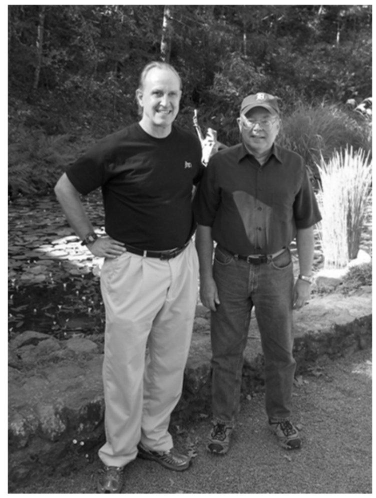
Design of Experiments: New Methods and How to Use Them

February 8, 2012
Doug Montgomery, Arizona State University
Brad Jones, SAS Institute



1

Design Optimality

- One of the truly great paradigm shifts in DOX
- Can create custom designs for almost any situation
- Modern software makes this easy for at least some optimality criteria
- What optimality criteria should we use?

2

Three Historical Optimality Criteria:

The *D*-criterion: **M** = $\frac{\mathbf{X'X}}{N}$

$$D_{eff} = \left[\frac{|\mathbf{X'X}|}{Max[|\mathbf{X'X}|]} \right]^{1/p}$$

Maximize the determinant of **M**

The *G*-criterion: Maximize the *SPV* over *R*

$$SPV = \frac{NVar[\hat{y}(\mathbf{x})]}{\sigma^2} = N\mathbf{x}'^{(m)} (\mathbf{X'X})^{-1} \mathbf{x}^{(m)}$$

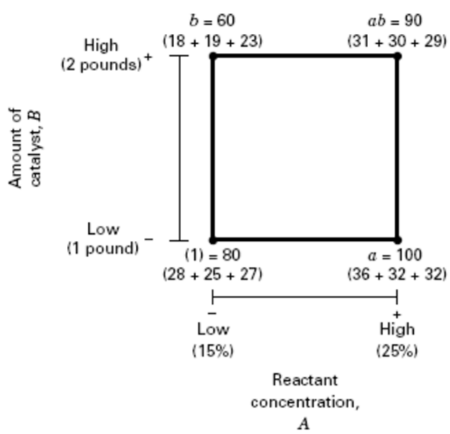
$$G_{eff} = \frac{p}{\max_{x \in R}(SPV)}$$

The *I*-criterion: Minimize the average prediction variance over *R*

$$I = \frac{1}{A} \int_R \frac{NVar[\hat{y}(\mathbf{x})]}{\sigma^2} d\mathbf{x}$$

$$I_{eff} = \frac{Min(APV(D))}{APV(D)}$$

3



$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \text{ where } \mathbf{y} = \begin{bmatrix} (1) \\ a \\ b \\ ab \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix}, \text{ and } \boldsymbol{\epsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

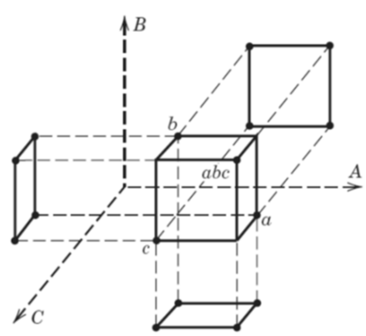
4

Why do Regular Fractional Factorial Designs Work for Factor Screening?

- The **sparsity of effects** principle
  - There may be lots of factors, but few are important
  - System is dominated by main effects, low-order interactions
- The **projection** property
  - Every fractional factorial contains full factorials in fewer factors
- **Optimality** properties
  - Regular fractional factorial designs are optimal by every criterion for every model they can fit.

5

Projection of Fractional Factorials



■ **FIGURE 8.2 Projection of a 2_{III}^{3-1} design into three 2^2 designs**

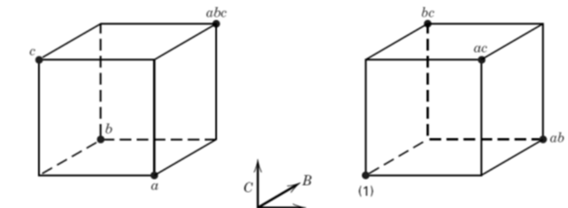
Every fractional factorial contains full factorials in fewer factors

The "flashlight" analogy

A one-half fraction will project into a full factorial in any $k-1$ of the original factors

