Most Robust BIBDs

J. P. Morgan and Valentin Parvu
Department of Statistics, Virginia Tech

14$^{\text{th}}$ June 2006; revised 14$^{\text{th}}$ October 2006

Summary

Intense combinatorial study of balanced incomplete block designs since the time of Fisher and Yates has led to a great many designs with the same numbers of treatments, blocks, and block size. While the basic analysis does not differentiate among different BIBDs with the same parameters, they do differ in their capacity to withstand loss of experimental material. Competing BIBDs are compared here for their robustness in terms of average loss and worst loss. A table of most robust BIBDs is compiled. Two useful criteria are *minimum intersection aberration* and *minimum efficiency aberration*.

1 Introduction

Balanced incomplete block designs, or BIBDs, are popular as experimental designs for several reasons. Like the randomized complete block designs, they are variance balanced, that is, every normalized contrast is estimated with the same variance. Amongst the binary block designs with block size k less than number of treatments v, they are the only designs which are variance balanced (Rao, 1958). BIBDs are universally optimal in the sense of Kiefer (1975), and M-optimal in the sense of Bagchi and Bagchi (2001), and so, for instance, minimize the average variance of elementary treatment contrasts over all competitors with the same v, k, and number of blocks b. Because there is only one within-blocks efficiency factor, recovery of interblock information is straightforward (e.g. Houtman and Speed, 1983). BIBDs are optimal designs under the mixed effects model (Mukhopadhyay, 1981).

Recently there has been interest in investigating the robustness properties of BIBDs. Should experimental difficulties arise causing loss of plots or entire blocks, does the resulting residual design still behave well in a statistical sense? Ghosh (1982) established that a BIBD remains connected after removal of any $r-1$ observations, or of any $r-1$ blocks, where $r = \frac{bk}{v}$ is the replication number for the original BIBD. The paper by Lal, Gupta, and Bhar (2001) gives a comprehensive study on the A-efficiency of residual designs from several classes of block designs. Tables 1 and 3 in that paper summarize results for residual efficiencies when two plots (respectively, two blocks) are lost from any of the designs in three published tables of BIBDs; relatively few cases show residual efficiency less than 80% of the original BIBD. Other robustness results can be found in Baksalary and Tabis (1987), Das and Kageyama (1992) and Dey (1993), including for designs other than BIBDs. A fundamental result due to Bhaumik and Whittinghill (1991) will be explored later in this paper. The collective conclusion of these papers is yes, BIBDs as a whole are fairly characterized as robust. The case for BIBDs presented in the first paragraph is correspondingly strengthened.

Nonetheless, the robustness work to now fails to address a very pragmatic issue. Loss of experimental material in designed experiments is rarely expected; rather, it is a somewhat low probability event that will not be the primary driver of design selection. Given low probability for loss of units, the first objective is to determine an optimal design for the selected b and k, irrespective of potential losses. If there are then several designs meeting the optimality criteria, robustness becomes the second objective: *amongst the optimal designs, select the member that is maximally robust.*

Thus, having decided for many reasons to employ a BIBD with given v, b, and k, which of the potentially large number of nonisomorphic competitors is best? Neither theoretical results for BIBDs as a class, nor calculated results for the individual BIBDs listed in particular tables, both available in the papers cited above, tell an experimenter which BIBD for her parameter set is most robust. This paper investigates that question from several angles, both theoretical and computational, under loss of entire blocks. In accordance with the idea that fewer losses have higher likelihood, and since all BIBDs for fixed (v, b, k) are statistically equivalent for loss of one block, focus is first on loss of two blocks. Symmetric ($v = b$) BIBDs for fixed (v, b, k) are statistically equivalent for loss of two blocks, so in this case the primary study is for loss of three blocks. The most robust BIBDs are tabled for up to 15 treatments and 15 replicates. The entirety of this investigation is for the standard additive, homoscedastic model.

2 Information matrices and other preliminaries

Efficiency comparisons of block designs are made in terms of the treatments information matrix in the within-blocks analysis (e.g. Shah and Sinha, 1989, chapter 1). For an arbitrary design d this is

$$C_d = Diag(r_{d1}, r_{d2}, \ldots, r_{dv}) - N_d Diag(k_1, k_2, \ldots, k_b)^{-1} N_d'$$

where r_{di} is the replication number for treatment i, k_j the number of experimental units in block j, and $N_d = (n_{dij})$ is the treatment×block incidence matrix with entries the number of units in block j receiving treatment i. For a BIBD, $C_d = rI - \frac{1}{k}[(r - \lambda)I + \lambda J]$ where λ is the number of blocks common to every pair of treatments. Of concern here is the information matrix for a residual design of a BIBD d under loss of t blocks:

$$C_{d(t)} = C_d - C_{d_t} \tag{1}$$

where C_{d_t} is the $v \times v$ information matrix for the t lost blocks when viewed as a design for v treatments. While $C_{d(t)}$ certainly depends on the t blocks lost, its trace does not.

Let $\Phi : C_d \to \Re$ be a measure of design quality, for which smaller is better. Many such measures are based on the $v-1$ positive eigenvalues $z_{d1} \leq z_{d2} \leq \cdots \leq z_{d,v-1}$ of C_d. For instance, the average variance over the pairwise treatment contrasts is proportional to $\Phi_A(C_d) = \sum_{i=1}^{v-1} \frac{1}{z_{di}}$, called the A-value of the design. The worst variance over all normalized contrasts is proportional to $\Phi_E(C_d) = \frac{1}{z_{d1}}$, called the E-value. A general class of eigenvalue-based criteria is $\Phi_f(C_d) = \sum_{i=1}^{v-1} f(z_{di})$ where f is convex. An important criterion not a function of the eigenvalues is the worst (scaled) variance over the pairwise treatment contrasts, $\Phi_{MV}(C_d) = \max_{i \neq i'} \mathrm{var}(\widehat{\tau_i - \tau_{i'}})/\sigma^2$, called the MV-value. All of these criteria are scale free, that is, do not depend on the underlying error variance σ^2.

Definition 1. The $\Phi_{(t)}$-efficiency of a residual of a BIBD is

$$\Phi_{(t)}\text{-}eff = \frac{\Phi(C_d)}{\Phi(C_{d(t)})}.$$

Any sensible criterion Φ will obey $\Phi(C_1) \leq \Phi(C_2)$ for $C_1 - C_2$ non-negative definite so that $\Phi_{(t)}$-*eff* is bounded above by 1. It is sometimes convenient to think in terms of the *risk* associated with a loss of t blocks. Risk is defined by $1 - \Phi_{(t)}$-*eff*. $\Phi_{(t)}$-efficiency depends on the set of t blocks lost.

A portion of the results in later sections will focus on the criteria A, E, and MV. Evaluation of $\Phi_{(t)}$-*eff* thus requires the values of these criteria for BIBDs. For a BIBD d, the z_{di} are identically equal to $\frac{v\lambda}{k}$ so that $\Phi_A(C_d) = \frac{k(v-1)}{v\lambda}$ and $\Phi_E(C_d) = \frac{k}{v\lambda}$. The MV-value is $\Phi_{MV}(C_d) = \frac{2k}{v\lambda}$. Obtaining criterion values for $C_{d(t)}$ requires a bit more work. A useful technical result is given here, with computation of residual information matrices and associated values following in section 3.

Definition 2. A symmetric matrix X is *generalized block-diagonal* if it can be partitioned as

$$\begin{pmatrix} A_{n_1\times n_1} & c_{12}J_{n_1\times n_2} & \cdots & c_{1p}J_{n_1\times n_p} \\ c_{21}J_{n_2\times n_1} & A_{n_2\times n_2} & \cdots & c_{2p}J_{n_2\times n_2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1}J_{n_p\times n_1} & c_{p2}J_{n_p\times n_2} & \cdots & A_{n_p\times n_p} \end{pmatrix} \tag{2}$$

where the matrices $A_{n_i\times n_i} = x_iI + y_iJ$ are completely symmetric.

Residual information matrices $C_{d(t)}$ for BIBDs are generalized block-diagonal. Efficiency calculations require the eigenvalues of these matrices, and for the MV-criterion, a generalized inverse. It is evident that $\sum_{i=1}^{p}(n_i - 1)$ eigenvalues of (2) are $n_i - 1$ copies of x_i for $i = 1, \ldots, p$. Removing these eigenvalues and corresponding (contrast) eigenvectors has the effect of replacing $A_{n_i\times n_i}$ by $\frac{a_i}{n_i}J_{n_i\times n_i}$ for $a_i = (x_i + n_iy_i)$. From there the remaining eigenvalues can be extracted by noting that the remaining eigenvectors must be of the form $(w_1 1'_{n_1}, w_2 1'_{n_2}, \ldots, w_p 1'_{n_p})$ for some $w_1, w_2, \ldots, w_p$.

