Biometrika (2019), **00**, 0, *pp.* 1–12
Printed in Great Britain
doi: 10.1093/biomet/asz025

Design of order-of-addition experiments

BY JIAYU PENG

Department of Statistics, The Pennsylvania State University, 331B Thomas Building, University Park, Pennsylvania 16802, U.S.A.
jup250@psu.edu

RAHUL MUKERJEE

Indian Institute of Management Calcutta, Joka, Diamond Harbour Road, Kolkata 700104, India
rmuk0902@gmail.com

AND DENNIS K. J. LIN

Department of Statistics, The Pennsylvania State University, 317 Thomas Building, University Park, Pennsylvania 16802, U.S.A.
dkl5@psu.edu

SUMMARY

In an order-of-addition experiment, each treatment is a permutation of m components. It is often unaffordable to test all the $m!$ possible treatments, and thus the design problem arises. We consider a flexible model that incorporates the order of each pair of components and can also account for the distance between the two components in every such pair. Under this model, the optimality of the uniform design measure is established, via the approximate theory, for a broad range of criteria. Coupled with an eigenanalysis, this result serves as a benchmark that paves the way for assessing the efficiency and robustness of any exact design. The closed-form construction of a class of robust optimal fractional designs that can also facilitate model selection is explored and illustrated.

Some key words: Approximate theory; Association algebra; Optimality; Pairwise order; Robustness; Signed permutation; Tapered model.

1. INTRODUCTION

This paper deals with a problem where the output of a process depends on the order of adding m different components into the system, and interest is in understanding this dependence. Order-of-addition experiments have wide applications in chemistry and related areas. For example, Ryberg (2008) studied a reaction in which the order of adding the reagents had a strong effect on the performance of the reaction, and a series of experiments was conducted to evaluate different orders of addition. More order-of-addition chemical experiments have been described by Fuleki & Francis (1968), Shinohara & Ogawa (1998), Jiang & Ng (2014), Song et al. (2014) and Ding et al. (2015), among others.

It is often unaffordable to test all the $m!$ possible orders, and the design problem is then to choose a subset of orders for comparison. The naïve design which randomly chooses orders is often adopted in practice. Naturally, an optimal design, with a solid theoretical foundation, will

be preferable to this naïve design. However, the statistical literature on this topic is limited. The work of van Nostrand (1995) appears to be the first statistical reference on the order-of-addition design; he suggested considering a design to detect the pairwise order effects that are often of great interest, with practitioners wishing to know, for example, whether adding component 1 before component 2 or vice versa significantly influences the response. Voelkel (2017) studied design criteria and design construction under a pairwise order model defined in § 2. A key feature of his work is the idea of extending orthogonal arrays to designs that are naturally restricted, of which the order-of-addition orthogonal array is an important example. He showed that such an array leads to the same value of the D-criterion as the full design, but he did not report optimality results.

We aim to initiate the systematic development of an optimal design theory for order-of-addition experiments. This is done under a model which is more flexible than the pairwise order model and includes it as a special case. In particular, our results provide firm justification for the work of Voelkel (2017) from the perspective of optimality and robustness. The tools that we employ include approximate theory via the use of signed permutation matrices and an eigenanalysis motivated by association algebra. All proofs appear in the Appendix.

2. MODEL FORMULATION

Suppose there are $m \geqslant 3$ components $1, \ldots, m$, which can be ordered in $m!$ ways. Any such ordering, say $a = a_1 \ldots a_m$, which is a permutation of $1, \ldots, m$, is a treatment. As in van Nostrand (1995) and Voelkel (2017), our model incorporates the order of every pair of components in any treatment. A new feature of the model is that it allows for possible tapering, as may happen in practice, of the impact of any such pairwise order with an increase in the distance between the components in the pair. Thus, the effect of component i preceding component j when they are at the two extremes of a treatment can be smaller than when they are adjacent.

To present the model formally, we write A for the set of the $m!$ treatments. For $a \in A$, let $\tau(a)$ be the treatment mean of a, that is, the expectation of the response from a. It is assumed that the responses have equal variance and are uncorrelated. Let S be the set of all pairs ij $(1 \leqslant i < j \leqslant m)$. For $a = a_1 \ldots a_m \in A$ and $ij \in S$, write $h(ij, a)$ for the distance between i and j in a; that is, if $a_k = i$ and $a_l = j$, then $h(ij, a) = |k - l|$, so that $h(ij, a) \in \{1, \ldots, m-1\}$. Then, according to our model,

$$\tau(a) = \beta_0 + \sum\nolimits_{ij} z_{ij}(a)\beta_{ij} \quad (a \in A), \tag{1}$$

where $\sum_{ij}$ denotes summation over $ij \in S$, the β_{ij} and β_0 are unknown parameters and, for each $ij \in S$,

$$z_{ij}(a) = \begin{cases} c_{h(ij,a)}, & i \text{ precedes } j \text{ in } a, \\ -c_{h(ij,a)}, & j \text{ precedes } i \text{ in } a. \end{cases} \tag{2}$$

Thus, if $m = 3$, then $\tau(312) = \beta_0 + c_1\beta_{12} - c_1\beta_{13} - c_2\beta_{23}$. In (2), the c_h $(h = 1, \ldots, m-1)$ are known quantities such that $1 = c_1 \geqslant \cdots \geqslant c_{m-1} \geqslant 0$. These are intended to capture the possible tapering as indicated above. For example, one can take $c_h = 1/h$ or $c_h = c^{h-1}$, with known c such that $0 < c < 1$. On the other hand, if no such tapering is anticipated, then one could also take $c_h = 1$ for all h, in which case (1) and (2) reduce to the pairwise order model of van Nostrand (1995) and Voelkel (2017). As an example of a practical situation which is

well represented by model (1) with an appropriate choice of the c_h in (2), we refer to the order-of-addition experiment of Jiang & Ng (2014), who used the full design for $m = 4$. The usual pairwise order model, given by $c_1 = c_2 = c_3 = 1$, conforms well with their data on fluorescence outputs at 445 nm, accounting for over 90% of the total variation. Moreover, if this dataset is truncated to the 12-run optimal fractional design in § 5, then the same model accounts for over 93% of the total variation. Indeed, our findings indicate a high degree of robustness of the optimal designs with respect to the specific choice of the c_h. This facilitates model selection, as will be seen in § 5.

