# An Introduction to Optimality Criteria and Some Results on Optimal Block Design 

Ashish Das<br>Theoretical Statistics and Mathematics Unit, Indian Statistical Institute New Delhi, India

## 1. An Introduction to Optimality Criteria

## Linear Model Set-up (fixed-effects linear model):

Observation-vector $Y_{n \times 1}$ follows a standard linear model.

$$
E(Y)=X \theta, \quad \operatorname{cov}(Y)=\sigma^{2} I_{n}
$$

where, $X_{n \times t}$ is the design matrix, $\theta_{t \times 1}$ the unknown parameter, and $\sigma^{2}$ the constant error variance.

Normal equations for obtaining the best linear unbiased estimator (BLUE) of $\theta$ is

$$
X^{\prime} X \theta=X^{\prime} Y
$$

Here, $X^{\prime} X$ is called the "information matrix" of $\theta$ and $\operatorname{var}(\hat{\theta})=\sigma^{2}\left(X^{\prime} X\right)^{-1}$ (provided $\operatorname{Rank}(X)=t)$. Let $l^{\prime} \theta$ be an estimable linear parametric function. Then $\operatorname{var}\left(l^{\prime} \hat{\theta}\right)=$ $\sigma^{2} l^{\prime}\left(X^{\prime} X\right)^{-} l$. We choose a design $d$, with design matrix $X_{d}$, whose information matrix $X_{d}^{\prime} X_{d}$ is "large" (equivalently, $\left(X_{d}^{\prime} X_{d}\right)^{-}$is "small") in some sense.

Now, suppose we are interested in a component $\theta_{1}$ of $\theta$. We write

$$
\theta=\binom{\theta_{1}}{\theta_{2}}
$$

and accordingly partition $X$ as $X=\left(\begin{array}{ll}X_{1} & X_{2}\end{array}\right)$ so that the model can be written as

$$
E(Y)=X_{1} \theta_{1}+X_{2} \theta_{2}, \operatorname{cov}(Y)=\sigma^{2} I_{n}
$$

The non-negative definite information matrix of $\theta_{1}$ is

$$
I\left(\theta_{1}\right)=X_{1}^{\prime} X_{1}-X_{1}^{\prime} X_{2}\left(X_{2}^{\prime} X_{2}\right)^{-} X_{2}^{\prime} X_{1} .
$$

Example of specific design models:
I. One-way Model (Completely Randomized Design)

$$
\begin{aligned}
\theta_{1} & \equiv \tau_{v \times 1} \\
\theta_{2} & \equiv \mu_{1 \times 1}
\end{aligned}
$$

Given a design $d$

$$
\begin{aligned}
X_{1 d} & \equiv\left(\left(x_{1 i j}\right)\right)_{n \times v} \text { observation versus treatment matrix } \\
x_{1 i j} & = \begin{cases}1 & \text { if } i \text { th observation arise out of application of } j \text { th treatment } \\
0 & \text { otherwise }\end{cases} \\
X_{2 d} & =1_{n}, \\
I_{d}\left(\theta_{1}\right) & =R_{d}-n^{-1} r_{d} r_{d}^{\prime}\left(=C_{d}\right), \\
R_{d} & =\operatorname{diag}\left(r_{d 1}, r_{d 2}, \ldots, r_{d v}\right), \\
r_{d} & =\left(\begin{array}{c}
r_{d 1} \\
r_{d 2} \\
\vdots \\
r_{d v}
\end{array}\right), \\
r_{d i} & =\text { Number of times treatment } i \text { is replicated in } d .
\end{aligned}
$$

II. Two-way Model (Block Design)

A block design is an arrangement of $v$ treatments in $b$ blocks each of size $k_{d 1}, k_{d 2}, \ldots, k_{d b}$ respectively. The replication of treatment $i$ is $r_{d i}, i=1, \ldots, v$.

$$
\begin{aligned}
\theta_{1} & \equiv \tau_{v \times 1} \\
\theta_{2} & \equiv\binom{\mu}{\beta}_{(b+1) \times 1}
\end{aligned}
$$

Given a design $d$

$$
\begin{aligned}
X_{1 d} & \equiv\left(\left(x_{1 i j}\right)\right)_{n \times v} \text { observation versus treatment matrix, } \\
X_{2 d} & \equiv\left(1_{n} X_{\beta d}\right)_{n \times(b+1)}, \\
X_{\beta d} & \equiv\left(x_{\beta i j}\right) \text { observation versus block matrix, } \\
x_{\beta i j} & = \begin{cases}1 & \text { if ith observation is from jth block } \\
0 & \text { otherwise }\end{cases} \\
I_{d}\left(\theta_{1}\right) & =R_{d}-N_{d} K_{d}^{-1} N_{d}^{\prime}\left(=C_{d}\right),
\end{aligned}
$$

where $R_{d}=\operatorname{diag}\left(r_{d 1}, r_{d 2}, \ldots, r_{d v}\right), K_{d}=\operatorname{diag}\left(k_{d 1}, k_{d 2}, \ldots, k_{d b}\right), N_{d}=\left(\left(n_{d i j}\right)\right)_{v \times b}$ is the treatment-block incidence matrix and $n_{d i j}$ is the number of times treatment $i$ appears in block $j$.
III. Three-way Model (Row-Column Design)

A row-column design $d$ is an arrangement of $v$ treatments in a $k \times b$ array of $k$ rows and $b$ columns.

$$
\begin{aligned}
& \theta_{1} \equiv \tau_{v \times 1} \\
& \theta_{2} \equiv\left(\begin{array}{c}
\mu \\
r \\
c
\end{array}\right)_{(k+b+1) \times 1} \\
& I_{d}\left(\theta_{1}\right)=R_{d}-\frac{1}{k} N_{d} N_{d}^{\prime}-\frac{1}{b} M_{d} M_{d}^{\prime}+\frac{1}{b k} r_{d} r_{d}^{\prime}\left(=C_{d}\right)
\end{aligned}
$$

where

$$
R_{d}=\operatorname{diag}\left(r_{d 1}, \ldots, r_{d v}\right)
$$

$$
\begin{aligned}
& N_{d}=\left(\left(n_{d i j}\right)\right)_{v \times b} \text { treatment-column incidence } \\
& M_{d}=\left(\left(m_{d i j}\right)\right)_{v \times k} \text { treatment-row incidence }
\end{aligned}
$$

Here, $n_{d i j}\left(m_{d i j}\right)$ is the number of times treatment $i$ appears in column (row) $j$.

## Inference problem:

As we are interested in the treatment effects, the problem of inference may be specified as

$$
\Pi: \eta=L \tau
$$

where $L$ is a $p \times v$ matrix with $L 1=0$. Thus, $\eta$ contains $p$ treatment contrasts. (In fact, only treatment contrasts are estimable).

With reference to $\Pi$, we call a design $d$ as acceptable if all components of $\eta$ are estimable using $d$. Let $D_{\pi}$ be the class of all acceptable designs with reference to the problem $\pi$. Problem $\pi$ is referred to as
(1) non-singularly estimable iff $\operatorname{Rank}(L)=p$
(2) non-singularly estimable full rank iff $\operatorname{Rank}(L)=p=v-1$.

For a full rank problem $\pi, D_{\pi}$ consists only of such designs $\{d\}$ for which $\operatorname{Rank}\left(C_{d}\right)=$ $v-1$. Such designs are called connected designs.

In all the three types of designs considered, $C_{d} 1=0$ and $l^{\prime} \tau$ is estimable iff $l$ belongs to the column-space of $C$. Thus, $l^{\prime} \tau$ need to be a treatment contrast in order to be estimable. Rank $\left(C_{d}\right)=v-1$ iff all treatment contrasts are estimable and in that case, the underlying design is said to be connected.

## Optimality criterion to select good designs:

Suppose $\hat{\eta}_{d}$ is the BLUE of $\eta$ using a design $d$ with $\operatorname{var}\left(\hat{\eta}_{d}\right)=V_{d}$. It is reasonable to define an optimality criterion as a meaningful function of $V_{d}$.
$\frac{A \text {-optimality }}{\text { A design } d^{*}} \in D$ is said to be $A$-optimal in $D$ iff

$$
\operatorname{tr}\left(V_{d^{*}}\right) \leq \operatorname{tr}\left(V_{d}\right) \text { for any other design } d \in D
$$

The trace of $V_{d}$ is minimized in the $A$-optimality criterion, which implies the minimization of the average variance of the BLUE of the components of $\eta$.
$D$-optimality
A design $d^{*} \in D$ is said to be $D$-optimal in $D$ iff

$$
\operatorname{det}\left(V_{d^{*}}\right) \leq \operatorname{det}\left(V_{d}\right) \text { for any design } d \in D
$$

The $D$-optimality criterion has the following statistical significance.
Let the observation vector $Y$ follow a multivariate normal distribution. Then, $\hat{\eta}_{d}$ also follows a multivariate normal distribution. with mean $\eta$ and dispersion matrix $V_{d}$.

