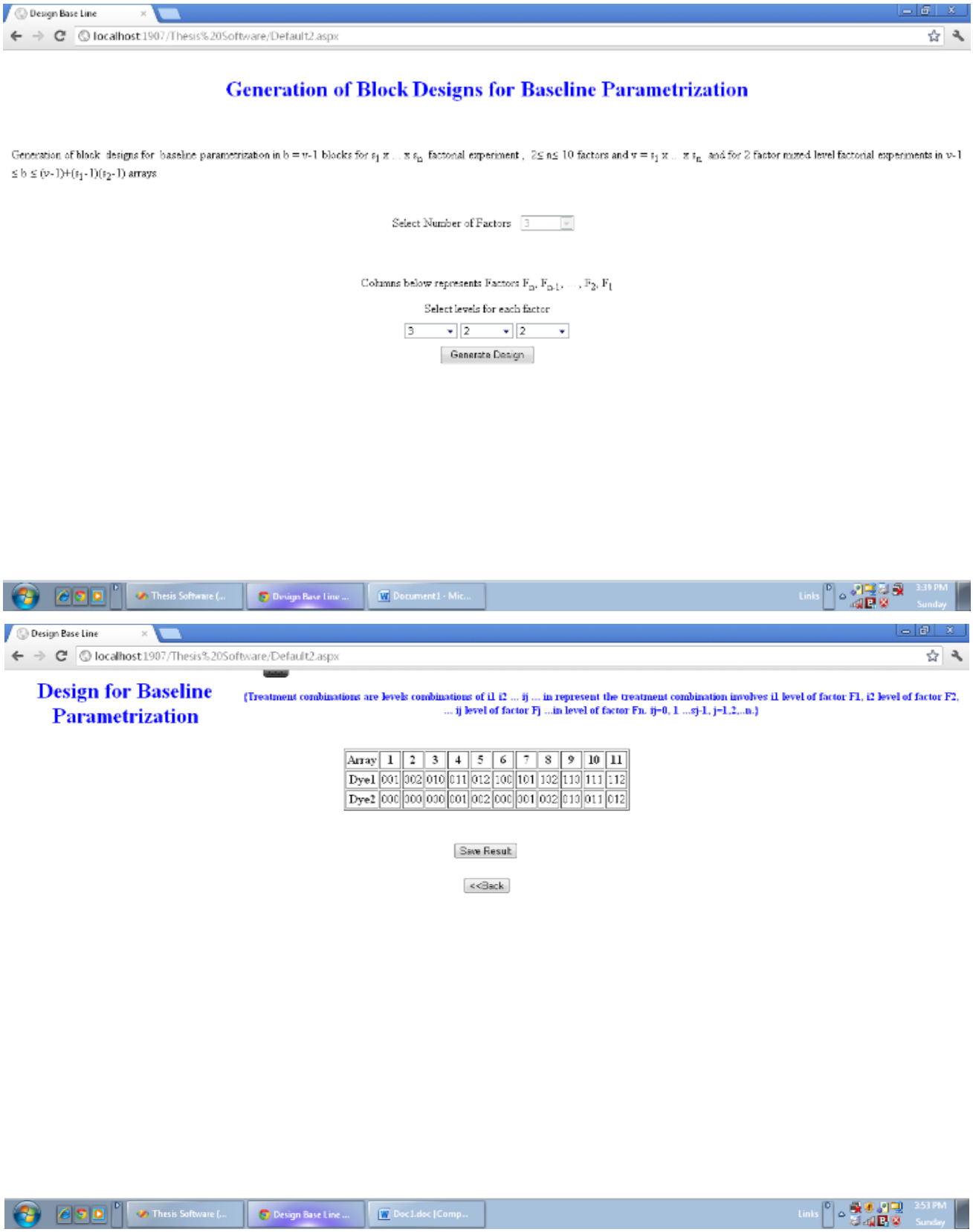

Fig. 3. A block design for 2×2×3 factorial experiment in 11 blocks for estimation of factorial effects under baseline parameterization

If one selects, any number of factors greater than 2, and their levels ($2 \le s_j \le 10, 1 \le j \le n$), then optimal saturated block design in block size 2 in $b = v - 1$ blocks for estimation of factorial effects under baseline parameterization gets generated. One such design in 11 blocks for 2×2×3 factorial experiment is given in Fig. 3.

3. UNSATURATED BLOCK DESIGNS FOR *N* FACTORS

For $n \ge 3$ factors, we have generated optimal saturated block designs in block size 2 for estimation of factorial effects under baseline parameterization. As discussed earlier in a saturated design, there is no error degree of freedom. If the interest is in only main effects and two factor interactions, then it leaves some degrees of freedom for error under the assumption of absence of 3 factor interactions. In this section, we extend the procedure of obtaining block designs for three factors in any number of blocks satisfying

$$v-1 \le b \le (v-1) + \sum_{j=2}^{n} \left\{ (j-1) \sum_{i_1 \ne i_2 \ne \cdots \ne i_j = 1}^{n} (s_{i_1} - 1)(s_{i_2} - 1) \cdots (s_{i_j} - 1) \right\}$$

Example 3: In an experiment it is desired to study the three factors, represented by F_1, F_2, and F_3. The factors F_1 and F_2 have 2 levels {0 and 1} each and factor F_3 has three levels (0, 1 and 2). The cell population corresponds to 12 treatment combinations 000, 001, 002, 010, 011, 012, 100, 101, 102, 110, 111, 112. Let $\tau_{000}, \tau_{001}, \tau_{002}, \tau_{010}, \tau_{011}, \tau_{012}, \tau_{100}, \tau_{101} \tau_{102}, \tau_{110}, \tau_{111}$ and τ_{112} denote the expected log intensities, that is, the effects of these treatment combinations. We focus on the situation where, there is a null state or baseline level, say, 0, of each factor. Then $\theta_{000} = \tau_{000}$ stands for the baseline effect. We consider the baseline parameterization according to which the main effects of F_1, F_2, and F_3 and their two factor interactions respectively are