6

The Regular One-Half Fraction of the 2^3



(a) The principal fraction, $I = +ABC$ (b) The alternate fraction, $I = -ABC$

■ **FIGURE 8.1** The two one-half fractions of the 2^3 design

■ **TABLE 8.2**
The Two One-Half Fractions of the 2^3 Design

	Full 2^2 Factorial (Basic Design)		2^{3-1}_{III}, $I = ABC$			2^{3-1}_{III}, $I = -ABC$		
Run	A	B	A	B	C = AB	A	B	C = − AB
1	−	−	−	−	+	−	−	−
2	+	−	+	−	−	+	−	+
3	−	+	−	+	−	−	+	+
4	+	+	+	+	+	+	+	−

Suppose we believe in the model:

$y = X_1\beta_1 + \varepsilon$

But the true model is:

$y = X_1\beta_1 + X_2\beta_2 + \varepsilon$

Then, the Alias matrix is:

$A = (X_1^t X_1)^{-1} X_1^t X_2$

The model for our specific example on the next slide is:

$y = \beta_0 + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + \varepsilon$

In the notation defined above.

$$\boldsymbol{\beta}_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \quad \text{and} \quad \mathbf{X}_1 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Suppose that the true model contains all the two-factor interactions, so that

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \epsilon$$

and

$$\boldsymbol{\beta}_2 = \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix}, \quad \text{and} \quad \mathbf{X}_2 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Now

$$\mathbf{X}_1'\mathbf{X}_1 = 4\,\mathbf{I}_4 \quad \text{and} \quad \mathbf{X}_1'\mathbf{X}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix}$$

9

Therefore,

$$(\mathbf{X}_1'\mathbf{X}_1)^{-1} = \frac{1}{4}\mathbf{I}_4$$

and

$$E(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\beta}_1 + \mathbf{A}\,\boldsymbol{\beta}_2$$

$$E\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \frac{1}{4}\mathbf{I}_4 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix}$$

$$= \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{12} \\ \beta_{13} \\ \beta_{23} \end{bmatrix}$$

$$= \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} 0 \\ \beta_{23} \\ \beta_{13} \\ \beta_{12} \end{bmatrix}$$

$$= \begin{bmatrix} \beta_0 \\ \beta_1 + \beta_{23} \\ \beta_2 + \beta_{13} \\ \beta_3 + \beta_{12} \end{bmatrix}$$

Main effects aliased with the 2fis

10

The Alias Matrix:

Columns represent the additional parameters in the full model

Rows represent the parameters in the model that you fit

Effect	**A*B**	**A*C**	**B*C**
Intercept	0	0	0
A	0	0	1
B	0	1	0
C	1	0	0

Note that every element in the alias matrix is either a 1 or a 0.
More generally, all regular fractional factorial designs have alias matrices having entries of -1, 0 or 1.

11

Is designing for either complete aliasing or no aliasing a good idea?

Suppose we have six factors and we suspect that many two-factor interactions may be active. What are our choices?

1. Full factorial – 64 runs – no aliasing
2. Regular half fraction – 32 runs – resolution VI
3. Regular quarter fraction – 16 runs – resolution IV

These are the only regular fractional factorial design choices that allow any ability to estimate two-factor interactions.

Are there other choices?

How about using a 24 run Plackett-Burman design?

Demonstration 1

VIFs for a 24 run Plackett-Burman design and a main effects model

Variance Inflation Factors

Parameter	VIF
Intercept	1
X1	1
X2	1
X3	1
X4	1
X5	1
X6	1

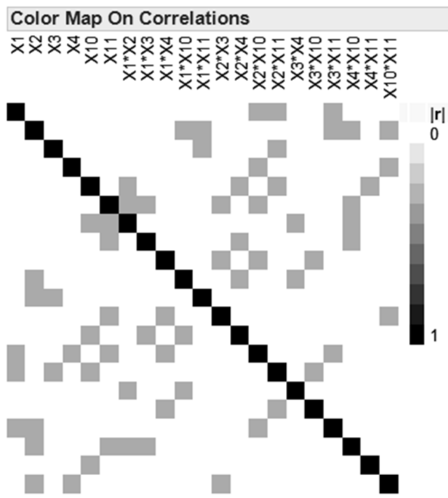
First 6 columns

Variance Inflation Factors

Parameter	VIF
Intercept	1
X1	1
X2	1
X3	1
X4	1
X10	1
X11	1

Best 6 columns

Column correlation cell plot for a 24 run Plackett-Burman design



Absolute correlations are either zero or one-third.

