"Dispersion" effects are considered in addition to "Location" effects of factors in the inferential procedure of sequential factor screening experiments with m factors each at two levels under search linear models. Search designs in measuring "Dispersion" and "Location" effects of factors are presented for both stage one and stage two of factor screening experiments with 4 < m < 10.
NON-ORTHOGONAL DESIGNS FOR MEASURING DISPERSION EFFECTS IN SEQUENTIAL FACTOR SCREENING EXPERIMENTS USING SEARCH LINEAR MODELS

Subir Ghosh
University of California, Riverside

ABSTRACT

"Dispersion" effects are considered in addition to "Location" effects of factors in the inferential procedure of sequential factor screening experiments with m factors each at two levels under search linear models. Search designs in measuring "Dispersion" and "Location" effects of factors are presented for both stage one and stage two of factor screening experiments with $4 \leq m \leq 10$.

1. INTRODUCTION

In a factor screening experiment the problem is to find a few effective factors from a list of a large number of possible factors influencing the response and screen out all non-effective factors. Watson (1961), Patel (1962) and others considered only "location" effects in their study of influence of factors on response. Taguchi and Wu (1985), Box and Meyer (1986) and others
considered both "location" and "dispersion" effects in the study of influence of factors on response. The concept of "dispersion" effects in factor screening experiments was introduced in Taguchi and Wu (1985) for replicated fractional factorial designs and in Box and Meyers (1986) for unreplicated fractional factorial designs. Srivastava (1975, 1976) introduced the search linear model and showed that factor screening problems can be solved using designs under the search linear model (called search designs) which have smaller numbers of treatments than that in comparable designs in the literature. Ghosh (1979), Ghosh and Avila (1985) presented many such search designs. However, those search designs were constructed to measure the "location" effects only and most of them are unable to measure the "dispersion" effects of factors. In this paper we define the "dispersion" effects of factors under the search linear model and present search designs to measure both the "location" and "dispersion" effects for factor screening experiments.

Throughout this paper, we consider the sequential factor screening experiments with m factors each at two levels under search linear models. At the first stage of the experimentation, the problem is to estimate the dispersion effects at two levels of factors under search linear models, to use the dispersion effects in finding the most effective factor out of m factors and finally to determine the optimum level combination of the most effective factor using the "signal to noise" ratio. At the second stage, the problem is to estimate again the dispersion effects at four level combinations of every two factors under search linear models, to find two most effective factors out of m factors in presence of interactions and finally to determine the optimum level combinations of two most effective factors. The process continues until we find all effective factors. In section 2 of this paper we discuss the models and inferential procedures. In
In a $2^m$ factorial experiment, the treatments are denoted by $(a_1, \ldots, a_m)$, $a_i = 0, 1$; the general mean, main effects and two factor interactions are denoted by $\mu$, $F_i$, and $F_i F_j$, respectively; the observation corresponding to the treatment $(a_1, \ldots, a_m)$ is denoted by $y(a_1, \ldots, a_m)$. The expectation form of the model is

$$E(y(a_1, \ldots, a_m)) = \mu + \sum_{i=1}^{m} b_i F_i + \sum_{i=1}^{m} \sum_{j<i} b_i b_j F_i F_j + \ldots,$$

where $b_i = 1$ if $a_i = 1$ and $b_i = -1$ if $a_i = 0$.

**Stage 1**

At the stage 1 of the experimentation we consider $m$ different models and the $i$th model is

$$E(y(a_1, \ldots, a_i, \ldots, a_m)) = \mu + b_i F_i,$$  

and the observations are uncorrelated, where the variance $\sigma^2_i(0)$ and $\sigma^2_i(1)$, $i = 1, \ldots, m$, are unknown constants, called the "dispersion" effects for the factor $i$. We now consider a fractional factorial design with $N_1$ treatments. (Treatments may or may not be all distinct.) We denote for $u = 0, 1$,

