Statistica Sinica Preprint No: SS-2023-0357	
Title	Complete Consecutive Order-Pairing (CCOP) Design and Its Distance-based Linear Model: Design Construction and Analysis for Order-of-Addition (OofA) Experiments
Manuscript ID	SS-2023-0357
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202023.0357
Complete List of Authors	Jing-Wen Huang and Frederick Kin Hing Phoa
Corresponding Authors	Frederick Kin Hing Phoa
E-mails	fredphoa@stat.sinica.edu.tw

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Complete Consecutive Order-Pairing (CCOP) Design and Its Distance-based Linear Model: Design Construction and Analysis for Order-of-Addition (OofA) Experiments

Jing-Wen Huang and Frederick Kin Hing Phoa

Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract: An order-of-addition (OofA) experiment investigates how the sequence of input factors influences the experimental response. This type of experiment has recently gain significant interest among practitioners in clinical trials and industrial processes. In this work, we introduce a new cost-efficient design called the Complete Consecutive Order-Pairing (CCOP) design. The CCOP design not only considers the effects of the component order on the response but also simultaneously accounts for the effects due to the component levels. We also propose a new statistical model associated with the CCOP design for identifying the optimal settings of both component order and levels. The CCOP design method evaluates the effects of two successive treatments by using the minimal number of runs, as each pair of level settings for two different components appears exactly once. Compared to recent studies on OofA experiments, our design effectively handles pure order experiments and multi-level experiments with a relatively

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

small run size.

Key words and phrases: Clinical Trials, Cost-Efficiency, Order-of-Addition Experiments

1. Introduction

The concept of order-of-addition (OofA) experiments arises in fields such as biochemistry, industry, agriculture, among many others. Different arrangements of components in an experiment can lead to varying effects on the response. The optimal order results in the best possible use of the components, yielding better prognoses, more stable products, or higher profits in different applications. Recent examples of using order effects in prognosis experiments are provided by Ding et al. (2015) and Bashkirtseva et al. (2021), both concluding that the sequential use of certain therapies contributed to the response.

To incorporate order effects in experimental analyses, OofA experiments have resurfaced among researches in the design and analysis of experiments. A major application of OofA experiments is to investigate how different orders of medicine intakes significantly affect the response value, enabling the determination of the optimal order of medicine intakes. However, this treatment involves not only the order of medicine intakes but also the dosage.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Among all researches on OofA design and analysis, only Rios et al. (2022), Xiao et al. (2022), and Tsai (2023) considered both order and level. Their designs exhibit good statistical properties only when the run size is affordable for the experimenter. However, in many real-world cases in biomedical sciences and clinical trials, the number of observations is often limited, necessitating a design that requires a small run size to identify the optimal component orders and levels.

The main contribution of this work is the introduction of a new class of cost-efficient plans and the corresponding analysis procedure for OofA experiments. This enables the collection and summarization of information on all possible successive component-ordered pairs, regardless of the number of component levels involved, using minimal experimental resources. The rest of the paper is arranged as follows. Section 2 reviews all existing methods for OofA experiments and introduces the definitions, notation, and mathematical tools underlying our work. Section 3 provides a model that considers all possible successive component pairs. Section 4 presents the main results, including the general solution of starting sequences for constructing our designs for a number of levels $k > 1$, depending on the numbers of components (even or odd) and the numbers of levels (even or odd). Section 5 compares the simulations and data analysis results for

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

$k > 1$ and $k = 1$, demonstrating that our design can handle both cases involving components and their levels as well as pure order arrangements. A summary, along with future extensions, is provided in Section 6. All proofs and large-scale examples are included in the supplementary material.

2. Reviews and Preliminaries

2.1 Literature Reviews

Generally speaking, there are three main modeling approaches for OofA experiments: pairwise-order (PWO), component-position (CP), and Gaussian Process (GP). The PWO approach considers the relative positions of components in the ordering sequence, while the CP approach focuses on the absolute positions of components. The GP approach can be generalized from both PWO and CP factors. Optimal designs are derived for each model, with a comprehensive review available in Stokes (2021). Below, we provide a brief overview of these models and their favorable designs.

A Review of PWO Modeling and Its Favorable Design. Following the pseudo-factor approach introduced in Van Nostrand (1995), we recall the general framework of a PWO favorable design. Suppose there are m components (e.g, drugs, reactants). Each pair of components (i, j) is considered a factor z_{ij} in the design. We denote its value for a run in an

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

experiment as 1 if j is applied after i and -1 if i is applied after j. This results in a two-level factorial design in which each factor indicates the order of two components. Peng, Mukerjee and Lin (2019) provided a sufficient and necessary condition for a (fractional) design to be optimal under PWO model.

