Optimal regular graph designs

Sera Aylin Cakiroglu

Abstract

A typical problem in optimal design theory is finding an experimental design that is optimal with respect to some criteria in a class of designs. The most popular criteria include the $A$- and $D$-criteria. Regular graph designs occur in many optimality results, and if the number of blocks is large enough, an $A$-optimal (or $D$-optimal) design is among them (if any exist). To explore the landscape of designs with a large number of blocks, we introduce extensions of regular graph designs. These are constructed by adding the blocks of a balanced incomplete block design repeatedly to the original design. We present the results of an exact computer search for the best regular graph designs and the best extended regular graph designs with up to 20 treatments $v$, block size $k \leq 10$ and replication $r \leq 9$.

Keywords $A$-optimality · $D$-optimality · Incomplete block design · Regular graphs · Regular graph design

1 Introduction

Suppose we are to design the following statistical experiment: there are $v$ treatments to be compared on a number of experimental units that can be partitioned into $b$ blocks of size $k$ with $k < v$. Typically, the blocks might differ systematically but all units in a block are assumed to be alike. An example may be a sensory experiment to quantify the perceptions of a product by letting a panel of judges score several products by their preference (for example by making a mark on a line). When the judges are confronted with too many products, sensory fatigue can occur and may produce biases. In taste testing, this is sometimes called 'palate fatigue' or 'palate paralysis' and it can be avoided when only a subset (four to five) products are presented to each panellist. If the judges are not available to return for further sessions to evaluate all of the products, the products represent the treatments and the judges are incomplete blocks of treatments.

For fixed $v$, $b$ and $k$, how should the treatments be allocated to the units to get as much information as possible from the available data? This often means the estimate of the unknown parameters with the least possible variance. If there are several parameters, this is a multidimensional problem and the design of the experiment can be ‘good’ in different ways. We will give the basic definitions in the following, and the reader is referred to Bailey and Cameron (2009), where a good overview and more details on the application of combinatorics can be found. Formally, a block design $d$ is an assignment of $v$ treatments or varieties to a set of experimental units that have been partitioned into $b$ blocks of size $k$. The statistical model to analyse the data obtained from $d$ is assumed to be a linear model. Let $f(\omega m)$ be the function that specifies which treatment is allocated to unit $\omega$ in block $m$. For each yield $Y_{\omega m}$ of unit $\omega$ in block $m$, the model is

$$Y_{\omega m} = \tau f(\omega m) + \beta_m + \epsilon_{\omega m},$$

where $\beta_m$ is the effect of the block $m$, $\tau_i$ is the effect of treatment $i$, and $\epsilon_{\omega m}$ are uncorrelated random errors with expectation 0 and variance $\sigma^2$.

Electronic supplementary material

The online version of this article (https://doi.org/10.1007/s11222-016-9720-8) contains supplementary material, which is available to authorized users.

Sera Aylin Cakiroglu
aylin.cakiroglu@crick.ac.uk

1 School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

2 Present Address: Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK
The design is said to be connected if all pairwise differences \( \tau_i - \tau_j \) are estimable. If every treatment occurs at most once per block, the design is called binary. In the following, we always assume the designs to be binary and connected and further that \( k < v \). The replication \( r_i \) of a treatment \( i \) is the total number of units that have been assigned treatment \( i \). If the replications are all equal to a constant \( r \), the design is called equireplicate.

Let \( N_d \) be the \( v \times b \) treatment-block incidence matrix of the design \( d \), i.e. the \( im \)-entry of \( N_d \) is the number of units in block \( m \) that have been assigned treatment \( i \). The \( ij \)-entry of the product \( N_d N_d^T \) will be denoted by \( \lambda_{ij} \) and is called the concurrence of treatments \( i \) and \( j \). Note that for binary designs \( \lambda_{ii} = r_i \) and \( \lambda_{ij} \) is the number of blocks containing both treatments \( i \) and \( j \). The information matrix for the estimation of the treatment effects is

\[
C_d = \text{diag}(r_1, \ldots, r_v) - \frac{1}{k} N_d N_d^T.
\]

Since \( C_d \) has row sums 0, the all-one vector is an eigenvector with eigenvalue 0. All other eigenvalues are called the non-trivial eigenvalues. If the design is connected, the rank of the information matrix is \( v - 1 \), the eigenvalue 0 has multiplicity 1, and all non-trivial eigenvalues are strictly positive. We will use another matrix associated with the design: the Laplacian matrix of a binary connected design, which is

\[
L_d = k \text{diag}(r_1, \ldots, r_v) - N_d N_d^T = kC_d.
\]

The smallest eigenvalue of \( L_d \) is 0, and if the design is connected, the other \( v - 1 \) non-trivial Laplacian eigenvalues are all strictly positive.