Lemma 1. *Let X be a generalized block-diagonal matrix as given in (2). The eigenvalues of X are x_i with multiplicity $n_i - 1$ for $i = 1, \ldots, p$, and the eigenvalues of the (not necessarily symmetric) matrix:*

$$X_0 = \begin{pmatrix} a_1 & c_{12}\cdot n_2 & \cdots & c_{1p}\cdot n_p \\ c_{21}\cdot n_1 & a_2 & \cdots & c_{2p}\cdot n_p \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1}\cdot n_1 & c_{p2}\cdot n_2 & \cdots & a_p \end{pmatrix}$$

3 Residual information matrices

Residual information matrices for BIBDs, and associated values, will be investigated here for the cases of primary interest. The eigenvalues of a residual design formed by removing two blocks from a BIBD as determined by Bhaumik and Whittinghill (1991, lemma 6) are listed in Table 1. Here, q denotes the number of treatments common to the two blocks.

Study of criteria based on pairwise variances, including Φ_{MV}, requires a generalized inverse of $C_{d(2)}$. When $0 < q < k$, $C_{d(2)}$ has form

$$C_{d(2)} = \begin{pmatrix} A_1 & \frac{1}{k}J_{q\times(k-q)} & \frac{1}{k}J_{q\times(k-q)} & 0_{q\times(v-2k+q)} \\ \frac{1}{k}J_{(k-q)\times q} & A_2 & 0_{(k-q)\times k-q} & 0_{(k-q)\times(v-2k+q)} \\ \frac{1}{k}J_{(k-q)\times q} & 0_{(k-q)\times(k-q)} & A_3 & 0_{(k-q)\times(v-2k+q)} \\ 0_{(v-2k+q)\times q} & 0_{(v-2k+q)\times(k-q)} & 0_{(v-2k+q)\times(k-q)} & A_4 \end{pmatrix} - \frac{\lambda}{k}J_{v\times v}, \qquad (3)$$

where

$$A_1 = (\frac{v\lambda}{k} - 2)I_q + \frac{2}{k}J_q, \quad A_2 = A_3 = (\frac{v\lambda}{k} - 1)I_{(k-q)} + \frac{1}{k}J_{(k-q)}, \quad A_4 = \frac{v\lambda}{k}I_{(v-2k+q)}.$$

One generalized inverse of (3) is $G_{d(2)} = [C_{d(2)} + \frac{\lambda}{k}J]^{-1}$. This matrix (see Parvu, 2004) is

$$\begin{pmatrix} \frac{k}{v\lambda-2k}I_q - \frac{2k(v\lambda-k)}{v\lambda(v\lambda-2k)(v\lambda-2k+q)}J_q & \frac{-k}{v\lambda(v\lambda-2k+q)}J_{q\times(k-q)} & \frac{-k}{v\lambda(v\lambda-2k+q)}J_{q\times(k-q)} & 0_{q\times(v-2k+q)} \\ \frac{-k}{v\lambda(v\lambda-2k+q)}J_{(k-q)\times q} & M_{(k-q)\times(k-q)} & \frac{kq}{v\lambda(v\lambda-q)(v\lambda-2k+q)}J_{(k-q)} & 0_{(k-q)\times(v-2k+q)} \\ \frac{-k}{v\lambda(v\lambda-2k+q)}J_{(k-q)\times q} & \frac{kq}{v\lambda(v\lambda-q)(v\lambda-2k+q)}J_{(k-q)} & M_{(k-q)\times(k-q)} & 0_{(k-q)\times(v-2k+q)} \\ 0_{(v-2k+q)\times q} & 0_{(v-2k+q)\times(k-q)} & 0_{(v-2k+q)\times(k-q)} & \frac{k}{\lambda v}I_{v-2k+q} \end{pmatrix}$$

where $M_{(k-q)\times(k-q)} = \frac{k}{v\lambda-k}I_{(k-q)} - \frac{k[v^2\lambda^2-k(2v\lambda-q)]}{v\lambda(v\lambda-k)(v\lambda-q)(v\lambda-2k+q)}J_{(k-q)}$. The variance of a pairwise treatment comparison is $Var[\widehat{\tau_i - \tau_{i'}}]/\sigma^2 = G_{d(2)ii} + G_{d(2)i'i'} - 2G_{d(2)ii'}$ as displayed for the seven possible cases in Table 2. For some q not all cases can occur (some counts are zero) .

For a symmetric BIBD (or SBIBD), every two blocks intersect in λ treatments. It follows that for fixed (v, b, k), every SBIBD d has the same residual information matrix $C_{d(2)}$ regardless of the pair of blocks lost, and thus the same $\Phi_{(2)}$-*eff* values for any Φ. So for SBIBDs primary interest is in loss of three blocks. Let n_a be the number of treatments occurring in a of three lost blocks for $a = 0, 1, 2, 3$. Take treatments $1, \ldots, n_3$ to be those occurring three times, treatments $n_3 + 1, \ldots, n_3 + n_2$ to be those occurring twice, and so on. Then $n_a \geq 0$, $n_3 + n_2 + n_1 + n_0 = v$, $3n_3 + 2n_2 + n_1 = 3k$, and $3n_3 + n_2 = 3\lambda$. Now setting $n_3 = q$, the n_a's in terms of q and the design parameters are

$$n_3 = q, \quad n_2 = 3\lambda - 3q, \quad n_1 = 3k - 6\lambda + 3q, \quad n_0 = v - 3k + 3\lambda - q \tag{4}$$

Because $n_a \geq 0$, a necessary condition for q is $\max(0, 2\lambda - k) \leq q \leq \min(\lambda, v - 3k + 3\lambda)$. The form of C_{d_3} depends on which n_a's are strictly positive, and (4) implies that of the 16 possibilities for choosing $n_a = 0$ or $n_a > 0$, only eight can actually occur (for instance $n_3 = n_2 = 0$ would imply $q = \lambda = 0$). These cases are:

Case	1	2	3	4	5	6	7	8
n_3	+	0	0	+	+	+	+	+
n_2	+	+	+	0	0	+	+	+
n_1	+	+	0	+	+	0	0	+
n_0	+	+	+	+	0	+	0	0

Not all eight cases can occur with any SBIBD. Case 3 can occur in SBIBDs with $k = 2\lambda$, case 5 when $v = 3k - 2\lambda$, and case 7 when $v = 2k - \lambda$.

Let r_{d_3} be the vector $r_{d_3} = (31'_{n_3}, 21'_{n_2}, 1'_{n1}, 0'_{n_o})'$. The information matrix for the three deleted blocks is $C_{d_3} = Diag(r_{d_3}) - \frac{1}{k} N_{d_3} N_{d_3}$. If all $n_a > 0$ (case 1) then $N_{d3} N'_{d3}$ is