Let $\tilde{\beta} = (\beta_{12}, \beta_{13}, \ldots, \beta_{m-1\,m})^{\mathrm{T}}$, where the superscript T denotes transpose. Then $\beta = (\beta_0, \tilde{\beta}^{\mathrm{T}})^{\mathrm{T}}$ represents the parameter vector of interest. Similarly, for any $a \in A$, let $z(a) = [z_{12}(a), \ldots, z_{m-1\,m}(a)]^{\mathrm{T}}$ and $x(a) = [1, z(a)^{\mathrm{T}}]^{\mathrm{T}}$. Then (1) can be expressed as

$$\tau(a) = \beta_0 + z(a)^{\mathrm{T}}\tilde{\beta} = x(a)^{\mathrm{T}}\beta. \tag{3}$$

Write $q = m(m-1)/2$ and $p = q + 1$. Then $\tilde{\beta}$ and $z(a)$ are $q \times 1$, while β and $x(a)$ are $p \times 1$. Also,

$$x(a)^{\mathrm{T}}x(a) = 1 + (m-1)c_1^2 + \cdots + c_{m-1}^2 \tag{4}$$

for each a, because by (2), $m-h$ of the $z_{ij}(a)$ $(ij \in S)$ have absolute value c_h, for $h = 1, \ldots, m-1$.

3. Optimality of the uniform design measure

Let d_0 be the full design which replicates each treatment once. Theorem 1 establishes the optimality of d_0 for inference on β under a broad range of criteria, among all designs with the same number, $m!$, of runs. This result holds irrespective of the c_h in (2). Although d_0 is impractical for large m, Theorem 1 provides a useful benchmark for assessing smaller designs.

One may anticipate the optimality of d_0 because, with run size $m!$, a design that replicates some treatments more than once while omitting others altogether is intuitively unappealing. However, unlike in similar situations such as traditional full factorial experiments, the moment matrix of d_0, obtained in § 4, is rather involved. For instance, it need not be completely symmetric (Kiefer, 1975) in the sense of having all diagonal elements equal and all off-diagonal elements equal. The c_h in (2) further complicate matters. As a result, a direct combinatorial proof of the optimality of d_0 is quite challenging. An approach based on approximate theory is found to yield a subtle noncomputational proof that does not require explicit evaluation of the moment matrix of d_0.

To motivate the ideas, consider an N-run exact design d, where any treatment $a \in A$ is replicated $r(a)$ times and the integers $r(a) \geqslant 0$ sum to N. By (3), d has per-run moment matrix

$$M(w) = \sum\nolimits_a w(a)x(a)x(a)^{\mathrm{T}}, \tag{5}$$

where $w(a) = r(a)/N$ and $\sum_a$ denotes summation over $a \in A$. In approximate theory, the requirement that the design weights $w(a)$ be integer multiples of $1/N$ is relaxed, and they are allowed to be any nonnegative quantities, subject to $\sum_a w(a) = 1$. Then $w = \{w(a) : a \in A\}$ is called a design measure, having moment matrix $M(w)$. In particular, the full design d_0 corresponds to the uniform design measure w_0 over A, which has moment matrix

$$M_0 = M(w_0) = (1/m!)\sum\nolimits_a x(a)x(a)^{\mathrm{T}}. \tag{6}$$

Let $\mathscr{M}$ denote the class of $p \times p$ nonnegative-definite matrices. Recall that a signed permutation matrix is a square matrix having exactly one nonzero entry in each row and column, where each nonzero entry is either 1 or -1. We consider optimality criteria $\phi(\cdot)$ that are concave over $\mathscr{M}$ and signed permutation invariant, that is, $\phi(R^{\mathrm{T}}MR) = \phi(M)$ for every signed permutation matrix R of order p and every $M \in \mathscr{M}$. Given any such $\phi(\cdot)$, a design measure is said to be ϕ-optimal if it maximizes $\phi\{M(w)\}$ among all design measures. The commonly used D-, A- and E-criteria correspond to $\phi(M) = \log\det(M)$, $-\operatorname{tr}(M^{-1})$ and $\lambda_{\min}(M)$, respectively, where $\log\det(M)$ and $-\operatorname{tr}(M^{-1})$ are interpreted as $-\infty$ for singular M, and $\lambda_{\min}$ stands for the smallest eigenvalue. As all design measures here have the same $\operatorname{tr}\{M(w)\}$ by (4) and (5), our framework also covers the MS-optimality criterion of Eccleston & Hedayat (1974), with $\phi(M) = -\operatorname{tr}(M^2)$. This is equivalent to the $E(s^2)$ criterion for two-level supersaturated factorial designs (Booth & Cox, 1962) when $c_h = 1$ for all h, in which case each $x(a)$ has elements ± 1. All these criteria are concave and signed permutation invariant.

Theorem 1 establishes the optimality of the uniform design measure or, equivalently, that of the full design. Its proof exploits the concavity and signed permutation invariance of the optimality criteria along with the fact, proved in the Appendix, that the action of any permutation of $1, \ldots, m$ on the set of treatments induces an action of a signed permutation matrix on $\{x(a) : a \in A\}$.

THEOREM 1. *The uniform design measure w_0 is ϕ-optimal for every optimality criterion $\phi(\cdot)$ which is concave and signed permutation invariant.*

Theorem 1 has several important implications.

(i) It shows the D-, A-, E- and MS-optimality of the uniform design measure w_0. By the equivalence theorem (Kiefer & Wolfowitz, 1960), D-optimality of w_0 also implies its G-optimality; that is, for every design measure w,

$$\max\{x(a)^{\mathrm{T}}M^{-1}(w)x(a) : a \in A\} \geqslant p = \max\{x(a)^{\mathrm{T}}M_0^{-1}x(a) : a \in A\}.$$

Thus, w_0 minimizes the maximum variance of the estimated responses at $a \in A$.