$y_i^u = \text{the vector of } N_1^u \text{ observations corresponding to the treatments with the } i \text{th factor at the level } u,$

$\bar{y}_i^u = \text{the simple arithmetic mean of } N_1^u \text{ observations in } y_i^u,$

$$s_i^2(u) = \frac{(\bar{y}_i^u - \bar{y}_1^u)}{(N_1^u - 1)}.$$
\[ s^2_i = \frac{(N_i^0 - 1) s_i^2(0) + (N_i^1 - 1) s_i^2(1)}{(N_i^0 - 1) + (N_i^1 - 1)}. \]

Note that \( N_i = N_i^0 + N_i^1 \). We now write (2) and (3) as

\[
E \left( \begin{pmatrix} y^0_1 \\ y^1_1 \\ \vdots \end{pmatrix} \right) = \begin{pmatrix} \frac{1}{N_1} N_i^0 - \frac{1}{N_1} N_i^1 \\ \frac{1}{N_1} N_i^1 \end{pmatrix} \begin{pmatrix} \mu \\ F_i \end{pmatrix} = \begin{pmatrix} \frac{1}{N_1} N_i^0 \\ \frac{1}{N_1} N_i^1 \end{pmatrix} \begin{pmatrix} \mu \\ F_i \end{pmatrix},
\]

\[
V \left( \begin{pmatrix} y^0_1 \\ y^1_1 \end{pmatrix} \right) = \begin{pmatrix} \sigma_i^2(0) I \sigma_i^2(1) \\ \sigma_i^2(0) I \sigma_i^2(1) \end{pmatrix},
\]

where \( j \) is a vector with all elements unity. It can be checked that \( E(s_i^2(u)) = \sigma_i^2(u) \) and the generalized least squares (and also the ordinary least squares) estimators of \( (\mu-F_i) \) and \( (\mu+F_i) \) are \( \bar{y}_i^0 \) and \( \bar{y}_i^1 \), respectively. We select the \( i \)th factor as the most effective if \( s_i^2 \) is a minimum for \( i \in \{1, \ldots, m\} \). We select the \( u \)th level of the factor \( i \) as the optimum level if the (square of) "signal to noise" ratio \( [y_i^u/s_i(u)]^2 \) is a maximum.

**Stage 2**

At the stage 2 of the experimentation we consider \( \binom{m}{2} \) different models and the \( (i,j) \)th model is

\[
E(y(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_m)) = \mu + b_i F_i + b_j F_j + b_i b_j F_i F_j,
\]

\[ i, j = 1, \ldots, m, i \neq j, \]

\[
V(y(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_m)) = \sigma_{ij}^2(a_i a_j),
\]

and the observations are uncorrelated, where the variances \( \sigma_{ij}(00), \sigma_{ij}(01), \sigma_{ij}(10) \) and \( \sigma_{ij}(11) \) are unknown constants, called the "dispersion" effects for the factors \( i \) and \( j \). We now consider a fractional factorial design with \( N_2 \) treatments. Note
again that a particular treatment may or may not be replicated in
the design. We denote for \( u, v = 0, 1 \),
\[
\bar{y}_{ij}^{uv} = \text{the vector of } N_{ij}^{uv} \text{ observations corresponding to the treat-
mements with the levels of factors } i \text{ and } j \text{ are } u \text{ and } v, \]
respectively,
\[
y_{ij}^{uv} = \text{the simple arithmetic mean of } N_{ij}^{uv} \text{ observations in } \bar{y}_{ij}^{uv},
\]
\[
s_{ij}^{2}(uv) = \frac{\left( y_{ij}^{uv} - \bar{y}_{ij}^{uv} \right) \left( y_{ij}^{uv} - \bar{y}_{ij}^{uv} \right)}{N_{ij}^{uv} - 1},
\]
\[
s_{ij}^{2} = \sum_{u=0}^{1} \sum_{v=0}^{1} \frac{1}{N_{ij}^{uv} - 1} s_{ij}^{2}(uv)
\]