VIFs for a 24 run Plackett-Burman design and a two-factor interactions model

Variance Inflation Factors

Parameter	VIF
Intercept	1
X1	29.2
X2	3.161
X3	5.363
X4	7.323
X5	14.41
X6	18.75
X1*X2	66.65
X1*X3	12.81
X1*X4	13.38
X1*X5	21.16
X1*X6	10.81
X2*X3	3.645
X2*X4	5.411
X2*X5	3.282
X2*X6	39.27
X3*X4	19.4
X3*X5	4.032
X3*X6	8.75
X4*X5	19.4
X4*X6	34.06
X5*X6	21.57

First 6 columns

Variance Inflation Factors

Parameter	VIF
Intercept	1
X1	2.108
X2	3.284
X3	1.641
X4	1.695
X10	2.818
X11	3.978
X1*X2	2.17
X1*X3	2.871
X1*X4	1.962
X1*X10	1.878
X1*X11	1.693
X2*X3	1.836
X2*X4	2.17
X2*X10	3.161
X2*X11	2.865
X3*X4	1.448
X3*X10	1.722
X3*X11	1.718
X4*X10	2.882
X4*X11	1.313
X10*X11	2.016

Best 6 columns

Are there other choices?

How about using a 24 run D-optimal design?

Demonstration 2

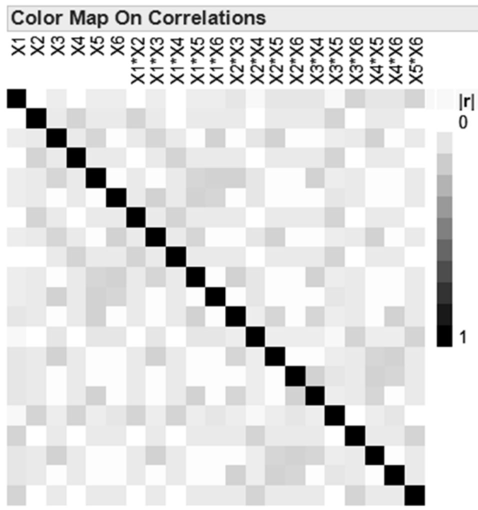
VIFs for a 24 run D-optimal design and a main effects model

Variance Inflation Factors

Parameter	VIF
Intercept	1.043
X1	1.043
X2	1.043
X3	1.047
X4	1.043
X5	1.047
X6	1.025

Note that the design was created to fit the two-factor interactions model.

Column correlation cell plot for a 24 run D-optimal design



Maximum absolute correlation is less than 0.2.

VIFs for a 24 run D-optimal design and a two-factor interactions model

Variance Inflation Factors

Parameter	VIF
Intercept	1.215
X1	1.171
X2	1.165
X3	1.191
X4	1.165
X5	1.191
X6	1.158
X1*X2	1.165
X1*X3	1.191
X1*X4	1.165
X1*X5	1.191
X1*X6	1.158
X2*X3	1.191
X2*X4	1.171
X2*X5	1.191
X2*X6	1.158
X3*X4	1.191
X3*X5	1.207
X3*X6	1.165
X4*X5	1.191
X4*X6	1.158
X5*X6	1.165

New Optimality Criterion Development

Consider two kinds of effects:

Primary effects are ones you are sure you want to estimate. There are p_1 of these.

Potential effects are ones you are afraid to ignore. There are p_2 of these.

Usually, for sample size, n,

$p_1 < n < p_1 + p_2$

21

Alias Optimality Criterion Development

The full model containing both primary and potential terms is:

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

Suppose you fit the model for just the primary terms. Let *b*, be the least squares estimate of β_1. Then the expected value of *b* is

$$E(b) = \beta_1 + A\beta_2$$

$$where$$

$$A = (X_1'X_1)^{-1}X_1'X_2$$

22

Alias Optimality Criterion Development

We use d to denote a design, X(d) to denote the model matrix corresponding to design d, A(d) to denote the alias matrix for design d, and d^* to denote a D-optimal design for the primary model. The D-efficiency of a design d is:

$$D_e(d) = \left[\frac{|\mathbf{X}_1(d)'\mathbf{X}_1(d)|}{|\mathbf{X}_1(d^*)'\mathbf{X}_1(d^*)|}\right]^{1/p_1}$$

23

Alias Optimality Criterion – Constrained Version

$$\min_d \mathrm{Tr}[\mathbf{A}(d)'\mathbf{A}(d)], \text{ subject to } D_e(d) \geq l_D$$

where l_D denotes the experimenter's lower bound for D efficiency and $0 < l_D < 1$.