$$\begin{pmatrix}
3 \cdot J_{n_3 \times n_3} & 2 \cdot J_{n_3 \times \frac{n_2}{3}} & 2 \cdot J_{n_3 \times \frac{n_2}{3}} & 2 \cdot J_{n_3 \times \frac{n_2}{3}} & J_{n_3 \times \frac{n_1}{3}} & J_{n_3 \times \frac{n_1}{3}} & J_{n_3 \times \frac{n_1}{3}} & 0_{n_3 \times n_0} \\
2 \cdot J_{\frac{n_2}{3} \times n_3} & 2 \cdot J_{\frac{n_2}{3} \times \frac{n_2}{3}} & J_{\frac{n_2}{3} \times \frac{n_2}{3}} & J_{\frac{n_2}{3} \times \frac{n_2}{3}} & J_{\frac{n_2}{3} \times \frac{n_1}{3}} & J_{\frac{n_2}{3} \times \frac{n_1}{3}} & 0_{\frac{n_2}{3} \times \frac{n_1}{3}} & 0_{\frac{n_2}{3} \times n_0} \\
2 \cdot J_{\frac{n_2}{3} \times n_3} & J_{\frac{n_2}{3} \times \frac{n_2}{3}} & 2 \cdot J_{\frac{n_2}{3} \times \frac{n_2}{3}} & J_{\frac{n_2}{3} \times \frac{n_2}{3}} & J_{\frac{n_2}{3} \times \frac{n_1}{3}} & 0_{\frac{n_2}{3} \times \frac{n_1}{3}} & J_{\frac{n_2}{3} \times \frac{n_1}{3}} & 0_{\frac{n_2}{3} \times n_0} \\
2 \cdot J_{\frac{n_2}{3} \times n_3} & J_{\frac{n_2}{3} \times \frac{n_2}{3}} & J_{\frac{n_2}{3} \times \frac{n_2}{3}} & 2 \cdot J_{\frac{n_2}{3} \times \frac{n_2}{3}} & 0_{\frac{n_2}{3} \times \frac{n_1}{3}} & J_{\frac{n_2}{3} \times \frac{n_1}{3}} & J_{\frac{n_2}{3} \times \frac{n_1}{3}} & 0_{\frac{n_2}{3} \times n_0} \\
J_{\frac{n_1}{3} \times n_3} & J_{\frac{n_1}{3} \times \frac{n_2}{3}} & J_{\frac{n_1}{3} \times \frac{n_2}{3}} & 0_{\frac{n_1}{3} \times \frac{n_2}{3}} & J_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times n_0} \\
J_{\frac{n_1}{3} \times n_3} & J_{\frac{n_1}{3} \times \frac{n_2}{3}} & 0_{\frac{n_1}{3} \times \frac{n_2}{3}} & J_{\frac{n_1}{3} \times \frac{n_2}{3}} & 0_{\frac{n_1}{3} \times \frac{n_1}{3}} & J_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times n_0} \\
J_{\frac{n_1}{3} \times n_3} & 0_{\frac{n_1}{3} \times \frac{n_2}{3}} & J_{\frac{n_1}{3} \times \frac{n_2}{3}} & J_{\frac{n_1}{3} \times \frac{n_2}{3}} & 0_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times \frac{n_1}{3}} & J_{\frac{n_1}{3} \times \frac{n_1}{3}} & 0_{\frac{n_1}{3} \times n_0} \\
0_{n_0 \times n_3} & 0_{n_0 \times \frac{n_2}{3}} & 0_{n_0 \times \frac{n_2}{3}} & 0_{n_0 \times \frac{n_2}{3}} & 0_{n_0 \times \frac{n_1}{3}} & 0_{n_0 \times \frac{n_1}{3}} & 0_{n_0 \times \frac{n_1}{3}} & 0_{n_0 \times n_0}
\end{pmatrix} \tag{5}$$

For the other seven cases, delete rows and columns of the partitioned matrix (5). For instance, if $n_2 = 0$, the second, third, and fourth row and column partitions are deleted.

From (1), the information matrix of the residual design of a SBIBD d when three blocks are removed is $C_{d(3)} = (k - 1 + \frac{\lambda}{k}) I_v - Diag(r_{d_3}) - \frac{\lambda}{k} J_v + \frac{1}{k} N_{d_3} N'_{d_3}$. It is apparent that C_{d_3} is generalized block-diagonal with number of partitions $p \leq 8$, and hence so too is $C_{d(3)}$. Thus lemma 1 can be used to find its eigenvalues. With the notation of definition 2, $C_{d(3)}$ has the following sub-matrices on the diagonal when $n_a > 0$ for all a:

$$\begin{aligned}
A_1 &= (k-4+\tfrac{\lambda}{k})I_{n_3} + \tfrac{3-\lambda}{k}J_{n_3}, \quad a_1 = k-4+3\tfrac{n_3}{k} - (n_3-1)\tfrac{\lambda}{k} \\
A_2 = A_3 = A_4 &= (k-3+\tfrac{\lambda}{k})I_{n_2/3} + \tfrac{2-\lambda}{k}J_{n_2/3}, \quad a_2 = a_3 = a_4 = k-3+2\tfrac{n_2}{3k} - (n_2/3-1)\tfrac{\lambda}{k} \\
A_5 = A_6 = A_7 &= (k-2+\tfrac{\lambda}{k})I_{n_1/3} + \tfrac{1-\lambda}{k}J_{n_1/3}, \quad a_5 = a_6 = a_7 = k-2+\tfrac{n_1}{3k} - (n_1/3-1)\tfrac{\lambda}{k} \\
A_8 &= (k-1+\tfrac{\lambda}{k})I_{n_0} + \tfrac{-\lambda}{k}J_{n_0}, \quad a_8 = k-1-(n_0-1)\tfrac{\lambda}{k}
\end{aligned} \tag{6}$$

The eigenvalues of $C_{d(3)}$ arising from the individual matrices in (6) are $\frac{v\lambda}{k}-3$, $\frac{v\lambda}{k}-2$, $\frac{v\lambda}{k}-1$, and $\frac{v\lambda}{k}$ with multiplicities n_3-1, n_2-3, n_1-3, and n_0-1 respectively; not all appear in cases 2-8. The other eigenvalues of $C_{d(3)}$ are computed from the reduced matrix displayed in lemma 1:

$$\frac{1}{k}\begin{pmatrix}
k\cdot a_1 & (2-\lambda)\frac{n_2}{3} & (2-\lambda)\frac{n_2}{3} & (2-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_1}{3} & (1-\lambda)\frac{n_1}{3} & (1-\lambda)\frac{n_1}{3} & -\lambda n_0 \\
(2-\lambda)n_3 & k\cdot a_2 & (1-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_1}{3} & (1-\lambda)\frac{n_1}{3} & -\lambda\frac{n_1}{3} & -\lambda n_0 \\
(2-\lambda)n_3 & (1-\lambda)\frac{n_2}{3} & k\cdot a_3 & (1-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_1}{3} & -\lambda\frac{n_1}{3} & (1-\lambda)\frac{n_1}{3} & -\lambda n_0 \\
(2-\lambda)n_3 & (1-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_2}{3} & k\cdot a_4 & -\lambda\frac{n_1}{3} & (1-\lambda)\frac{n_1}{3} & (1-\lambda)\frac{n_1}{3} & -\lambda n_0 \\
(1-\lambda)n_3 & (1-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_2}{3} & -\lambda\frac{n_2}{3} & k\cdot a_5 & -\lambda\frac{n_1}{3} & -\lambda\frac{n_1}{3} & -\lambda n_0 \\
(1-\lambda)n_3 & (1-\lambda)\frac{n_2}{3} & -\lambda\frac{n_2}{3} & (1-\lambda)\frac{n_2}{3} & -\lambda\frac{n_1}{3} & k\cdot a_6 & -\lambda\frac{n_1}{3} & -\lambda n_0 \\
(1-\lambda)n_3 & -\lambda\frac{n_2}{3} & (1-\lambda)\frac{n_2}{3} & (1-\lambda)\frac{n_2}{3} & -\lambda\frac{n_1}{3} & -\lambda\frac{n_1}{3} & k\cdot a_7 & -\lambda n_0 \\
-\lambda n_3 & -\lambda\frac{n_2}{3} & -\lambda\frac{n_2}{3} & -\lambda\frac{n_2}{3} & -\lambda\frac{n_1}{3} & -\lambda\frac{n_1}{3} & -\lambda\frac{n_1}{3} & k\cdot a_8
\end{pmatrix}$$

Again, some rows/columns of this matrix are removed depending on the case. What remains is purely computational, so the eigenvalues for $C_{d(3)}$ are stated (Table 3) without further comment.

4 Comparison of nonisomorphic designs

With the basic tools for computing efficiencies in place, paradigms for design comparison will now be developed. Using the GAP Design package (Soicher, 2003) one can, within bounds, obtain the collection $\mathcal{D}(v, b, k)$ of all nonisomorphic BIBDs for given (v, b, k). The bounds here are in time and computational effort: the isomorphism problem is notoriously hard, even computationally, and for some parameter sets the number of nonisomorphic designs is extremely large. For instance, it is known that there are more than 1.25×10^8 designs in $\mathcal{D}(9, 36, 4)$. The theory of this section provides a route for avoiding full enumeration, and in doing so makes addressing the robustness problem feasible for a much greater parameter range than would otherwise be possible. A table of BIBD parameters including number of isomorphism classes where known is in Mathon and Rosa (2006).

4.1 Robustness of non-symmetric BIBDs to loss of two blocks

As discussed in section 3, the information matrix $C_{d(2)}$ of a residual design depends only on q, the intersection number of the two lost blocks. There are $\binom{b}{2}$ pairs of blocks that could be lost, and $k+1$ possible values for their intersections.

Definition 3. The $\Phi_{(t)}$ *distribution* of BIBD d is the collection of distinct $\Phi_{(t)}$-*eff* values over all $\binom{b}{t}$ possible losses of t blocks, and their corresponding frequencies.

The problem is to compare the $\Phi_{(2)}$ distributions of nonisomorphic designs with the same (v, b, k). How one chooses to do this not only determines which BIBD is most robust, but defines how the robustness concept is implemented. One choice, corresponding to one extreme in what one expects for *how* the blocks may be lost, is to compare distributions through their *maximum* value. This is appropriate if one may assign which two blocks are most likely to be lost. Though not formulated in the same way, this idea was discussed by Bhaumik and Whittinghill (1991, page 406).