A $(1-\alpha) \%$ joint confidence region for $\eta$ is the ellipsoid

$$
\begin{equation*}
\left(\eta-\hat{\eta}_{d}\right)^{\prime} V_{d}^{-1}\left(\eta-\hat{\eta}_{d}\right) \leq \sigma^{2} \chi_{\alpha(v-1)}^{2} \tag{1.1}
\end{equation*}
$$

where $\sigma^{2}=$ per observation variance (known), $\chi_{\alpha(v-1)}^{2}=(1-\alpha)$ percentile of a central $\chi^{2}$ with $(v-1)$ d.f.
or is the ellipsoid

$$
\begin{equation*}
\left(n-\hat{\eta}_{d}\right)^{\prime} V_{d}^{-1}\left(n-\hat{\eta}_{d}\right) \leq v s^{2} F_{\alpha}\left(v-1, \eta_{e}\right) \tag{1.2}
\end{equation*}
$$

where $s^{2}=$ unbiased estimator of $\sigma^{2}$ (unknown), $F_{\alpha}\left(v-1, n_{e}\right)=(1-\alpha)$ percentile of $F$ with $(v-1)$ and $n_{e}$ d.f.. Here, $n_{e}$ is the error degrees of freedom.

The volume of (1.1) (expected volume in (1.2)) is proportional to the square root of $\operatorname{det}\left(V_{d}\right)$.

Thus, the $D$-optimality criterion chooses that design as the "best" for which the volume (expected volume) of the joint confidence ellipsoid is least.

## E-optimality

$\overline{\text { A design } d^{*}} \in D$ is said to be $E$-optimal in $D$ iff for all normalized treatment contrasts $l^{\prime} \tau$ with BLUE $l^{\prime} \hat{\tau}$,

$$
\max _{l: l^{\prime} l=1}\left(\operatorname{var}_{d^{*}}\left(l^{\prime} \hat{\tau}\right)\right) \leq \max _{l: l^{\prime} l=1}\left(\operatorname{var}_{d}\left(l^{\prime} \hat{\tau}\right)\right)
$$

for any other design $d \in D$.
Let $0=\mu_{d 0}<\mu_{d 1} \leq \cdots \leq \mu_{d, v-1}$ be the eigenvalues of $C_{d}$. Then

$$
\operatorname{var}_{d}\left(l^{\prime} \hat{\tau}\right)=\sigma^{2} l^{\prime} C_{d}^{+} l
$$

where $C_{d}^{+}$is the Moore-Penrose inverse of $C_{d}$. Also,

$$
\mu_{d, v-1}^{-1} \leq \frac{l^{\prime} C_{d}^{+} l}{l^{\prime} l} \leq \mu_{d 1}^{-1}
$$

Thus, if $l^{\prime} l=1$, we have

$$
\left.\sigma^{-2} \max _{\operatorname{var}_{d}\left(l^{\prime} \hat{\tau}\right)}\right)=\mu_{d 1}^{-1}
$$

Hence, we can say, a design $d^{*} \in D$ is $E$-optimal in $D$ iff $\mu_{d^{*} 1} \geq \mu_{d 1}$ where $d$ is any other competing design in $D$ (i.e., $d^{*}$ minimizes the maximum variance of all normalized treatment contrasts over $d$ ).
$M V$-optimality
A design $d^{*} \in D$ is said to be $M V$-optimal iff

$$
\max _{i \neq j} \operatorname{var}_{d^{*}}\left(\hat{\tau}_{i}-\hat{\tau}_{j}\right) \leq \max _{i \neq j} \operatorname{var}_{d}\left(\hat{\tau}_{i}-\hat{\tau}_{j}\right)
$$

where $d$ is any other competing design in $D$.
Here our interest is only on elementary treatment contrasts and accordingly the $M V$ optimality criteria is based on only such specific contrasts.

Optimality Criterion as a Functional
$\overline{\text { Let } P \tau \text { represent a complete set of orthonormal treatment contrasts with BLUE } P \hat{\tau} \text {. Thus, }}$ $P$ is of order $(v-1) \times v$ and $\operatorname{Rank}(P)=v-1$ with $P 1=0, P P^{\prime}=I_{v-1}$. Also,

$$
\sigma^{-2} \operatorname{var}_{d}(P \hat{\tau})=P C_{d}^{+} P^{\prime}=\left(P C_{d} P^{\prime}\right)^{-1}
$$

Consider

$$
A=\left[\begin{array}{c}
v^{-1 / 2} \\
P
\end{array}\right] \text { with } A A^{\prime}=A^{\prime} A=I_{v}
$$

Then, since $C_{d}^{+} 1=0$

$$
A C_{d}^{+} A^{\prime}=\left(\begin{array}{cc}
0 & 0^{\prime} \\
0 & P C_{d}^{+} P^{\prime}
\end{array}\right)
$$

and $\operatorname{det}\left(C_{d}^{+}-\lambda I_{v}\right)=\operatorname{det}\left(A C_{d}^{+} A^{\prime}-\lambda I_{v}\right)=-\lambda \operatorname{det}\left(P C_{d}^{+} P^{\prime}-\lambda I_{v-1}\right)$.
Thus, the non-zero eigenvalues of $C_{d}^{+}$and the eigenvalues of $P C_{d}^{+} P^{\prime}$ are the same.
It follows that instead of minimizing a function of the eigenvalues of $P C_{d}^{+} P^{\prime}$ to arrive at an optimal design, one may as well minimize the same function of the non-zero eigenvalues of $C_{d}^{+}$.

One may thus think of an optimality criterion as a function on the set of n.n.d. symmetric matrices of order $v$ with zero row sums.

Let, $B_{v, 0} \rightarrow$ set of all n.n.d. symmetric matrices of order $v$ with zero row sums.
An optimality criterion $\phi$ is a function $\phi: B_{v, 0} \rightarrow(-\infty, \infty]$.
A design $d$ is $\phi$-optimal if it minimizes $\phi\left(C_{d}\right)$. Note that $C_{d} \in B_{v, 0}$. Thus we have,

$$
\begin{aligned}
A-\text { optimality : } \phi_{A}\left(C_{d}\right) & =\sum_{i=1}^{v-1} \mu_{d i}^{-1} \\
D-\text { optimality : } \phi_{D}\left(C_{d}\right) & =\prod_{i=1}^{v-1} \mu_{d i}^{-1} \equiv-\sum_{i=1}^{v-1} \log \left(\mu_{d i}\right), \\
E-\text { optimality }: \phi_{E}\left(C_{d}\right) & =\max _{i} \mu_{d i}^{-1}=\mu_{d 1}^{-1}
\end{aligned}
$$

$\phi_{p^{-}}$optimality
Kiefer (1974, Ann. Statist.)
Let, $\phi_{p}\left(C_{d}\right)=\left[(v-1)^{-1} \sum_{i=1}^{v-1} \mu_{d i}^{-p}\right]^{1 / p}, \quad 0<p<\infty$. A design $d^{*}$ is $\phi_{p}$-optimal if for $d^{*}, \phi_{p}\left(C_{d}\right)$ is minimum over $D$ for all $p$.

The $\phi_{p}$ family of optimality criteria has $A-, D-, E-$ criterion as particular cases.

$$
\begin{aligned}
\phi_{p=1}\left(C_{d}\right) & =\phi_{A}\left(C_{d}\right), \\
\phi_{p \rightarrow 0}\left(C_{d}\right) & =\phi_{D}\left(C_{d}\right), \\
\phi_{p \rightarrow \infty}\left(C_{d}\right) & =\phi_{E}\left(C_{d}\right) .
\end{aligned}
$$

Universal optimality
Kiefer (1975, Survey Design, J. N. Srivasava ed., North Holland)
(A strong family of optimality criteria which includes $A$-, $D$ - and $E$-criteria as perticular cases.)
$d^{*} \in D$ is universally optimal over $D$ if $d^{*}$ minimizes $\phi\left(C_{d}\right), d \in D$ for any $\phi: B_{v, 0} \rightarrow$ $(-\infty, \infty]$ satisfying

1. $\phi$ is matrix convex, i.e., $\phi\left\{a C_{1}+(1-a) C_{2}\right\} \leq a \phi\left(C_{1}\right)+(1-a) \phi\left(C_{2}\right)$ for $C_{i} \in$ $B_{v, 0}(i=1,2)$ and $0 \leq a \leq 1$,
2. $\phi(b C)$ is nonincreasing in the scalar $b \geq 0$ for each $C \in B_{v, 0}$,
3. $\phi$ is invariant under each simultaneous permutation of rows and columns of $C$ in $B_{v, 0}$.