The average variance of the set of the best linear unbiased estimators of the pairwise differences of the treatment effects is proportional to the reciprocal of the harmonic mean of the non-trivial eigenvalues of \( C_d \). The volume of the confidence ellipsoid for the estimate of \( (\tau_1, \ldots, \tau_v) \) (in the hyperplane \( \sum \tau_i = 0 \)) is proportional to the reciprocal of the product of the non-trivial eigenvalues of \( C_d \) (for example Shah and Sinha 1989). These facts give us different criteria against which a design can be considered ‘good’: a design maximizing the harmonic mean of the non-trivial eigenvalues of \( C_d \) is called \( A \)-optimal, and a design maximizing the product of the non-trivial eigenvalues of \( C_d \) is called \( D \)-optimal. There are many more optimality criteria; another popular example is the \( E \)-criterion which is the maximization of the smallest non-trivial eigenvalue of \( C_d \) and is equivalent with minimizing the largest variance of the estimators (see Shah and Sinha 1989).

There is no general answer to the question of which design is to be chosen for given \( (v, b, k) \), but there are several partial results. A famous class of block designs are the balanced incomplete block designs (BIBDs); these are binary equireplicate incomplete block designs with replications \( bk/v \), where \( k < v \) and any pair of treatments is contained in exactly \( \lambda \) blocks for some \( \lambda > 0 \). BIBDs with these parameters are also called \( 2-(v, k, \lambda) \)-designs. Kiefer (1975) proved that BIBDs are optimal with regard to a wide range of criteria, in particular the \( A \)- and \( D \)-criteria. However, it is not clear which designs to choose if no BIBD exists. Regular graph designs (RGDs) are another class of designs that have been suggested to be good candidates. RGDs are equireplicate binary designs that are ‘close’ to BIBDs in the sense that any pair of treatments occurs in either \( \lambda \) or \( \lambda + 1 \) blocks for some integer \( \lambda \geq 0 \). The Laplacian matrix of an RGD \( d \) with \( v \) treatments, replication \( r \) and block size \( k \) can be written as

\[
L_d = \{r(k - 1) + \lambda\} I_v - T_d - \lambda J_v,
\]

where \( I_v \) is the \( v \times v \) identity matrix, \( J_v \) denotes the \( v \times v \) all-one matrix, and \( T_d \) is a symmetric \( v \times v (0, 1) \)-matrix with 0’s on the diagonal and exactly \( r(k - 1) - \lambda(v - 1) \) number of 1’s in each row and each column. This means the Laplacian matrix of any RGD is fully determined by the parameters \( v, k, r \) and the matrix \( T_d \). Note that if \( \lambda = 0 \), the fact that the design is connected is equivalent to the matrix \( T_d \) being irreducible (i.e. the matrix cannot be transformed into a block upper-triangular matrix by row or column permutations).

The reference to graphs in their name is due to a simple equivalence. A graph \( G \) is given by a set of \( v \) vertices with edges that connect vertices \( i \neq j \). The \( v \times v \) matrix, whose \( ij \)-entry is the number of edges joining vertices \( i \) and \( j \), is the adjacency matrix of the graph \( G \). If \( G \) contains no multiple edges, we say that \( G \) is simple. A graph is regular if its adjacency matrix has constant row and column sums, called the degree. If any vertex of the graph can be reached from any other vertex by going along edges, then the graph is called connected. In this case, its adjacency matrix is irreducible. All regular simple graphs of degree \( \delta \) with \( v \) vertices correspond to all symmetric \( v \times v \)-matrices with \((0, 1)\)-entries, zero diagonal and row and column sum \( \delta \). The matrix \( T_d \) is precisely the adjacency matrix of a simple regular graph of degree \( r(k - 1) - \lambda(v - 1) \). If \( \lambda = 0 \), the design and the graph being connected are equivalent.