24

Are there other choices?

How about using a 24 run Alias optimal design?

Demonstration 3

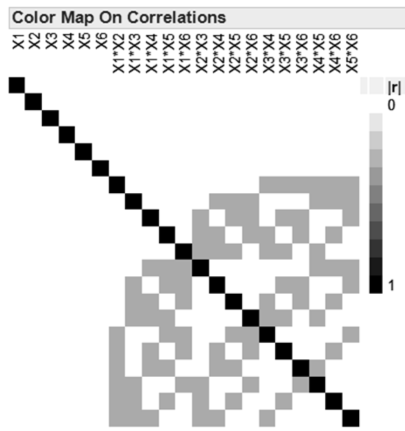
VIFs for a 24 run Alias optimal design and a main effects model

Variance Inflation Factors

Parameter	VIF
Intercept	1
X1	1
X2	1
X3	1
X4	1
X5	1
X6	1

Note that the design was created to fit the main effect model but de-alias two-factor interactions.

Column correlation cell plot for a 24 run Alias optimal design



Main effects and two-factor interactions are uncorrelated.

Two-factor interactions are correlated either zero or one-third.

VIFs for a 24 run Alias optimal design and a two-factor interactions model

Variance Inflation Factors

Parameter	VIF
Intercept	1
X1	1
X2	1
X3	1
X4	1
X5	1
X6	1
X1*X2	2
X1*X3	2
X1*X4	4.5
X1*X5	2
X1*X6	2
X2*X3	2
X2*X4	2
X2*X5	2
X2*X6	5
X3*X4	2
X3*X5	5

Only 11 of the 15 two-factor interactions are estimable but the VIFs are reasonably small.

Introducing Definitive Screening

Jones and Nachtsheim (2011) provide a class of screening designs for 5 factors and up.

These are three level designs for numeric factors.

We will discuss their properties, show how to construct them and give an example.

29

Motivation: Problems with Standard Screening Designs

Resolution III designs confound main effects and two-factor interactions.

Plackett-Burman designs have "complex aliasing" of the main effects by two-factor interactions.

Resolution IV designs confound two-factor interactions with each other, so if one is active, you usually need further runs to resolve the active effects.

Center runs give an overall measure of curvature but you do not know which factor(s) are causing the curvature.

Even the nonregular orthogonal designs we just discussed have aliasing of +0.5 or -0.5 of some main effects by some two-factor interactions. Plus they have no way of resolving quadratic effects.

Screening Conundrum – Two Models

The full model containing both 6 first-order and 15 second-order terms is:

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$$

But n = 12, so we can only fit the intercept and the main effects:

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}^*$$

Standard result: some main effects estimates are biased:

$$\mathrm{E}(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\beta}_1 + \mathbf{A}\boldsymbol{\beta}_2$$

where the "alias" matrix is: $\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$

31

Alias Matrix of Plackett-Burman design

Alias Matrix

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0.333	-0.33	0.333	0.333	0.333	-0.33	0.333	-0.33	0.333	-0.33
B	0	0.333	-0.33	0.333	0.333	0	0	0	0	-0.33	-0.33	-0.33	-0.33	0.333	0.333
C	0.333	0	0.333	-0.33	0.333	0	-0.33	-0.33	-0.33	0	0	0	-0.33	0.333	-0.33
D	-0.33	0.333	0	-0.33	0.333	-0.33	0	-0.33	0.333	0	-0.33	0.333	0	0	-0.33
E	0.333	-0.33	-0.33	0	-0.33	-0.33	-0.33	0	0.333	-0.33	0	-0.33	0	-0.33	0
F	0.333	0.333	0.333	-0.33	0	-0.33	0.333	0.333	0	0.333	-0.33	0	-0.33	0	0

PB (non-regular) design has "complex aliasing"

If only there were another six factor 12 run design with this alias matrix:

Alias Matrix

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Screening Design – Wish List

1. Orthogonal main effects.
2. Main effects uncorrelated with two-factor interactions and quadratic effects.
3. Estimable quadratic effects – three-level design.
4. Small number of runs – order of the number of factors.
5. Good projective properties.

