QUALITY ENGINEERING
2019, VOL. 31, NO. 1, 49–59
https://doi.org/10.1080/08982112.2018.1548021

Order-of-addition experiments: A review and some new thoughts

Dennis K. J. Lin and Jiayu Peng

Department of Statistics, Pennsylvania State University, State College, Pennsylvania

ABSTRACT

The study of order-of-addition (OofA) experiments is prevalent in many scientific and industrial areas. The statistical design of experiments (DOE) will considerably improve the efficiency of OofA experiments. Designing and modeling the OofA experiments have increasingly received a great deal of attention. In this article, we review the latest work on the design and model of OofA experiments, and introduce some new thoughts. We believe that this article will motivate fruitful applications in real OofA experiments as well as future developments in the methodology.

KEYWORDS

constraints; design construction; exchange algorithm; higher-order; optimal designs; pairwise-order; symmetry

Introduction

In Fisher (1937), a lady was able to distinguish (by tasting) whether the tea or the milk was first added to the cup. This is perhaps the first order-of-addition (OofA) experiment in the literature. Two orders, "milk then tea" or "tea then milk," yield different responses (the taste). A general OofA experiment involves m (≥ 2) components, and the $m!$ different orders of adding these components into the system yield different responses. The OofA experiment is prevalent in chemistry-related areas. In many chemical experiments, a number of reactants are added into the apparatus sequentially, rather than simultaneously. The formation (the amount/size/purity, etc.) of the reaction product depends on the sequence of adding reactants. For example, Fuleki and Francis (1968) studied an experiment for extracting anthocyanins from cranberries. They found that "The order of addition of the lead acetate (before or after the pH adjustment) had a definite influence on the reaction. Higher recoveries were obtained by adjusting the pH after lead acetate addition." During the past decades, the OofA effect is frequently mentioned in the area of bio-chemistry (Shinohara and Ogawa 1998), food science (Jourdain et al. 2009), nutritional science (Karim, McCormick, and Kappagoda 2000), and pharmaceutical science (Rajaonarivony et al. 1993), just to name a few. As another kind of application in genomics, in the algorithmic construction of "maximum-likelihood phylogenetic trees" from DNA sequences, the likelihood of the fitted tree relies on the order of adding different taxa into the computer program (Olsen et al. 1994; Stewart et al. 2001). There are typically more than 10, or even hundreds of taxa involved.

Ultimately, an OofA experiment is to find the optimal addition order. For this purpose, one needs to compare a number of different orders to learn the dependence of the response on the order. With m components to add, an exhaustive search of all permutations requires $m!$ runs of experiments, which is usually not affordable (e.g., for $m = 10$, there will be $m! = 10! \approx 3.6 \times 10^6$ orders). Often in practice, a number of randomly selected orders are tested (see, e.g., Stewart et al. 2001). Of course, such a random design is a convenient choice as it does not exploit any pattern that the order-of-addition drives the variation of responses. Systematic designs are desired to learn the dependence of the response on the order-of-addition.

The statistical design of experiments (DOE) is a powerful tool to collect informative data under the control of experimental costs. A successful application of DOE in the OofA experiments will help experimenters identify the important order effects and find out the optimal addition order, with substantially fewer experimental runs and higher reliability. Van Nostrand (1995) proposed a model which assumes that an order-of-addition affects the response via pairwise-order (PWO) effects. Voelkel (2017) studied the orthogonality of designs and suggested a number of

CONTACT Jiayu Peng pengjy08@gmail.com Department of Statistics, Pennsylvania State University, Thomas Building, University Park, State College, PA 16802, USA.

This article was presented at the Sixth Stu Hunter Research Conference in Roanoke, Virginia, March 2018.

design criteria under Van Nostrand's PWO model. Via computer search, Voelkel (2017) found some OofA designs which have the same correlation structures as the full OofA designs, for small number of components (m). While the designs Voelkel found are intuitively optimal under, say, the D-criterion, a theoretical support on their optimality remains absent.

While the OofA experiments are prevalent in practice, the statistical methods for its design and model remain primitive. In this article, we introduce up-to-date developments in the OofA problem. We review the latest literature and propose some new thoughts. Many of these results can be directly used in scientific areas, such as chemistry and biology. We believe that these results form a solid basis for the future work. Moreover, the OofA problems provide a golden opportunity for interdisciplinary studies. Some research topics in group theories or combinatorics may arise therein. A fruitful application of computer science is also anticipated, because the computational complexity for finding efficient OofA designs tremendously increases as m becomes large.

The remainder of this article is organized as follows. The PWO model section introduces the PWO model for OofA effects. Optimality theory of PWO design section summarizes the theory of optimal PWO designs. Systematic construction of favorable PWO designs and Algorithmic construction of efficient PWO designs with small number of runs sections are devoted to the design construction under the PWO model. Systematic construction of favorable PWO designs section is focused on the systematic construction methods while Algorithmic construction of efficient PWO designs with small number of runs section proposes a new algorithmic construction method. More order-of-addition models section discusses on a number of new models beyond the PWO model. Concluding remarks are given in the last section.

The PWO model

Following Van Nostrand (1995) and Voelkel (2017), we first study the OofA designs under the pairwise-order (PWO) model defined below. Suppose there are m (≥ 2) components $1, 2, ..., m$, there will be a total of $m!$ possible orders. An order is represented by a vector $\boldsymbol{a} = [a_1, a_2, ..., a_m]^T$, where $a_1, a_2 ..., a_m$ is a permutation of $1, ..., m$. Here, an order defines a treatment. Denote the set of all the $m!$ treatments by $\mathcal{A}$. In a PWO model, the features of an order are represented by the precedence patterns between all the $\binom{m}{2}$ pairs of components. Explicitly, let $\mathcal{S}$ be set of all pairs (j, k) for $1 \leq j<k \leq m$, and for each $jk \in \mathcal{S}$, define the PWO indicator between j and k as:

$$z_{jk}(\boldsymbol{a}) = \begin{cases} 1 & \text{if } j \text{ precedes } k \text{ in } \boldsymbol{a}, \\ -1 & \text{if } k \text{ precedes } j \text{ in } \boldsymbol{a}. \end{cases} \quad [1]$$

As an illustration, for $m = 4$ and $\boldsymbol{a} = 2143$: we have $\mathcal{S} = \{12, 13, 14, 23, 24, 34\}$, then $z_{12}(\boldsymbol{a}) = -1, z_{13}(\boldsymbol{a}) = +1$, ..., and $z_{34}(\boldsymbol{a}) = -1$.