John and Mitchell (1977) conjectured that if an incomplete block design is \( D \)-optimal (or \( A \)-optimal or \( E \)-optimal), then it is an RGD (if any RGDs exist). In the same paper, they provided a list of the best RGDs with \( v \leq 12 \), \( r \leq 10 \) and \( v \leq b \) which they found using numerical methods. Jones and Eccleston (1980) found counterexamples for \( A \)-optimality for \( k = 2, b = v \in \{10, 11, 12\} \) for this conjecture (they were dismissed as of academic but not of practical interest) and Constantine (1986) disproved the conjecture for \( E \)-optimality. In all of their 209 cases, the matrices \( T_d \) of the \( A \)- and \( D \)-best RGDs are the same and the matrix is only
different in 14 cases for the $E$-criterion. This observation inspired the conjecture that among RGDs the same design is $A$-optimal and $D$-optimal (John and Williams 1982). These conjectures seemed to be in accordance with existing results for good designs for the sizes needed in practice. However, computers are growing more powerful, which has made searches for optimal designs without the restriction to RGDs possible. Kerr and Churchill (2001) and Wit et al. (2005) reported the results of a computer search for $A$-optimal designs (not restricted to RGDs) with $k = 2$, $b = v$ supporting the results of Jones and Eccleston (1980); for $v \geq 9$ the $A$-optimal designs (which are not RGDs) differ from the $D$-optimal designs but are consistent with the results in Kerr and Churchill (2001) and theoretical results for these cases followed (Bailey 2007). Morgan (2007) provided a near-complete solution for the $E$-criterion and a downloadable catalogue of $E$-optimal designs for up to 15 treatments.

However, the conjecture by John and Mitchell (1977) holds if the number of blocks is large enough (Cheng 1992). In this context, we were interested in: (i) repeating and expanding their computer search with exact methods to confirm that those RGDs are indeed $A$- and $D$-best and (ii) investigating the effect of adding the blocks of a BIBD repeatedly to an RGD on the performance on the $A$- and $D$-criteria. In the spirit of Morgan (2007), we call these designs BIBD-extended RGDs. We can write the Laplacian matrix of the BIBD-extended RGD in terms of the RGD and the BIBD as follows: suppose, the RGD has Laplacian matrix $L_d$ and $\tilde{d}$ is a $2-(v, k, \lambda)$-design on $v$ treatments and block size $k$ with $\tilde{b}$ blocks and Laplacian matrix $L_{\tilde{d}}$. Then for $y \in \mathbb{N}$ the matrix $L_d + yL_{\tilde{d}}$ is the Laplacian matrix of an RGD on $v$ treatments, replication $r + y\lambda(v - 1)/(k - 1)$ and $b + y\hat{b}$ blocks of size $k$. We can write $L_d + yL_{\tilde{d}}$ with $\lambda = [r(k - 1)/(v - 1)]$ and $\delta = r(k - 1) - \lambda(v - 1)$ as

$$L_d + yL_{\tilde{d}} = (\delta + yv\lambda + v\lambda)I_v - (y\lambda + \lambda)J_v - T_d.$$  

Since any differences between RGDs (and their BIBD extensions) is captured in the matrix $T_d$, we take a similar approach to John and Mitchell (1977) (as described in John and Mitchell 1976) and use the matrices $T_d$ to identify the $A$- and $D$-best RGDs. We use exact methods to compute the $A$- and $D$-values as polynomials in $y$, and this allows us also to report the $A$- and $D$-best BIBD-extended RGDs for large $y$.

We found the direct approach of computing the Laplacian eigenvalues of $L_d + yL_{\tilde{d}}$ and then computing the $A$- and $D$-value as their harmonic mean and product led to long computation times even in small cases. Instead, we will show in Sect. 2 how we can represent the $A$- and $D$-values for BIBD-extended RGDs to make their computation with exact methods feasible. Section 3 describes the method we used to identify the best RGDs and BIBD-extended RGDs. Section 4 discusses some aspects of the catalogue, including the description of a counterexample to the above-mentioned conjecture in John and Williams (1982), and the computation. The summary discussion is found in Sect. 5, and the catalogue of all best RGDs that are also the best BIBD-extended RGDs for large $y$ can be found in the appendix.

2 A representation of the $A$-value of an BIBD-extended RGD

Different expressions have been found for the $A$- and $D$-value. For example, the $D$-value is given as the product of the non-trivial eigenvalues of the Laplacian matrix which can be computed as its cofactor (Kirchhoff 1847) exactly and efficiently (Wolfram Research 2012). The goal of this section is to find an expression that makes the computation of the $A$-value of an BIBD-extended RGD just as computationally tractable.