If $d^{*}$ is universally optimal in $D$ then $\operatorname{tr}\left(C_{d^{*}}\right) \geq \operatorname{tr}\left(C_{d}\right)$ for any other $d \in D$ i.e., maximization of $\operatorname{tr}\left(C_{d}\right)$ is a necessary condition for universal optimality.
$S$ - and $M S$-optimality
Shah (1960, Ann. Statist.); Eccleston and Hedayat (1974, Ann. Statist.)
A design $d^{*} \in D$ is $S$-optimal if $d^{*}$ minimizes $\operatorname{tr}\left(C_{d}^{2}\right)=\sum_{i=1}^{v-1} \mu_{d i}^{2}$ for all $d \in D$.
Let $\operatorname{tr}\left(C_{d}\right)=A$, a constant, for all $d \in D$. Then a balanced design has all its eigenvalues equal to $A /(v-1)$. It is possible that such a balanced design do not exist. In that situation an $S$-optimal design would be one which is "closet" to a balanced design. For this the Euclidian distance between $\left(\mu_{d 1}, \mu_{d 2}, \ldots, \mu_{d, v-1}\right)$ and $(A /(v-1), \ldots, A /(v-1))$ is minimized and an $S$-optimal design $d^{*}$ is obtained. Here,

$$
\text { Distance }=\left\{\sum_{i=1}^{v-1} \mu_{d i}^{2}-A^{2} /(v-1)\right\}^{1 / 2}
$$

A design $d^{*} \in D$ is said to be $M S$-optimal if

$$
\max _{d \in D} \operatorname{tr}\left(C_{d}\right)=\operatorname{tr}\left(C_{d^{*}}\right)
$$

and

$$
\min _{d \in D^{\prime}} \operatorname{tr}\left(C_{d}^{2}\right)=\operatorname{tr}\left(C_{d^{*}}^{2}\right)
$$

where $D^{\prime}$ is the sub-class of all designs $d \in D$ for which $\operatorname{tr}\left(C_{d}\right)$ is maximum.

$$
\text { Distance }=\left\{\sum \mu_{d i}^{2}-\left(\sum \mu_{d i}\right)^{2} /(v-1)\right\}^{1 / 2}
$$

Generalized Optimality Criteria
Cheng (1978, Ann. Statist.)
Consider a class of optimality functionals $\psi_{f}\left(C_{d}\right)=\sum_{i=1}^{v-1} f\left(\mu_{d i}\right)$ where $f$ is defined over $\left(0, \max _{d \in D} \operatorname{tr}\left(C_{d}\right)\right)$ and satisfy

1. $f$ is continuously differentiable on $\left(0, \max \operatorname{tr}\left(C_{d}\right)\right)$
2. $f^{\prime}<0, f^{\prime \prime}>0, f^{\prime \prime \prime}<0$ on $\left(0, \max \operatorname{tr}\left(C_{d}\right)\right)$
3. $f$ is continuous at 0 or $f(0)=\lim _{\mu_{d} \rightarrow 0+} f\left(\mu_{d}\right)=+\infty$

A design $d^{*} \in D$ is $\psi_{f}$-optimal if $d^{*}$ minimizes $\psi_{f}\left(C_{d}\right)$ for any $d \in D$.
Special Cases:

| $A$ - criteria | $f(x)=\frac{1}{x}$ |
| :--- | :--- |
| $D$ - criteria | $f(x)=-\log x$ |
| $S$ - criteria/MS- criteria | $f(x)=x^{2}$ |
| $\phi_{p^{-}}$criteria | $f(x)=x^{-p}$ |

## Distance Criterion

Sinha (1970, Calcutta Bull.)
Specific optimality criterion based on concept of Distance.
Choose a design $d^{*} \in D$ such that

$$
\operatorname{Pr}\left[\left|\hat{\eta}_{d^{*}}-\eta\right|<\epsilon\right] \geq \operatorname{Pr}\left[\left|\hat{\eta}_{d}-\eta\right|<\epsilon\right]
$$

for all $\epsilon>0$ and $d$ is any other design in $D$.

## 2. Optimal Block Designs : Some Results

## Kiefer (1975)

Suppose $d^{*} \in D$ and $C_{d^{*}}$ satisfies
(a) $C_{d^{*}}$ is completely symmetric, i.e., $C_{d^{*}}=\alpha I_{v}+\beta J_{v}$ ( $d^{*}$ is variance balanced)
(b) $\operatorname{tr}\left(C_{d^{*}}\right)=\max _{d \in D} \operatorname{tr}\left(C_{d}\right)$
then $d^{*}$ is universally optimal in $D$.
Universal optimality of CRD $d \in D(v, n)$ with $n / v=r$, an integer

$$
\begin{aligned}
\operatorname{tr}\left(C_{d}\right) & =\operatorname{tr}\left(R_{d}-n^{-1} r_{d} r_{d}^{\prime}\right) \\
& =n-n^{-1} \sum_{i=1}^{v-1} r_{d i}^{2}
\end{aligned}
$$

Here minimization of $\sum r_{d i}^{2}$ subject to $\sum r_{d i}=n$ is attained when $r_{d i}=n / v=r$.
So, a design $d^{*}$ with $r_{d^{*} i}=r$ for all $i$ is universally optimal in $D(v, n)$.
Universal optimality of Block Design in $D(v, b, k)$
Let $d$ be a block design with $v \geq 3$ treatments and $b$ blocks, each of size $k \geq 2$. Then $d$ is called a balanced block design (BBD) if
(i) $\sum_{j=1}^{b} n_{d i j} n_{d m j}=\lambda$, for $i \neq m$, (i.e., $d$ is variance balanced)
(ii) $\left|n_{d i j}-k / v\right|<1$, for all $i, j$.

Condition (ii) $\Rightarrow n_{d i j}=[k / v]$ or $[k / v]+1$ and (i) $\Rightarrow$ variance balance. Also, conditions (i) and (ii) $\Rightarrow \mathrm{BBD}$ is an equireplicate design. If $k<v$, then $n_{d i j}=0$ or 1 and a BBD reduces to a balanced incomplete block design (BIBD).

A BBD $d^{*} \in D(v, b, k)$ is universally optimal in $D(v, b, k)$. We show that $d^{*}$ maximizes $\operatorname{tr}\left(C_{d}\right)$.

$$
\begin{aligned}
\operatorname{tr}\left(C_{d}\right) & =\operatorname{tr}\left(R_{d}-k^{-1} N_{d} N_{d}^{\prime}\right) \\
& =b k-k^{-1} \sum_{i} \sum_{j} n_{d i j}^{2} .
\end{aligned}
$$

Minimization of $\sum \sum n_{d i j}^{2}$ subject to $\sum \sum n_{d i j}=b k$ is attained when $n_{d i j}$ 's are as nearly equal as possible, i.e., when $n_{d i j}=[k / v]$ of $[k / v]+1$.

For a design $d$, it can be shown that the sum of the variances of the estimates of all elementary treatment contrasts is proportional to the sum of the reciprocals of the nonzero eigenvalues of $C_{d}$. Thus, a design which is $A$-optimal for inferring on a full set of orthonormalized treatment contrast is optimal for the estimation of the overall elementary treatment contrasts.

When the class of designs does not contain any completely symmetric $C$-matrix with maximum trace, Kifer's sufficient condition of universal optimality cannot be used.

## Yeh (1986, Biometrika)

A generalization of Kiefer's result on universal optimality.
Suppose a class $\mathcal{C}=\left\{C_{d}: d \in D\right\}$ of matrices in $B_{v, 0}$ contain a $C_{d^{*}}$ such that
i) for any $d \in D$ and $C_{d} \neq 0$, there exist scalars $a_{d i} \geq 0(i=1, \ldots, m)$ satisfying

$$
C_{d^{*}}=\sum_{i=1}^{m} a_{d i} P_{i} C_{d} P_{i}^{\prime}
$$

ii) $\operatorname{tr}\left(C_{d^{*}}\right)=\max _{d \in D} \operatorname{tr}\left(C_{d}\right)$
where $m=v$ ! and $P_{i}$ is the $i$ th permutation matrix. Then $d^{*}$ is universally optimal in $D$.
A class of universally optimal binary block designs, identified by Yeh (1988, JSPI) is given below.