(ii) In view of (i), the full design d_0 is D-, A-, E-, MS- and G-optimal among all designs having $m!$ runs. An exact design, with a smaller number of runs but having the same moment matrix M_0 as d_0, also enjoys these optimality properties. More generally, such a design is ϕ-optimal for every $\phi(\cdot)$ as in Theorem 1. The matrix M_0 will be examined in more detail in § 4 with a view to assessing the efficiencies of any given exact design under the aforementioned criteria.

(iii) While w_0 may not be the unique design measure having the optimality properties mentioned in (i), any other design measure which is D-, A- or MS-optimal must have the same moment matrix M_0 as w_0. This is because these three criteria are strictly concave.

(iv) In particular, for the usual pairwise order model given by $c_h = 1$ for all h, the order-of-addition orthogonal arrays of Voelkel (2017) have, by definition, moment matrix M_0, and therefore, as noted in (ii) above, these are indeed optimal in the sense of Theorem 1. Conversely, in this situation, an exact design that corresponds to a D-, A- or MS-optimal design measure must have moment matrix M_0, in view of (iii) above. Hence it is not hard to see that such a design has to be an order-of-addition orthogonal array. This settles an issue left open by Voelkel (2017).

4. EIGENANALYSIS AND EFFICIENCY ASSESSMENT

Although the proof of Theorem 1 does not require explicit knowledge of M_0, we need to find M_0 and its eigenvalues to assess the efficiencies of a given design measure or a given exact design

under various optimality criteria. Write I for the identity matrix of order q, and define V as the $q \times q$ matrix with rows and columns indexed by the elements of S such that for $ij, kl \in S$, the (ij, kl)th element of V is

$$V(ij,kl) = \begin{cases} 1, & i = k, j \neq l \text{ or } i \neq k, j = l, \\ -1, & i = l \text{ or } j = k, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

For instance, if $m = 4$, then $S = \{12, 13, 14, 23, 24, 34\}$ and

$$V = \begin{pmatrix} 0 & 1 & 1 & -1 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ -1 & 1 & 0 & 0 & 1 & -1 \\ -1 & 0 & 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & -1 & 1 & 0 \end{pmatrix},$$

where $V(12, 12) = V(12, 34) = 0$, $V(13, 23) = 1$, $V(23, 12) = -1$, and so on. Let

$$b_0 = 2\{(m-1)c_1^2 + \cdots + c_{m-1}^2\}/\{m(m-1)\}, \tag{8}$$

$$b_1 = 2\sum\nolimits_h \{m - h(1) - h(2)\}c_{h(1)}\{2c_{h(1)+h(2)} - c_{h(2)}\}/\{m(m-1)(m-2)\}, \tag{9}$$

where $\sum_h$ denotes summation over all positive integers $h(1)$ and $h(2)$ such that $h(1) + h(2) \leqslant m - 1$. In general, (8) and (9) do not permit further simplification, although when $c_h = 1$ for all h, they reduce to

$$b_0 = 1, \quad b_1 = 1/3. \tag{10}$$

THEOREM 2. *The moment matrix of the uniform design measure w_0 is $M_0 = \mathrm{diag}(1, b_0 I + b_1 V)$, having eigenvalues 1, $b_0 + (m-2)b_1$ and $b_0 - 2b_1$ with multiplicities 1, $m-1$ and $(m-1)(m-2)/2$, respectively.*

If every -1 in V were 1 instead, V would equal an association matrix of the triangular association scheme (Raghavarao, 1971, Ch. 8), and the related association algebra could be useful in studying the eigenvalues of M_0 in Theorem 2. Somewhat in the same spirit, the proof in the Appendix obtains and uses an expression for V^2 as a linear combination of I and V.

Theorem 2 enables us to assess the efficiencies of any design measure w under various criteria, relative to the uniform design measure w_0 seen to be optimal in Theorem 1. In particular, the D- and A-efficiencies of w, defined as $[\det\{M(w)\}/\det(M_0)]^{1/p}$ and $\mathrm{tr}(M_0^{-1})/\mathrm{tr}\{M^{-1}(w)\}$, are

$$\text{D-eff}(w) = \left[\frac{\det\{M(w)\}}{\{b_0 + (m-2)b_1\}^{m-1}(b_0 - 2b_1)^{(m-1)(m-2)/2}}\right]^{1/p}, \tag{11}$$

$$\text{A-eff}(w) = \frac{1 + (m-1)\big[\{b_0 + (m-2)b_1\}^{-1} + \{(m-2)/2\}(b_0 - 2b_1)^{-1}\big]}{\text{tr}\{M^{-1}(w)\}}. \tag{12}$$

The design measure w in (11) and (12) can correspond to an exact design.

Equations (11) and (12) are pivotal in efficiency calculations for a given choice of the c_h in (2), as well as for studying robustness across such choices. For illustration, we study the performance of the order-of-addition orthogonal arrays of Voelkel (2017) under tapered models given by (i) $c_h = 1/h$ and (ii) $c_h = (1/2)^{h-1}$ $(h = 1, \ldots, m-1)$. In § 3 these arrays were found to be optimal under the usual pairwise order model. By (11) and (12), for $m = 5$, the 12-run array in Table 3 of Voelkel (2017) has D- and A-efficiencies of 0.985 and 0.972 under (i), and of 0.989 and 0.980 under (ii). Similarly, for $m = 6$, the 24-run arrays in Table 7 of Voelkel (2017) are also quite robust, the best being the leftmost one in that table, having D- and A-efficiencies of 0.992 and 0.984 under (i) and of 0.994 and 0.988 under (ii). Indeed, even these figures can be conservative, as under (i) or (ii) there may not exist any exact design of the same run size as these arrays that has moment matrix M_0.