Let $x = \lambda + y\tilde{\lambda}$. We can write the Laplacian matrix $L_d + yL_{\tilde{d}}$ of the BIBD-extended RGD as

$$L(x, T_d) = (\delta + vx)I_v - T_d - xJ_v.$$  

For the matrix $L(x, T_d)$ to be the Laplacian matrix of an existing design, only some values for $x$ will be admissible depending on $\lambda$ and $\tilde{\lambda}$. However, this is not relevant for the ordering of the matrices according to their performance on $A$- and $D$-value; in fact we can even allow non-integral values for $x$. For the ease of the notation, we will write $x$ as a variable and not as a function of $\lambda$, $\tilde{\lambda}$ and $y$.

Let $\psi_1^{T_d} \geq \psi_2^{T_d} \geq \ldots \geq \psi_v^{T_d}$ denote the eigenvalues of $T_d = J_v$. Then, $\psi_v^{T_d} = 0$ since $T_d$ has row sum $\delta$. The non-trivial eigenvalues of $L(x, T_d)$ are $vx + \psi_1^{T_d}, \ldots, vx + \psi_v^{T_d} - 1$. We can write the $A$- and $D$-value as functions that only depend on $x$ and $T_d$ as

$$A(x, T_d) := \frac{v - 1}{\sum_{i=1}^{v-1} \frac{1}{vx + \psi_i^{T_d}}}$$  

and

$$D(x, T_d) := \prod_{i=1}^{v-1} (vx + \psi_i^{T_d}),$$  

respectively.

For $\psi^{T_d} = (\psi_1^{T_d}, \ldots, \psi_v^{T_d} - 1)$ and $I = \{1, \ldots, v - 1\}$, let

$$S_j(\psi^{T_d}) := \sum_{\mathcal{J} \subseteq I} \prod_{i \in \mathcal{J}} \psi_i^{T_d},$$  

and

$$S_{j,j}(\psi^{T_d}) := \sum_{\mathcal{J} \subseteq I \setminus \{j\}} \prod_{i \in \mathcal{J}} \psi_i^{T_d}.$$

\@ Springer
With the relation \( \sum_{i=1}^{v-1} S_{j,i}(\psi T_d) = (v - 1 - j)S_j(\psi T_d) \) (for example, Beckenbach and Bellman 1965, p. 34) we can express the \( A \)-value as a rational function whose coefficients are completely determined by the coefficients of the \( D \)-value:

\[
A(x, T_d) = \frac{(v - 1)D(x, T_d)}{\sum_{i=1}^{v-1} \prod_{j=1}^{v-1}(v x + \psi_i T_d)} = \frac{(v - 1)D(x, T_d)}{\sum_{i=1}^{v-1} \sum_{j=0}^{v-2}(v x)^{v-2-j}S_j(\psi T_d)} = \frac{v(v - 1)D(x, T_d)}{D_x(x, T_d)}, \tag{3}
\]

where \( D_x(x, T_d) = \frac{d}{dx} D(x, T_d) \) denotes the derivative of \( D(x, T_d) \) in \( x \). It is therefore sufficient to compute the coefficients of \( D(x, T_d) \). The Laplacian matrix \( L_d \) of the RGD is \( L_d = [\lambda, T_d] \), and \( D(\lambda, T_d) \) and \( A(\lambda, T_d) \) are the product and harmonic mean of its non-trivial eigenvalues, respectively.

3 Determining the \( A \)- and \( D \)-best RGDs and BIBD-extended RGDs

We will denote the set of all symmetric irreducible \( v \times v \) (0, 1)-matrices with 0’s on the diagonal and row and column sum \( \delta \) by \( \mathcal{M}(v, \delta) \). We can compute all matrices in \( \mathcal{M}(v, \delta) \) by generating the adjacency lists of all connected \( \delta \)-regular graphs on \( v \) vertices with the program genreg (Meringer 1999) and then the matrices \( L[x, T_d] \) as given in Eq. 2.

We ordered the matrices in decreasing order of \( A(x, T) \) for \( x = 0, 1, 2, 3, ... \) until we found a value \( x^A_0 \) where the order stayed the same for \( x^A_0 + 1 \). For all \( i = 1, \ldots, |\mathcal{M}(v, \delta)| - 1 \) we computed the exact roots (using built-in algorithms in Wolfram Research 2012) of the polynomials

\[
P_i(x) = D(x, T_i)D_x(x, T_{i+1}) - D(x, T_{i+1})D_x(x, T_i)
\]

and found that \( P_i \) never had any roots bigger than \( x^A_0 \), that is if \( P_i(x) = 0 \), then \( x < x^A_0 \). This means that this order of the matrices according to the \( A \)-value is valid for all \( x \geq x^A_0 \).