Let $D_{1}$ be the class of binary designs i.e. for $d \in D, n_{d i j}=[k / v]$ or $[k / v]+1$. Consider $v \geq 3, b=v m+n \geq 2, k=v-1$ where $m, n$ are integers with $m \geq 0$ and $1 \leq n<v$. Suppose $d^{*}$ is a $\operatorname{BIBD}(v, \bar{b}=v m, r=(v-1) m, k=v-1, \lambda=(v-2) m)$ plus the last $n$ distinct binary blocks of size $v-1$. Then $d^{*}$ is universally optimal in $D_{1}(v, b, k)$.
Example: $v=5, m=1, n=2, b=7, k=4$. The following design is universally optimal in $D_{1}(5,7,4)$.

| 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 2 | 2 | 3 | 3 | 2 | 2 |
| 3 | 3 | 4 | 4 | 4 | 3 | 3 |
| 4 | 5 | 5 | 5 | 5 | 4 | 5 |
| $b_{1}$ | $b_{2}$ | $b_{3}$ | $b_{4}$ | $b_{5}$ | $b_{6}$ | $b_{7}$ |

Yeh's result can be extended for $k>v$.
Let $v \geq 3, b=v m+n \geq 2, k=v x \pm 1$ with $x \geq 0, m \geq 0$, and $1 \leq n<v$. Suppose $d^{*}$ is a $\operatorname{BBD}(v, \bar{b}=v m, k=v x \pm 1)$ plus the last $n$ distinct binary blocks of size $v x \pm 1$. Then $d^{*}$ is universally optimal over $D_{1}(v, b, k)$.

Example: $v=5, x=1, m=1, n=2, b=7, k=6$. The following design is universally optimal in $D_{1}(5,7,6)$.

| 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 1 | 2 | 3 | 4 | 5 | 1 | 2 |
| $b_{1}$ | $b_{2}$ | $b_{3}$ | $b_{4}$ | $b_{5}$ | $b_{6}$ | $b_{7}$ |

We note that if $d^{*}$ is a universally optimal design over $D_{1}(v, b, k)$ then if there exists a universally optimal design in $D(v, b, k)$ then $d^{*}$ is universally optimal over $D(v, b, k)$.

## Cheng (1978, Ann. Statist.)

A design $d$ is said to be Most Balanced group divisible design (MB-GDD) if
i) $n_{d i j}=[k / v]$ or $[k / v]+1$,
ii) $r_{d i}$ 's are all equal,
iii) the treatments can be divided into $m$ groups of $n$ each such that $\lambda_{i i^{\prime}}=\lambda_{1}$ if $i$ and $i^{\prime}$ are in the same group, and $\lambda_{i i^{\prime}}=\lambda_{2}$ otherwise. Here, $N N^{\prime}=\left(\left(\lambda_{i i^{\prime}}\right)\right), i, i^{\prime}=1, \ldots, v$.
iv) $\lambda_{2}=\lambda_{1} \pm 1$.

Cheng (1978) showed that MB-GDD with $m=2$ and $\lambda_{2}=\lambda_{1}+1$ is $\psi_{f-}$ optimal in $D(v, b, k)$.

When $m=2$, a MB-GDD has 2 distinct eigenvalues with multiplicities 1 and $v-2$.
Earlier Takeuchi (1961, Rep. Statist. Japan Engrs.) has shown that any GDD with $\lambda_{2}=\lambda_{1}+1$, is E-optimal in $D(v, b, k)$ and later Cheng (1980, JRSS "B") showed that MB-GDD with $n=2$ and $\lambda_{2}=\lambda_{1}-1$ is $E$-optimal in $D(v, b, k)$.

## Takeuchi (1961, Rep. Statist. Japan Engrs.)

Lemma 2.1: For $d \in D(v, b, k)$ let $F=k C_{d}+a I+b J$ where $a, b$ are integers chosen so that $a+b v>0$. Then, if $F$ is not strictly positive definite ( $p . d$. ), then

$$
\mu_{d 1} \leq-a / k
$$

Proof: Let $\lambda_{1}, \ldots, \lambda_{v}$ be the eigenvalues of $F$. Using the fact that $k C_{d}$ and $a I+b J$ commute,

$$
\begin{aligned}
\lambda_{1} & =k \mu_{d 0}+a+b v=a+b v>0 \quad \text { (given) } \\
\lambda_{2} & =k \mu_{d 1}+a \\
& \vdots \\
\lambda_{2} & =k \mu_{d 1}+a \\
\lambda_{v} & =k \mu_{d_{1}, v-1}+a
\end{aligned}
$$

where $0=\mu_{d 0}<\mu_{d 1} \leq \mu_{d 2} \leq \cdots \leq \mu_{d_{1}, v-1}$ are the eigenvalues of $C_{d}$.
Now since $F$ is not strictly $p$.d., therefore $\min _{i} \lambda_{i} \leq 0$, i.e., $k \mu_{d 1}+a \leq 0$ or, $\mu_{d 1} \leq-a / k$.
Theorem 2.1: For $d \in D(v, b, k)$,

$$
\mu_{d 1} \leq \frac{1}{k}\{r(k-1)+m\}
$$

where $r=b k / v$ is an integer and $m=\left[\frac{r(k-1)}{(v-1)}\right]$. Here, $[x]$ represents the largest integer contained in $x$.
Proof: Put $F=k C_{d}-\{r(k-1)+m\} I+(m+1) J$ and let $\lambda=r(k-1) /(v-1)$.
Since $-r(k-1)-m+v(m+1)=-\lambda(v-1)+(m+1)(v-1)+1=(v-1)(m+1-\lambda)+1>0$, in view of Lemma 2.1, it is sufficient to show that the minimum eigenvalue of $F$ is not positive, or equivalently, $F$ is not strictly $p . d$.

We now show that $F=\left(\left(f_{i j}\right)\right)$ is not strictly p.d. Note that all elements of $F$ are integers. Now

$$
f_{i j}=m+1-\sum_{u=1}^{b} n_{i u} n_{j u} \quad(i \neq j) .
$$

We shall show that $f_{i j} \geq 1$ for some $i \neq j$. If possible suppose $f_{i j}<1$ for all $i \neq j$. Since $f_{i j}$ is an integer, it follows that $f_{i j} \leq 0$ for all $i \neq j$.
$\Rightarrow m+1 \leq \sum_{u=1}^{b} n_{i u} n_{j u}$ for all $i \neq j$

$$
\begin{align*}
& \Rightarrow(m+1) v(v-1) \leq \sum_{i \neq j=1}^{v} \sum_{u=1}^{b} n_{i u} n_{j u}=\sum_{j=1}^{v} \sum_{u=1}^{b}\left(\sum_{\substack{i=1 \\
i \neq j}}^{v} n_{i u}\right) n_{j u}=\sum_{j=1}^{v} \sum_{u=1}^{b}\left(k-n_{j u}\right) n_{j u} \\
& \qquad=k \sum_{j=1}^{v} \sum_{u=1}^{b} n_{j u}-\sum_{j=1}^{v} \sum_{u=1}^{b} n_{j u}^{2} \leq(k-1) \sum_{j=1}^{v} \sum_{u=1}^{b} n_{j u}=(k-1) k b=(k-1) v r . \\
& \Rightarrow(m+1)(v-1) \leq r(k-1) \\
& \Rightarrow m+1 \leq r(k-1) /(v-1) \\
& \Rightarrow\left[\frac{r(k-1)}{(v-1)}\right]+1 \leq \frac{r(k-1)}{(v-1)} \text { which is impossible. } \\
& \text { Therefore, } \\
& \qquad f_{i j} \geq 1 \quad \text { for some } \quad i \neq j \tag{2.1}
\end{align*}
$$

Also,

$$
\begin{aligned}
f_{i i} & =k\left(r_{i}-\frac{1}{k} \sum_{u=1}^{b} n_{i u}^{2}\right)-\{r(k-1)+m\}+(m+1) \\
& =k r_{i}-\sum_{u=1}^{b} n_{i u}^{2}-r(k-1)+1 \text { for all } i
\end{aligned}
$$

We shall show that if $f_{i i}>0$ for all $i$ then $f_{i i}=1$ for all $i$.
Let $f_{i i}>0$ for all $i$. Since $f_{i i}$ is an integer for all $i$, we have $f_{i i} \geq 1$ for all $i$.
If possible let $f_{i i}>1$ for some $i$.
Then $1 \leq f_{i i}$ for all $i$ with strict inequality for some $i$.

$$
\begin{aligned}
\Rightarrow v<\sum_{i=1}^{v} f_{i i} & =k b k-\sum_{i=1}^{v} \sum_{u=1}^{b} n_{i u}^{2}-(k-1) r v+v \\
& \leq b k^{2}-\sum \sum n_{i u}-(k-1) r v+v \\
& =b k^{2}-b k-(k-1) r v+v \\
& =b k(k-1)-b k(k-1)+v \\
& =v
\end{aligned}
$$

i.e., $v<v$ which is impossible. Hence

$$
\begin{equation*}
f_{i i}=1 \text { for all } i \text { if } f_{i i}>0 \text { for all } i . \tag{2.2}
\end{equation*}
$$

Case 1. If $f_{i i} \leq 0$ for some $i$, then letting $x^{\prime}=(0,0, \ldots, 1,0, \ldots, 0)=e_{i}$, $\Rightarrow x^{\prime} F x=f_{i i} \leq 0$
$\Rightarrow F$ is not p.d.

Case 2. If $f_{i i}>0$ for all $i$, then from (2.1) and (2.2) $\exists i, j$ for which $f_{i j} \geq 1$ and $f_{i i}=f_{j j}=1$.