For $\boldsymbol{a} \in \mathcal{A}$, write $\tau(\boldsymbol{a})$ for the treatment mean of $\boldsymbol{a}$, that is, $\tau(\boldsymbol{a})$ is the expectation of any observation arising from treatment $\boldsymbol{a}$. As usual, it is assumed that the observations have equal variance and are uncorrelated. The PWO model assumes that the expected response is determined by summing the effects of all z_{jk}'s, namely:

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{jk \in \mathcal{S}} z_{jk}(\boldsymbol{a})\beta_{jk}, \quad [2]$$

where β_{jk}'s and β_0 are unknown parameters. Let $\tilde{\boldsymbol{\beta}} = (\beta_{12}, \beta_{13}, ..., \beta_{(m-1)m})^T$, where ^{T} denotes the transpose. Then $\boldsymbol{\beta} = (\beta_0, \tilde{\boldsymbol{\beta}}^T)^T$ represents the parametric vector of interest. Similarly, for any $\boldsymbol{a} \in \mathcal{A}$, let $\boldsymbol{z}(\boldsymbol{a}) = (z_{12}(\boldsymbol{a}), z_{13}(\boldsymbol{a}), ..., z_{(m-1)m}(\boldsymbol{a}))^T$ and $\boldsymbol{x}(\boldsymbol{a}) = (1, \boldsymbol{z}(\boldsymbol{a})^T)^T$. Then Eq. [2] can be expressed as

$$\tau(\boldsymbol{a}) = \beta_0 + \boldsymbol{z}(\boldsymbol{a})^T\tilde{\boldsymbol{\beta}} = \boldsymbol{x}(\boldsymbol{a})^T\boldsymbol{\beta}. \quad [3]$$

Throughout, write $q = \binom{m}{2}$ and $p = q + 1$. Then $\tilde{\boldsymbol{\beta}}$ and $\boldsymbol{z}(\boldsymbol{a})$ are $q \times 1$ while $\boldsymbol{\beta}$ and $\boldsymbol{x}(\boldsymbol{a})$ are $p \times 1$. The PWO model is appealing since the PWO effects β_{jk}'s are often great interest in many experiments. Practitioners often expect conclusions like "Adding A before or after B has a significant influence on the response" (see the aforementioned example in Fuleki and Francis 1968).

The design of OofA experiments is to choose a collection of orders, represented by an OofA design matrix $\mathbf{D} = [\boldsymbol{a}_1^T, \boldsymbol{a}_2^T, ..., \boldsymbol{a}_N^T]^T$, where $\boldsymbol{a}_i \in \mathcal{A}$ for each i and N indicates the number of runs. We represent each $\mathbf{D}$ by a PWO design, defined as $\mathbf{Z} = [\boldsymbol{z}(\boldsymbol{a}_1), \boldsymbol{z}(\boldsymbol{a}_2), ..., \boldsymbol{z}(\boldsymbol{a}_N)]^T$. Then $\mathbf{Z}$ is the design matrix in model [1]. Let $\mathbf{D}_f$ be the full design which replicates each treatment in $\mathcal{A}$ once, and $\mathbf{Z}_f$ be the full PWO design, that is, the PWO design of $\mathbf{D}_f$. For illustration, Table 1 shows the matrices $\mathbf{D}_f$ and $\mathbf{Z}_f$ for $m = 3$.

Remark 1. Due to the transitive property of order, the region of PWO design points does not include all the level combinations. For example, when $m = 3$, the points (–, +, –) and (+, –, +) are not valid for a PWO design. That is, if component 2 precedes component 1 (z_{12} = –) and component 1 precedes

Table 1. The full OofA design and full PWO design with three components.

(a) Full OofA design $\mathbf{D}_f$.		
1	2	3
1	3	2
2	1	3
2	3	1
3	1	2
3	2	1
(b) Full PWO design $\mathbf{Z}_f$		
z_{12}	z_{13}	z_{23}
+	+	+
+	+	−
−	+	+
−	−	+
+	−	−
−	−	−

component 3 ($z_{13} = +$), then component 2 *must* precede component 3 (z_{23} must be $+$, not $-$).

The PWO model is an economic model which requires small number of runs ($p = \binom{m}{2} + 1$) compared with the total number of candidate runs ($m!$). For example, when $m = 8$, $p/m! \approx 0.0007$. In addition, this model is easy to interpret: the effect β_{jk} indicates the difference between the average treatment mean of all the possible orders in which j precedes k and the average treatment mean of all the orders where k precedes j. Furthermore, the analytical results of PWO model are easy to utilize. Fitting this model identifies the sign and significance of each pairwise-order effect $\hat{\beta}_{jk}$. One can identify the significant effects via t-tests or variable selection, etc. Its analysis results can be represented in a directed graph—a positively significant $\hat{\beta}_{jk}$ suggests an edge from j to k, a negatively significant β_{jk} suggests an edge from k to j, and a non-significant $\hat{\beta}_{ij}$ suggests no edge between j and k. A topological sorting on this directed graph can give optimal order(s) that are consistent with each inferred preference of pairwise precedence. In summary, the PWO model serves as a convenient initial model for identifying the important components, and is especially favorable when each run of OofA experiment is expensive or when the number of components is large.

Remark 2. A topological sorting exists if and only if the directed graph is acyclic. For example, j preceding k, k preceding l, and l preceding j cannot be simultaneously identified as favorable. In the presence of cycles, more advanced algorithms are needed to speculate the optimal order(s). This will be investigated in the future. A possible approach is to "break" the cycles using existing tools in computer science, such as Tarjan (1972)'s strongly connected components algorithm.

We evaluate the property of any n-run PWO design $\mathbf{Z}$ via its moment matrix, defined as $\mathbf{M} = \mathbf{X}^T\mathbf{X}/n$, where $\mathbf{X} = [1, \mathbf{Z}]$. Note that under the PWO model [3], the variance-covariance matrix of the least squares estimator of $\boldsymbol{\beta}$ is proportional to $\mathbf{M}^{-1}$. Thus it is desirable to maximize the matrix $\mathbf{M}$ under some criteria. The popular criteria in the design field include the D-criterion $\det(\mathbf{M})$, the A-criterion $\mathrm{tr}(\mathbf{M}^{-1})$, the $M.S.$-criterion $\mathrm{tr}(\mathbf{M}^2)$, and the E-criterion which is the minimum eigenvalue of $\mathbf{M}$. We will discuss when the PWO design attains optimum under these criteria in "Optimality theory of PWO design" section and how to construct (nearly-)optimal PWO designs in "Systematic construction of favorable PWO designs" and "Algorithmic construction of efficient PWO designs with small number of runs" sections.