Although the properties of PWO favorable designs are good from a theoretical perspective, the desired optimal design properties are lost when there are not enough budget. To overcome this obstacle, Zhao, Lin and Liu (2021) introduced another approach via the minimal-point (MP) design. It reduces the experimental cost by using only $\binom{m}{2}+1$ runs. Wang and Mee (2022) also proposed a supersaturated Bayesian D-optimal design constructed by the Federov exchange algorithm.

For cases involving both component order and component level, Voelkel (2019) and Mee (2020) considered an additive model for the order effects and level effects; the model is $\hat{Y} = \beta_0 + \sum \beta_{ij} Z_{ij} + \sum \beta_k X_k$, where the Z_{ij}s are PWO effects and the X_ks are component-level effects. Tsai (2023) further proposed dual-orthogonal arrays (DOAs), which combined cross arrays of order-of-addition orthogonal arrays (OofA-OAs) and the classical orthogonal array to construct optimal designs for such additive models.

A Review of CP Modeling and Its Favorable Design. Yang, Sun

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

and Xu (2021), Stokes (2021) and Stokes and Xu (2022) followed nominal CP and position-based models to address the effects of positions. The effects $X_c^{(j)}$ in the CP model are coded by 1 if component c is at position j, and the effects in the position-based model are considered to be evenly spaced time points, which is a quantitative method. Position-based models include 1st-order, quadratic and 2nd-order models with orthogonal polynomials of the time points to apply components. The corresponding favorable designs for nominal CP, the component orthogonal arrays (COAs), have strength 2 so that for any two columns, all pairs of components appear exactly once. The corresponding favorable design for position-based models, $F_{n,m}$, where n is the run size and m is the number of components, is a generalized minimum aberration design for n runs where $m \leq n \leq m(m-1)$. This approach is useful when the absolute positions of components are important.

A Review of GP-type Modeling and Design. In addition to PWO and CP factors, there is another class of model that depends on GP regression. Xiao and Xu (2021) first considered the generalization of both PWO and CP factors to GP regression. Xiao et al. (2022) further generalized the models on the basis of both order effects and dose-level effects and called them MaGP models. Order effects can be of PWO or CP type, and dose-level effects are considered in the traditional quantitative setting. They

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

called the factors, which have both order and level effects on the response, the quantitative-sequence (QS) factors, and proposed a QS design that has a certain best space-filling property.

Problems. These approaches are useful for interpretation, but most of them need a large run size to preserve desirable design properties and estimation stability. As an alternative to existing models, we start from the steps of the original data analysis scheme: defining important information conceptually, collecting data, modeling important information, analyzing data, and making decisions. We define the important information of the order to be the consecutive pieces in the order. To reduce the run size, we focus on the smallest pieces, which are the consecutive pairs. For example, the order (1,2,4,3) has the smallest pieces (1,2), (2,4) and (4,3). Therefore, the problem becomes finding a design that has the smallest run size and covers all possible smallest paired pieces. We propose a design with a small run size of mk^2, where m is the number of components and k is the number of component levels. This design is supersaturated under all existing models, so we also propose a new model to analyze such small experiments based on a Distance-based Linear Model (DBLM) Cuadras and Arenas (1990) that will be introduced later in the next section.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

2.2 Definitions, Notation and Existing Related Tools

We provide the definitions required for this problem. In the remaining of the paper, m and k are the number of components and the number of their levels, respectively. We assume that all components have same number of levels. The mixed-level cases are outside the scope of our paper and will be considered as a future extension. We define the equivalence class $F_i = \{i, m+i, \cdots, (k-1)m+i\}$ for each k-level component i, $i = 1, ..., m$. Any two elements a and b in the same equivalence class are said to be congruent modulo m and denoted as $a \equiv b \pmod{m}$. If $a \neq b$ but $a \equiv b \pmod{m}$, then they represent the same component but are at different levels.

We define a run of an experiment as an m-ordered sequence formed by elements in $\cup_{i=1}^{m} F_i$. A design is valid for OofA experiments only if all its runs go through each component exactly once; that is, each run is a permutation of the m components. Mathematically, we can say that for any run $(a_1, \cdots, a_m)$, there is a bijection from $\{a_i, i = 1, \cdots, m\}$ to $\{F_i, i = 1, \cdots, m\}$ such that each a_i belongs to exactly one equivalence class. Given a run $\vec{a} = (a_1, \cdots, a_m)$, we denote a run difference vector of length $m-1$ as $\Delta\vec{a} = (a_2 - a_1, a_3 - a_2, \cdots, a_m - a_{m-1})$. Rephrased in terms of the above notation, this paper aims to construct a class of valid designs

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

$D = [D_{i,j}]$ that satisfies the following definition.