Let $x^{\prime}=(0,0,0, \ldots, 1,0, \ldots,-1,0, \ldots, 0)=e_{i j}$. Then,

$$
\begin{aligned}
x^{\prime} F x & =f_{i i}+f_{j j}-2 f_{i j} \quad\left(\text { since } F^{\prime}=F\right) \\
& =2\left(1-f_{i j}\right) \leq 0
\end{aligned}
$$

$\Rightarrow F$ is not p.d. Hence $\mu_{d 1} \leq \frac{1}{k}\{r(k-1)+m\}$.
The above theorem gives an improved bound of $\mu_{d 1}$ than $\mu_{d 1} \leq \frac{b(k-1)}{(v-1)}$ obtained earlier which lead to E-optimality of BIBD.

Theorem 2.2: A group divisible (GD) PBIB design $d^{*}$ with $\lambda_{2}=\lambda_{1}+1$ is $E$-optimal in $D(v, b, k)$.
Proof: Let $N_{d^{*}}$ be the incidence matrix of a $G D P B I B D$ with parameters $v=m n, b, r$, $k, \lambda_{1}, \lambda_{2}=\lambda_{1}+1$.
The eigenvalues of $N_{d^{*}} N_{d^{*}}^{\prime}$ are
$\theta_{0}=r k$ with multiplicity $\alpha_{0}=1$,
$\theta_{1}=r-\lambda_{1}$ with multiplicity $\alpha_{1}=m(n-1)$,
$\theta_{2}=r k-v \lambda_{2}$ with multiplicity $\alpha_{2}=m-1$.
Recall that,
$r=\lambda_{1}$ for a Singular GDD,
$r>\lambda_{1}, r k=v \lambda_{2}$ for a Semi-regular GDD,
$r>\lambda_{1}, r k>v \lambda_{2}$ for a Regular GDD.
Therefore

$$
\begin{aligned}
\mu_{d_{0}^{*}} & =r-\frac{r k}{k}=0 \\
\mu_{d_{i}^{*}} & =r-\frac{1}{k}\left(r-\lambda_{1}\right)=\frac{1}{k}\left\{r(k-1)+\lambda_{1}\right\}, \quad i=1, \ldots, m(n-1) \\
\mu_{d_{i}^{*}} & =r-\frac{1}{k}\left(r k-v \lambda_{2}\right)=\frac{v \lambda_{2}}{k}, \quad i=m(n-1)+1, \ldots, v-1
\end{aligned}
$$

Thus $\mu_{d^{*} 1}=\frac{1}{k}\left\{r(k-1)+\lambda_{1}\right\}=\frac{1}{k}\{r(k-1)+m\}$ (since $\frac{n_{1} \lambda_{1}+n_{2} \lambda_{2}}{v-1}=\frac{r(k-1)}{v-1}$ and $\lambda_{1} \leq \frac{r(k-1)}{v-1} \leq \lambda_{2}$ implies that for $\lambda_{2}=\lambda_{1}+1, v \lambda_{2}-\left(r(k-1)+\lambda_{1}\right)=(v-1) \lambda_{2}-r(k-1)+1>$ $0)$.

Hence $d^{*}$ is $E$-optimal in $D(v, b, k)$.
Optimality of Dual Designs in $D(v, b, k, r)$
Let $d \in D(v, b, k)$ with $b k / v=r$ be an equireplicate design with incidence matrix $N_{d}$, then $\bar{d}$ is said to be dual of $d$ if

$$
N_{\bar{d}}=N_{d}^{\prime}
$$

and then $\bar{d} \in D(\bar{v}=b, \bar{b}=v, \bar{k}=r)$ and is an equireplicate design with $\bar{r}=k$.
For $d \in D(v, b, k)$,

$$
C_{d}=r I-k^{-1} N_{d} N_{d}^{\prime},
$$

and for $\bar{d} \in D(\bar{v}, \bar{b}, \bar{k})$,

$$
\begin{aligned}
C_{\bar{d}} & =\bar{r} I-\bar{k}^{-1} N_{\bar{d}} N_{\bar{d}}^{\prime} \\
& =k I-r^{-1} N_{d}^{\prime} N_{d}
\end{aligned}
$$

Also, $\mu_{\bar{d} i}=k \mu_{d i} / r, 1 \leq i \leq v-1$, and if $v<b, \mu_{\bar{d} i}=k, v \leq i \leq b-1$.
Let $D(v, b, k, r)$ be the sub-class of designs in $D(v, b, k)$ with constant replication $r$. Thus, if $d$ is $\psi_{f}$-optimal in $D(v, b, k, r)$ then $\bar{d}$ is $\psi_{f}$-optimal in $D(\bar{v}=b, \bar{b}=v, \bar{k}=r, \bar{r}=$ $k)$. Hence duals of BBD's are $\psi_{f}$-optimal in the equireplicate class of designs.

Cheng (1980, JRSS 'B'; 1992, R.C.Bose Proc.) later established that duals of BBD's are $E$ - and $D$ - optimal in the unrestricted class $D(v, b, k)$.

Constantine (1981, Ann. Statist.) used "averaging technique" to show optimality of certain designs in situations when $b k / v$ is not an integer.

## Averaging technique

For any $C_{d}$, Let

$$
\begin{aligned}
\bar{C}_{d} & =(1 / n) \sum_{i=1}^{n} C_{d}^{\sigma_{i}} \\
& =(1 / n) \sum_{i=1}^{n} P_{i} C_{d} P_{i}^{\prime}
\end{aligned}
$$

where $\left\{\sigma_{i}\right\}$ denotes a collection of $n$ permutations on the symbols $1, \ldots, v$; and $P_{i}$ is the $v \times v$ (permutation) matrix representation of $\sigma_{i}$.

Then

$$
\sum_{i=1}^{v-1} f\left(\mu_{d i}\right) \geq \sum_{i=1}^{v-1} f\left(\bar{\mu}_{d i}\right) .
$$

If we are able to express the r.h.s. of the above inequality in terms of the parameters $v, b, k$, we would then get a lower bound for $\sum f\left(\mu_{d i}\right)$. We then try to identify a design $d^{*}$ which attains this bound. For example, for $d \in D(v, b, k)$

$$
\begin{gathered}
\sum_{i=1}^{v-1} \mu_{d i}^{-1} \geq \frac{(v-1)^{2}}{b(k-1)} \\
\mu_{d 1}^{-1} \geq \frac{(v-1)}{b(k-1)}
\end{gathered}
$$

The following result is due to Constantine (1981, Ann. Statist.).
Let $d^{*} \in D(v, b=\bar{b}+x, k)$ be a design obtained by adding $x$ disjoint blocks to a BIB $(v, \bar{b}, k)$ or a GDD with $\lambda_{2}=\lambda_{1}+1$. Then $d^{*}$ is $E$-optimal in $D(v, b=\bar{b}+x, k)$ provided $x<v / k$ (in case of BIBD) and provided $x<(v-m) / k$ (in case of GDD).

Similarly, $d^{*} \in D(v, b=\bar{b}-x, k)$ obtained by deleting $x$ disjoint blocks from a BIB $(v, \bar{b}, k)$ is $E$-optimal in $D(v, b, k)$ provided $v / k^{2} \leq x \leq v / k$.

Later, Sathe and Bapat (1985, Calcutta Bull.), showed that deleting any $x$ blocks, $x \leq\left(\frac{v-\sqrt{v}}{v-k}\right)$ from a BIB $(v, \bar{b}, k)$ yield an $E$-optimal design in $D(v, b=\bar{b}-x, k)$.

Jacroux (1983, Sankhya), obtained an upper bound for $\mu_{d 1}$ which resulted in indentifying new $E$-optimal designs.

Let $D(v, b, k)$ with $b k=v \bar{r}+s, 0 \leq s<v, \bar{r}(k-1)=(v-1) \bar{\lambda}+t, 0 \leq t<v-1$.
Then for $d \in D(v, b, k)$

$$
\mu_{d 1} \leq(\bar{r}(k-1)+\bar{\lambda}) / k
$$

provided $v \leq(v-s)(v-t)$.
Das (1993, Sankhya), gave the following result.
Let $d^{*} \in D(v=\bar{v}-p, b=\bar{b}+x, k)$ be a design obtained from $\operatorname{BIB}(\bar{v}, \bar{b}, r, k, \lambda)$ by adding $x$ arbitrary blocks and collapsing $p+1$ treatments into one. Then $d^{*}$ is $E$-optimal in $D(v, b, k)$ provided

$$
\begin{gathered}
v-p r-x k \geq 1, \\
v-p \lambda \geq 2 \\
v<(v-p r-x k)(v-p \lambda) .
\end{gathered}
$$

Example : Let $v=12, b=14, k=4, p=1, x=1$. The design obtained by adding 1 arbitrary block and collapsing 2 treatments into one in a BIB (13,13,4,4,1) we get an $E$-optimal in $D(12,14,4)$ as shown below.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 12 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 12 | 1 | 2 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 12 | 1 | 2 | 3 | 3 |
| 10 | 11 | 12 | 12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 4 |

Here, we have added the last block $\{1,2,3,4\}$ and collapsed treatment numbers 13 and 12 into one, and call it treatment number 12.