Optimality theory of PWO design

Among all the possible choices of PWO designs with certain number of runs, consider which one(s) attain the optimum in terms of the aforementioned criteria. Peng, Mukerjee, and Lin (2017) proved a non-trivial result: the optimal correlation structure of PWO design is the one given by the full PWO design. Explicitly, let $\mathbf{M}_f$ be the moment matrix of the full PWO design, Peng, Mukerjee, and Lin (2017) showed:

Theorem 1.

i. *Any PWO design with a moment matrix of $\mathbf{M}_f$ is ϕ-optimal for every optimality criterion $\phi(\cdot)$ which is concave and signed permutation invariant.*
ii. *For strictly concave and signed permutation invariant criterion $\phi(\cdot)$, any PWO design is ϕ-optimal if and only if it attains a moment matrix of $\mathbf{M}_f$.*

In particular, the conventional D-, A-, E-, and $M.S.$-criteria are all signed permutation invariant and concave, and strictly concave except the E-criterion, so Theorem 1 applies. In view of the equivalence theorem, the D-optimality of full design also implies its G-optimality, that is,

$$\max_{\boldsymbol{a}\in\mathcal{A}} \boldsymbol{x}(\boldsymbol{a})^T\mathbf{M}^{-1}\boldsymbol{x}(\boldsymbol{a}) \geq p = \max_{\boldsymbol{a}\in\mathcal{A}} \boldsymbol{x}(\boldsymbol{a})^T\mathbf{M}_f^{-1}\boldsymbol{x}(\boldsymbol{a}) \quad [4]$$

for the moment matrix ($\mathbf{M}$) of any PWO design; see, for example, Silvey (2013), Chapter 3 for more details. By Eq. [4], a moment matrix of $\mathbf{M}_f$ minimizes the maximum variance of the estimated responses at $\boldsymbol{a} \in \mathcal{A}$.

In practice, the explicit optimal values of different criteria are needed for the purpose of comparing different designs. Peng, Mukerjee, and Lin (2017) studied the properties of the full (thus optimal) PWO design by identifying its eigenstructure. Their results are as below.

Theorem 2. $\mathbf{M}_f$ has eigenvalues 1, $(m+1)/3$ and 1/3, with multiplicities 1, $m-1$ and $\binom{m-1}{2}$, respectively. Then for the full (thus optimal) design:

$$D\text{−criterion} = \left[\det(\mathbf{M}_f)\right]^{1/p} = \left[\frac{(m+1)^{m-1}}{3^q}\right]^{\frac{1}{p}},$$

$$A\text{−criterion} = \operatorname{tr}\left(\mathbf{M}_f^{-1}\right) = 1 + \frac{3m(m-1)^2}{2(m+1)},$$

$$E\text{−criterion} = \lambda_{\min}(\mathbf{M}_f) = \frac{1}{3}, \text{ and}$$

$$\text{M.S.}\text{−criterion} = \operatorname{tr}\left(\mathbf{M}_f^2\right) = 1 + \frac{(m-1)m(2m+5)}{18},$$

with $q = \binom{m}{2}$ and $p = \binom{m}{2} + 1$.

Remark 3. One of the most important features of the order-of-addition problem is that the design points under the PWO model, or in general any OofA model, are subject to heavy constraints. The constraints, essentially imposed by the transitive property of orders, result in inevitable correlations in any PWO design. Note that even the optimal PWO design is not orthogonal; from Theorem 2, as m approaches infinity, the D-criterion of optimal design approaches 1/3. These constraints further raise substantial algebraic difficulties in proving the optimality theorems under the PWO, or other OofA models.

Remark 4. Theorems 1 and 2 provide benchmarks for evaluating all PWO designs, and enable us to define the D-, A-, E-, and M.S.-efficiencies for the purpose of searching efficient, fractional PWO designs. For example, for any PWO design $\mathbf{Z}$ with moment matrix $\mathbf{M}$, its D-efficiency is

$$d(\mathbf{Z}) = [\det(\mathbf{M})]^{1/p} \Big/ \left[\frac{(m+1)^{m-1}}{3^q}\right]^{\frac{1}{p}}.$$

Throughout this article, the D-efficiency is relative to the optimal design (hence the full design). Likewise, one can define the A-, E-, or M.S.-efficiency. For simplicity, only the D-efficiency is considered in the remainder of this article. The larger D-efficiency, the better, and an optimal design has a D-efficiency of 1.

Systematic construction of favorable PWO designs

In this section, we consider the construction of PWO designs. The design construction problem is based upon two objectives: (i) maximizing the design efficiency and (ii) minimizing the experimental cost. We first review the construction of optimal (but not necessarily minimal-point) PWO designs, and then review the construction of minimal-point (but not necessarily optimal) PWO designs.

Table 2. Optimal fractional PWO design given by Peng, Mukerjee, and Lin (2017) with $m = 4$.

1	2	3	4
2	1	4	3
4	3	1	2
3	4	2	1
1	3	2	4
3	1	4	2
4	2	1	3
2	4	3	1
1	4	2	3
4	1	3	2
3	2	1	4
2	3	4	1

A class of optimal fractional PWO designs

Recall that a fractional PWO design is D-optimal if and only if it has the same moment matrix as the full design. However, such designs are very rare: for $m = 6$, 7, 8, respectively, we randomly generated 100,000 half-fractional PWO designs (with $m!/2$ runs), and none of them attains optimum.

Peng, Mukerjee, and Lin (2017) obtained a class of optimal fractional designs with closed-form construction. Their construction is based on a new blocking scheme and a rowwise-reverse technique. Table 2 shows their design with $m = 4$. This design can be divided (by the horizontal lines in Table 2) into three blocks, with different blocks sharing a common structure. Each block has a form of $\begin{bmatrix} A & B \\ \tilde{B} & A \end{bmatrix}$: in Table 2, $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix}$, and $\tilde{B} = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$. One block can be obtained from any other block by relabeling the components. For instance, if we replace each "2" by "3" and replace each "3" by "2" in the first block, then we obtain the second block. All these features guarantee that the design is symmetric about different components and is thus optimal. Detailed mathematical descriptions and proofs are given in Peng, Mukerjee, and Lin (2017). These obtained designs are $1/r!$ fractions of the full design, where r equals $m/2$ for even m, and $(m-1)/2$ for odd m. Such designs are useful for small m, but may not be affordable for larger m (see Table 3). For example when $m = 6$, their design is a (1/6)-fractional design which has 120 runs; while when $m = 10$, their design requires 30,240 runs, which is often not affordable. Despite that their results may not be practically useful for large m,

Table 3. The number of runs ($m!/r!$) for the obtained optimal fractional PWO design versus the number of runs ($m!$) for the full PWO design.

m	$m!$	$m!/r!$
4	24	12
5	120	60
6	720	120
7	5,040	840
8	40,320	1,680
9	3.6×10^5	15,120
10	3.6×10^6	30,240

they proved a theoretically important fact that the optimal fractional PWO design does exist (for $m \geq 4$).