Design Structure

1. Scaled values of each element of the design are either +1, -1 or 0.
2. Each even numbered row is a mirror image of the previous row.
3. For the *kth* column, rows *2k* and 2*k*−*1* contain zeros.
4. The last row contains all zero elements.

Run	A	B	C	D	E	F
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0
13	0	0	0	0	0	0

How did we find this design? – we used an optimization algorithm

$$D_e(d) = \left[\frac{|\mathbf{X}_1(d)'\mathbf{X}_1(d)|}{|\mathbf{X}_1(d_D)'\mathbf{X}_1(d_D)|}\right]^{1/p_1}.$$

The optimization of interest is

$$\min_{d\in\mathcal{D}} \mathrm{Tr}[\mathbf{A}(d)'\mathbf{A}(d)], \quad \text{subject to } D_e(d) \geq l_D,$$

Problem: Finding orthogonal main effects plans.

Algorithmic approach gave orthogonal main effects plans for:

- 6 factors
- 8 factors
- 10 factors
- but not 12 factors. ☹

Orthogonal Main Effects Design Construction

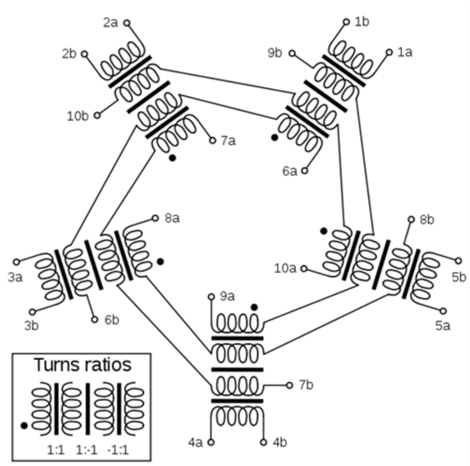
Conference matrix

From Wikipedia, the free encyclopedia

In mathematics, a **conference matrix** (also called a **C**-**matrix**) is a square matrix C with 0 on the diagonal and +1 and −1 off the diagonal, such that $C^T C$ is a multiple of the identity matrix I. Thus, if the matrix has order n, $C^T C = (n-1)I$.

Conference matrices first arose in connection with a problem in telephony.[3] They were first described by Vitold Belevitch who also gave them their name. Belevitch was interested in constructing ideal telephone conference networks from ideal transformers and discovered that such networks were represented by conference matrices, hence the name.[4] Other applications are in statistics,[5] and another is in elliptic geometry.[6]

Conference Matrices &Telephony



From Wikipedia article on Conference Matrices

Conference Matrix of Order 6

$$\begin{pmatrix} 0 & +1 & +1 & +1 & +1 & +1 \\ +1 & 0 & +1 & -1 & -1 & +1 \\ +1 & +1 & 0 & +1 & -1 & -1 \\ +1 & -1 & +1 & 0 & +1 & -1 \\ +1 & -1 & -1 & +1 & 0 & +1 \\ +1 & +1 & -1 & -1 & +1 & 0 \end{pmatrix}$$

The definitive screening design can be constructed by

$$D = \begin{pmatrix} C \\ -C \\ \mathbf{0} \end{pmatrix},$$

where C is an $m \times m$ conference matrix and $\mathbf{0}$ is a $1 \times m$ zero matix .

Statistical Properties

1. Orthogonal for the main effects.
2. The number of required runs is only one more than twice the number of factors. ***
3. Unlike resolution III designs, main effects are independent of two-factor interactions.
4. Unlike resolution IV designs, two-factor interactions are not completely confounded with other two-factor interactions, although they may be correlated
5. Unlike resolution III, IV and V designs with added center points, all quadratic effects are estimable in models comprised of any number of linear and quadratic main effects terms.
6. Quadratic effects are orthogonal to main effects and not completely confounded (though correlated) with interaction effects.
7. If there are more than six factors, the designs are capable of efficiently estimating all possible full quadratic models involving three or fewer factors

Example 1

6 factors and 13 runs.

Run	A	B	C	D	E	F
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0
13	0	0	0	0	0	0

Note that even numbered row mirrors the previous row.

D-efficiency is 85.5% and the design is orthogonal for the main effects.

Alias Matrices

The D-optimal design with one added center point has substantial aliasing of each main effect with a number of two-factor interactions.