An analogous procedure is applied for finding the orders of the matrices in \( \mathcal{M}(v, \delta) \) with regard to the \( D \)-value. Table 1 shows the values of \( x^A_0 \) and \( x^D_0 \) for different \( v \) and \( \delta \).

| \( v \) | \( \delta \) | \( x^A_0 \) | \( x^D_0 \) |
|------|------|------|------|
| 5    | 4    | 0    | 0    |
| 6    | 3    | 0    | 0    |
| 4    | 0    | 0    | 0    |
| 7    | 4    | 0    | 0    |
| 6    | 0    | 0    | 0    |
| 8    | 3    | 0    | 0    |
| 5    | 0    | 0    | 0    |
| 6    | 0    | 0    | 0    |
| 9    | 4    | 1    | 1    |
| 6    | 0    | 0    | 0    |
| 10   | 3    | 1    | 1    |
| 4    | 1    | 1    | 1    |
| 5    | 1    | 1    | 1    |
| 6    | 1    | 1    | 1    |
| 7    | 0    | 0    | 0    |
| 8    | 0    | 0    | 0    |
| 11   | 4    | 1    | 1    |
| 11   | 6    | 1    | 1    |
| 8    | 0    | 0    | 0    |
| 12   | 3    | 1    | 1    |
| 4    | 2    | 2    | 2    |
| 5    | 3    | 3    | 3    |
| 6    | 3    | 2    | 2    |
| 7    | 2    | 2    | 2    |
| 8    | 1    | 1    | 1    |
| 9    | 0    | 0    | 0    |
| 13   | 4    | 3    | 3    |
| 6    | 4    | 3    | 3    |
| 8    | 5    | 3    | 3    |
| 14   | 3    | 1    | 1    |
| 4    | 4    | 3    | 3    |
| 5    | 6    | 5    | 5    |
| 15   | 4    | 5    | 4    |
| 16   | 3    | 1    | 1    |
| 18   | 3    | 2    | 1    |
| 20   | 3    | 2    | 2    |

The input to GAP DESIGN is \( \lambda \) and the concurrence structure given by the matrix; given sufficient time the package will either find a design or conclude that no such design exists. Unfortunately, the performance decreases with growing \( v \), \( b \) or \( k \) and in some cases either the number of matrices in \( \mathcal{M}(v, \delta) \) or the computational complexity to decide whether there exists a design with given matrix in GAP was too high to perform a complete search in a reasonable time. In these cases we had to restrict the search to the first 100–20,000 best matrices (depending on the computational complexity);
Table 2 The A-best RGD (y = 0) for v = 14, r = 5, k = 2

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

The blocks are given as pairs of treatments in the rows of the columns.

this may explain why we failed to find any designs for some choices of r and k.

4 The design catalogue

The A- and D-best RGDs (y = 0) also produce A- and D-best BIBD-extended RGDs for large y, except in the case of v = 14, r = 5, k = 2; here, only the D-best RGD is the A- and D-best BIBD-extended RGD. These designs have been catalogued in downloadable form as XML files in external representation format (Soicher 2006) as an electronic appendix. For the exceptional case v = 14, r = 5, k = 2, we report separately the A-best RGD in Table 2.

For v ≤ 12 and δ ≤ 5, the best RGDs are isomorphic to the designs presented by John and Mitchell except for the cases v = 11, k = r = 3 and k = r = 8. This is particularly remarkable since we made the additional restriction on all matrices to be irreducible. In the cases v = 11, k = r = 3 and k = r = 8 we found an A- and D-best RGD that John and Mitchell failed to find. The A- and D-best RGDs correspond in both cases to the same matrix in \( M(11, 6) \).

The A-best and D-best matrices for x = 0 were the same and remained A- and D-best with growing x for v < 14. Interestingly, we found that the asymptotic orders (the order corresponding to the A-values evaluated at \( x = x_0^A \) and the order corresponding to the D-values evaluated at \( x = x_0^D \) on the matrices are exactly the same. That means that asymptotically not only one design dominates all others, but that they can be ordered so that any design beats all successive designs on both the A- and D-value. We found in all cases that \( x_0^A \cdot x_0^D \leq \delta + 1 \) indicating that when we say ‘asymptotically’, we might not mean very large values for y.