Definition 1. A valid design D is called a k-layer complete consecutive order-pairing (CCOP) design if

(i) for all (a, b), where $1 \leq a, b \leq km$ and $a \not\equiv b \pmod{m}$, there is exactly the same number of pairs of (i, j) such that $(a, b) = (D_{i,j}, D_{i,j+1})$;

(ii) the run size of D is the minimum among all valid designs satisfying property (i).

Example 1. The 1-layer CCOP design

$$\begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 3 & 1 & 4 \\ 3 & 4 & 2 & 1 \\ 4 & 1 & 3 & 2 \end{pmatrix}$$

contains all possible different pairs $(1,2), (1,3), (1,4), (2,1) \ldots, (4,3)$ exactly once.

Remark. The first condition in the definition when $k = 1$ is identical to the balanced crossover design in Dean et al. (2015) and the pair-balanced design in Xiao et al. (2022). In addition, some COA's are unions of $\frac{m}{n}$ 1-layer designs, where n is the run size. However, a k-layer CCOP design in which

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

$k > 1$ is different from these because of the structure of equivalence classes that use the same components. The feature of such layer-type structures results in a new constraint on these numbers to make a valid run. For example, when $m = 4$ and $k = 2$, we will have 8 possible ways, labelled $1, ..., 8$, to assign components. In particular, $1 \equiv 5 \pmod 4$ belong to component 1, $2 \equiv 6 \pmod 4$ belong to component 2, and so on. If 1 is in a run then 5 can never appear in any position of that run, that is, component 1 can appear only once in a valid run. The same argument can be applied to components 2, 3 and 4.

Our main purpose is to give a simple way to construct a k-layer CCOP design for any number of components m and number of levels k. Such constructions for $k = 1$ are provided in Williams (1949) and Tillson (1980). For $k > 1$, we propose some k-layer CCOP designs inspired by Williams (1949). For notational simplicity in the lemmas and theorems in the later sections, we introduce a function, denoted as $f(i, a, b)$, that returns the value a if i is odd and b if i is even.

2.3 1-layer CCOP Designs: Construction

Williams (1949) provided a construction method for 1-layer CCOP designs, applicable to pure order problems. Before stating the construction from

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Williams (1949), we define the Latin square and the cyclic addition procedure.

Definition 2. A Latin square A of order m is an $m \times m$ array in which every entry x in A is an image of a mapping $L : \{1, \ldots, m\} \times \{1, \ldots, m\} \to \{1, \ldots, m\}$. In this mapping, any row or column in A is a permutation of $\{1, \ldots, m\}$, where the first and second dimensions of the domain represent the row and column positions of x in A respectively.

Definition 3. A matrix is said to be constructed via cyclic addition procedure if each row is obtained from its preceding row $+1$ under modulo m. The first row is commonly called the base block.

Example 2. The 1-layer CCOP design in Example 1 can be constructed via a cyclic addition procedure. In specific, the second row $2 \to 3 \to 1 \to 4$ is obtained from the first row $1 \to 2 \to 4 \to 3$ by adding one to every component under modulo 4. The third and forth rows are obtained in the same way.

To construct a 1-layer CCOP design for even m, Williams (1949) suggested using a base block $\vec{r_1} = (1, 2, m, 3, m-1, 4, m-2, \cdots, \frac{m}{2}+1)$. The remaining rows are then obtained via a cyclic addition procedure $r_1 + a \times (1, \cdots, 1) \pmod{m}$, $1 \leq a \leq n-1$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Due to its cyclic addition structure, if we subtract one column of the design matrix from another, the resulting difference will be a constant vector under modulo m. For example, in the 1-layer CCOP design of order 4, the difference between the second and first column is $(1, 1, 1, 1)$. TTherefore, if the base block consists of all possible differences induced by all possible pairs of components, the design generated by that base block must be a 1-layer CCOP design. We describe all base blocks in the form of difference vectors and restate the above result in the following lemma.