Observe that \( N_{2} = N_{00}^{ij} + N_{01}^{ij} + N_{10}^{ij} + N_{11}^{ij} \). We write (6) and (7) as
\[
E = \begin{pmatrix}
\bar{y}_{ij}^{00} & \bar{y}_{ij}^{01} & \bar{y}_{ij}^{10} & \bar{y}_{ij}^{11}
\end{pmatrix}
= \begin{pmatrix}
\frac{1}{N_{ij}^{00}} & \frac{1}{N_{ij}^{01}} & \frac{1}{N_{ij}^{10}} & \frac{1}{N_{ij}^{11}}
\end{pmatrix}
\begin{pmatrix}
\frac{1}{N_{ij}^{00}} & \frac{1}{N_{ij}^{01}} & \frac{1}{N_{ij}^{10}} & \frac{1}{N_{ij}^{11}}
\end{pmatrix}
\begin{pmatrix}
\mu \\
F_{1} \\
F_{j} \\
F_{1j}
\end{pmatrix}
\]
\[
= \begin{pmatrix}
\frac{1}{N_{ij}^{00}} & 0 & 0 & 0 \\
0 & \frac{1}{N_{ij}^{01}} & 0 & 0 \\
0 & 0 & \frac{1}{N_{ij}^{10}} & 0 \\
0 & 0 & 0 & \frac{1}{N_{ij}^{11}}
\end{pmatrix}
\begin{pmatrix}
\mu - F_{1j} \\
F_{1} + F_{1j} \\
F_{j} - F_{1j} \\
F_{1j}
\end{pmatrix}
\]
It can be seen that \( E(s_{ij}^2(\text{uv})) = \sigma_{ij}^2(\text{uv}) \) and the generalized least squares (and also the ordinary least squares) estimators of 
\[
\begin{align*}
(y_{i1}^{00})^2 &= \sigma_{ij}^2(00)I_{00}^{ij} \\
(y_{i1}^{01})^2 &= \sigma_{ij}^2(01)I_{01}^{ij} \\
(y_{i1}^{10})^2 &= \sigma_{ij}^2(10)I_{10}^{ij} \\
(y_{i1}^{11})^2 &= \sigma_{ij}^2(11)I_{11}^{ij}
\end{align*}
\]

are \( y_{ij}^{00}, y_{ij}^{01}, y_{ij}^{10}, \) and \( y_{ij}^{11}, \) respectively. We select the factors \( i \) and \( j \) as the most effective factors if \( s_{ij}^2 \) is a minimum for \( i,j \) in \{1,\ldots,m\}. We select the level \( (u,v) \) of the factors \( i \) and \( j \) as the optimum level if the (square of) "signal to noise" ratio
\[
\frac{y_{ij}^2}{s_{ij}^2(\text{uv})^2}
\]
is a maximum.

We stop at the stage 2 if Minimum \( s_{ij}^2 \) is very close to Minimum \( s_{ij}^2 \); otherwise we go to the stage 3. The stage 3 inferential procedure is similar to those in stages 1 and 2.

3. DESIGN

The orthogonal fractional factorial designs are indeed efficient but have restriction on the number of treatments in the design. The designs used in Taguchi and Wu (1985), Box and Meyer (1986) are all classical orthogonal designs. In this section we construct designs to measure the "dispersion" effects defined in section 2 for the stage 1 and stage 2 of the experiment, relaxing the restriction on the number of treatments. We want to have enough observations (at least two!) in measuring the "dispersion" effects, the number of treatments to be small and furthermore the designs to be near orthogonal.
3.1 DESIGNS FOR THE STAGE 1 EXPERIMENT

The design condition obtained in Srivastava (1975) requires that for all \(i\) and \(j\), \(i \neq j\), \(i, j \in \{1, \ldots, m\}\), \(\nu, F_i\) and \(F_j\) are unbiasedly estimable (u.e.) under the model (1) assuming all other parameters to be zero. We denote a design by a treatment matrix \(T(N \times m)\) with rows as treatments and columns as factors.