We also report an counterexample to the conjecture that among RGDs, an A-optimal (D-optimal) RGD is also D-optimal (A-optimal) in John and Williams (1982): for v = 14, r = 5, k = 2 the A-best RGD is the design \( d_A \) described in Table 2. The A- and D-best BIBD-extended RGD for y > 0 and D-best RGD for y = 0 are #12 in the appendix; we denote it by \( d_0 \) here. To compare the performances of the designs on the A- and D-value, we use the fraction of the values: \( A(0, T_{d_A})/A(0, T_{d_D}) \) = 1.000004 that means \( d_D \) performs worse on the A-value than \( d_A \), and \( D(0, T_{d_A})/D(0, T_{d_D}) \) = 0.999904 that means \( d_D \) performs better on the D-value than \( d_A \). However, this changes when the blocks of a BIBD are added to the designs. For the A-value to reach the asymptotic order we need to add \( x_0^A = 6 \) times the blocks of a BIBD and for the D-value \( x_0^D = 5 \) times: \( A(6, T_{d_A})/A(6, T_{d_D}) = 0.999999998 \) and \( D(5, T_{d_A})/D(5, T_{d_D}) = 0.9999999986 \), showing that the BIBD-extended \( d_D \) performs better on both the D- and the A-value than the BIBD-extended \( d_A \).

Let \( K_{m, \ldots, m} \) denote the regular complete multipartite graph, that is a graph whose vertex set can be partitioned into groups of size \( m \) such that any pair of vertices is joined by an edge if and only if they are in different groups. An RGD with a multipartite concurrence graph and block size 2 is a group divisible design. Cheng (1981) proved that regular complete bipartite graphs give rise to the concurrence graphs of the unique A- and D-optimal designs for all \( y \geq 0 \) (not necessarily only among RGDs). He extended his result to complete regular multipartite graphs for \( y = 0 \). We found that in the cases listed, the RGDs corresponding to \( K_{2,2,2,2} \), \( K_{2,2,2,2,2} \), \( K_{2,2,2,2,2,2} \), \( K_{3,3,3} \) and \( K_{3,3,3,3} \) are also the A- and D-best BIBD-extended RGDs for all \( y \geq 0 \).

5 Discussion

The online catalogue provides design plans for experiments with high replication. Examples may include sensory evaluation trials or preference studies in consumer and food research: when the judges cannot assess all products, it is recommended to use BIBDs to design the study (for example Gacula et al. 2009; Amerine and Roessler 1979). If the products are presented in a randomized order to each panelist, and no judge scores a product more than once and evaluates only one block of samples, and there are no carryover effects, then we can assume a linear model as given in Eq. 1 and the RGDs in the online catalogue are valid A- and D-efficient designs. The BIBD-extended RGDs provide highly A- and D-efficient (if not optimal) design plans for sensory trials when the number of blocks (i.e. judges) is large enough to surpass the bounds in Table 1; this is the case in many studies where more than 100 judges are used (for example Lockshin et al. 2011 used 420 consumers). RGDs have been recommended in sensory trials but not explicitly (Best et al. 2011). Instead, partially balanced incomplete
block designs (PBIBDs) have been suggested (Cochran and Cox 1957; Best et al. 2011); these are a natural extension of BIBDs and include some RGDs (but not all), and the reader is referred to Bose and Nair (1939) and Bose and Shimamoto (1952) for an introduction. PBIBDs are not always optimal and often inferior to RGDs (Cheng 1978); in fact many of the existing optimality results on PBIBDs apply only to those PBIBDs that are also RGDs (for example Constantine 1982; Cheng and Bailey 1991; Cheng 1978; Takeuchi 1961).

A difficulty in using RGDs in practice is that most existing catalogues are paper bound and spread over different journals. One major achievement of this paper is the online catalogue of the $A$- and $D$-best RGDs (except for the case $v = 14$, $r = 5$, $k = 2$) in the electronic appendix that extends the current catalogues of designs to up to 20 treatments and fills in the two cases for $v = 11$ that are missing in John and Mitchell (1977). In the exceptional case $v = 14$, $r = 5$, $k = 2$ we include the $D$-best RGD in the catalogue and report the $A$-best RGD separately. Because of the exact computation all of the catalogued designs are indeed $A$- or $D$-best RGDs and $A$- and $D$-best BIBD-extended RGDs for large $y$. Our results therefore confirm that all the designs listed in John and Mitchell (1977) are indeed the best RGDs. We add to the previous computer searches (Jones and Eccleston 1980; Kerr and Churchill 2001; Wit et al. 2005) by not restricting to cases with $b = v$ and by presenting, to our knowledge, the first counterexample to the conjecture in John and Williams (1982) that among RGDs, an $A$-optimal ($D$-optimal) RGD is also $D$-optimal ($A$-optimal). However, in this counterexample no blocks of a BIBD have been added ($y = 0$) and we did not find a counterexample for BIBD-extended RGDs.