A class of minimal-point PWO designs

In some OofA experiments, each run of experiment is rather expensive or time-consuming. It is then desirable to minimize the number of runs in a PWO design. The PWO model has $p = \binom{m}{2} + 1$ parameters to be estimated and therefore a minimal-point PWO design shall have p runs.

Zhao, Lin, and Liu (2017) studied the construction of minimal-point PWO designs (Table 4). They first found that if we randomly generate a PWO design with p runs, the obtained design is mostly singular and often has a very low D-efficiency. Zhao, Lin, and Liu (2017) then systematically constructed a class of minimal-point designs (for any m) which are reasonably-efficient. Basically, given an efficient design with small m, they developed a method to augment this small design so that a minimal-point PWO design with larger m is obtained. This can serve as a "lower-bound design" (as compared to the full $m!$-run design as a "upper-bound design").

Explicitly, let $\mathbf{Q}$ be an efficient minimal-point design with $m = 5$. (Such a $\mathbf{Q}$ is found by exhausted computer search.) Then for any larger m, let $\mathbf{H}_1 = [\mathbf{Q}, 1]$ with 1 being a vector of 1's. Let $\mathbf{H}_2$ be an $11 \times (q-11)$ matrix of 1's (recall that $q = \binom{m}{2}$), $\mathbf{H}_2$ be a $(q-10) \times 11$ matrix of 1's, and $\mathbf{H}_4 = (h_{ij})$ be a $(q-10) \times (q-11)$ matrix with

$$h_{ij} = \begin{cases} 1 & \text{when } i \leq j \\ -1 & \text{otherwise.} \end{cases}$$

Zhao, Lin, and Liu (2017) showed the following result.

Theorem 3. $\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{H}_3 & \mathbf{H}_4 \end{bmatrix}$ is a minimal-point PWO design with m components, with a D-criterion of

$$[\det(\mathbf{M})]^{1/p} = \frac{1}{p}\left(4^{q-10}|\mathbf{H}_1^T\mathbf{H}_1|\right)^{1/p},$$

$\mathbf{M}$ being the information matrix of $\mathbf{H}$.

Table 4. The D-efficiency and the number of runs (p) of the minimal-point designs obtained in Zhao, Lin, and Liu (2017).

m	$m!$	$p = \binom{m}{2} + 1$	D-efficiency
3	6	4	0.810
4	24	7	0.897
5	120	11	0.837
6	720	16	0.532
7	5,040	22	0.375

Note that the quantity $|\mathbf{H}_1^T\mathbf{H}_1|$ is determined by the D-efficiency of $\mathbf{Q}$. Using Theorem 3, Zhao, Lin, and Liu (2017) constructed a series of minimal-point PWO designs. The size and efficiency of their obtained designs are tabulated as follows. These designs are highly economic and are relatively efficient among the minimal-point designs. Yet for large m, the D-efficiency of their design is far below 1. A tradeoff between the D-efficiency and the design economicity is necessary, and this will be studied in the next section.

Algorithmic construction of efficient PWO designs with small number of runs

Systematic construction of favorable PWO designs section gives a series of PWO designs either with the highest efficiency or with the minimal number of runs. In this section, we consider the tradeoff between these two objectives. Consider, for practical needs, how to construct designs with (i) a D-efficiency close to 1 (e.g., greater than 0.9) and (ii) a relatively small number of runs (e.g., $n<2p$). For this purpose, we propose a new algorithm called the "Bubblesort-exchange" algorithm.

Note that the problem of finding optimal order appears similar to the traditional "sorting" problem: given an array of m numbers, how to arrange it in an ascending order? As a simple sorting algorithm, "bubblesort" compares a pair of adjacent items at each step and swaps them if they are in the wrong order. The comparison of adjacent items starts from the bottom of list, so that the largest item will move up to the top of array, like a bubble rising to the surface. After the comparison of adjacent items reaches the top of array, the largest element in this array is then determined. We next restart the comparison of adjacent items from the beginning, and the next bubble which rises up to the top is the second largest element, and so on; finally, the correct order is obtained.

The idea of bubblesort is applicable to the construction of OofA designs. As a toy example, suppose $m = 3$ and the first run of an OofA design, $\mathbf{D}$, is $[a_1, a_2, a_3]$.

Write $\mathbf{D}$ as $\begin{bmatrix} a_1, a_2, a_3 \\ \mathbf{D}_1 \end{bmatrix}$. Via the following steps, the updated (a_1, a_2, a_3) tends to be (although is not 100%) the best run given the rest of design, $\mathbf{D}_1$. **if** $\begin{bmatrix} a_2, a_1, a_3 \\ \mathbf{D}_1 \end{bmatrix}$ has a higher efficiency than $\begin{bmatrix} a_1, a_2, a_3 \\ \mathbf{D}_1 \end{bmatrix}$ **then** $(a_1, a_2) \leftarrow (a_2, a_1)$; **end** **if** $\begin{bmatrix} a_1, a_3, a_2 \\ \mathbf{D}_1 \end{bmatrix}$ has a higher efficiency than $\begin{bmatrix} a_1, a_2, a_3 \\ \mathbf{D}_1 \end{bmatrix}$ **then** $(a_2, a_3) \leftarrow (a_3, a_2)$;**endif** $\begin{bmatrix} a_2, a_1, a_3 \\ \mathbf{D}_1 \end{bmatrix}$ has a higher efficiency than $\begin{bmatrix} a_1, a_2, a_3 \\ \mathbf{D}_1 \end{bmatrix}$ **then**$(a_1, a_2) \leftarrow (a_2, a_1)$;**end**

Given any initial OofA design $\mathbf{D}$, the proposed Bubblesort-exchange algorithm updates the first run via the above approach, and then updates the second run, …, until the last run. The procedure is conducted for several iterations. At each step, bubble-sort swaps only an adjacent pair of components, and thereby the PWO design is updated by only one element—thus the coordinate-exchange algorithm (Meyer and Nachtsheim 1995) can be applied to accelerate our algorithm.