Similar results hold when starting from a GDD with $\lambda_{2}=\lambda_{1}+1$.
Minimally connected designs (MCD) belong to $D_{0}(v, b, k)$ with $b k=v+b-1$. Also, one more than MCD belong to $D_{1}(v, b, k)$ with $b k=v+b$.

Mandal, Shah, Sinha (1991, Calcutta Bull.), showed that for a design $d^{*} \in D_{0}(v, b, k)$ such that any one treatment appears in each of the binary blocks is $A$-optimal in $D_{0}(v, b, k)$. Such a design was also shown to be $D$ - and $E$-optimal by Bapat and Dey (1991, Prob. Letters).

Example: Let $v=19, b=6, k=4$. Then the following design is $A$-, $D$ - and $E$-optimal in $D(19,6,4)$.

| 1 | 1 | 1 | 1 | 1 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 5 | 8 | 11 | 14 | 17 |
| 3 | 6 | 9 | 12 | 15 | 18 |
| 4 | 7 | 10 | 13 | 16 | 19 |

Balasubramaniam and Dey (1996. JSPI) used graph-theoretic methods to establish $D$-optimality of certain types of designs in $D_{1}(v, b, k)$.
Example: Let $v=15, b=5, k=4$. Then the following design is $D$-optimal in $D(15,5,4)$.

| 1 | 4 | 7 | 10 | 13 |
| :---: | :---: | :---: | :---: | :---: |
| 2 | 5 | 8 | 11 | 14 |
| 3 | 6 | 9 | 12 | 15 |
| 4 | 7 | 10 | 13 | 1 |

For further results on optimal block designs with minimal and nearly minimal number of units, one may refer Dey, Shah and Das (1995, Statist. Sinica).

## 3. Optimal Block Designs for Combinatorially Aberrant Settings

What follows is based on a talk by Dr. John P. Morgan at the Dekalb IISA Conference2002.

Consider the usual experimental setting of $v$ treatments to be compared using $b$ blocks of experimental material such that there are $k$ experimental units in each block. The problem then is to find the best assignment of the treatments to the units. In other words, the problem is to find the optimal design $d \in D(v, b, k)$ where $D(v, b, k)$ is the class of all possible assignments of treatments into the experimental units.

We have already defined the optimality criteria in terms of functions of $C_{d}$, the Cmatrix. $C_{d}$ is determined by the replications $r_{d i}$ and the concurrences $\lambda_{d i i^{\prime}}$. Define the associated parameters for the setting $(v, b, k)$ as

- $r=\operatorname{int}\left(\frac{b k}{v}\right) \quad$ target replication
- $\lambda=\operatorname{int}\left(\frac{r(k-1)}{v-1}\right) \quad$ target concurrence

We consider the optimality criteria based on symmetric functions of the nonzero eigenvalues $\mu_{d 1} \leq \mu_{d 2} \leq \cdots \leq \mu_{d, v-1}$ of $C_{d}$. Recall that,

$$
\begin{aligned}
A-\text { criterion } & =\sum_{i} \mu_{d i}^{-1} \\
D-\text { criterion } & =-\sum_{i} \log \left(\mu_{d i}\right) \\
E-\text { criterion } & =\mu_{d 1}^{-1}
\end{aligned}
$$

All of these are included in the family of type I optimality criterion as defined by Cheng (1978).

One well known class of block designs is the BIBDs. A balanced incomplete block design $d \in D(v, b, k)$ is a block design satisfying the following conditions: (i) each $n_{d i j}=0$ or 1 , (ii) each $r_{d i}=r$, (iii) each $\lambda_{d i i^{\prime}}=\lambda$.

As seen earlier BIBDs are optimal with respect to every type I criterion (and many others as well - Kiefer, 1975). Also, BIBDs are binary and perfectly on target.

Example : $v=7, b=7, k=4$. Targets are $r=\operatorname{int}\left(\frac{28}{7}\right)=4$ and $\lambda=\operatorname{int}\left(\frac{12}{6}\right)=2$. Consider the BIBD $d$ as given below.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 2 | 3 | 4 | 5 | 6 | 7 | 1 |
| 3 | 4 | 5 | 6 | 7 | 1 | 2 |
| 5 | 6 | 7 | 1 | 2 | 3 | 4 |

Here, all $r_{d i}=4$ and all $\lambda_{d i i^{\prime}}=2$.
Example: $v=9, b=11, k=5$.

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 | 4 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 2 | 2 | 2 | 2 | 3 | 3 | 5 | 3 | 4 | 5 | 5 |
| 3 | 3 | 4 | 5 | 4 | 4 | 6 | 6 | 6 | 6 | 7 |
| 4 | 5 | 6 | 8 | 5 | 7 | 7 | 7 | 7 | 8 | 8 |
| 9 | 7 | 8 | 9 | 6 | 8 | 9 | 8 | 9 | 9 | 9 |

$\lambda_{\max }=4=\lambda_{12}=\lambda_{13}=\lambda_{14}=\lambda_{15}=\lambda_{59}$
$\lambda_{\text {min }}=2=\lambda_{25}=\lambda_{45}=\lambda_{49}$
$r_{1}=7, r_{2}=r_{3}=\cdots=r_{9}=6$
The target values are
$r=\operatorname{int}\left(\frac{b k}{v}\right)=\operatorname{int}\left(\frac{55}{6}\right)=6$,
$\lambda=\operatorname{int}\left(\frac{r(k-1)}{v-1}\right)=\operatorname{int}\left(\frac{24}{8}\right)=3$.
Note that:
BIBDs are the preferred designs whenever they can be found.

Necessary conditions for existence of a BIBD are (i) $v$ divides $b k$ and (ii) $v(v-1)$ divides $b k(k-1)$.
The class $D(v, b, k)$ is called a $B I B D$ setting if the necessary conditions are satisfied. BIBDs are the optimal designs in these settings (provided they exist).

Two questions that arise are:
I. If the divisibility conditions do not hold (non-BIBD settings), then what designs will be good?
II. If the divisibility conditions do hold but BIBDs do not exist, then what designs will be good?

Known theory points to getting the $r_{d i}$ and the $\lambda_{d i i^{\prime}}$ as close as possible to the targets $r$ and $\lambda$, i.e., get as close as possible to a BIBD. We motivate a possible general theory through two examples.
Definition 3.1: A nearly balanced incomplete block design $d$ in $D(v, b, k)$ with concurrence range $l$, or $\operatorname{NBBD}(l)$, is an incomplete block design satisfying the following conditions:
(i) each $n_{d i j}=0$ or 1 ,
(ii) each $r_{d i}=r$ or $r+1$
(iii) $\max _{i \neq i^{\prime}, j \neq j^{\prime}}\left|\lambda_{d i i^{\prime}}-\lambda_{d j j^{\prime}}\right|=l$, and
(iv) $d$ minimizes $\operatorname{tr}\left(C_{d}^{2}\right)$ over all designs satisfying (i)-(iii).

If $b k / v$ is an integer and $l=0$, this gives the BIBDs. If $b k / v$ is an integer and $l=1$, this is the regular graph design (RGD) class of John and Mitchell (1977) for which certain optimality properties are known. If $l=1$, this is the strongly regular graph design (SRGD) class of Jacroux (1985) for which again certain optimality properties are known. Note that $l=1$ is the best combinatorial approximation to a BIBD.

Now for given $v, b, k$, suppose $l \leq 1$ is not achievable. In such a situation we shall call it a combinatorially aberrant setting. The need to extend the "nearly balanced" notion arises in settings where neither BIBDs nor $\operatorname{NBBD}(1)$ s exist.

Example: $v=9, b=11, k=5$

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 | 4 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 2 | 2 | 2 | 2 | 3 | 3 | 5 | 3 | 4 | 5 | 5 |
| 3 | 3 | 4 | 5 | 4 | 4 | 6 | 6 | 6 | 6 | 7 |
| 4 | 5 | 6 | 8 | 5 | 7 | 7 | 7 | 7 | 8 | 8 |
| 9 | 7 | 8 | 9 | 6 | 8 | 9 | 8 | 9 | 9 | 9 |

$\lambda_{\max }=4=\lambda_{12}=\lambda_{13}=\lambda_{14}=\lambda_{15}=\lambda_{59}$
$\lambda_{\text {min }}=2=\lambda_{25}=\lambda_{45}=\lambda_{39}$
$r_{1}=7, r_{2}=r_{3}=\cdots=r_{9}=6$
This is a $\operatorname{NBBD}(2)$. It can be shown that $\lambda_{\max }-\lambda_{\min }=1$ cannot be achieved. For this design, three $\lambda_{\text {dii }}{ }^{\prime}$ fall short of the target $\lambda=3$.