More common are experiments in which one wishes to guard against possible losses with no prior information as to which blocks may be lost. In broad terms there are two further choices. The opposite extreme from that just discussed is to compare $\Phi_{(2)}$ distributions through their *minimum* value. This approach casts robustness in terms of protecting against a worst case loss for the experiment at hand. Between the two extremes, comparison could be made through the *mean* values of the distributions, formulating robustness via long-run average loss over many experiments. Both the minimum and the mean will be pursued here. The ideas are first illustrated with an example.

Example 1. There are four nonisomorphic BIBDs with parameters $(v, b, k) = (7, 14, 3)$ and thus $(r, \lambda) = (6, 2)$, distinguished in Table 4 by the distribution of intersection numbers for the $\binom{14}{2} = 91$ pairs of blocks in each design. Average and minimum $\Phi_{(2)}$-efficiencies over the 91 possible pairs of lost blocks are also displayed in Table 4. When averaging $\Phi_{(2)}$ values, the robustness ordering for these designs is $1 > 2 > 3 > 4$. Design 4 is best for minimum A-efficiency; the other two "worst case" calculations do not distinguish among the competitors.

One feature of example 1 is the small range of the various efficiencies. This is a masking of relevant behavior due to repeated, structurally fixed values, an issue addressed following example 2 below.

Another prominent feature of example 1, the essentially reversed design ordering produced by the average and the minimum efficiencies, will be seen again. Theorem 1 gives some clarification of this phenomenon. Let $C_{d(2)}^{(q)}$ denote the information matrix of a design obtained by removing two blocks with intersection number q from a BIBD. Also, let $\mu\{C_{d(2)}^{(q)}\}$ denote the ordered vector of positive eigenvalues of the matrix $C_{d(2)}^{(q)}$. For vectors x and y, the notation "$x \prec y$" means "y majorizes x" (also read "x is majorized by y"). Refer to Bhaumik and Whittinghill (1991) or more generally Bhatia (1997) for definitions and details related to majorization.

Theorem 1 (Bhaumik and Whittinghill, 1991)**.** *The vectors of eigenvalues of residual designs obtained by removing 2 blocks from a BIBD satisfy:*

$$\mu\{C_{d(2)}^{(0)}\} \prec \mu\{C_{d(2)}^{(1)}\} \prec \cdots \prec \mu\{C_{d(2)}^{(k)}\} \tag{7}$$

and thus

$$\Phi(C_{d(2)}^{(0)}) \leq \Phi(C_{d(2)}^{(1)}) \leq \cdots \leq \Phi(C_{d(2)}^{(k)}) \tag{8}$$

for any criterion Φ preserving the majorization partial order on vectors.

Theorem 1 orders the residual designs of a BIBD in terms of the intersection numbers of the removed blocks. Since Φ_A and Φ_E both preserve the majorization ordering, $\Phi_A(C_{d(2)}^{(0)}) \leq \Phi_A(C_{d(2)}^{(1)}) \leq \cdots \leq \Phi_A(C_{d(2)}^{(k)})$ and $\Phi_E(C_{d(2)}^{(0)}) \leq \Phi_E(C_{d(2)}^{(1)}) \leq \cdots \leq \Phi_E(C_{d(2)}^{(k)})$. This explains why design 4 in example 1 maximizes the minimum Φ_A-*eff* and Φ_E-*eff*. Since it contains no pair of identical blocks, its minimum A-efficiency is $\frac{\Phi_A(C_d)}{\Phi_A(C_{d(2)}^{(2)})}$, while for the other designs the minimum A-efficiency is $\frac{\Phi_A(C_d)}{\Phi_A(C_{d(2)}^{(3)})}$. In this example, the strict inequality does not hold with respect to minimum E-efficiency (more on this later).

Theorem 1 suggests a method for differentiating among designs on majorization criteria without computing efficiencies. Let $\eta(d) = (\eta_0(d), \eta_1(d), \ldots, \eta_k(d))$ be the vector of block intersection counts for the BIBD d, that is, $\eta_s(d)$ is the number of pairs of blocks that intersect in s treatments. For instance, in d_1 of example 1, 84 of the 91 pairs of blocks intersect in one treatment (see Table 4), while the other seven pairs each intersect in 3 treatments, so $\eta(d_1) = (0, 84, 0, 7)$.

Definition 4. Let d_1 and d_2 be two balanced incomplete block designs, and s be the largest integer such that $\eta_s(d_1) \neq \eta_s(d_2)$. Then design d_1 is said to have *less intersection aberration* than d_2 if $\eta_s(d_1) < \eta_s(d_2)$. A design $d \in \mathcal{D}(v, b, k)$ has *minimum intersection aberration* (MIA), if no other design in $\mathcal{D}(v, b, k)$ has less intersection aberration than d.

Theorem 2. *Let $d \in \mathcal{D}(v, b, k)$ have minimum intersection aberration. Then d maximizes the minimum $\Phi_{(2)}$-eff over $\mathcal{D}$ for every majorization criteria Φ.*

In short, MIA designs offer the best protection against worst-case loss of two blocks, for all reasonable criteria that can be expressed in terms of eigenvalues. Theorem 2 is immediate from Theorem 1 and definition 4. Though the concepts here are quite different, the "aberration" terminology parallels usage in the fractional factorial literature (e.g. Wu and Hamada, 2000, section 4.2), with MIA sequentially minimizing η_s beginning with largest s.

For an experimenter focused on pairwise comparisons $\tau_i - \tau_{i'}$, eigenvalue-based criteria are not necessarily the proper basis for evaluation of competing designs. Theorem 2 does guarantee that worst-case loss in terms of average pairwise variance is minimized, due to equivalence with Φ_A. It does not, however, address Φ_{MV}. Theorem 3 does this and much more. Write $v_{ii'}$ for $\text{var}(\widehat{\tau_i - \tau_{i'}})/\sigma^2$.

Theorem 3. *Let $d \in \mathcal{D}(v, b, k)$ have minimum intersection aberration. For any increasing, convex g consider the criterion $\Phi(C_d) = \sum_i \sum_{i'>i} g(v_{ii'})$. Then d maximizes the minimum $\Phi_{(2)}$-eff over $\mathcal{D}$ for every such g.*

Proof. For any BIBD d let $\nu\{C_{d(2)}^{(q)}\}$ with coordinates ν_{jq} denote the vector of pairwise variances $v_{ii'}$ in nonincreasing order. It is sufficient to show, akin to Theorem 1, that $\nu\{C_{d(2)}^{(q)}\} \prec_w \nu\{C_{d(2)}^{(q+1)}\}$ (Bhatia, 1997, p. 40; this is weak submajorization, as $\sum\sum v_{ii'}$ depends on the blocks lost), that is,

$$\sum_{j=1}^{h} \nu_{j,q+1} - \sum_{j=1}^{h} \nu_{jq} \geq 0 \tag{9}$$

for each $h = 1, \ldots, \binom{v}{2}$ and any $0 \leq q \leq k-1$. Since all designs have the same $\Phi_{(2)}$-distribution if $\lambda = 1$ or $b = v$, only the asymmetric case with $\lambda \geq 2$ needs to be considered. It can then be checked that the seven variances displayed in Table 2 satisfy $V_{1q} \geq V_{2q} \geq \cdots \geq V_{7q}$. Denote the corresponding counts by $c_{1q}, \ldots, c_{7q}$ and write $s_{jq} = \sum_{i=1}^{j} c_{iq}$. Then it is sufficient to establish (9) for h at each s_{jq} and each $s_{j,q+1}$, of which only a few need actually be computed.

First, $s_{1q} < s_{1,q+1}$, $V_{1q} = V_{1,q+1}$ and $V_{2,q+1} \geq V_{2q} \Rightarrow$ (9) holds for $h \leq s_{2,q+1}$. Depending on k and q, s_{2q} could be larger than $s_{j,q+1}$ for $j = 2, 3$ or 4, in which case it is possible that (9) fails at $h = s_{2q}$. If so, then it also fails at $h = s_{3,q+1} < s_{3q}$. Moreover, $V_{4,q+1} < V_{3,q+1} = V_{3q}$ so the difference in (9) is decreasing for $h \in (s_{3,q+1}, s_{3,q})$. Thus the first h for which (9) need be checked is s_{3q}. Removing the positive denominator terms $\delta = v\lambda(v\lambda-2k+q)(v\lambda-q)(v\lambda-2k+q+1)(v\lambda-q-1)(v\lambda-k)(v\lambda-2k)$, the relevant difference is

$$W_1(v,k,\lambda,q) = \delta[\sum_{i=1}^{3} c_{i,q+1}V_{i,q+1} + (s_{3q} - s_{3,q+1})V_{4,q+1} - \sum_{i=1}^{3} s_{iq}c_{iq}]$$

which is a lengthy expression best left to an algebraic manipulator such as Maple. The task is to show that $W_1(v,k,\lambda,q) \geq 0$ for $v \geq k+1$, $k \geq q+1$, and $\lambda \geq 2$, or equivalently that $W_1(v+(k+q+1)+1, k+q+1, \lambda+2, q) \geq 0$ for $v, k, \lambda, q \geq 0$. It is trivial (with a program such as Maple) to see that all coefficients for the latter polynomial in v, k, λ and q are positive.