Lemma 1. *(Williams, 1949). Let m be even. An $m \times m$ Latin square constructed by the cyclic addition procedure is a 1-layer CCOP design if the difference vector is defined by $\Delta\vec{r_1}[i] = f(i, i, m - i)$.*

For odd m greater than 5, Tillson (1980) proved the existence of a 1-layer CCOP design and provided an exact construction without the cyclic addition structure. For example, we can express a design matrix for any

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

$m = 4p + 3$ with positive integer p as shown below.

$$\begin{pmatrix} 4p+2 & 1 & 4p+1 & \cdots & 2p+2 & 2p+1 & 4p+3 \\ p+1 & \cdots & 2p+2 & 4p+3 & 1 & \cdots & 3p+4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ p-1 & \cdots & 2p & 4p+3 & 4p+2 & \cdots & 3p+2 \\ 4p+3 & 4p+2 & 2p-1 & 4p-2 & \cdots & 3p+4 & p+1 \end{pmatrix}.$$

This design does not have a cyclic addition structure, so it is outside the scope of this work. Readers interested in such 1-layer CCOP designs are referred to Tillson (1980) for details. In the rest of this paper, all 1-layer CCOP designs are constructed via the cyclic addition procedure.

When m is odd, it is impossible to construct a 1-layer CCOP design using the cyclic addition procedure. Recall in the lemma below, Williams (1949) proposed a base block that can have every possible pair of components appear exactly twice. The resulting designs are no longer 1-layer CCOP designs, but they are useful for constructing designs when $k > 1$.

Lemma 2. *(Williams, 1949). Let m be odd. A $2m \times m$ Latin rectangle constructed by the cyclic addition procedure satisfies the first condition of 1-layer CCOP design if the difference vectors of two non-overlapping $m \times m$ submatrices are $\Delta\vec{r_1}[i] = f(i, i, m-i)$ and $\Delta\vec{r_2}[i] = f(i, m-i, i)$.*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

3. New Model for k-layer CCOP Designs

The model matrix of 1-layer CCOP designs can be recast as a supersaturated variant of both PWO and CP models, but there always exist pairs of candidate components that are fully aliased with each other due to the mathematical structure of our designs. For example, the design in Example 1 corresponds to the model matrix

$$\begin{pmatrix} Z_{12} & Z_{13} & Z_{23} & Z_{14} & Z_{24} & Z_{34} \\ 1 & 1 & 1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \end{pmatrix}$$

in PWO modeling and corresponds to

$$\begin{pmatrix} X_2^{(1)} & X_2^{(2)} & X_2^{(3)} & X_3^{(1)} & X_3^{(2)} & X_3^{(3)} & X_4^{(1)} & X_4^{(2)} & X_4^{(3)} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

in CP modeling. Similar to Z_{12}, Z_{13}, and Z_{34} in the PWO model matrix, or $X_2^{(1)}$ and $X_3^{(2)}$ in the CP model matrix, they are both fully aliased with each other in groups. It can be proved that full aliasing is inevitable, so we propose a new model suitable for our k-layer CCOP designs with such small run sizes.

Since our goal is to predict the largest or smallest response, we use the Distance-based Linear Model (DBLM) as formulated in Cuadras and Arenas (1990). The procedure is as follows: (1) Metric Definition: For the set of all permutations S, define a metric $d : (S, S) \to \mathbb{R}^+ \cup \{0\}$. (2) Model Matrix Construction: For any d, the columns of the model matrix are defined as $d(x_i, x)$ for $i = 1, \cdots, n$, where the x_is are the design points selected from S and $x \in S$. Dimension reduction on the columns can be performed at this step if needed. (3) Model Fitting: Fit the model using $\hat{\beta} = (X^T X)^{-1} X^T Y$ $\hat{Y} = X^* \beta$. Here, X is the distance matrix (or reduced-dimensional distance matrix) of the n observations, and X^* represents the distance measurement between the points we want to predict and the n observations $x_1, \ldots, x_n$.

A major concern with this model class is defining a suitable d to measure the discrepancy between elements in S. For $k = 1$, we define $d(x, y) = (m - 1) - s(x, y)$, where $s : (S, S) \to \mathbb{R}^+ \cup \{0\}$ counts the number of pairs

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

that the two sequences have in common within their orders. For example, if 1243 is a design point, then $s(1243, 1234) = 1$ because they both have the pair 12, and $s(1243, 3124) = 2$ because they both have the pairs 12 and 24. We consider only pair similarities because consecutive pairs are the smallest units that affect the response. Here is an example of a model matrix X:

$$\begin{pmatrix} 0 & 3 & 3 & 3 \\ 3 & 0 & 3 & 3 \\ 3 & 3 & 0 & 3 \\ 3 & 3 & 3 & 0 \end{pmatrix}$$