Alias Matrix

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	-0.33	0.333	-0.33	0.333	-0.33	-0.33	0.333	0.333	0.333	0.333
B	0	-0.33	0.333	-0.33	0.333	0	0	0	0	-0.33	-0.33	0.333	-0.33	-0.33	0.333
C	-0.33	0	-0.33	-0.33	0.333	0	-0.33	-0.33	0.333	0	0	0	0.333	-0.33	0.333
D	0.333	-0.33	0	0.333	0.333	-0.33	0	-0.33	-0.33	0	0.333	-0.33	0	0	0.333
E	-0.33	-0.33	0.333	0	0.333	-0.33	-0.33	0	0.333	0.333	0	0.333	0	0.333	0
F	0.333	0.333	0.333	0.333	0	0.333	-0.33	0.333	0	-0.33	0.333	0	0.333	0	0

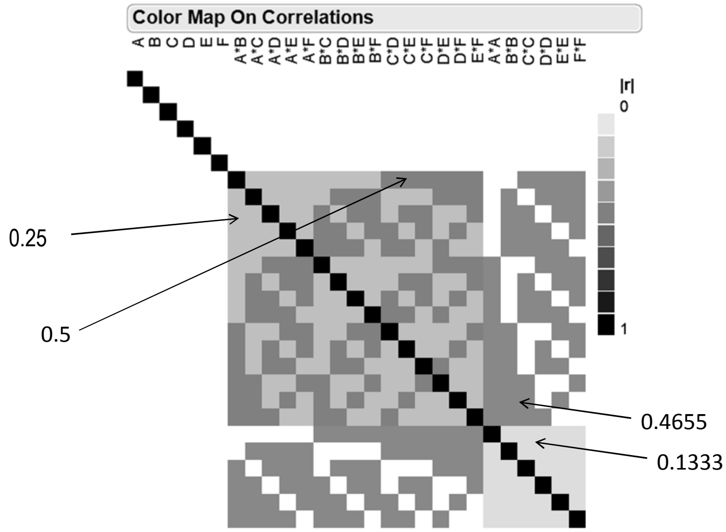
For our design there is no aliasing between main effects and two-factor interactions.

Alias Matrix

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Column Correlations

Color Map On Correlations

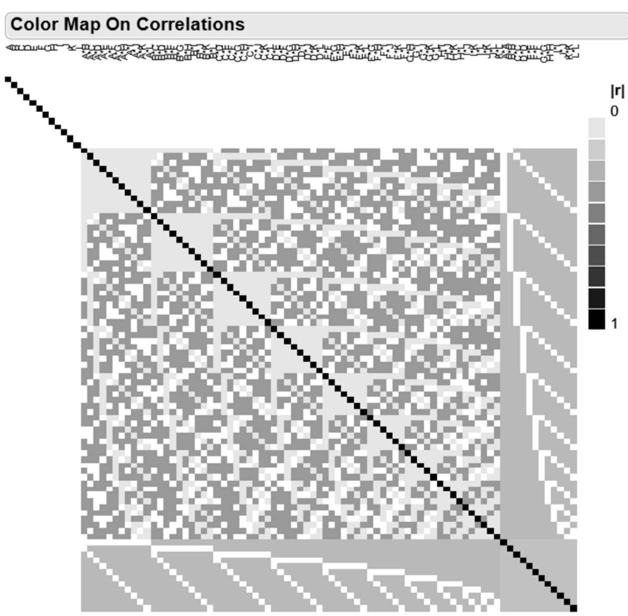


Example 2

12 factors – 25 runs.

A	B	C	D	E	F	G	H	I	J	K	L
0	1	1	1	1	1	1	1	1	1	1	1
0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
1	0	-1	1	-1	-1	-1	1	1	1	-1	1
-1	0	1	-1	1	1	1	-1	-1	-1	1	-1
1	1	0	-1	1	-1	-1	-1	1	1	1	-1
-1	-1	0	1	-1	1	1	1	-1	-1	-1	1
1	-1	1	0	-1	1	-1	-1	-1	1	1	1
-1	1	-1	0	1	-1	1	1	1	-1	-1	-1
1	1	-1	1	0	-1	1	-1	-1	-1	1	1
-1	-1	1	-1	0	1	-1	1	1	1	-1	-1
1	1	1	-1	1	0	-1	1	-1	-1	-1	1
-1	-1	-1	1	-1	0	1	-1	1	1	1	-1
1	1	1	1	-1	1	0	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	0	1	-1	1	1	1
1	-1	1	1	1	-1	1	0	-1	1	-1	-1
-1	1	-1	-1	-1	1	-1	0	1	-1	1	1
1	-1	-1	1	1	1	-1	1	0	-1	1	-1
-1	1	1	-1	-1	-1	1	-1	0	1	-1	1
1	-1	-1	-1	1	1	1	-1	1	0	-1	1
-1	1	1	1	-1	-1	-1	1	-1	0	1	-1
1	1	-1	-1	-1	1	1	1	-1	1	0	-1
-1	-1	1	1	1	-1	-1	-1	1	-1	0	1
1	-1	1	-1	-1	-1	1	1	1	-1	1	0
-1	1	-1	1	1	1	-1	-1	-1	1	-1	0
0	0	0	0	0	0	0	0	0	0	0	0