Next $s_{4,q+1} > s_{4q} \geq s_{3q}$ and $V_{4,q+1} > V_{4q} > V_{5q}$ imply that the difference in (9) is nondecreasing for $h \in (s_{3q}, s_{4,q+1}]$. But $V_{5,q+1} < V_{5q}$, so it is possible that (9) fails at $h = s_{5q}$ if $s_{4,q+1} < s_{5q}$, that is, if $(v - 2k + q) \leq (k-q)^2 - 2$.

Moreover, $s_{6,q+1} \leq s_{6q}$, $V_{7,q+1} = V_{7q} < V_{6q}$, and (9) holds at $h = s_{7q} = s_{7,q+1}$ by Theorem 1 and equivalence with Φ_A, so (9) must hold at $h = s_{6,q+1}$. For $h \in (s_{5q}, s_{6,q+1}]$ the difference in (9) is not monotonic only if $s_{5,q} < s_{5,q+1} < s_{6,q+1}$, but since $V_{6,q+1} < V_{5q}$, it is in this case increasing for h up to $s_{5,q+1}$ and then monotonic at least until $h = s_{6,q+1}$. Thus if (9) holds at $h = s_{5q}$, it holds for all h.

It remains to check (9) at $h = s_{5q}$. There are two cases. If $2(k-q-1) \leq v - 2k + q \leq (k-q)^2 - 2$ then $s_{4,q+1} < s_{5,q} \leq s_{5,q+1}$ and the relevant difference is

$$W_2(v,k,\lambda,q) = \delta[\sum_{i=1}^{4} c_{i,q+1}V_{i,q+1} + (s_{5q} - s_{4,q+1})V_{5,q+1} - \sum_{i=1}^{5} s_{iq}c_{iq}].$$

If $v - 2k - q < 2(k - q - 1)$ then $s_{4,q+1} \leq s_{5,q+1} < s_{5q}$ and the relevant difference is

$$W_3(v, k, \lambda, q) = \delta[\sum_{i=1}^{5} c_{i,q+1} V_{i,q+1} + (s_{5q} - s_{5,q+1}) V_{6,q+1} - \sum_{i=1}^{5} s_{iq} c_{iq}].$$

Establishing that W_2 and W_3 are each positive can be done similarly as described for W_1 above: shift function arguments in accordance with lower bound restrictions so that each may be taken as nonnegative, and observe that the resulting "shifted" functions are polynomials all of whose coefficients are positive. This is again a task for an algebraic manipulator program, which is also the simplest way to check statements above regarding orderings of the c_{jq}, s_{jq}, and V_{jq}. □

While not explicitly addressed in the statement of Theorem 3, the proof by majorization implies that MIA designs minimize maximum MV-risk. Theorem 3 essentially says that MIA produces the same properties for estimating pairwise comparisons as it does for estimating all contrasts.

Example 2. Robustness against the loss of two blocks for the four nonisomorphic BIBDs in $\mathcal{D}(8, 14, 4)$ is summarized in Table 6. Design 4 has no blocks intersecting in more than two treatments, while designs 1, 2, and 3 contain blocks intersecting in 3 treatments. Design 4 has MIA, and thus maximizes the minimum efficiency of its residuals with respect to all criteria covered by Theorems 2 and 3.

Loss of two blocks disturbs relatively few of the eigenvalues, and of those that do change, some values must be the same for all BIBDs and all losses. Consequently, as already seen in Tables 4 and 6, shared multiplicities tend to swamp efficiency calculations, in effect diluting assessment of changes for the treatments affected. For instance, since the smallest nonzero eigenvalue of a residual design is $\frac{v\lambda}{k} - 2$ unless $q \leq 1$, the minimum E-efficiency is common to all designs in each example. For the purpose of assessing the relative effect of losses in one design versus another, sharper focus is gained by comparing their lists of eigenvalues only where different.

Definition 5. For designs d_1 and d_2 whose information matrices have lists of positive eigenvalues $\mu(d_1)$ and $\mu(d_2)$ sharing c common values, let $\mu^-(d_1)$ and $\mu^-(d_2)$ denote the corresponding reduced lists of $v - c$ non-common eigenvalues. The Φ-*impact* of d_2 relative to d_1 is the ratio of Φ-value of design d_1 to that of design d_2, where Φ is any eigenvalue-based criterion as discussed in section 2 for the $(v - c)$-dimensional vectors $\mu^-(d_1)$ and $\mu^-(d_2)$.

The MV-impact is similarly computed as the ratio of the two maximum and non-equal variances of pairwise treatment contrasts. Whether for eigenvalues or pairwise variances, eliminating common values is statistically equivalent to comparing only variances of contrasts which are indeed different.

Returning to example 1, an analysis of worst case losses for $(7, 14, 3)$ BIBDs is presented in Table 5. The worst case residuals for designs 1, 2 and 3 have $q = 3$, while the worst case residual for design 4, the MIA design, has $q = 2$. Thus, the table presents a comparison of the non-common eigenvalues and non-common pairwise variances arising from $C^{(2)}_{d(2)}$ and $C^{(3)}_{d(2)}$ for $(7, 14, 3)$ designs. The advantage enjoyed by design 4 is now obvious. For instance, even though the largest variance of pairwise treatment comparisons (0.571429) is common to all four designs, the ratio of the largest non-equal variances is 0.816667, a loss of roughly 20% relative to the MIA design. Impact relative to the MIA design in $\mathcal{D}(8, 14, 4)$ is shown in Table 7.

If one adopts worst case protection as the formulation of robustness, then the worst case judgement can also be posed in terms of aberration of efficiencies. Let $\phi_{(t)}(d) = (\phi_{(t)1}(d), \phi_{(t)2}(d), \ldots, \phi_{(t)\binom{b}{t}}(d))$ be the vector of $\Phi_{(t)}$-*eff* values for d arranged in nondecreasing order.

Definition 6. Let d_1 and d_2 be two BIBDs in $\mathcal{D}(v, b, k)$, and s be the smallest integer such that $\phi_{(t)s}(d_1) \neq \phi_{(t)s}(d_2)$. Then design d_1 is said to have *less efficiency aberration* than d_2 if $\phi_{(t)s}(d_1) > \phi_{(t)s}(d_2)$. A design $d \in \mathcal{D}(v, b, k)$ has *minimum efficiency aberration* (MEA), if no other design in $\mathcal{D}(v, b, k)$ has less efficiency aberration than d.

Efficiency aberration better distinguishes exposure to worst case risk than does intersection aberration, in the sense of being valid for evaluating loss with respect to any fixed number t of blocks. The advantage of intersection aberration is computational efficiency for the important $t = 2$ case, as one works solely with the distribution of block intersection counts rather than functions of the corresponding lists of eigenvalues or pairwise variances. For $t = 2$, MEA and MIA are equivalent for all Theorem 2 and 3 criteria satisfying $\Phi(C^{(q)}_{d(2)}) < \Phi(C^{(q+1)}_{d(2)})$ for each q. When there is more than one MIA design for $t = 2$, they can be distinguished based on MEA for $t = 3$ (or higher).

Both aberration criteria address worst case risk, and generally do not minimize average risk. A "perfectly robust" BIBD, say d^*, with respect to a criterion Φ, would satisfy $\phi_{(t)}(d^*) \prec^w \phi_{(t)}(d)$ for all d (this is weak supermajorization; see Bhatia, 1997, page 30), and thus would minimize average risk, maximum risk, and indeed would minimize $\sum_s f(\phi_{(t)s}(d))$ for any decreasing, convex f. Examples 1 and 2 above make clear that perfectly robust BIBDs do not generally exist. Other similarly detailed examples in Parvu (2004) show likewise. One more example, this one producing two MIA designs, is given here.

Example 3. There are 21 nonisomorphic BIBDs in $\mathcal{D}(10, 18, 5)$. Of these, 17 have maximum pairwise block intersection $q = 4$, while two have identical η-vectors $(0, 9, 81, 63, 0, 0)$ and thus are MIA designs. The A-, E-, and MV-impacts for the 17 non-MIA designs are 0.991, 0.909, and 0.914, respectively.