Of course, as a local-greedy-search method, the bubblesort-exchange algorithm cannot guarantee achieving the global optimum. Roughly speaking, the proposed algorithm identifies the local optimum with high chance in computationally efficient manner. As such, it is able to find a variety of new, efficient PWO designs, with m being large (up to 30). In addition, the proposed algorithm is flexible in the run size: n can be any number between p and $m!$. When $n = 2q$, the obtained design efficiencies are displayed in Table 5. With $m \leq 10$, our obtained designs have D-efficiencies greater than 95%. With $m \leq 25$, the obtained designs have D-efficiencies greater than 90%.

Voelkel (2017) searched for small-run efficient PWO design using the Fedorov (1972) exchange algorithm. The Fedorov exchange algorithm obtains slightly better results than our Bubblesort-exchange algorithm for small m, say, $m = 5$, 6, 7. But for larger m, say, $m > 10$, Fedorov exchange is computationally infeasible due to both time and space complexity. The Fedorov exchange algorithm requires to store the full PWO design, which will be a matrix with 4×10^7 runs and 55 columns when $m = 11$. We were unable to store such a large matrix in the memory (under an environment of Intel i7-6700, 3.40 GHz, RAM 16.0 GB, Matlab 2014a). On the contrary, the proposed bubblesort-exchange algorithm is able to handle large m (up to at least $m = 30$).

Table 5. Some D-efficiencies obtained by the Bubblesort-exchange algorithm (with $n = 2q$).

m	$p = q + 1$	$n = 2q$	D-efficiency
4	7	12	1
5	11	20	0.971
6	16	30	0.970
7	22	42	0.970
8	29	56	0.962
9	37	72	0.958
10	46	90	0.953
11	56	110	0.947
12	67	132	0.943
13	79	156	0.940
14	92	182	0.934
15	106	210	0.932
16	121	240	0.929
17	137	272	0.927
18	154	306	0.921
19	172	342	0.918
20	191	380	0.917
25	301	600	0.903
30	485	870	0.891

More order-of-addition models

The PWO model is a parsimonious OofA model. Recall that a minimal-point PWO design requires $p = \binom{m}{2} + 1$ runs. If one can afford considerably more than p experimental runs, a more complicated OofA model would be desired.

An important direction of OofA modeling is to extend the PWO model to general "s-way" models. The PWO design is capable of describing the effect of the precedence pattern between any pair of components. In some cases it is desirable to capture the effect of the precedence pattern among triplets, quadruplets, or in general, s-way tuplets of components ($2 \leq s \leq m$). The PWO model is merely the simplest two-way model. The m-way, that is, the "full" OofA model, is naturally defined. The m-way model assumes that each of the $m!$ orders has a treatment mean and the $m!$ treatment means are not necessarily related in any sense. Between the two-way (PWO) and the m-way models, the definition of any s-way model is not that straightforward. Mee (2017) is the first to discuss the modeling of s-way effects. This section reviews the contribution of Mee (2017) and proposes modifications as well as new frameworks for the s-way model.

In The interaction model section we review a simple interaction model, which describes the three-way effects. The symmetric interaction model section proposes a "symmetric interaction model" which is theoretically more favorable than the simple interaction model. The s-way Oofa model under the constrained linear regression framework section proposes a new series of s-way models under the framework of constrained linear regression. These models deserve further study.

Table 6. Interaction terms of the simple interaction model (5) under the full design ($m = 3$).

Order $\boldsymbol{a}$	$z_{12}(\boldsymbol{a})z_{13}(\boldsymbol{a})$	$z_{12}(\boldsymbol{a})z_{23}(\boldsymbol{a})$
123	+	+
132	+	−
213	−	−
231	+	−
312	−	−
321	+	+

The interaction model

The conventional wisdom of constructing a higher-order model from a lower-order model is to add interactions. Intuitively, the three-way pattern, that is, the precedence pattern among any triplet of components, can be represented by the interactions of PWO factors. Recall that z_{jk} is the PWO indicator between j and k. For any triplet (j, k, l), the relevant PWO interactions include $z_{jk} \cdot z_{jl}, z_{jk} \cdot z_{kl}$, and $z_{jl} \cdot z_{kl}$. A non-trivial concern is: which of the three interactions shall we add into the three-way model? Mee (2017) noticed that the three interactions cannot be added simultaneously. He showed the equation $z_{jk} \cdot z_{jl} - z_{jk} \cdot z_{kl} + z_{jl} \cdot z_{kl} = 1$; and therefore a model is not estimable, if it incorporates all the three interactions (as well as the intercept). Mee (2017) defined an estimable three-way model by flexibly adding "any full-rank set of $df_{3|2} = m(m-1)(m-2)/3$ two-factor [PWO] interactions that share a common factor" into the PWO model. For instance, one can incorporate the first two interactions of each triplet, thus the model

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{1 \le j < k \le m} \beta_{jk} z_{jk}(\boldsymbol{a}) + \sum_{1 \le j < k < l \le m} \left[\beta^{(1)}_{jkl} z_{jk}(\boldsymbol{a}) z_{jl}(\boldsymbol{a}) + \beta^{(2)}_{jkl} z_{jk}(\boldsymbol{a}) z_{kl}(\boldsymbol{a}) \right]. \quad [5]$$

Again, $\tau(\boldsymbol{a})$ denotes the expected response from order $\boldsymbol{a}$. The unknown parameters include β_0, β_{jk}'s for any $1 \le j < k \le m$, as well as $\beta^{(1)}_{jkl}$'s and $\beta^{(2)}_{jkl}$'s for any $1 \le j < k < l \le m$, thus $1 + \binom{m}{2} + 2\binom{m}{3}$ parameters in total.

As a well-defined three-way model, model [5] solves the issue of estimability by incorporating two out of the three interactions for each triplet. However, the choice of only the first two interactions causes some "asymmetry" on different components. Namely, why do we keep $z_{12}z_{13}$ and $z_{12}z_{23}$, while dropping $z_{13}z_{23}$? What if we incorporate another two interactions? The asymmetry of model [5] results in some potential concerns. For one thing, any interaction term, say, $z_{12}z_{13}$, is unbalanced even in the full design. See Table 6: when $m = 3$, $z_{12}z_{13}$ is at the "+" level for four runs while at the "−" level for two runs. This imposes some intrinsic deficiency in estimating the interaction effects. The asymmetry of model [5] raises another potential issue regarding the labeling of different components. The m different components, for example, reagents in chemical experiments, are not naturally ordinal. Ideally, the way of labeling components should not be mattered. However, relabeling the components will alter the design efficiency. As an illustrative example, consider an OofA design $\mathbf{D}$ with $m = 4$ obtained by excluding the 3rd, 9th, 15th, and 21st runs, namely, 1,324, 2,314, 3,214, and 4,213 from the full OofA design. Under model [5], $\mathbf{D}$ has a A-criterion of 39.9. Now consider relabeling the components via 4,132, that is, components 1, 2, 3, 4 are now called components 4, 1, 3, 2, respectively. The relabeled design has a different A-criterion that equals 46.3. Consequently, practitioners have to be cautious about the labeling when using a design under model [5]. Such inconvenience of labeling is even more serious when fitting a reduced model (after screening out the unimportant components).