Column Correlations



Note the zero correlations among main effects

Example from JQT paper

TABLE 2. Three-Level Definitive Screening Design for Six Factors with a Simulated Response Vector

Run (i)	$x_{i,1}$	$x_{i,2}$	$x_{i,3}$	$x_{i,4}$	$x_{i,5}$	$x_{i,6}$	y_i
1	0	1	−1	−1	−1	−1	21.04
2	0	−1	1	1	1	1	10.48
3	1	0	−1	1	1	−1	17.89
4	−1	0	1	−1	−1	1	10.07
5	−1	−1	0	1	−1	−1	7.74
6	1	1	0	−1	1	1	21.01
7	−1	1	1	0	1	−1	16.53
8	1	−1	−1	0	−1	1	20.38
9	1	−1	1	−1	0	−1	8.62
10	−1	1	−1	1	0	1	7.80
11	1	1	1	1	−1	0	23.56
12	−1	−1	−1	−1	1	0	15.24
13	0	0	0	0	0	0	19.91

Where did the data come from?

We simulated it using the model below…

$$y_i = 20 + 4x_{i,1} + 3x_{i,2} - 2x_{i,3} - x_{i,4} + 5x_{i,2}x_{i,3} + 6x_{i,1}^2 + \varepsilon_i,$$

Stepwise Results

Current Estimates

Lock	Entered	Parameter	Estimate	nDF	SS	"F Ratio"	"Prob>F"
		Intercept	20.5782979	1	0	0.000	1
		X1	3.408	2	197.8935	177.357	4.6e-6
		X2	2.748	2	241.8807	216.780	2.54e-6
		X3	-1.309	2	183.5004	164.458	5.75e-6
		X4	-0.851	1	7.24201	12.981	0.01133
		X5	0	1	0.26896	0.437	0.53788
		X6	0	1	0.43264	0.742	0.42834
		X1*X2	0	1	2.082135	8.228	0.03506
		X1*X3	0	1	0.058259	0.089	0.77798
		X1*X4	0	1	0.1764	0.278	0.62047
		X1*X5	0	2	0.744728	0.572	0.60454
		X1*X6	0	2	0.515108	0.364	0.71591
		X2*X3	5.15244681	1	166.3656	298.202	2.42e-6
		X2*X4	0	1	0.580345	1.049	0.35278
		X2*X5	0	2	0.636038	0.469	0.65608
		X2*X6	0	2	0.493121	0.346	0.72707
		X3*X4	0	1	1.181568	2.728	0.15953
		X3*X5	0	2	0.616614	0.452	0.66552
		X3*X6	0	2	1.876522	2.552	0.19308
		X4*X5	0	2	0.420923	0.288	0.76432
		X4*X6	0	2	0.908408	0.745	0.53089
		X5*X6	0	3	1.905	1.321	0.4123
		X1*X1	-6.7247872	1	81.74884	146.531	0.00002
		X2*X2	0	1	0.076481	0.117	0.74631
		X3*X3	0	1	0.062247	0.095	0.77064
		X4*X4	0	1	2.165907	9.166	0.02916
		X5*X5	0	2	1.793049	2.307	0.21561
		X6*X6	0	2	0.442558	0.305	0.75306

49

Recapitulation – Definitive Screening Design

1. Orthogonal main effects plans.
2. Two-factor interactions are uncorrelated with main effects.
3. Quadratic effects are uncorrelated with main effects.
4. All quadratic effects are estimable.
5. The number of runs is only one more than twice the number of factors.
6. For six factors or more, the designs can estimate all possible full quadratic models involving three or fewer factors