4.2 A catalog of MIA balanced incomplete block designs

The MIA criterion differentiates between designs so to guarantee minimax risk, with clear computational advantage over efficiency aberration. The GAP Design program allows enumeration of BIBDs with specified block intersection counts η, so that determination of all designs can often be avoided. Table 8 is a catalog of MIA BIBDs mostly compiled in this way. Initial blocks for cyclic constructions are listed, including fixed blocks where necessary. In a few cases, MIA designs have been found by trial and error.

Where there is more than one MIA design for given (v, b, k), the MEA design for loss of three blocks has been reported. In a few such cases, there is more than one MEA-best design, in which case ties have been broken by inspecting MEA for $t = 4$, and in one case, $t = 5$. MEA is calculated with respect to the A-criterion. Thus among designs minimizing maximum risk with respect to all criteria in Theorems 2 and 3 for loss of two blocks, the tabled design minimizes maximum A-risk for loss of three blocks. If $\mathcal{D}(v, b, k)$ has $\lambda = 1$, all BIBDs have the same distribution of block intersection counts, and so all are MIA designs.

Only parameter sets with more than one nonisomorphic design are included in Table 8, which covers $v \leq 15$ and $r \leq 15$. For d in any $\mathcal{D}(v, b, k)$, if d is MIA, then the complement of d is MIA in $\mathcal{D}(v, b, v - k)$; consequently $k > \frac{v}{2}$ is excluded from the table, with one exception. There is no guarantee that the MEA-best of MIA competitors remains so under complementation. For $v \leq 15$, $r \leq 15$ this occurs only for $\mathcal{D}(10, 15, 6)$, and so this parameter set is included in Table 8.

In a few cases all MIA designs are not yet known; here the best in terms of MEA of those we do know is listed. In a few other cases, due to the current impossibility of determining even all designs with fixed block intersection counts, we have been unable to establish the MIA η. In these cases the best design currently known to us in terms of intersection aberration is listed, and the number of MIA designs is reported as a question mark. A lower bound for the largest pairwise block intersection number, achieved by all of the "?" designs, is $\lfloor \frac{k(r-1)}{b-1} \rfloor$.

4.3 Robustness of SBIBDs to the removal of three blocks

For given $(v, b = v, k)$, any two blocks in a SBIBD intersect in $\lambda = k(k-1)/(v-1)$ treatments, so that all nonisomorphic competitors are equally robust against the loss of two blocks. For SBIBDs, the study of robustness begins with $t = 3$. As in section 2, let q denote the number of treatments common to the three blocks removed. As shown there, eigenvalues of the residual information matrix $C_{d(3)}^{(q)}$ depend only on q and the setting parameters. Intuitively it would seem that the majorization order in (7) and (8) would hold here as well: $\mu\{C_{d(3)}^{(q)}\} \prec \mu\{C_{d(3)}^{(q+1)}\}$ and so $\Phi\{C_{d(3)}^{(q)}\} \leq \Phi\{C_{d(3)}^{(q+1)}\}$ for each q. Unfortunately, this is not the case, as can be seen by working with the eigenvalues in Table 3. With no 3-block residual inferior to another with respect to all majorization criteria, there is no undisputed worst-case loss. Maximal E-risk, however, can be minimized.

The minimum eigenvalue of a residual when removing three blocks with intersection $q > 1$ from a SBIBD is $\frac{v\lambda}{k} - 3$ with multiplicity $q - 1$. Thus $C_{d(3)}^{(q+1)}$ has a count one larger for the minimum eigenvalue it shares with $C_{d(3)}^{(q)}$. In statistical terms, a residual obtained by removing three blocks intersecting in $q + 1$ treatments, relative to a residual with intersection q, has one additional dimension for contrasts of maximal variance. This justifies the MIA criterion for differentiating among SBIBDS under loss of three blocks.

Example 4. There are five nonisomorphic SBIBDs in $\mathcal{D}(15, 7, 7)$. For each design there are $\binom{15}{3} = 455$ possible residuals. The η-vectors for counts of sizes $q \in \{0, 1, 2, 3\}$ of 3-block intersections are $(0, 420, 0, 35)$, $(16, 372, 48, 19)$, $(24, 348, 72, 11)$, and (for two of the designs, which are both therefor MIA) $(28, 336, 84, 7)$. The E-impact for the three non-MIA designs is $\frac{3.4286}{3.8179} = \frac{0.5333}{0.5939} = 0.898$. Notably, none of the four distinct eigenvalue lists majorizes any of the others.

The designs of example 4 are the only SBIBDs falling within the range of Table 8 for which there is more than one nonisomorphic alternative. Other detailed examples for SBIBDs can be found in Parvu (2004). E-impacts are easily calculated for all BIBDs using the values in Tables 1 or 3.

5 Robustness of designs to loss of data

Minimum intersection aberration and minimum efficiency aberration have been introduced as methods for identifying the most robust BIBD amongst all BIBDs with the same numbers of treatments, blocks, and experimental units per block. If concern is with protecting against a worst-case loss of two blocks, then MIA allows evaluation in terms of simple intersection counts while providing the best design with respect to all reasonable criteria. MEA is a more general criterion for evaluating robustness to the loss of any number t of blocks, that also protects against a worst-case scenario. For SBIBDs, MIA applies to loss of three blocks, guaranteeing minimax E-risk. With a few open cases, MIA BIBDs with $v \leq 15$ and $r \leq 15$ have been identified. Numeric assessment of loss is appropriately measured by efficiency impact. For any BIBD class, with loss of three or more blocks, there need not be a majorization ordering of the possible residual designs in the sense of (7) and (8).

MIA and MEA are not the only feasible routes for evaluating robustness. The underlying problem is that of comparing distributions of $\Phi_{(t)}$-*eff* values for competing designs, and as with any collection of distributions, there are a variety of ways in which they could be compared and ordered. It has been shown that MIA and MEA designs, while minimizing maximum risk, do not generally minimize average risk. In proposing MIA/MEA we have taken the view we believe to be most widely applicable to general statistical practice: that experimenter goals are focused squarely on the experiment at hand, so that choice of design mitigating effects of loss *for this experiment* is the most relevant implementation of robustness. Should other implementations be preferred, the concepts of this paper still provide a framework for design choice. Average efficiency is certainly a reasonable alternative should risk over many experiments be the focus.

Importantly, the MEA approach is restricted neither to BIBDs nor to loss of blocks. It is suitable for robustness evaluation of any collection of competing block designs $\mathcal{D}$ having common v, b, and k. Should (unlike for BIBDs) values $\Phi(C_d)$ not be constant for $d \in \mathcal{D}$, simply modify definition 1 to $\Phi_{(t)}$-$\mathit{eff} = \frac{\max_{d\in\mathcal{D}} \Phi(C_d)}{\Phi(C_{d(t)})}$ and proceed as in definition 6. If concern is with potential loss of t units rather than t blocks, then redefining $C_{d(t)}$ of (1) accordingly, the vector $\phi_{(t)}(d)$ will contain $\binom{bk}{t}$ values, and MEA again offers a rational basis for choice of robust design. Capacity to withstand loss of experimental material is a pragmatic standard for selecting among competing, optimal or highly efficient designs, for which the ideas introduced in this paper offer a rigorous schema.

ACKNOWLEDGEMENT

J. P. Morgan was supported by National Science Foundation grants DMS01-04195 and 06-04997.

REFERENCES

Bagchi, B. and Bagchi, S. (2001). Optimality of partial geometric designs. *Ann. Statist.* **29**, 577–594.

Baksalary, J. K. and Tabis, Z. (1987). Conditions for the robustness of block designs against the unavailability of data. *J. Statist. Plann. Inf.* **16**, 49–54.

Bhatia, R. (1997). *Matrix Analysis.* Springer-Verlag, New York.

Bhaumik, D. K. and Whittinghill, D. C. (1991). Optimality and robustness to the unavailability of blocks in block designs. *J. Roy. Statist. Soc. B* **53**, 399–407.

Das, A. and Kageyama, S. (1992). Robustness of BIB and extended BIB designs against the unavailability of any number of observations in a block. *Comput. Statist. Data Anal.* **14**, 343–358.

Dey, A. (1993). Robustness of block designs against missing data. *Statist. Sinica* **3**, 219–231.

Houtman, A. M. and Speed, T. P. (1983). Balance in designed experiments with orthogonal block structure. *Ann. Statist.* **11**, 1069–1085.