**Theorem 1.** A necessary and sufficient condition that a treatment matrix \(T(N \times m)\) satisfies the design condition is that for every submatrix \(T_i(N \times 2)\) of \(T\), at least three pairs out of the four pairs \((00), (01), (10)\) and \((11)\) appear as rows in \(T_i\).

Proof. To estimate the three parameters \(\nu, F_i\) and \(F_j\), we need three independent equations in parameters from (1) under the assumption that all other parameters are exactly equal to zero. For the submatrix \(T_i(N \times 2)\) of \(T\) corresponding to the factors \(i\) and \(j\), any three rows out of four possible distinct rows \((00), (01), (10)\) and \((11)\) will give three independent equations in \(\nu, F_i\) and \(F_j\) under (1). This completes the proof.

The characterization in Theorem 1 is so simple that the checking can even be done by eye inspection. A result similar to Theorem 1 is also available in Srivastava (1975).

We now present some designs for \(4 \leq m \leq 10\) satisfying the condition in Theorem 1. In the Table I, for brevity, we indicate a treatment (i.e., a row) in \(T\) by the positions where the level 1 is occurring. The treatment \((0, \ldots, 0)\) will be denoted by \(\theta\). To illustrate this for \(m=4\), \(T(5 \times 4)\) matrix is given below.

\[
T = \begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0
\end{bmatrix}
\]
We denote this \( T \) in our representation as \((1234, 0, 12, 34, 13)\). Notice that in every column of \( T \), \( u(=0,1) \) is appearing either twice or thrice.

**TABLE I.**

**DESIGNS \((4 < m < 10)\) FOR THE STAGE 1 EXPERIMENT**

<table>
<thead>
<tr>
<th>( m )</th>
<th>Designs</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1234, 0, 12, 34, 13</td>
</tr>
<tr>
<td>5</td>
<td>12345, 12, 13, 24, 35</td>
</tr>
<tr>
<td>6</td>
<td>3456, 1256, 1234, 146, 235</td>
</tr>
<tr>
<td>7</td>
<td>127, 3456, 1234, 567, 1357, 1234567</td>
</tr>
<tr>
<td>8</td>
<td>1278, 3456, 3478, 1256, 1357, 1234</td>
</tr>
<tr>
<td>9</td>
<td>1278, 3456, 34789, 1256, 13579, 1234</td>
</tr>
<tr>
<td>10</td>
<td>1278910, 3456, 3478910, 1256, 13579, 1234, 24689</td>
</tr>
</tbody>
</table>

### 3.2 DESIGNS FOR THE STAGE 2 EXPERIMENT

The design condition obtained in Srivastava (1975) requires for all \( i, j (i \neq j), u, v (u \neq v), (i, j) \neq (u, v), i, j, u, v \in \{1, \ldots, m\} \), \( u, F_i, F_j, F_{ij}, F_u, F_v \) and \( F_{uv} \) are u.e. under the model (1) assuming all other parameters to be zero.

Theorem 2. Suppose a design \( T(N_2 \times m) \) is such that \( u, F_i, F_j, F_u \) are u.e. for every distinct \( i, j \) and \( u \) in \( \{1, \ldots, m\} \) under (1) assuming all other parameters except \( F_{ij} \) and \( F_{iu} \) to be zero. Then \( F_{ij} \) and \( F_{iu} \) are also u.e.

Proof. It can be seen using Theorem 1 that \( F_{ij} + F_{iu} \) and \( -F_{ij} + F_{iu} \) are u.e. This completes the proof.

Theorem 3. Suppose a design \( T(N_2 \times m) \) is such that \( u, F_i, F_j, F_u \) and \( F_v \) are u.e. for every distinct \( i, j, u \) and \( v \) in \( \{1, \ldots, m\} \)
under (1) assuming all other parameters except \( F_i F_j \) and \( F_u F_v \) to be zero. Then \( F_i F_j \) and \( F_u F_v \) are u.e. if and only if for \( T_i(N_2x2) \) corresponding to factors \((i,j)\) and \( T_2(N_2x2) \) corresponding to factors \((u,v)\) there are two distinct rows in \( T \) with the number of 1's in \( T_1 \) and \( T_2 \) as follows.