Remark 5. The A-, E-, and M.S.-criteria are typically altered after relabeling. Interestingly, the D-criterion is invariant under relabeling in all the examples we have studied.

Last but not least, as can be expected, the asymmetry causes some mathematical difficulty in establishing the conditions as well as properties of optimal designs under this model. The knowledge of optimal designs is important because it provides benchmarks for evaluating any design. The symmetric model proposed below, as a reparametrization of model [5], will facilitate the theory of optimal designs.

The symmetric interaction model

To transform the two interactions in model [5], so that the transformed interaction terms are symmetric about different components, a proposed model is as follows.

$$\tau(\boldsymbol{a}) = \beta_0 + \sum_{1 \le j < k \le m} \beta_{jk} z_{jk}(\boldsymbol{a}) + \sum_{1 \le j < k < l \le m} \left[\beta^{(1)}_{jkl} T^{(1)}_{jkl}(\boldsymbol{a}) + \beta^{(2)}_{jkl} T^{(2)}_{jkl}(\boldsymbol{a}) \right], \quad [6]$$

where

$$\left[T^{(1)}_{jkl}(\boldsymbol{a}), T^{(2)}_{jkl}(\boldsymbol{a}) \right] = \begin{cases} (1, 0) & \text{if } jkl \text{ or } lkj \text{ in } \boldsymbol{a}, \\ (-1/2, \sqrt{3}/2) & \text{if } jlk \text{ or } klj \text{ in } \boldsymbol{a}, \\ (-1/2, -\sqrt{3}/2) & \text{if } kjl \text{ or } ljk \text{ in } \boldsymbol{a}. \end{cases}$$

Here "jkl in $\boldsymbol{a}$" means "j precedes k and k precedes l in $\boldsymbol{a}$". It is easy to show that $T^{(1)}_{jkl} = 1/4 + 3/4 z_{jk} z_{kl}$

Table 7. Interaction terms of the symmetric interaction model (6) under the full design ($m = 3$).

Order $\boldsymbol{a}$	$T^{(1)}_{jkl}(\boldsymbol{a})$	$T^{(2)}_{jkl}(\boldsymbol{a})$
123	1	0
132	$-1/2$	$\sqrt{3}/2$
213	$-1/2$	$-\sqrt{3}/2$
231	$-1/2$	$\sqrt{3}/2$
312	$-1/2$	$-\sqrt{3}/2$
321	1	0

and that $T^{(2)}_{jkl} = -\sqrt{3}/4 + \sqrt{3}/2 z_{jk}z_{jl} - \sqrt{3}/4 z_{jk}z_{kl}$. Hence model [6] is a *linear coding* of model [5]. The modified model is symmetric in many senses. First, it is relabeling-invariant. For any design under model [6], its moment matrix changes by only a orthogonal transformation after any relabeling. Thus, conventional design criteria such as the *D*-, *A*-, *E*, M.S.-criteria are invariant under any relabeling. A detailed proof is not given here—the main idea follows the proof of Lemma A1 in Peng, Mukerjee, and Lin (2017).

Compared with model [5], model [6] also improves the overall estimation/prediction accuracy in some sense. For one thing, the "symmetric interaction" terms $T^{(1)}_{jkl}$ and $T^{(2)}_{jkl}$ are now balanced under the full design (cf. the unbalancedness of interactions in Table 6). Here a factor is called balanced if the inner product equals 0 between the column of this factor and the intercept column in the design. See Table 7 for illustration. Such a balancedness is favorable for the statistical inference. More properties of model [6] in the estimation/prediction efficiency shall be explored in the near future.

The symmetry of model [6] also facilitates the theory of optimal design. Under this model, we have obtained the same optimality result as Theorem 1. Explicitly, the full OofA design attains the *D*-, *A*-, *E*, *and M.S.*-optimum under model [6]. Meanwhile, a fractional design is *D*-, *A*-, *E*, and M.S.-optimal under model [6] only if it attains the same moment matrix as the full design. These can be verified using the method of proving Theorem 1 in Peng, Mukerjee, and Lin (2017) together with the relabeling-invariance property of this model as mentioned above. Note that the *D*-optimality is invariant under coding schemes (i.e., linear transformations of variables). Thus our result in turn proves that the full design is *D*-optimal under model [5].

Model [5] allows straightforward interpretations on the coefficients; see the example in Section 2.4 of Mee (2017).[AQ4] On the contrary, model [6] appears not easy to interpret, since the variables $T^{(1)}_{jkl}$ and $T^{(2)}_{jkl}$ are artificially defined. Nevertheless, this is not a critical concern in practice, as explained below. A major purpose of using the three-way model is to detect "which triplets indicate useful interaction terms". When analyzing model [5], we shall first detect overall significance of $[\beta^{(1)}_{jkl}, \beta^{(2)}_{jkl}]$ for each triplet *jkl*. There is no need to interpret the single coefficient $\beta^{(1)}_{jkl}$ or $\beta^{(2)}_{jkl}$ itself. Of course, after identifying any significant triplet, say *jkl*, one can conduct a post-hoc analysis to evaluate the linear contrasts regarding difference precedence patterns among the triplet *jkl*. These contrasts should be estimated via a linear combination of $\beta^{(1)}_{jkl}$ and $\beta^{(2)}_{jkl}$.

As a summary, a symmetric interaction model [6] is proposed as a modification of the simple three-way model [5]. Compared to model [5], model [6] has several advantages as it (i) possesses the relabeling-invariance, (ii) improves the estimation/prediction efficiency in some aspects, and (iii) facilitates the design theory.