Ghosh, S. (1982). Robustness of BIBD against the unavailability of data. *J. Statist. Plann. Inf.* **6**, 29–32.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models*, pp. 333-53. Ed. J. N. Srivastava, North Holland: Amsterdam.

Lal, K., Gupta, V. K., and Bhar, L. (2001). Robustness of designed experiments against missing data. *J. Appl. Stat.* **28**, 63–79.

Mathon, R. and Rosa, A. (2006). 2-(v, k, λ) designs of small order. In *The CRC Handbook of Combinatorial Designs, 2nd ed.*, pp. 25–58. Ed. C. J. Colbourn and J. H. Dinitz, CRC Press: Boca Raton.

Mukhopadhyay, S. (1981). On the optimality of block designs under mixed effects model. *Calcutta Statist. Assoc. Bull.* **30**, 171–185.

Parvu, V. (2004). *Optimal blocking for three treatments and BIBD robustness: two problems in design optimality.* Ph.D. Dissertation, Virginia Polytechnic Institute & State University, Blacksburg.

Rao, V. A. (1958). A note on balanced designs. *Ann. Math. Statist.* **29**, 290–294.

Soicher, L. H. (2003). The DESIGN package for GAP. *designtheory.org/software/gap_design/*

Shah, K. R. & Sinha, B. K. (1989). *Theory of Optimal Designs*. Springer-Verlag: Berlin.

Wu, C. F. & Hamada, M. (2000). *Experiments*. John Wiley & Sons: New York.

Table 1: Eigenvalues for residuals of BIBDs under loss of two blocks

Case 1: $q = 0$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k}$	$v - 2k + 1$
$\frac{v\lambda}{k} - 1$	$2(k-1)$

Case 2: $0 < q < k$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k}$	$v - 2k + q$
$\frac{v\lambda}{k} - \frac{q}{k}$	1
$\frac{v\lambda}{k} - 1$	$2(k-q-1)$
$\frac{v\lambda}{k} - 2 + \frac{q}{k}$	1
$\frac{v\lambda}{k} - 2$	$q - 1$

Case 3: q=k

Eigenvalue	Multiplicity
$\frac{v\lambda}{k}$	$v - k$
$\frac{v\lambda}{k} - 2$	$k - 1$

Table 2: Variances of pairwise treatment contrasts when two blocks are removed from a BIBD

Case	$Var[\widehat{\tau_i - \tau_{i'}}]/\sigma^2$	Count
i and i' occur in both blocks	$V_{1q} = \frac{2k}{v\lambda-2k}$	$\binom{q}{2}$
i occurs in both blocks, i' occurs in one block	V_{2q}	$2q(k-q)$
i and i' both occur once, in the same block	$V_{3q} = \frac{2k}{v\lambda-k}$	$2\binom{k-q}{2}$
i occurs in both blocks, i' does not occur in either block	V_{4q}	$q(v-2k+q)$
i and i' both occur once, in different blocks	$V_{5q} = \frac{2k(v\lambda-q-1)}{(v\lambda-k)(v\lambda-q)}$	$(k-q)^2$
i occurs in one block, i' does not occur in either block	V_{6q}	$2(k-q)(v-2k+q)$
i and i' do not occur in the removed blocks	$V_{7q} = \frac{2k}{v\lambda}$	$\binom{v-2k+q}{2}$

$$V_{2q} = \frac{k(q-1)}{q(v\lambda-2k)} + \frac{k(k-q-1)}{(k-q)(v\lambda-k)} + \frac{k}{2(k-q)(v\lambda-q)} + \frac{k(2k-q)}{2q(k-q)(v\lambda-2k+q)}$$
$$V_{4q} = \frac{k(2k-q+1)}{v\lambda(2k-q)} + \frac{k(q-1)}{q(v\lambda-2k)} + \frac{2k(k-q)}{q(2k-q)(v\lambda-2k+q)}$$
$$V_{6q} = \frac{k(2k-q+1)}{v\lambda(2k-q)} + \frac{k(k-q-1)}{(k-q)(v\lambda-k)} + \frac{k}{2(k-q)(v\lambda-q)} + \frac{kq}{2(k-q)(2k-q)(v\lambda-2k+q)}$$

Table 3: Eigenvalues for residuals of SBIBDs under loss of three blocks

Case 1 : $q \notin \{0, \lambda, 2\lambda - k, v - 3k + 3\lambda\}$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 3$	$q - 1$
$\frac{v\lambda}{k} - 2$	$3\lambda - 3q - 3$
$\frac{v\lambda}{k} - 1$	$3k - 6\lambda + 3q - 3$
$\frac{v\lambda}{k}$	$v - 3k + 3\lambda - q$
$\frac{v\lambda}{k} - 2 - \frac{k-2\lambda+\sqrt{(k+2\lambda)^2-8kq}}{2k}$	1
$\frac{v\lambda}{k} - 2 - \frac{k-2\lambda-\sqrt{(k+2\lambda)^2-8kq}}{2k}$	1
$\frac{v\lambda}{k} - 1 - \frac{\lambda+\sqrt{(2k-\lambda)^2-4k(\lambda-q)}}{2k}$	2
$\frac{v\lambda}{k} - 1 - \frac{\lambda-\sqrt{(2k-\lambda)^2-4k(\lambda-q)}}{2k}$	2

Case 2: $q = 0,\ k \neq 2\lambda$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 2$	$3\lambda - 3$
$\frac{v\lambda}{k} - 1$	$3k - 6\lambda - 3$
$\frac{v\lambda}{k}$	$v - 3k + 3\lambda$
$\frac{v\lambda}{k} - 2 + \frac{2\lambda}{k}$	1
$\frac{v\lambda}{k} - 1 - \frac{\lambda+\sqrt{(2k-\lambda)^2-4k\lambda}}{2k}$	2
$\frac{v\lambda}{k} - 1 - \frac{\lambda-\sqrt{(2k-\lambda)^2-4k\lambda}}{2k}$	2

Case 3: $q = 0,\ k = 2\lambda$

Eigenvalue	Multiplicity
$\frac{v-4}{2}$	$3\lambda - 3$
$\frac{v-3}{2}$	2
$\frac{v}{2}$	$v - 3\lambda$

Case 4: $q = \lambda, v \neq 3k - 2\lambda$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 3$	$\lambda - 1$
$\frac{v\lambda}{k} - 3 + \frac{2\lambda}{k}$	1
$\frac{v\lambda}{k} - 1$	$3k - 3\lambda - 3$
$\frac{v\lambda}{k} - \frac{\lambda}{k}$	2
$\frac{v\lambda}{k}$	$v - 3k + 2\lambda$

Case 5: $q = \lambda,\ v = 3k - 2\lambda$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 3$	$\lambda - 1$
$\frac{v\lambda}{k} - 3 + \frac{2\lambda}{k}$	1
$\frac{v\lambda}{k} - 1$	$v - \lambda - 3$
$\frac{v\lambda}{k} - \frac{\lambda}{k}$	2

Case 6: $q = 2\lambda - k > 0,\ v \neq 2k - \lambda$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 3$	$2\lambda - k - 1$
$\frac{v\lambda}{k} - 4 + \frac{2\lambda}{k}$	1
$\frac{v\lambda}{k} - 2$	$3k - 3\lambda - 3$
$\frac{v\lambda}{k} - 1 - \frac{\lambda}{k}$	2
$\frac{v\lambda}{k}$	$v - 2k + \lambda$

Case 7: $q = 2\lambda - k > 0,\ v = 2k - \lambda$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 3$	$2\lambda - k - 1$
$\frac{v\lambda}{k} - 4 + \frac{2\lambda}{k}$	1
$\frac{v\lambda}{k} - 2$	$3k - 3\lambda - 3$
$\frac{v\lambda}{k} - 1 - \frac{\lambda}{k}$	2

Case 8: $0 < q = v - 3k + 3\lambda < \lambda,\ v \neq 2k - \lambda$

Eigenvalue	Multiplicity
$\frac{v\lambda}{k} - 3$	$v - 3k + 3\lambda - 1$
$\frac{v\lambda}{k} - 2$	$9k - 3v - 6\lambda - 3$
$\frac{v\lambda}{k} - 1$	$3v - 6k + 3\lambda - 3$
$\frac{v\lambda}{k} - 2 - \frac{k-2\lambda+\sqrt{(5k-2\lambda)^2-8kv}}{2k}$	1
$\frac{v\lambda}{k} - 2 - \frac{k-2\lambda-\sqrt{(5k-2\lambda)^2-8kv}}{2k}$	1
$\frac{v\lambda}{k} - 1 - \frac{\lambda+\sqrt{-8k^2+4kv+4k\lambda+\lambda^2}}{2k}$	2
$\frac{v\lambda}{k} - 1 - \frac{\lambda-\sqrt{-8k^2+4kv+4k\lambda+\lambda^2}}{2k}$	2