5. OPTIMAL FRACTIONAL DESIGNS

The literature on order-of-addition design lacks a systematic procedure for the construction of optimal fractional designs. To this end, we now propose a method that yields optimal fractions much smaller than the full design d_0 and has potential for further improvement, as discussed in § 6. Our construction is based on the fact, noted in § 3, that an N-run exact design d having the same moment matrix M_0 as d_0 is optimal among all N-run designs, for every criterion $\phi(\cdot)$ as in Theorem 1. Throughout this section, we work under the usual pairwise order model given by $c_h = 1$ for all h. Then, by (10) and Theorem 2,

$$M_0 = \text{diag}\{1, I + (1/3)V\}. \tag{13}$$

The resulting designs are found to be very efficient under tapered models as well. We first illustrate the structure of the proposed designs, and then present the general construction procedure.

Let $m = 4$. Consider the 12-run half fractional design d, given in transposed form by

$$\begin{array}{cccccccccccc} 1 & 2 & 4 & 3 & 1 & 3 & 4 & 2 & 1 & 4 & 3 & 2 \\ 2 & 1 & 3 & 4 & 3 & 1 & 2 & 4 & 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 & 2 & 4 & 1 & 3 & 2 & 3 & 1 & 4 \\ 4 & 3 & 2 & 1 & 4 & 2 & 3 & 1 & 3 & 2 & 4 & 1 \end{array}$$

This can be obtained from the first design in Table 2 of Voelkel (2017) by relabelling the components and row permutation. The design d has moment matrix M_0 and hence enjoys the optimality properties of d_0. To see the structure of d, let

$$B_1 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \bar{B}_1 = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}, \quad \bar{B}_2 = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix},$$
$$B_3 = \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix}, \quad \bar{B}_3 = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}.$$

Then d can be expressed as $[D_1^{\mathrm{T}}\ D_2^{\mathrm{T}}\ D_3^{\mathrm{T}}]^{\mathrm{T}}$ where

$$D_u = \begin{pmatrix} B_u & \bar{B}_u \\ \sim\bar{B}_u & B_u \end{pmatrix} \quad (u = 1, 2, 3), \tag{14}$$

with $\sim$ representing the operator of column reversal of any matrix; for example, $\sim\bar{B}_1$ has first column $(4, 3)^{\mathrm{T}}$ and second column $(3, 4)^{\mathrm{T}}$.

We now show how the above design structure allows an extension to the case of general even $m \geqslant 4$. Let $s = m/2$, $L = m!/\{2(s!s!)\}$ and $\Gamma = \{1, \ldots, m\}$. Consider lexicographically arranged distinct sets $C_1, \ldots, C_L$, where each C_u consists of 1 and $s-1$ other elements of Γ. For each u, let $\bar{C}_u$ be the complement of C_u in Γ. In both C_u and $\bar{C}_u$, the elements are arranged in ascending order. For $u = 1, \ldots, L$, let B_u be the $s! \times s$ array with rows formed by all permutations of the elements of C_u, and define $\bar{B}_u$ similarly with reference to $\bar{C}_u$. Thus, if $m = 4$, then $C_1 = \{1, 2\}$, $\bar{C}_1 = \{3, 4\}$, $C_2 = \{1, 3\}$, $\bar{C}_2 = \{2, 4\}$, $C_3 = \{1, 4\}$ and $\bar{C}_3 = \{2, 3\}$, entailing B_u and $\bar{B}_u$ $(u = 1, 2, 3)$ as shown above. Similarly, if $m = 6$, then, for example, $C_5 = \{1, 3, 4\}$ and $\bar{C}_5 = \{2, 5, 6\}$, and hence

$$B_5 = \begin{pmatrix} 1 & 1 & 3 & 3 & 4 & 4 \\ 3 & 4 & 1 & 4 & 1 & 3 \\ 4 & 3 & 4 & 1 & 3 & 1 \end{pmatrix}^{\mathrm{T}}, \quad \bar{B}_5 = \begin{pmatrix} 2 & 2 & 5 & 5 & 6 & 6 \\ 5 & 6 & 2 & 6 & 2 & 5 \\ 6 & 5 & 6 & 2 & 5 & 2 \end{pmatrix}^{\mathrm{T}}.$$

For $u = 1, \ldots, L$, now define D_u as in (14). Along the lines of the 12-run design shown above for $m = 4$, let d^* be the design consisting of the $m!/s!$ treatments given by the rows of $D = (D_1^{\mathrm{T}}\ \ldots\ D_L^{\mathrm{T}})^{\mathrm{T}}$.

THEOREM 3. *For every even $m \geqslant 4$, the design d^* has moment matrix M_0 and hence is ϕ-optimal among designs with the same number of runs, for every optimality criterion $\phi(\cdot)$ which is concave and signed permutation invariant.*

The above construction readily yields ϕ-optimal designs for odd $m = 2s + 1$ $(s \geqslant 2)$ as well. To get such a design in $(2s + 1)!/s!$ runs, one has to stack $2s + 1$ copies of D. Then it suffices to insert a column consisting only of $2s+1$ just before the lth column of the lth copy $(l = 1, \ldots, 2s)$ and also after the last column of the last copy.

In order to explore the performance of the above fractional designs under tapered models, we again consider (i) $c_h = 1/h$ and (ii) $c_h = (1/2)^{h-1}$ $(h = 1, \ldots, m-1)$. Quite reassuringly, the D- and A-efficiencies of our fractional designs for $m = 4, \ldots, 10$, calculated from (11) and (12), turn out to be over 0.99 under both (i) and (ii). For the reason indicated at the end of § 4, even these figures may be conservative. Moreover, the same pattern is seen for other choices of the c_h. This robustness facilitates the use of our fractional designs in model selection, allowing a wide range of models to choose from. One does not need to prespecify the c_h in (2), but rather can make a post-experimentation data-driven choice of these quantities with the assurance that the design will remain highly efficient under the model so reached. For illustration, let $m = 4$ and suppose that the design d^* in Theorem 3 leads to observations