<table>
<thead>
<tr>
<th>Row</th>
<th>The number of 1's in ( T_1 )</th>
<th>The number of 1's in ( T_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>even</td>
<td>even</td>
</tr>
<tr>
<td>2</td>
<td>odd</td>
<td>odd</td>
</tr>
</tbody>
</table>

Proof. It can be seen that from the rows 1 and 2 that \( F_i F_j F_i F_j \) and \( -F_i F_j F_i F_j \) are u.e. This completes the proof.

**3.2.1 CONSTRUCTION OF DESIGNS \( m=4 \)**

The treatment matrix consists of 9 treatments as rows; the first eight treatments are solutions of \( x_1+x_2+x_3+x_4=0 \) over the finite field \( GF(2) \) and the last treatment is \((1000)\). The treatment matrix is represented as \((1234,0,12,34,13,24,14,23,1)\).

The first eight treatments form an orthogonal resolution IV plan. It follows from Theorem 2 that \( \nu, F_i, F_j, F_i F_j, F_i u, F_i u \) are u.e. for any three distinct \( i,j,u \) in \([1,\ldots,m]\). It can be checked that \( \nu, F_i, F_j, F_i F_j, F_i u, F_j u, F_i F_j F_i u, F_i F_j F_j u \) are u.e. for any four distinct integers \( i,j,u \) and \( v \). It can be seen that for all rows in the treatment matrix the numbers of 1's in submatrices corresponding to \((i,j)\) and \((u,v)\) are either (even, even) or (odd, odd) and therefore the treatment matrix does not satisfy the conditions in Theorem 3. When the treatment \((1000)\) is added to the treatment matrix, it follows from Theorem 3 that \( \nu, F_i, F_j, F_i F_j, F_i u, F_j u, F_i F_j F_i u, F_i F_j F_j u \) are u.e.
m=5

The treatment matrix consists of 11 treatments as all \((1\times5)\) vectors with two elements are unity and the other elements are zero and furthermore the treatment with the levels of all factors are unity.

This is a balanced array of full strength. There is a complete \(2^3\) factorial experiment w.r.t. any three factors and there is a resolution V plan for a \(2^4\) factorial experiment w.r.t. any four factors. Thus the design satisfies the design condition.

m=6

Design 1: The treatment matrix consists of 16 treatments as follows:

The 14 treatments are solutions of \(x_1+x_2+x_3+x_4=0\) and 
\(x_3+x_4+x_5+x_6=0\) over the finite field \(GF(2)\) excluding the treatments \((000000)\) and \((111111)\); and the other two treatments are \((10\ 00\ 00)\) and \((00\ 00\ 01)\).

It can be seen that all \(2^3\) distinct treatments are present in rows of \(T\) w.r.t. any three factors \(F_1, F_j\) and \(F_u\). It thus follows \(u, F_1, F_j, F_1^F_j, F_u\) and \(F_1^F_u\) are u.e. for every three distinct factors \(i, j\) and \(u\). For four factors of the type \((F_1,F_2,F_3,F_4),(F_1,F_2,F_5,F_6)\) and \((F_3,F_4,F_5,F_6)\), we get basically the design for \(m=4\) with replications. We get a resolution V plan for a \(2^4\) factorial experiment w.r.t. any four factors other than \((F_1,F_2,F_3,F_4),(F_1,F_2,F_5,F_6)\) and \((F_3,F_4,F_5,F_6)\). Therefore, \(w, F_1,F_j,F_1^F_j,F_u,F_v\) and \(F_1^F_uF_v\) are u.e. for every four distinct factors \(i, j, u\) and \(v\).