$\theta_{100} = \tau_{100} - \tau_{000}$, $\theta_{010} = \tau_{010} - \tau_{000}$, $\theta_{001} = \tau_{001} - \tau_{000}$, $\theta_{002} = \tau_{002} - \tau_{000}$, $\theta_{110} = \tau_{110} - \tau_{100} - \tau_{010} + \tau_{000}$,

$\theta_{101} = \tau_{101} - \tau_{100} - \tau_{001} + \tau_{000}$, $\theta_{102} = \tau_{102} - \tau_{100} - \tau_{002} + \tau_{000}$,

$\theta_{011} = \tau_{011} - \tau_{010} - \tau_{001} + \tau_{000}$, $\theta_{012} = \tau_{012} - \tau_{010} - \tau_{002} + \tau_{000}$

The optimal block design for 2×2×3 factorial experiment with baseline parameterization with $b = v - 1 = 11$ blocks obtained through Case I of Section 2 is

Array	1	2	3	4	5	6	7	8	9	10	11
Dye 1	001	002	010	011	012	100	101	102	110	111	112
Dye 2	000	000	000	001	002	000	001	002	010	011	012

Using Steps of construction 5 to 10 given above, we can get 9 more blocks in the following order for getting a block design in any number of blocks from 12 to 20.

Array	12	13	14	15	16	17	18	19	20
Dye 1	011	012	101	102	110	111	111	112	112
Dye 2	010	010	100	100	100	101	110	102	110

4. DISCUSSION

The efficient block designs in blocks of size 2 for estimation of factorial effects under baseline parameterization have been obtained in Sections 2 and 3. These designs have been obtained under the assumption of absence of dye effects. In the presence of dye effects, we need to obtain row-column designs. For obtaining a row-column design, consider the blocks of a block design as columns and use the simplest approach of dye swap arrangement recommended in the literature for each of the columns to make the treatment combinations balanced with respect to dye effects, *i.e.*, if a column consists of treatment combinations in the ordered pair as ($i_1 \cdots i_n$, $j_1 \cdots j_n$), add one more column with the order pair ($j_1 \cdots j_n$, $i_1 \cdots i_n$).Banerjee and Mukerjee (2008) have shown that the row-column design with $b = 2(v-1)$ columns obtained by dye-swap arrangement of a optimal saturated block design in $b = v - 1$ arrays is optimal in the sense that all factorial effects are estimated with minimum variance . In this approach number of arrays becomes twice that of required in block design set up, which may not be feasible due to constraints of resources. Therefore, it is required to see whether it is possible to get a row-column design in same or few extra columns as that of a block design and still the treatment combinations are orthogonal with respect to dyes. For 2×2 factorial, such a design in 4 columns is possible and given as

Array	1	2	3	4
Dye 1	01	00	11	10
Dye 2	00	10	01	11

The procedure of obtaining ω-efficiency of a block design for 2 or 3-factor experiment obtained through augmentation in $b > v - 1$ blocks, the minimum value

of T_1 in (1) or T_2 in (2) can only be obtained through generating all possible $\binom{v(v-1)/2}{b}$ designs. Therefore, it is required to obtain a lower bound to T_1 in (1) or T_2 (2). Further, the method of construction of optimal/efficient designs in number of arrays equal to or more than the number of treatment combinations for more than 3-factor factorial experiments needs to be studied.

ACKNOWLEDGEMENTS

The authors are grateful to the anonymous reviewer for valuable suggestions that have led to improvement of the manuscript. The research work is a part of Ph.D. thesis of first author submitted to PG School of IARI, New Delhi. The IASRI fellowship received is gratefully acknowledged.

REFERENCES

Banerjee, T. and Mukerjee, R. (2008). Optimal factorial designs for CDNA microarray experiments, *Ann. Appl. Stat.*, **2(1)**, 366-385.

Churchill, G.A. (2002). Fundamentals of experimental design for cDNA microarrays. *Nature Genetics (Suppl.)*, **3**, 490-495.

Glonek, G.F.V. and Solomon, P.J. (2004). Factorial and time course designs for cDNA microarray experiments. *Biostatistics*, **5**, 89-111.

Grossman, H. and Schwabe, R. (2008). The relationship between optimal designs for microarray and paired comparison experiments. Preprint

Gupta, S. (2006). Balance factorial design for cDNA microarrray experiments. *Comm. Stat. Theory and Methods*. **35**, 1469-1476.

Kerr, K.F. (2006). Efficient 2^k factorial designs for blocks of size two with microarray applications, *Journal of Quality Tech.*, **38(4)**, 309-318.

Mukerjee, R. and Tang, B. (2012). Optimal fractions of two-level factorials under a baseline parameterization *Biometrika*, **99(1)**, 71-84.

Wang, P.C. (2004). Designing two-level fractional factorial experiments in blocks of size two. *Sankhya*, **66**, 327-342.

Yang, Y.J. and Draper, N.R. (2003). Two-level factorial and fractional factorial designs in blocks of size two, *Journal of Quality Tech.*, **35(3)**, 294-305.

Yang, Y.H. and Speed, T. (2002). Design issues for cDNA microarray experiments. *Nature and Rev. in Genetics.* **3**, 579-588.