Table 4: Comparison of $(7, 14, 3)$ designs for robustness to loss of 2 blocks

Design	Int. numbers				A-efficiency		E-efficiency		MV-efficiency	
no.	0	1	2	3	average	min	average	min	average	min
1	0	84	0	7	0.833134	0.8	0.637363	0.571429	0.69172	0.571429
2	4	72	12	3	0.833073	0.8	0.637363	0.571429	0.683955	0.571429
3	6	66	18	1	0.833042	0.8	0.637363	0.571429	0.680072	0.571429
4	7	63	21	0	0.833027	0.820046	0.637363	0.571429	0.678131	0.571429

Table 5: Worst case comparisons of $(7, 14, 3)$ designs for loss of 2 blocks

Maximum		Impact relative to MIA design		
q	Designs	A-impact	E-impact	MV-impact
2	4 (MIA)	1.0	1.0	1.0
3	1-3	0.933333	0.8	0.816667

Table 6: Comparison of $(8, 14, 4)$ designs for robustness to loss of 2 blocks

Design	Int. numbers					A-efficiency		E-efficiency		MV-efficiency	
no.	0	1	2	3	4	average	min	average	min	average	min
1	3	12	72	4	0	0.841905	0.832737	0.677656	0.666667	0.684874	0.666667
2	1	18	66	6	0	0.841908	0.832737	0.67674	0.666667	0.687568	0.666667
3	0	21	63	7	0	0.84191	0.832737	0.676282	0.666667	0.688915	0.666667
4	7	0	84	0	0	0.841897	0.840917	0.679487	0.666667	0.679487	0.666667

Table 7: Worst case comparisons of $(8, 14, 4)$ designs for loss of 2 blocks

Maximum		Impact relative to MIA design		
q	Designs	A-impact	E-impact	MV-impact
2	4 (MIA)	1.0	1.0	1.0
3	1-3	0.983332	0.888889	0.874242

Table 8: MIA-BIBDs with $v \leq 15$, $r \leq 15$

v	b	k	r	λ	#noniso	#MIA	design
6	20	3	10	4	4	1	$(\infty,0,1),(\infty,0,2),(0,1,2),(0,1,3)$ (mod 5)
6	30	3	15	6	6	1	$(\infty,0,1),(\infty,0,1),(\infty,0,2),(0,1,2),(0,1,3),(0,1,3)$ (mod 5)
7	14	3	6	2	4	1	$(0,1,3),(0,2,3)$ (mod 7)
7	21	3	9	3	10	1	$(0,1,2),(0,2,4),(0,3,4)$ (mod 7)
7	28	3	12	4	35	1	$(0,1,2),(0,1,3),(0,2,4),(0,3,4)$ (mod 7)
7	35	3	15	5	109	1	$(0,1,2),(0,1,3),(0,2,3),(0,2,4),(0,3,4)$ (mod 7)
8	14	4	7	3	4	1	$(\infty,0,1,3),(0,2,3,4)$ in (mod 7)
8	28	4	14	6	2310	1	$(\infty,0,1,3),(\infty,0,2,3),(0,2,3,4),(0,2,5,6)$ (mod 7)
9	24	3	8	2	36	13	$(\infty,0,4),(0,1,3),(0,2,3)$ (mod 8)
9	36	3	12	3	22521	332	$(0,1,2),(0,2,4),(0,3,6),(0,4,8)$ (mod 9)
9	18	4	8	3	11	1	$(0,1,4,6),(0,1,2,4)$ (mod 9)
10	30	3	9	2	960	394	$(1,5,\infty),(1,2,3),(1,3,6)$ (mod 9), $(1,4,7)$ (partial cycle)
10	15	4	6	2	3	3	$(1_1,2_1,4_1,2_2),(1_1,2_1,1_2,3_2),(1_1,3_2,4_2,5_2)$ (mod 5)
10	30	4	12	4	$> 1.7 \times 10^6$	1	$(1,2,3,7),(1,2,4,5),(1,3,5,8)$ (mod 10)
10	18	5	9	4	21	2	$(1,2,5,7,\infty),(0,1,2,4,8)$ (mod 9)
10	15	6	9	5	3	3	$(1_1,2_1,2_2,3_2,2_3,3_3)$, $(1_1,1_2,3_2,1_3,3_3,\infty)$, $(1_1,2_1,2_2,1_3,3_3,\infty)$ $(1_1,2_1,1_2,3_2,2_3,\infty)$ (mod 3), $(1_1,2_1,3_1,1_2,2_2,3_2)$, $(1_1,2_1,3_1,1_3,2_3,3_3)$, $(1_2,2_2,3_2,1_3,2_3,3_3)$ (fixed)
11	55	3	15	3	≥ 436800	≥ 6	$(1,2,4),(1,2,5),(1,2,8),(1,3,6),(1,3,8)$ (mod 11)
11	33	5	15	6	≥ 127	?	$(1,2,3,5,6),(1,2,3,7,9),(1,2,4,6,9)$ (mod 11)
12	44	3	11	2	$\geq 10^6$	≥ 1	$(0,1,\infty),(0,1,4),(0,2,5),(0,2,6)$ (mod 11)
12	33	4	11	3	≥ 17172470	≥ 2	$(1,3,7,\infty),(1,2,3,6),(1,2,4,9)$ (mod 11)
12	22	6	11	5	11603	1	$(1,2,3,5,8,\infty)$, $(1,2,3,7,9,10)$ (mod 11)
13	26	3	6	1	2	2	$(1_1,2_1,1_2),(1_1,2_2,1_3),(1_1,2_3,2_4),(1_1,3_3,1_4),(1_1,3_4,\infty),(1_2,1_3,3_4)$ $(1_2,2_3,\infty),(1_2,1_4,2_4)$ (mod 3), $(1_2,2_2,3_2),(1_3,2_3,3_3)$ (fixed)
13	52	3	12	2	≥ 92714	≥ 7	$(1,2,5),(1,2,11),(1,3,8),(1,3,9)$ (mod 13)
13	39	4	12	3	≥ 3702	≥ 5	$(1,2,4,10),(1,2,5,7),(1,3,4,8)$ (mod 13)
13	39	5	15	5	≥ 30	?	$(1,2,3,7,10),(1,2,4,6,7),(1,2,4,8,12)$ (mod 13)
13	26	6	12	5	≥ 2572156	≥ 1	$(1,2,3,4,7,11)$, $(1,2,4,6,8,9)$ (mod 13)
14	26	7	13	6	≥ 17896	?	$(1,2,3,4,6,0,\infty)$, $(1,2,3,5,8,10,11)$ (mod 13)
15	70	3	14	2	$\geq 685{,}521$	≥ 1	$(1,4,\infty),(1,2,3),(1,3,9),(1,5,8),(1,5,10)$ (mod 14)
15	35	3	7	1	80	80	$(1_1,2_1,1_2),(1_2,2_2,1_3),(1_1,2_3,2_4),(1_1,3_3,2_5),(1_1,1_4,3_5),(1_1,3_4,1_5)$, $(1_2,1_3,2_4),(1_2,2_3,1_4),(1_2,3_4,3_5),(1_2,1_5,2_5),(1_3,2_3,2_5)$ (mod 3), $(1_2,2_2,3_2),(1_4,2_4,3_4)$ (fixed)
15	42	5	14	4	≥ 207	≥ 1	$(1_1,2_1,4_1,1_2,\infty),(1_1,3_2,4_2,6_2,\infty),(1_1,2_1,3_1,3_2,5_2)$, $(1_1,2_1,5_1,6_2,7_2),(1_1,1_2,2_2,5_2,7_2),(1_1,3_1,5_1,1_2,4_2)$ (mod 7)
15	35	6	14	5	≥ 117	?	$(1_1,2_1,3_1,5_1,4_2,\infty),(1_1,1_2,4_2,5_2,6_2,\infty),(1_1,2_1,4_1,1_2,2_2,6_2)$ $(1_1,2_1,6_1,2_2,4_2,5_2),(1_1,2_1,6_1,1_2,3_2,7_2)$ (mod 7)
15	15	7	7	3	5	2	$(1_1,2_1,4_1,1_2,2_2,6_2,\infty),(1_1,2_1,3_1,6_1,2_2,4_2,5_2)$ (mod 7), $(1_2,2_2,3_2,4_2,5_2,6_2,7_2)$ (fixed)
15	30	7	14	6	≥ 34	?	$(1,2,3,4,6,11,12),(1,2,4,5,8,12,14)$ (mod 15)