The s-way OofA model under the constrained linear regression framework

Consider the definition of *s*-way models with $s \geq 4$. Mee (2017) proposed a general framework which defines an *s*-way model by adding $df_{s|s-1}$ terms into the $(s-1)$-way model, where $df_{s|s-1}$ is from Rencontres series (Riordan 2012). The added terms are chosen from all the products of PWO indicators that involve *s* components in total. This framework is quite appealing as it decomposes $m!$ into the degrees of freedoms for the two-way (i.e., pairwise), three-way, up to the *m*-way effects, respectively. However, Mee's definition of the *s*-way model is not fully justified before the explicit choices of the $df_{s|s-1}$ add terms are given and the resulting model is proved to be estimable. The justification of this model is rather difficult, in view of the complicated confoundedness among the products of the PWO indicators. For example, the confoundedness relative to the four-way model includes $z_{12}z_{13}z_{14} - z_{12}z_{23}z_{24} + z_{13}z_{23}z_{34} - z_{14}z_{24}z_{34} = 0$ and $z_{12}z_{13}z_{14} - z_{12}z_{13}z_{24} + z_{13}z_{14}z_{24} - z_{13}z_{23}z_{24} + z_{13}z_{23}z_{34} - z_{13}z_{24}z_{34} = 0$. So far, we have not seen any explicit four-way or higher-way model in the literature.

While Mee's framework for the *s*-way model is definitely worth investigating, here we propose a different framework, which models the *s*-way effects in a constrained linear regression. We first introduce some concepts. For any $2 \leq s \leq m$ and any $1 \leq j_1, j_2, \ldots, j_s \leq m$, define the *s*-way groups

$$\mathcal{A}_{j_1 j_2 \ldots j_s} = \{\boldsymbol{a} \in \mathcal{A} : j_t \text{ precedes } j_{t+1} \text{ in } a \text{ for any } 1 \leq t \leq s-1\}. \quad [7]$$

Such groups are partitioned by the precedence patterns of a subset of *s* components. Define the *s*-way

group means

$$\mu(\mathcal{A}_{j_1 j_2 \ldots j_s}) = \sum_{\boldsymbol{a} \in \mathcal{A}_{j_1 j_2 \ldots j_s}} \tau(\boldsymbol{a}) / |\mathcal{A}_{j_1 j_2 \ldots j_s}| = \sum_{\boldsymbol{a} \in \mathcal{A}_{j_1 j_2 \ldots j_s}} \tau(\boldsymbol{a}) / (m!/s!). \quad [8]$$

The proposed s-way order-of-addition model can be written as

$$\tau(\boldsymbol{a}) = \mu(\mathcal{A}) + \sum_{1 \le j_1, j_2, \ldots, j_s \le m} I(\boldsymbol{a} \in \mathcal{A}_{j_1 j_2 \ldots j_s}) \cdot \left[\mu(\mathcal{A}_{j_1 j_2 \ldots j_s}) - \mu(\mathcal{A}) \right], \quad [9]$$

The parameters in [9] include $\mu(\mathcal{A}_{j_1 j_2 \ldots j_s})$'s for any mutually different $j_1, j_2, \ldots, j_s$, as well as $\mu(\mathcal{A})$–thus $P_s^m + 1$ parameters in total. These parameters are subject to some natural constraints, as described below.

The constraints on the parameters are from the partitioning structure of the s-way groups defined in [7]. For any $2 \le r < m$, we have:

$$\mathcal{A}_{k j_1 \ldots j_r} \cup \mathcal{A}_{j_1 k j_2 \ldots j_r} \cup \ldots \cup \mathcal{A}_{j_1 \ldots j_{r-1} k j_r} \cup \mathcal{A}_{j_1 \ldots j_r k} = \mathcal{A}_{j_1 \ldots j_r}, \quad [10]$$

for any mutually different $j_1, \ldots, j_r$ and k. This results in the following constraints on the group means:

$$\mu(\mathcal{A}_{k j_1 \ldots j_r}) + \mu(\mathcal{A}_{j_1 k j_2 \ldots j_r}) + \ldots + \mu(\mathcal{A}_{j_1 \ldots j_{r-1} k j_r}) + \mu(\mathcal{A}_{j_1 \ldots j_r k}) = (r+1)\mu(\mathcal{A}_{j_1 \ldots j_r}). \quad [11]$$

We call Eq. [11] as the rth hierarchy of constraints. For $1 \le j \le m$, we define $\mathcal{A}_j = \mathcal{A}$ for convenience (so that the 1st hierarchy of constraints is well-defined). Since Eq. [11] holds for any mutually different $j_1, \ldots, j_r, k$, there are P_{r+1}^m constraints in the rth hierarchy.

As an illustration, the first hierarchy of constraints says

$$\mu(\mathcal{A}_{jk}) + \mu(\mathcal{A}_{kj}) = 2\mu(\mathcal{A}), \quad [12]$$

and the second hierarchy says

$$\mu(\mathcal{A}_{ljk}) + \mu(\mathcal{A}_{jlk}) + \mu(\mathcal{A}_{jkl}) = 3\mu(\mathcal{A}_{jk}). \quad [13]$$

and

$$\mu(\mathcal{A}_{lkj}) + \mu(\mathcal{A}_{klj}) + \mu(\mathcal{A}_{kjl}) = 3\mu(\mathcal{A}_{kj}). \quad [14]$$

Note that by inserting Eqs. [13] and [14] into [12], any two-way group means will be eliminated. In other words, combining the first and second hierarchies yields a linear system that includes only the three-way group means $\mu(\mathcal{A}_{j_1 j_2 j_3})$'s and the grand mean $\mu(\mathcal{A})$ as variables. In general, combining the first up to $(s-1)$th hierarchies of constraints yields a linear system involving only the s-way group means and the grand mean—we define this as the s-way cumulative-constraints linear system.

In summary, the s-way order-of-addition model is a constrained linear regression defined by Eq. [9] and the s-way cumulative-constraints linear system. We have clearly identified all these constraints, so that the models can be fitted using the constrained least squares method. The proposed framework is interpretable as it incorporates the P_s^m group means and the grand mean as parameters to estimate. Of course, this model needs to be improved for practical use. For one thing, this model may be less useful for large m as the number of variables increases quickly as m increases. The main purpose of this subsection is to provide some insights into future theories. A potential merit of the proposed s-way models is that they are symmetric (i.e., relabeling-invariant) to different components. This may facilitate the theory as well as the construction of optimal designs. We believe, in addition, that the constraints described above reveal some essential features of the order-of-addition problem.

Concluding remarks

Order-of-addition (OofA) experiments have been widely used in many scientific and industrial studies. A successful OofA experiment can help experimenters find out an optimal order. Such experiments also help in identifying the significant order effects which reveal some underlying scientific mechanisms. Design of OofA experiments is needed to select representative orders, testing which yields efficient estimates of order effects under the control of experimental costs. This article introduces many latest results as well as some new thoughts, regarding the design and model of the OofA experiment. Specifically, we have reviewed the pairwise-order (PWO) model and its relevant optimality theorems. The construction methods of PWO designs have been thoroughly reviewed, including the construction of optimal PWO designs, minimal-point PWO designs, and nearly-optimal PWO designs with small number of runs. We have further introduced several extensions of the PWO model, including two forms of the three-way model and the general s-way model. Their pros and cons have been discussed. The current results/thoughts under each model have been summarized. We believe that this review can help practitioners understand, implement, and utilize OofA designs. This review also provides some suggestions for statistical research on this topic. For example, fruitful future work is anticipated in the design and analysis under higher-order models (see More order-of-addition models section).