To look for an optimal design in a combinatorially aberrant setting, let's guess that it will be binary. As earlier let $D(v, b, k)$ be the class of all possible designs, and $M(v, b, k)$ represent the class of binary designs.

Now measure the combinatorial asymmetry in a binary design by

$$
\delta_{d}=\sum \sum_{i<i^{\prime}} \max \left\{0, \lambda-\lambda_{d i i^{\prime}}\right\}=\text { the discrepancy of } d
$$

Also define,

$$
\delta=\min _{d \in M} \delta_{d}=\text { the minimum discrepancy } .
$$

Example: The $\operatorname{NBBD}(2)$ for $(v, b, k)=(9,11,5)$ has $\delta_{d}=3$.

From the definition of $\delta$ it follows that (i) $\delta>0 \Longleftrightarrow l \leq 1$ not achievable, (ii) Aberrant settings are exactly those for which $\delta>0$.
Lemma 3.1: In any aberrant setting, the $\operatorname{NBBD}(2) \bar{d}$ minimizes $\operatorname{tr}\left(C_{d}^{2}\right)$ over $M$ provided $\delta_{\bar{d}}=\delta$, i.e., provided it has minimum discrepancy over all binary designs.

- This provides an angle of attack on the optimality problem.
- A type-I criterion bound can be constructed based on the first two moments of the eigenvalues of $C_{d}$.

Let $\bar{d}$ be a minimum discrepancy $\operatorname{NBIBD}(2)$ as in Lemma 3.1. Let

$$
A=\operatorname{tr}\left(C_{d}\right) \text { and } B=\operatorname{tr}\left(C_{d}^{2}\right)+\frac{4}{k^{2}}
$$

Then for any design $d \in M$,

$$
\sum_{i=1}^{v-1} f\left(\mu_{\overline{d i}}\right)<f\left(\mu_{1}\right)+(v-3) f\left(\mu_{2}\right)+f\left(\mu_{3}\right)
$$

and for any nonbinary design,

$$
\sum_{i=1}^{v-1} f\left(\mu_{\bar{d} i}\right)<f\left(\mu_{1}^{*}\right)+(v-2) f\left(\mu_{4}\right)
$$

Here

$$
\begin{aligned}
\mu_{1} & =\text { an upper bound for } \mu_{d 1} \text { over } M \\
\mu_{1}^{*} & =\text { an upper bound for } \mu_{d 1} \text { over } D / M \\
P & =\left[\left(B-\mu_{1}^{2}\right)-\left(\left(A-\mu_{1}\right)^{2} /(v-2)\right)\right]^{1 / 2} \\
\mu_{2} & =\left[\left(A-\mu_{1}\right)-\sqrt{(v-2) /(v-3)} P\right] /(v-2) \\
\mu_{3} & =\left[\left(A-\mu_{1}\right)+\sqrt{(v-2)(v-3) P}\right] /(v-2) \\
\mu_{4} & =\left[A-(2 / k)-\mu_{1}^{*}\right] /(v-2)
\end{aligned}
$$

(For details one may refer Jacroux, 1985)

Theorem 3.1: Let $D(v, b, k)$ be a setting with $k \geq 3$, and let $\bar{d}$ be a $\operatorname{NBBD}(2)$ having minimum discrepancy $\delta_{\bar{d}}=\delta>0$. Then if $\mu_{1} \leq \mu_{2}$ and

$$
\begin{equation*}
\sum_{i=1}^{v-1} f\left(\mu_{\bar{d} i}\right)<f\left(\mu_{1}\right)+(v-3) f\left(\mu_{2}\right)+f\left(\mu_{3}\right) \tag{3.1}
\end{equation*}
$$

a $\psi_{f}$-optimal design in $M(v, b, k)$ must be an $\operatorname{NBBD}(2)$. If, moreover, $\mu_{1}^{*} \leq \mu_{4}$ and

$$
\begin{equation*}
\sum_{i=1}^{v-1} f\left(\mu_{\bar{d} i}\right)<f\left(\mu_{1}^{*}\right)+(v-2) f\left(\mu_{4}\right) \tag{3.2}
\end{equation*}
$$

then a $\psi_{f}$-optimal design in $D(v, b, k)$ must be an $\operatorname{NBBD}(2)$.

- Given any aberrant setting, one can apply this theorem after addressing the sometimes burdensome combinatorial problem of determining the exact value of $\delta$.
- Typically $\delta_{d}=\delta$ for some binary $d$ in the subclass with all $r_{d i} \geq r$, and moreover, $\delta_{d}>\delta$ for all $d$ not in that subclass.
- The diffculties arise in sorting through the possibilities for the "as equally replicated as possible" designs.

Example: This $\bar{d}$ for $(v, b, k)=(9,11,5)$ has $\delta_{d}=3$ :

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 3 | 4 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 2 | 2 | 2 | 2 | 3 | 3 | 5 | 3 | 4 | 5 | 5 |
| 3 | 3 | 4 | 5 | 4 | 4 | 6 | 6 | 6 | 6 | 7 |
| 4 | 5 | 6 | 8 | 5 | 7 | 7 | 7 | 7 | 8 | 8 |
| 9 | 7 | 8 | 9 | 6 | 8 | 9 | 8 | 9 | 9 | 9 |

$$
\begin{aligned}
& \lambda_{\max }=4=\lambda_{12}=\lambda_{13}=\lambda_{14}=\lambda_{15}=\lambda_{59} \\
& \lambda_{\min }=2=\lambda_{25}=\lambda_{45}=\lambda_{39} \\
& r_{1}=7, r_{2}=r_{3}=\cdots=r_{9}=6
\end{aligned}
$$

Calculations using $\bar{d}$ show that the inequalities (3.1) and (3.2) of the theorem hold for both the $A$ - and $D$ - criteria. Thus A- and D-optimal designs in $(9,11,5)$ must be NBBDs, provided that
(i) $\bar{d}$ itself is an $\operatorname{NBBD}(2)$, that is, that $\delta_{d}=3$ is the smallest achievable $\delta_{d}$ among binary designs with replicate range of 1 and concurrence range of 2 , and
(ii) $\delta_{\bar{d}}$ is in fact the minimum discrepancy $\delta$, the latter value being determined over all binary designs.
Now we have a combinatorial problem to solve. Enumeration proves that $\bar{d}$ is indeed an $\operatorname{NBBD}(2)$, and turns up one other, nonisomorphic $\operatorname{NBBD}(2) \tilde{d}$ :

\[

\]

- The design $\tilde{d}$ is inferior to $\bar{d}$ in terms of the A and D criteria.
- $\tilde{d}$ is superior to $\bar{d}$ in terms of the MV criterion (not eigenvalue based).
- A- and D-optimality of $\bar{d}$ is immediate if no design with larger concurrence range can have smaller discrepancy.

The rest of the points are as follows:

- The discrepancy value of 3 , though the smallest discrepancy for designs with concurrence range of 2 , is not the minimum discrepancy.
- This design $d^{*}$ with $\delta_{d^{*}}=2$ is the unique $\operatorname{NBBD}(3)$, and establishes that $\delta=2$ for $D(9,11,5)$ :

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 3 | 3 | 4 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 2 | 2 | 2 | 2 | 2 | 3 | 3 | 5 | 5 | 4 | 5 |
| 7 | 3 | 3 | 4 | 4 | 4 | 4 | 6 | 6 | 7 | 6 |
| 8 | 5 | 6 | 5 | 6 | 6 | 5 | 7 | 7 | 8 | 8 |
| 9 | 9 | 8 | 7 | 9 | 7 | 8 | 8 | 9 | 9 | 9 |

$\lambda_{\text {max }}=5=\lambda_{12}$
$\lambda_{\text {min }}=2=\lambda_{23}=\lambda_{24}$.

- Since $d^{*}$ is the unique design achieving the minimum discrepancy $\delta=\delta_{\bar{d}}-1$, it is the only design not ruled out by the Theorem.
- Calculation of the criteria values immediately proves that the $\operatorname{NBBD}(3) d^{*}$ is uniquely A- and D-optimal.
- All three designs have the same $\operatorname{tr}\left(C_{d}^{2}\right)$.

Let's return to the BIBD settings: we have $(v, b, k)$ satisfying the necessary BIBD conditions $v \mid b k$ and $v(v-1) \mid b k(k-1)$. This opens the possibility of a perfectly on-target design. But suppose no BIBD exists

- $D(v, b, k)$ is called an irregular $B I B D$ setting if the divisibility conditions are satisfied but no BIBD exists.
- What is the optimal design in an irregular BIBD setting?
- $(v, b, k)=(15,21,5)$ is a BIBD setting, but it is known that no BIBD exists.
- $(v, b, k)=(22,33,8)$ is a BIBD setting, but existence of a BIBD is not yet settled.