$$8.27, 3.09, 10.21, 7.26, 0.76, 12.15, 3.90, 5.09, 11.95, 1.62, 4.03, 7.77,$$

which correspond to treatments in the same order as in d^* that are generated by adding perturbations, each uniform over $[-1, 1]$, to a random permutation of $1, \ldots, 12$. Standard crossvalidation techniques may be used to compare various choices of the c_h in (2). For instance, one could employ leave-one-out crossvalidation and find the ratio of the predictive residual sum of squares for any choice of the c_h to that for a model involving only the general mean. A choice of the c_h

may be considered satisfactory if this ratio is sufficiently small. With observations as above, this ratio is found to be only 0.083 for $c_h = 1/h$ ($h = 1, 2, 3$). Thus, a model given by this choice of the c_h in (2) is reasonable and one remains assured of high efficiency of our fractional design under the model so reached. Indeed, as indicated in the next section, our designs remain highly efficient when the model incorporates some three-way orderings in addition to the pairwise terms, and thus they allow even more flexibility in model selection, if necessary.

The optimal design d^* in Theorem 3 permits a natural blocking, with treatments arising from each D_u constituting one block. In model (1) for the treatment mean $\tau(a)$, β_0 then has to be replaced by a block effect parameter depending on the block where any treatment a in d^* appears, and we focus on $\tilde{\beta} = (\beta_{12}, \ldots, \beta_{m-1\,m})^{\mathrm{T}}$. The proof of Lemma A4(a) in the Appendix shows that the blocking of d^* as envisaged above is a case of orthogonal blocking under the usual pairwise order model. Therefore, under this model, the resulting block design remains ϕ-optimal for inference on $\tilde{\beta}$, for every monotone criterion $\phi(\cdot)$ which is concave and signed permutation invariant, such as the D- and A-criteria. Moreover, for m up to 10, the D- and A-efficiencies of these blocked fractional designs, calculated from appropriate versions of (11) and (12), are again found to be over 0.99, under both of the tapered models given by (i) and (ii) in the previous paragraph.

6. DISCUSSION

Theorem 3 is a first step towards systematic construction of optimal fractional designs for order-of-addition experiments. There is a need to develop optimal or efficient designs in even smaller run sizes. Theorems 1 and 2, along with the resulting expressions (11) and (12) for D- and A-efficiencies which facilitate the study of robustness, are powerful tools for this purpose. Initial studies suggest that a refinement of the procedure in § 5, via the use of certain incomplete block designs and partial rather than full permutations, should work. Some progress has already been made in this direction, covering in particular the cases of $m = 7, \ldots, 12$. For example, with $m = 8$, we have found a design that requires only 168 runs and has a moment matrix M_0 when $c_h = 1$ for all h. Thus, under the usual pairwise order model, the design is optimal in the sense of Theorem 3. It is also very robust with respect to the choice of the c_h, and has D- and A-efficiencies of over 0.99 under the tapered models given by $c_h = 1/h$ and $c_h = (1/2)^{h-1}$ ($h = 1, \ldots, m-1$). Step-down and exchange algorithms are quite promising in the construction of still smaller efficient designs. Thus, with $m = 8$, from the 168-run design mentioned above, we could obtain an 84-run design having D- and A-efficiencies of 0.986 and 0.973 for $c_h = 1$; 0.973 and 0.946 for $c_h = 1/h$; and 0.972 and 0.945 for $c_h = (1/2)^{h-1}$ ($h = 1, \ldots, m-1$). As noted in § 4, these efficiency figures are conservative. Streamlining of these algorithms could yield even better results.

From the viewpoint of model robustness, it is of interest to examine how our designs behave when the model is allowed to include some three-way orderings in addition to the pairwise terms. For $m = 3$ and any fixed c_h, under model (1), the space of possible 6×1 vectors with elements $\tau(a)$ ($a \in A$) has dimension four, and its two-dimensional orthocomplement does not depend on the c_h. Hence, for $m = 3$, the three-way ordering of components 1, 2 and 3 can be taken care of via augmentation of (1) by two more terms that account for this orthocomplement. The same idea can be extended to general m through the inclusion of two additional terms for each triplet of components for which the three-way ordering is intended to be captured, with only a few such triplets being entertained from the perspective of model parsimony. We omit the details to save space but note that, for $m \geqslant 5$, our fractional designs in § 5 perform very well even under such augmented models. For example, the design there for $m = 7$ has a D-efficiency of at least 0.95 and an A-efficiency of at least 0.9, under every such augmented model that incorporates

three-way orderings of one or two triplets, this being true for various choices of the c_h including $c_h = 1, 1/h$ and $(1/2)^{h-1}$ $(h = 1, \ldots, m-1)$; in fact, the D- and A-efficiencies exceed 0.98 and 0.95, respectively, in a vast majority of cases. The smaller designs mentioned in the preceding paragraph also tend to perform quite well, particularly under the D-criterion and often under the A-criterion. This allows further flexibility in model selection along the lines indicated in § 5, perhaps now with a provision for penalty for the additional parameters that correspond to three-way orderings. As before, one remains assured of high design efficiency under the model that one arrives at.

Our results do not require normality of the responses, but do assume their homoscedasticity. This is violated if the response variance is treatment-dependent. While theoretical results are then hard to obtain, algorithms in approximate theory can be readily employed to find a D- or an A-optimal design measure, say w^*, at least for $m \leqslant 7$; see, for example, Torsney & Martín-Martín (2009). Numerical studies based on comparison with w^* show that our optimal designs retain high efficiency when the variances do not differ widely. In contrast, if such heteroscedasticity is more severe, then, as intuitively expected, w^* turns out to be far from uniform, and this leads to a pleasant surprise, making the conversion of w^* to an efficient exact design rather easy. One needs to include only those treatments where w^* assigns greater weights and exclude others. Such designs are quite robust, with high efficiency across various tapered models and over an appreciable range of the variances. Thus, with $m = 4$, if it is believed that responses from treatments beginning with component 1 may have a common variance greater than that of the rest, then under model (1) the 12-run design $\{2134, 2143, 2341, 2431, 3124, 3142, 3241, 3421, 4123, 4132, 4231, 4321\}$, obtained as above, has D-efficiency over 0.98 and A-efficiency over 0.95, for a variance ratio greater than 1.2 and for $c_h = 1, 1/h$ or $(1/2)^{h-1}$. While a good design here is expected to exclude treatments beginning with 1, approximate theory guides us as to which 12 of the remaining 18 should be included. Incidentally, these heuristics do not work under homoscedasticity, where, by Theorem 1, the uniform design measure is optimal. This gives no immediate clue to which treatments should be retained in a smaller design and necessitates the development of further theory. Of course, the optimal design algorithms can become unmanageable for larger m, and more work is needed in this regard.