Design 2: The treatment matrix consists of 15 treatments as follows.

The 13 treatments are solutions of \(x_1+x_2+x_3+x_4=0\) and 
\(x_3+x_4+x_5+x_6=0\) over the finite field \(GF(2)\) excluding the treatments \((01\ 01\ 01), (10\ 10\ 10)\) and \((00\ 00\ 00)\); and the other two treatments are \((10\ 00\ 00)\) and \((00\ 00\ 01)\).
The argument is similar to that in Design 1.

\( m=7 \)

The treatment matrix consists of 17 treatments as follows.

The 14 treatments are solutions of \( x_1+x_2+x_3+x_4=0, \)
\( x_3+x_4+x_5+x_6=0 \) and \( x_3+x_4+x_5+x_7=0 \) over the finite field GF(2)
excluding the treatments (1111111) and (0000000); and the other 3 treatments
are \((10 \ 00 \ 00 \ 0), \ (00 \ 00 \ 01 \ 0), \ (00 \ 00 \ 00 \ 1). \)[One may
keep (1111111) in the design although the design is all right
without it.]

\( m=8 \)

The treatment matrix consists of 18 treatments as follows.

The 14 treatments are solutions of \( x_1+x_2+x_3+x_4=0, \)
\( x_3+x_4+x_5+x_6=0, \ x_5+x_6+x_7+x_8=0 \) and \( x_1+x_3+x_5+x_7=0 \) over the finite
field GF(2) excluding the treatments \((00 \ 00 \ 00 \ 00) \) and
\((11 \ 11 \ 11 \ 11); \) and the other 4 treatments are \((10 \ 00 \ 00 \ 00), \)
\((00 \ 00 \ 10 \ 00), \ (00 \ 00 \ 00 \ 01). \)

\( m=9 \)

The treatment matrix consists of 28 treatments as follows.

The treatments are solutions of \( x_1+x_2+x_3+x_4=0, \)
\( x_3+x_4+x_5+x_6=0, \ x_5+x_6+x_7+x_8=0 \) and \( x_1+x_3+x_5+x_7=0 \) over the finite field GF(2)
excluding the treatments \((000000000), \ (000000001), \)
\((11 \ 11 \ 11 \ 11) \) and the other 5 treatments are \((10 \ 00 \ 00 \ 00), \)
\((00 \ 00 \ 10 \ 00), \ (00 \ 00 \ 00 \ 10), \ (00 \ 00 \ 00 \ 01). \)

\( m=10 \)

The treatment matrix consists of 35 treatments as follows.

The 30 treatments are solutions of \( x_1+x_2+x_3+x_4=0, \)
\( x_3+x_4+x_5+x_6=0, \ x_5+x_6+x_7+x_8=0, \ x_7+x_8+x_9+x_{10}=0 \) and \( x_1+x_3+x_5+x_7=0 \)
over the finite field GF(2) excluding the treatments \((00 \ 00 \ 00 \ 00), \ (11 \ 11 \ 11 \ 11), \) and the other five treatments
are \((10 \ 00 \ 00 \ 00), \ (00 \ 00 \ 10 \ 00), \ (00 \ 00 \ 00 \ 10), \ (00 \ 00 \ 00 \ 01). \)

The arguments for the cases \( m=7, \ 8, \ 9, \ 10 \) are similar to the
argument for \( m=6. \)
4. FINAL REMARKS

This paper touches the major developments in factor screening designs over the period of more than 25 years and deals with the important issue of measuring the "dispersion" effects in addition to the "location" effects of factors using search linear models. The influence of Professor R. C. Bose to this author is not easy to describe in words and this author takes pride in dedicating this work to honor Professor R. C. Bose.

ACKNOWLEDGEMENTS

The work of the author is sponsored by the Air Force Office of Scientific Research under Grant AFOSR-86-0048.

BIBLIOGRAPHY


Key Words and Phrases: Dispersion effects, Factor Screening, Linear models, Search designs, Sequential experiments.