A practical advantage of BIBD-extended RGDs is that if one can afford to add enough blocks of a BIBD, our list provides highly $A$- and $D$-efficient designs and no decision as to which criterion to use is necessary. This is important because one major weakness of these optimality criteria is that, when taken alone, they reduce a high-dimensional problem into a single number and fail to capture the richness of the information available. Furthermore, the designs in the catalogue can be extended with BIBDs while staying $A$- and $D$-best. This means that additional blocks can be added to the experiment to gain power, even if their becoming available was unknown in the design phase of the study.

We have only considered $A$- and $D$-optimality since Morgan (2007) published, together with the design construction, a catalogue of binary connected $E$-optimal designs up to 15 treatments without the restriction to RGDs which can be found on www.designtheory.org.

The matrices used for our computer search were generated with the program genreg (Meringer 1999) which is very efficient but restricted to irreducible matrices. However, all optimal designs in John and Mitchell (1977) with $δ ≥ 2$ correspond to irreducible matrices, and it seems like a reasonable compromise to be able to handle cases with a large number of matrices. We like to note however that RGDs corresponding to irreducible matrices might not perform better than the ones corresponding to reducible matrices, but we failed to find an example.

To order all possible matrices $T$ according to the performance of the corresponding $L_d + yL_{d'}$ on the $A$- and $D$-criteria for different values of $y$, we guessed the smallest $x_0^A$ or $x_0^D$ and verified that the $A$- and $D$-values of any successive pair of matrices in this order do not intersect. There is to our knowledge no proven result implying that these values for $x_0^A$ or $x_0^D$ exist, but interestingly we did find this to be true in all the cases we list.

This article extends the computational results of John and Mitchell (1977) to explore the landscape of asymptotically optimal designs in the results in Cheng (1992). The online catalogue improves the accessibility for statistical application and will be valuable for the practical design of experiments.

Acknowledgements The author would like to thank Peter J. Cameron and R. A. Bailey for the many conversations on optimal designs and is extremely grateful to I. P. Morgan for bringing the topic to her attention and the helpful discussions about RGDs and the arising computational difficulties. Joachim Kunert has been an invaluable source of information for preference designs. Parts of the computations were performed using the facilities at the London Research Institute, Cancer Research UK.

Appendix 1: Catalogue of $A$- and $D$-best RGDs

There follows a table of the $A$- and $D$-best RGDs (except for the case $v = 14$, $r = 5$, $k = 2$, where only the $D$-best RGD is included). These RGDs also produce optimal BIBD-extended RGDs for large $y$. We list $δ = r(k − 1) − λ(v − 1)$ and the smallest $δ$ such that a $2-(v, k, λ)$-design exists. There exists only one irreducible matrix for $δ = 2$ or $δ = v − 1$, and we therefore exclude these cases.

Most designs can be found in either Clatworthy (1973) or John et al. (1972). To follow the convention as in John and Mitchell (1977), we write in the reference column $P.XY$ for a design in Clatworthy (1973) with reference number $XY$ and $C.XY$ for a design in John et al. (1972) with reference number $XY$. If the design is not in either catalogues but can be found in John and Mitchell (1977) we give the reference number as $JM.XY$, and if it is cyclic, we give the initial blocks. If the design is the complement of a design in one of the catalogue, we add an (C) to the reference number. All other designs can be found in the appendix. However, some of these designs are possible to construct with known methods, such as the ones in John (1967). A full catalogue of all designs as XML files in external representation format is included in the electronic appendix. For more information on external representation see Soicher (2006).
### Appendix 2: Optimal regular graph designs

There follows a list of all designs not found in the used reference catalogues. Blocks are represented as \(k\)-tuples of treatments in the rows.

**#1** \(v = 7, k = 2, r = 4 + 6y\)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
3 & 6 & 7 & 10
\end{array}
\]

**#2** \(v = 7, k = 5, r = 10 + 15y\)

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
3 & 6 & 7 & 10
\end{array}
\]
**References**

Amerine, M.A., Roessler, E.B.: Wines: Their Sensory Evaluation. W. H. Freeman, San Francisco (1979)

Bailey, R.A., Cameron, P.J.: Combinatorics of optimal designs. In: Surveys in Combinatorics 2009, London Mathematical Society Lecture Notes, vol 365, pp 19–73. Cambridge University Press, Cambridge (2009)

Bailey, R.A.: Designs for two-colour microarray experiments. Appl. Stat. 56(4), 365–394 (2007)