As pointed out by a reviewer, our work here may have connections with other types of ordering problems, such as the crossover experiment and the choice of splitting order in decision trees. We conclude with the hope that the present endeavour will generate further interest in order-of-addition designs and related topics.

Acknowledgement

We thank the referees, associate editor and editor for very constructive suggestions. This research was supported by the U.S. National Science Foundation, the Indian Institute of Management Calcutta, and a J. C. Bose National Fellowship of the Government of India.

Appendix

Proof of Theorem 1

Let $\pi = \pi_1 \cdots \pi_m$ be a permutation of $1, \ldots, m$, and let

$$\pi a = \pi_{a_1} \cdots \pi_{a_m} \tag{A1}$$

for any treatment $a = a_1 \ldots a_m \in A$. Clearly, A is also the set of all such $m!$ permutations π.

LEMMA A1. *Given any $\pi \in A$, there exists a signed permutation matrix $R(\pi)$ of order p such that $x(\pi a)^{\mathrm{T}} = x(a)^{\mathrm{T}} R(\pi)$ for every $a \in A$.*

Proof. Given any $\pi = \pi_1 \cdots \pi_m$, for $ij \in S$ let $\bar{\pi}_i \bar{\pi}_j$ equal $\pi_i \pi_j$ if $\pi_i < \pi_j$ and $\pi_j \pi_i$ if $\pi_i > \pi_j$. By (A1), for every $ij \in S$ and every $a \in A$, (i) i precedes j in a if and only if π_i precedes π_j in πa, and (ii) the distance between i and j in a equals that between π_i and π_j in πa. Recalling (2), by (i), therefore, $z_{\bar{\pi}_i \bar{\pi}_j}(\pi a)$ and $z_{ij}(a)$ have the same sign if and only if $\pi_i < \pi_j$, while by (ii) they have the same absolute value. In other words, $z_{\bar{\pi}_i \bar{\pi}_j}(\pi a)$ equals $z_{ij}(a)$ if $\pi_i < \pi_j$ and $-z_{ij}(a)$ if $\pi_i > \pi_j$. Thus, $z(\pi a)^{\mathrm{T}} = z(a)^{\mathrm{T}} \tilde{R}(\pi)$ for every $a \in A$, where $\tilde{R}(\pi)$ is a signed permutation matrix of order q such that for each $ij \in S$, the $(ij, \bar{\pi}_i \bar{\pi}_j)$th element of $\tilde{R}(\pi)$ is 1 or -1, according to whether $\pi_i < \pi_j$ or $\pi_i > \pi_j$, respectively, and all other elements of $\tilde{R}(\pi)$ are zeros. Hence $x(\pi a)^{\mathrm{T}} = x(a)^{\mathrm{T}} R(\pi)$ for every $a \in A$, where $R(\pi) = \mathrm{diag}\{1, \tilde{R}(\pi)\}$ is a signed permutation matrix of order p. □

Consider a design measure $w = \{w(a) : a \in A\}$. For any $\pi \in A$, let πw be the design measure that assigns, for each $a \in A$, weight $w(a)$ to treatment πa. Because $\phi(\cdot)$ is concave, writing $\sum_\pi$ for summation over $\pi \in A$,

$$\phi\Big\{(1/m!) \sum\nolimits_\pi M(\pi w)\Big\} \geqslant (1/m!) \sum\nolimits_\pi \phi\big\{M(\pi w)\big\}. \tag{A2}$$

Now, $\{\pi a : \pi \in A\} = A$ for every fixed $a \in A$, so that by (6), $\sum_\pi x(\pi a) x(\pi a)^{\mathrm{T}} = \sum_a x(a) x(a)^{\mathrm{T}} = m! M_0$. Hence, by (5),

$$(1/m!) \sum\nolimits_\pi M(\pi w) = (1/m!) \sum\nolimits_\pi \sum\nolimits_a w(a) x(\pi a) x(\pi a)^{\mathrm{T}} = \sum\nolimits_a w(a) M_0 = M_0. \tag{A3}$$

Also, by (5) and Lemma A1, any π leads to a signed permutation matrix $R(\pi)$ such that

$$\begin{aligned} M(\pi w) &= \sum\nolimits_a w(a) x(\pi a) x(\pi a)^{\mathrm{T}} \\ &= R(\pi)^{\mathrm{T}} \Big\{\sum\nolimits_a w(a) x(a) x(a)^{\mathrm{T}}\Big\} R(\pi) = R(\pi)^{\mathrm{T}} M(w) R(\pi), \end{aligned}$$

and hence $\phi\{M(\pi w)\} = \phi\{M(w)\}$ because $\phi(\cdot)$ is signed permutation invariant. Consequently, in view of (A3), from (A2) we get $\phi(M_0) \geqslant \phi\{M(w)\}$, and the result follows.

Proof of Theorem 2

LEMMA A2.