The main result applied to irregular BIBD settings says
Theorem 3.2: Let $D(v, b, k)$ be an irregular BIBD setting. Let $\bar{d} \in D$ be a $\operatorname{NBBD}(2)$ with $\delta_{\bar{d}}=\delta \leq 4$. If conditions (3.1) and (3.2) hold, then a $\psi_{f}$-optimal design must be a NBBD (2).

In an irregular BIBD setting, a $\operatorname{NBBD}(2)$ will have all $r_{d i}=r$ and all $\lambda_{d i i^{\prime}} \in\{\lambda-1, \lambda, \lambda+1\}$.
Corollary 3.1: Let $D(v, b, k)$ be an irregular BIBD setting in which $r \leq 41$. If there exists a binary, equireplicate design $\bar{d}$ with concurrence range 2 and $\delta_{\bar{d}} \leq 4$, then an $A$-optimal design must be a $\operatorname{NBBD}(2)$, and a $D$-optimal design must be a $\operatorname{NBBD}(2)$.

- The table of BIBDs in the CRC Handbook of Combinatorial Design shows 497 cases with $r \leq 41$ for which either a BIBD does not exist, or is not known.
- How do we find $\operatorname{NBBD}(2) \mathrm{s}$ with $\delta_{d} \leq 4$ in irregular BIBD settings?
- Little has been done on design construction for these settings.
- They find a design with $\delta_{d}=4$ for $D(22,33,8)$.
- BIBD existence is still open in this setting - should a BIBD not exist, then we have proven that an $\operatorname{NBBD}(2)$ is A - and D -optimal.

The smallest irregular BIBD setting is $(v, b, k)=(15,21,5)$, for which $r=7$ and $\lambda=2$. Can we find an optimal design?

After introducing a concept of U-BIBDs, the related combinatorial problem reduces to the determination of certain non-isomorphic U-BIBDs and to enumerate all possible completions of these, hoping to find $\delta \leq 4$. These considerations lead to the following design

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 5 | 5 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 10 | 2 | 2 | 3 | 4 | 5 | 6 | 3 | 4 | 5 | 7 | 8 | 4 | 5 | 6 | 7 | 5 | 6 | 8 | 6 | 7 |
| 12 | 3 | 6 | 7 | 8 | 9 | 7 | 9 | 6 | 8 | 9 | 10 | 9 | 6 | 13 | 8 | 7 | 11 | 9 | 9 | 11 |
| 13 | 4 | 10 | 11 | 11 | 14 | 8 | 11 | 7 | 13 | 12 | 11 | 10 | 8 | 14 | 10 | 10 | 12 | 12 | 10 | 12 |
| 14 | 5 | 12 | 13 | 15 | 15 | 9 | 13 | 14 | 15 | 15 | 14 | 14 | 12 | 15 | 15 | 13 | 15 | 13 | 11 | 14 |

- The enumeration shows only one other discrepancy 4 design, and none with smaller discrepancy.
- This design is $A$ - and $D$-optimal.
- Is the above design $E$-optimal?
- Theorem doesn't help.
- Could $\delta_{d}>4$ be $E$-superior?
- Enumeration of all designs with $\delta_{d}=5$ found all were $E$-inferior.
- Is proof possible without enumeration of all possible cases?

A bound on $\mu_{d 1}$ by working with a submatrix of $C_{d}$ is useful.
Partition $C_{d}$ as

$$
C_{d}=\left(\begin{array}{cc}
C_{d 11} & C_{d 12} \\
C_{d 21} & C_{d 22}
\end{array}\right)
$$

Lemma 3.2: Let $\left(\lambda_{i}, z_{i}\right)$ be the eigenvalue/vector pairs for $C_{d 11}$. Write $x_{i}=z_{i}^{\prime} 1$. Then

$$
\mu_{d 1} \leq \min _{i}\left[\frac{v-x_{i}^{2}}{v}\right]^{-1} \lambda_{i}
$$

The lemma provides a tool for a computational attack leading to a design

| 1 | 1 | 2 | 4 | 5 | 2 | 1 | 5 | 1 | 4 | 3 | 2 | 1 | 3 | 4 | 1 | 3 | 3 | 2 | 2 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | 6 | 3 | 5 | 6 | 4 | 3 | 9 | 2 | 7 | 5 | 6 | 4 | 7 | 8 | 10 | 6 | 4 | 7 | 5 | 5 |
| 3 | 7 | 8 | 6 | 8 | 9 | 7 | 11 | 6 | 9 | 8 | 11 | 8 | 10 | 12 | 11 | 9 | 6 | 8 | 7 | 9 |
| 4 | 8 | 9 | 7 | 10 | 10 | 11 | 12 | 10 | 10 | 10 | 12 | 11 | 12 | 13 | 13 | 13 | 13 | 13 | 13 | 14 |
| 5 | 9 | 10 | 11 | 12 | 12 | 12 | 13 | 14 | 15 | 15 | 15 | 14 | 14 | 14 | 15 | 14 | 15 | 15 | 14 | 15 |
| $\lambda_{1,12}=\lambda_{1,13}=\lambda_{10,11}=\lambda_{10,13}=\lambda_{11,14}=\lambda_{12,15}=\lambda_{14,15}=1$ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

- This is a $\delta_{d}=7$ design.
- It is uniquely $E$-optimal.

To conclude,

- This work directs attention squarely to the complex combinatorial problem of determining what incidence structures actually do exist in these settings.
- From the optimality perspective, it is now clear that the attack on irregular BIBD settings should focus on which discrepancy patterns can actually be achieved, and their relative merits.
- E-optimal designs need not have minimum discrepancy.
- In no setting is found an $A$-optimal design that did not have minimum discrepancy.
- Conjecture : $A$-optimal designs must be minimum discrepancy designs.


## References

Balasubramaniam, K. and Dey, A. (1996). D-optimal designs with minimal and nearly minimal number of units. J. Statist Planning Inference, 52, 255-262.

Bapat, R.B. and Dey, A. (1991). Optimal block designs with minimal number of observations. Statist. Probab. Letters, 11, 399-402.

Cheng, C.S. (1978). Optimality of certain asymmetrical experimental designs. Ann. Statist., 6, 1239-1261.

Cheng, C.S. (1980). On the E-optimality of some block designs. J. Roy. Statist. Soc. Series B, 42, 199-204.

Cheng, C.S. (1992). D-optimality of linked block designs and some related results. In: Bahadur R.R. (Ed.) R.C. Bose Proc. Probability Statistics and Design of Experiments, Wiley, New Delhi.

Constantine, G. M. (1981). Some E-optimal block designs. Ann. Statist., 9, 886-892.
Das, A. (1993). E-optimal block and row-column designs with unequal number of replicates. Sankhya Series B, 55, 77-90.

Dey, A., Shah, K.R. and Das, A. (1995). Optimal block designs with minimal and nearly minimal number of units. Statist. Sinica, 5, 547-558.

Eccleston, J.A. and Hedayat, A.S. (1974). On the theory of connected designs: characterization and optimality. Ann. Statist., 2, 1238-1255.

Jacroux, M. (1983). On the E-optimality of block designs. Sankhya Series B, 45, 351-361.
Jacroux, M. (1985). Some sufficient conditions for the type 1 optimality of block designs. J. Statist Planning Inference, 11, 385-398.

John, J.A. and Mitchell, T.J. (1977). Optimal incomplete block designs. J. Roy. Statist. Soc. Series B, 39, 39-43.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). Ann. Statist., 2, 849-879.

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

Mandal, N.K., Shah, K.R. and Sinha, B.K. (1991). Uncertain resources and optimal designs: problems and perspectives. H.K. Nandi Memorial Volume. Calcutta Statist. Assoc Bull., 40, 267-282.

Morgan, J. P. (2002). Optimal block design for combinatorially aberrant settings. Dekalb IISA Conference handout.

Sathe, Y.S. and Bapat, R.B. (1985). On the E-optimality of truncated BIBD. Calcutta Statist. Assoc Bull., 34, 113-117.

Shah, K.R. (1960). Optimality criteria for incomplete block designs. Ann. Math. Statist., 31, 791-794.

Shah, K.R. and Sinha B.K. (1989). Theory of Optimal Designs. Lecture Notes \#54, Springer - Verlag, New York.

Sinha, B.K. (1970). On the optimality of some designs. Calcutta Statist. Assoc Bull., 19, 1-22.

Takeuchi, K. (1961). On the optimality of certain type of PBIB designs. Rep. Statist. appl. Res. Un. Japan sci. Engrs., 8, 140-145.

Yeh, C.M. (1986). Conditions for universal optimality of block designs. Biometrika, 73, 701-706.

Yeh, C.M. (1988). A class of universally optimal binary block designs. J. Statist Planning Inference, 18, 355-361.