Beckenbach, E.F., Bellman, R.: Inequalities, Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Berlin (1965)

Best, D.J., Rayner, J.C., Allingham, D.: A statistical test for ranking data from partially balanced incomplete block designs. J. Sensory Stud. 26, 81–84 (2011)

Bose, R.C., Nair, K.R.: Partially balanced incomplete block designs. Sankhya 4, 337–372 (1939)

Bose, R.C., Shimamoto, T.: Classification and analysis of partially balanced incomplete block designs with two associate classes. J. Am. Stat. Soc. 47, 151–184 (1952)

Cheng, C.S.: Optimality of certain asymmetrical experimental designs. Annal Stat. 6, 1239–1261 (1978)
Cheng, C.S.: Maximizing the total number of spanning trees in a graph: Two related problems in graph theory and optimum design theory. J. Comb. Theory 31, 240–248 (1981)
Cheng, C.S.: On the optimality of (M.S)-optimal designs in large systems. Sankhya 54, 117–125 (1992)
Cheng, C.S., Bailey, R.A.: Optimality of some two-associate-class partially balanced incomplete-block designs. Annal Stat. 19(3), 1667–1671 (1991)
Clatworthy, W.H.: Tables of two-associate-class partially balanced designs, NBS Applied Mathematics Series, vol 63. The United States Department of Commerce Publications, National Bureau of Standards (U.S.) (1973)
Cochran, W.G., Cox, G.: Experimental Designs. John Wiley and Sons, New York, NY (1957)
Constantine, G.M.: On the E-optimality of PBIB designs with a small number of blocks. Annal Stat. 10, 1027–1031 (1982)
Constantine, G.M.: On the optimality of block designs. Ann. Inst. Stat. Math. 38, 161–174 (1986)
Gacula, M.C., Singh, J., Bi, J., Altan, S.: Statistical Methods in Food and Consumer Research. Academic Press, New York, NY (2009)
John, J.A., Mitchell, T.J.: Optimal Incomplete Block Designs. ORNL/CSD-8 Available from the National Technical Information Service, The United States Department of Commerce, 5285 Port Royal Road, Springfield, VA (1976)
John, J.A., Wolock, F.W., David, H.A.: Cyclic designs, NBS Applied Mathematics Series, vol 62. The United States Department of Commerce Publications, National Bureau of Standards (U.S.) (1972)
John, J.A.: Reduced group divisible paired comparison designs. Ann. Math. Stat. 38, 1887–1893 (1967)
John, J.A., Mitchell, T.J.: Optimal incomplete block designs. J. R. Stat. Soc. 39B, 39–43 (1977)
John, J.A., Williams, E.R.: Conjectures for optimal block designs. J. R. Stat. Soc. 44B, 221–225 (1982)
Jones, B., Eccleston, J.A.: Exchange and interchange procedures to search for optimal designs. J. R. Stat. Soc. 42, 238–243 (1980)
Kerr, M.K., Churchill, G.A.: Experimental design for gene expression microarrays. Biostatistics 2, 183–201 (2001)
Kiefer, J.: Optimality and Construction of Generalized Youden Designs. A Survey of Statistical Designs and Linear Models, pp. 333–354. North-Holland, Amsterdam (1975)
Krechhoff, G.: Über die Auflösung der Gleichung, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. Ann. Phys. Chem. 72, 497–508 (1847)
Lockshin, L., Mueller, S., Louviere, J., Francis, L., Osidacz, P.: Development of a new method to measure how consumers choose wine. Aust. N. Z. Wine Ind. J. 24, 81–84 (2011)
Meringer, M.: Fast generation of regular graphs and construction of cages. J. Graph Theory 30, 137–146 (1999)
Morgan, J.P.: Optimal incomplete block designs. J. Am. Stat. Assoc. 102, 655–663 (2007)
Shah, K.R., Sinha, B.K.: Theory of Optimal Designs, Lecture Notes in Statistics, vol 54. Springer, Berlin (1989)
Soicher, L.H.: The DESIGN Package for GAP, Version 1.4. http://designtheory.org/software_design/ (2006)
Takeuchi, K.: On the optimality of certain type of PBIB designs. Rep. Stat. Appl. Res. Union Jpn. Sci. Eng. 8, 140–145 (1961)
Wit, E., Nobile, A., Khanin, R.: Near-optimal designs for dual channel microarray studies. Appl. Stat. 54, 817–830 (2005)
Wolfram Research, I.: Mathematica (2012)