(a) *For $ij \in S$, $\sum_a z_{ij}(a) = 0$ and $\sum_a z_{ij}^2(a) = (m!) b_0$.*
(b) *For i, j and k satisfying $1 \leqslant i < j < k \leqslant m$, $\sum_a z_{ij}(a) z_{ik}(a) = (m!) b_1$, $\sum_a z_{ik}(a) z_{jk}(a) = (m!) b_1$, and $\sum_a z_{ij}(a) z_{jk}(a) = -(m!) b_1$.*
(c) *For $ij, kl \in S$, if the sets $\{i, j\}$ and $\{k, l\}$ are disjoint, then $\sum_a z_{ij}(a) z_{kl}(a) = 0$.*

Proof. (a) This follows from (2) and (8), noting that among the $z_{ij}(a)$ $(a \in A)$, each of c_h and $-c_h$ occurs with frequency $(m-2)!(m-h)$ $(h = 1, \ldots, m-1)$.

(b) By (2), considering the positions, in increasing order, occupied by some permutation of i, j and k in any treatment,

$$\begin{aligned} &\sum\nolimits_a z_{ij}(a) z_{ik}(a) \\ &\quad = 2\{(m-3)!\} \sum\nolimits_u \big\{c_{u(2)-u(1)} c_{u(3)-u(1)} - c_{u(2)-u(1)} c_{u(3)-u(2)} + c_{u(3)-u(1)} c_{u(3)-u(2)}\big\}, \end{aligned}$$

where $\sum_u$ denotes summation over all integers $u(1), u(2)$ and $u(3)$ satisfying $1 \leqslant u(1) < u(2) < u(3) \leqslant m$. Given any positive integers $h(1)$ and $h(2)$ such that $h(1)+h(2) \leqslant m-1$, there are $m-h(1)-h(2)$ triplets $\{u(1), u(2), u(3)\}$ as above such that $u(2)-u(1) = h(1)$ and $u(3)-u(2) = h(2)$. Therefore

$$\sum\nolimits_a z_{ij}(a) z_{ik}(a) = 2\{(m-3)!\} \sum\nolimits_h \{m-h(1)-h(2)\} \left\{ c_{h(1)} c_{h(1)+h(2)} - c_{h(1)} c_{h(2)} + c_{h(2)} c_{h(1)+h(2)} \right\},$$

where $\sum_h$ is as in (9). Because

$$\sum\nolimits_h \{m-h(1)-h(2)\} c_{h(2)} c_{h(1)+h(2)} = \sum\nolimits_h \{m-h(1)-h(2)\} c_{h(1)} c_{h(1)+h(2)}$$

by symmetry, the first of the three identities in (b) is now evident from (9). The other two identities follow similarly.

(c) This is evident from (2), upon noting that the set of treatments A can be partitioned into disjoint pairs such that the two treatments within each pair have the positions of i and j interchanged and every other component is in the same position. □

LEMMA A3.

(a) *Let* $V^2 = 2(m-2)I + (m-4)V$;

(b) V *has eigenvalues* $m-2$ *and* -2, *with respective multiplicities* $m-1$ *and* $(m-1)(m-2)/2$.

Proof. (a) For any $ij, kl \in S$, write ρ for the (ij, kl)th element of V^2. It suffices to show that ρ equals $2(m-2)$ if $ij = kl$ and $(m-4)V(ij, kl)$ otherwise. As V is symmetric, ρ equals the scalar product of the ijth and klth rows of V. Thus $\rho = \mu_{++} - \mu_{+-} - \mu_{-+} + \mu_{--}$, where μ_{+-} is the number of positions with the ijth row having 1 and the klth row having -1, and so on. The assertion now follows from (7) via consideration of the cases (i)–(vi) below. In the following, $\mu = (\mu_{++}, \mu_{+-}, \mu_{-+}, \mu_{--})$.

(i) If $ij = kl$, then $\mu = (m-i+j-3, 0, 0, m+i-j-1)$; $\rho = 2(m-2)$.

(ii) If $i = k$ and $j \neq l$, then $i \leqslant m-2$ and $\mu = (m-i-2, 0, 1, i-1)$ or $(m-i-2, 1, 0, i-1)$ according to whether $j < l$ or $j > l$, respectively; $\rho = m-4$.

(iii) If $i \neq k$ and $j = l$, then $j \geqslant 3$ and $\mu = (j-3, 1, 0, m-j)$ or $(j-3, 0, 1, m-j)$ according to whether $i < k$ or $i > k$, respectively; $\rho = m-4$.

(iv) If $i = l$, then $i \geqslant 2$ and $\mu = (1, m-i-1, i-2, 0)$; $\rho = -(m-4)$.

(v) If $j = k$, then $j \leqslant m-1$ and $\mu = (1, j-2, m-j-1, 0)$; $\rho = -(m-4)$.

(vi) If none of the above five cases arises, then $\{i, j\}$ and $\{k, l\}$ are disjoint sets. In this situation, μ equals $(1,1,1,1)$ if $i > l$ or $k > j$, $(2,2,0,0)$ if $i < k < l < j$, $(2,0,2,0)$ if $k < i < j < l$, and $(2,1,1,0)$ if $i < k < j < l$ or $k < i < l < j$. So $\rho = 0$.

(b) By part (a), any eigenvalue λ of V satisfies $\lambda^2 = 2(m-2) + (m-4)\lambda$ and hence equals $m-2$ or -2. As $\mathrm{tr}(V) = 0$, these eigenvalues have multiplicities $m-1$ and $(m-1)(m-2)/2$, respectively. □

We now complete the proof of Theorem 2. The expression for M_0 follows from (6), (7) and Lemma A2, recalling that $x(a) = [1, z(a)^{\mathrm{T}}]^{\mathrm{T}}$. Hence, by Lemma A3, eigenvalues of M_0 are as stated.

Proof of Theorem 3

Let $X = (X_0, X_{12}, \ldots, X_{m-1\,m})$ be the model matrix of d^*, arising from (1) and (2), with $c_h = 1$ for all h. Thus X_0 is a column of ones and each X_{ij} corresponds to β_{ij} $(ij \in S)$. Write $N = (m!)/(s!)$ for the run size of d^*. Because d^* has moment matrix $(1/N)X^{\mathrm{T}}X$, the result follows from (7), (13) and Lemma A4 below, which is akin to Lemma A2 but has a different proof.

LEMMA A4.