We refer interested readers to Voelkel (2017), Mee (2017), and Yang, Sun, and Xu (2017) for more results (which we have not covered here) on the OofA problem. For instance, Voelkel (2017) introduced the χ^2 criterion for PWO designs to measure the design efficiency of estimating higher-order effects; Mee (2017) introduced a real example as well as a simulated example of OofA experiments; and Yang, Sun, and Xu (2017) established the connection between difference matrix and the OofA design.

The OofA problem has recently become popular in the field of design of experiment. Besides the research topics covered in this article, there have been several more directions of tackling this problem. One direction of future work is to explore and extend the "position-effect" model proposed by Yang, Sun, and Xu (2017). They define the predictors in an OofA model via the positions of all components, rather than the precedence patterns among them. It is notable that they have obtained, via a smart algorithm, designs that are (nearly)-optimal under the position effect model and highly efficient under the PWO model. Construction of designs that are robust to different models is worth studying. Another future work is on the simultaneous design and inference of both the OofA effects and the real-factor effects. This topic is strongly motivated by the practical needs in chemistry, food science, and related areas. In chemical experiments, the formation of product depends on both the addition order and the amounts of each reactants. Many designs (such as fractional factorial designs) are available to control the levels of different reactants. Then how to combine such designs with our OofA designs, so as to estimate the OofA effects and the real-treatment effects simultaneously? We look forward to new models and designs along this line.

About the authors

Dennis K.J. Lin is a university distinguished professor of supply chain and statistics at Penn State University. His research interests are quality assurance, industrial statistics, data mining, and response surface. He has published near 200 SCI/SSCI papers in a wide variety of journals. He currently serves or has served as associate editor for more than 10 professional journals and was co-editor for Applied Stochastic Models for Business and Industry. Dr. Lin is an elected fellow of ASA, IMS and ASQ, an elected member of ISI, a lifetime member of ICSA, and a fellow of RSS. He is an honorary chair professor for various universities, including Renmin University of China (as a Chang-Jiang Scholar), Fudan University, The University of Hong Kong, and National Chengchi University (Taiwan). His recent awards including, the recipient of the 2004 Faculty Scholar Medal Award (Penn State), the Youden Address (ASQ, 2010), the Shewell Award (ASQ, 2010), the Don Owen Award (ASA, 2011), the Loutit Address (SSC, 2011), the Hunter Award (ASQ, 2014), and the Shewhart Medal (2015).

Jiayu Peng is a PhD student at the Department of Statistics, Penn State University. His thesis advisor is Dr. Dennis K.J. Lin. Mr. Peng received a bachelor's degree in Mathematics from Tsinghua University, China. Mr. Peng's research interest includes design of experiments, non-parametric statistics, and quantile regression. He has received the Honorable Mention in JSM 2017 for the SPES Outstanding Presentation Awards, and the Student Scholarships for JSM 2015 SPES sponsored roundtable, and the 2011 First-class scholarship at Tsinghua University.

References

Fedorov, V. V. 1972. *Theory of optimal experiments*. Amsterdam, The Netherlands: Elsevier.

Fisher, R. A. 1937. *The design of experiments*. Edinburgh, London: Oliver and Boyd.

Fuleki, T., and F. J. Francis. 1968. Quantitative methods for anthocyanins. *Journal of Food Science* 33 (3):266–274. doi:10.1111/j.1365-2621.1968.tb01365.x.

Jourdain, L. S., C. Schmitt, M. E. Leser, B. S. Murray, and E. Dickinson. 2009. Mixed layers of sodium caseinate + dextran sulfate: Influence of order of addition to oil-water interface. *Langmuir* 25 (17):10026–10037. doi: 10.1021/la900919w.

Karim, M., K. McCormick, and C. T. Kappagoda. 2000. Effects of cocoa extracts on endothelium-dependent relaxation. *The Journal of Nutrition* 130 (8S Suppl):2105S–2108S.

Mee, R. W. 2017. *Order-of-addition modeling*. Submitted for publication.

Meyer, R. K., and C. J. Nachtsheim. 1995. The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* 37 (1):60–69. doi:10.1080/00401706.1995.10485889.

Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek. 1994. Fastdnaml: A tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Computer Applications in the Biosciences* 10 (1):41–48. doi:10.1093/bioinformatics/10.1.41.

Peng, J., R. Mukerjee, and D. K. Lin. 2017. Design of order-of-addition experiments. *Biometrika*, forthcoming.

Rajaonarivony, M., C. Vauthier, G. Couarraze, F. Puisieux, and P. Couvreur. 1993. Development of a new drug carrier made from alginate. *Journal of Pharmaceutical Sciences* 82 (9):912–917.

Riordan, J. 2012. *Introduction to combinatorial analysis*. Chelmsford, MA: Courier Corporation.

Shinohara, A., and T. Ogawa. 1998. Stimulation by rad52 of yeast rad51-mediated recombination. *Nature* 391 (6665): 404–407.

Silvey, S. 2013. *Optimal design: an introduction to the theory for parameter estimation*, vol. 1. Berlin, Germany: Springer Science & Business Media.

Stewart, C. A., D. Hart, D. K. Berry, G. J. Olsen, E. A. Wernert, and W. Fischer. 2001. Parallel implementation and performance of fastdnaml-a program for maximum likelihood phylogenetic inference. In *Supercomputing, ACM/IEEE 2001 Conference*, 32–32. IEEE.

Tarjan, R. 1972. Depth-first search and linear graph algorithms. *SIAM Journal on Computing* 1 (2):146–160. doi: 10.1137/0201010.

Van Nostrand, R. C. 1995. *Design of experiments where the order of addition is important. In ASA Proceedings of the Section on Physical and Engineering Sciences*, 155–160. Alexandria, VA: American Statistical Association.

Voelkel, J. G. 2017. The design of order-of-addition experiments. Submitted for publication.

Yang, J., F. Sun, and H. Xu. 2017. Component orthogonal arrays for order-of-addition experiments. Submitted for publication.

Zhao, Y., D. K. Lin, and M. Liu. 2017. Design of order of addition experiment. Submitted for publication.