(a) *For* $ij \in S$, $X_0^{\mathrm{T}}X_{ij} = 0$ *and* $X_{ij}^{\mathrm{T}}X_{ij} = N$. *Also,* $X_0^{\mathrm{T}}X_0 = N$.
(b) *For* i,j *and* k *satisfying* $1 \leqslant i < j < k \leqslant m$, $X_{ij}^{\mathrm{T}}X_{ik} = N/3$, $X_{ik}^{\mathrm{T}}X_{jk} = N/3$, *and* $X_{ij}^{\mathrm{T}}X_{jk} = -N/3$.
(c) *For* $ij, kl \in S$, *if the sets* $\{i,j\}$ *and* $\{k,l\}$ *are disjoint, then* $X_{ij}^{\mathrm{T}}X_{kl} = 0$.

Proof. (a) Clearly, $X_{ij}^{\mathrm{T}}X_{ij} = X_0^{\mathrm{T}}X_0 = N$ for $ij \in S$, as X has elements ± 1. Next, let $G_u(ij)$ be the contribution of D_u to $X_0^{\mathrm{T}}X_{ij}$. By (2) and (14), $G_u(ij) = 0$, irrespective of whether i is in C_u or $\bar{C}_u$ and whether j is in C_u or $\bar{C}_u$. Hence $X_0^{\mathrm{T}}X_{ij} = 0$ for any $ij \in S$.

(b) Write G_u for the row vector with elements $G_u(ij,ik)$, $G_u(ik,jk)$ and $G_u(ij,jk)$, where $G_u(ij,ik)$ is the contribution of D_u to $X_{ij}^{\mathrm{T}}X_{ik}$ and $G_u(ik,jk)$ and $G_u(ij,jk)$ are similarly defined. Because the rows of B_u and $\bar{B}_u$ in (14) are formed by all permutations of the elements of C_u and $\bar{C}_u$, respectively, in view of (7) and (13) the following hold:

(i) $G_u = (2/3)(s!)(1,1,-1)$ if either $i,j,k \in C_u$ or $i,j,k \in \bar{C}_u$;
(ii) $G_u = 2(s!)(1,0,0)$ if either $j,k \in C_u$ and $i \in \bar{C}_u$, or $i \in C_u$ and $j,k \in \bar{C}_u$;
(iii) $G_u = 2(s!)(0,1,0)$ if either $i,j \in C_u$ and $k \in \bar{C}_u$, or $k \in C_u$ and $i,j \in \bar{C}_u$;
(iv) $G_u = 2(s!)(0,0,-1)$ if either $i,k \in C_u$ and $j \in \bar{C}_u$, or $j \in C_u$ and $i,k \in \bar{C}_u$.

Because situation (i) corresponds to $(2s-3)!/\{s!(s-3)!\}$ choices of u, and each of (ii), (iii) and (iv) corresponds to $(2s-3)!/\{(s-1)!(s-2)!\}$ choices of u, part (b) follows after a little algebra.

(c) Let $G_u(ij,kl)$ be the contribution of D_u to $X_{ij}^{\mathrm{T}}X_{kl}$. Then, by (2) and (14),

(i) $G_u(ij,kl) = 2(s!)$ if either $i,k \in C_u$ and $j,l \in \bar{C}_u$, or $j,l \in C_u$ and $i,k \in \bar{C}_u$;
(ii) $G_u(ij,kl) = -2(s!)$ if either $i,l \in C_u$ and $j,k \in \bar{C}_u$, or $j,k \in C_u$ and $i,l \in \bar{C}_u$;

and $G_u(ij,kl) = 0$ in all other situations. Part (c) is now immediate, because each of (i) and (ii) corresponds to $(2s-4)!/\{(s-2)!(s-2)!\}$ choices of u. □

REFERENCES

BOOTH, K. H. V. & COX, D. R. (1962). Some systematic supersaturated designs. *Technometrics* **4**, 489–95.

DING, X., MATSUO, K., XU, L., YANG, J. & ZHENG, L. (2015). Optimized combinations of bortezomib, camptothecin, and doxorubicin show increased efficacy and reduced toxicity in treating oral cancer. *Anti-cancer Drugs* **26**, 547–54.

ECCLESTON, J. A. & HEDAYAT, A. (1974). On the theory of connected designs: Characterization and optimality. *Ann. Statist.* **2**, 1238–55.

FULEKI, T. & FRANCIS, F. J. (1968). Quantitative methods for anthocyanins. *J. Food Sci.* **33**, 266–74.

JIANG, X. J. & NG, D. K. P. (2014). Sequential logic operations with a molecular keypad lock with four inputs and dual fluorescence outputs. *Angew. Chem. Int. Ed.* **53**, 10481–4.

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

KIEFER, J. & WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Can. J. Math.* **12**, 363–6.

RAGHAVARAO, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments*. New York: Wiley.

RYBERG, P. (2008). Development of a mild and robust method for large-scale palladium-catalysed cyanation of aryl bromides: Importance of the order of addition. *Org. Process Res. Dev.* **12**, 540–3.

SHINOHARA, A. & OGAWA, T. (1998). Stimulation by rad52 of yeast rad51-mediated recombination. *Nature* **391**, 404–7.

SONG, Y., ZHU, S., XIANG, S., ZHAO, X., ZHANG, J., ZHANG, H., FU, Y. & YANG, B. (2014). Investigation into the fluorescence quenching behaviors and applications of carbon dots. *Nanoscale* **6**, 4676–82.

TORSNEY, B. & MARTÍN-MARTÍN, R. (2009). Multiplicative algorithms for computing optimum designs. *J. Statist. Plan. Infer.* **139**, 3947–61.

VAN NOSTRAND, R. C. (1995). Design of experiments where the order of addition is important. In *ASA Proc. Sect. Phys. Eng. Sci.* Alexandria, Virginia: American Statistical Association, pp. 155–60.

VOELKEL, J. G. (2017). The design of order-of-addition experiments. *arXiv:* 1701.02786v2.

[*Received on* 28 *July* 2017. *Editorial decision on* 17 *December* 2018]