An example of a model matrix X^* for all possible permutations can be characterized in a transposed form as follows:

$$\begin{pmatrix} 2 & 0 & 2 & 3 & 3 & 2 & 3 & 2 & 3 & 3 & 2 & 1 & 1 & 3 & 3 & 2 & 2 & 3 & 2 & 3 & 3 & 3 & 1 & 2 \\ 2 & 3 & 3 & 3 & 1 & 2 & 3 & 2 & 0 & 2 & 3 & 2 & 2 & 1 & 2 & 3 & 3 & 3 & 2 & 3 & 3 & 1 & 2 & 3 \\ 2 & 3 & 3 & 1 & 2 & 3 & 1 & 2 & 3 & 2 & 3 & 3 & 3 & 2 & 2 & 3 & 2 & 0 & 3 & 3 & 1 & 2 & 3 & 2 \\ 3 & 3 & 1 & 2 & 3 & 2 & 2 & 3 & 3 & 2 & 1 & 3 & 3 & 3 & 2 & 1 & 2 & 3 & 2 & 0 & 2 & 3 & 3 & 2 \end{pmatrix}.$$

If the effects of triples or larger groups of elements are considered important, our metric can be defined based on coinciding fragments of length 3 or other higher lengths. While reasonable metrics can be constructed to fit case-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

specific situations, it's essential to also develop other designs that ensure proper analysis properties according to the new model.

For the case where $k > 1$, in addition to assigning 1 for a fully similar pair and 0 for a fully dissimilar pair when $k = 1$, we define partial similarity. In this scenario, $s(x, y)$ may not be an integer. Partial similarity occurs when pairs consist of the same components but differ in their levels. A naive definition of partial similarity involves setting uniform intervals of points in $[0, 1]$. For example, when $m = 4$ and $k = 2$, we have $s(1234, 1243) = 1$, $s(1234, 5243) = 2/3$, $s(1234, 5643) = 1/3$ and $s(2134, 1243) = 0$. Although this definition combines two types of differences, numerical results support the usefulness of this model for differentiating the ranks of responses.

4. Construction of k-Layer CCOP Designs with $k > 1$

Inspired by the results in Williams (1949), this work presents a new and systematic method for constructing k-layer CCOP with $k > 1$. Several scenarios are considered based on the number of components m and the number of levels k. These scenarios include even numbers of components with any levels, odd numbers of components with even levels, and odd numbers of components with odd levels.

Regarding the relationship between the model and the design we pro-

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

pose, it should be noted that our k-layer CCOP design has the smallest run size necessary to capture all the information required by this model. In specific, the rows of the full-model matrix induced by the metric d will satisfy $d(x_1, x) + \cdots + d(x_n, x) = (m-1)^2$ along with n inequalities $d(x_1, x) \leq m-1, \cdots, d(x_n, x) \leq m-1$, where n is the run size of the design matrix, and $x_1, \cdots, x_n$ are the design points. The design points in a k-layer CCOP design correspond to the vertices of a polyhedron that bounds all possible points in the full design. Thus, our k-layer CCOP design is the smallest among all designs that avoid extrapolation predictions.

4.1 k-Layer CCOP Designs with Even Numbers of Components

The following theorem extends Lemma 1 to 1-layer CCOP designs with 2-level components when m is even.

Theorem 1. *When m is even, consider the difference vectors of two base blocks $\Delta\vec{r_1}[i] = f(i, i, m-i)$, and $\Delta\vec{r_2}[i] = f(i, m+i, 2m-i)$. Let 1 be the starting element of the base block. A $2m \times m$ submatrix M_1 is obtained from $\vec{r_1}$ while another $2m \times m$ submatrix M_2 is obtained from $\vec{r_2}$ under the cyclic addition procedure. Then $D = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$ is a $4m \times m$ 2-layer CCOP designs.*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Example 3. To illustrate this result, let us consider $m = 4$ and $k = 2$. Here, the numbers 1, 2, 3, and 4 represent the low-level settings of components 1, 2, 3, and 4, while the numbers 5, 6, 7, and 8 represent the high-level settings of these components. For example, both 1 and 5 represent component 1, which can be determined by checking if their remainders under modulo 4 are the same. Following Theorem 1, the difference vectors of the two base blocks are $\Delta\vec{r_1} = (1, 2, 3)$ and $\Delta\vec{r_2} = (5, 6, 7)$. The base blocks are $\vec{r_1} = (1, 2, 4, 7)$ and $\vec{r_2} = (1, 6, 4, 3)$. These two base blocks generate two 8×4 submatrices

$$M_1 = \begin{pmatrix} 1 & 2 & 4 & 7 \\ 2 & 3 & 5 & 8 \\ 3 & 4 & 6 & 1 \\ 4 & 5 & 7 & 2 \\ 5 & 6 & 8 & 3 \\ 6 & 7 & 1 & 4 \\ 7 & 8 & 2 & 5 \\ 8 & 1 & 3 & 6 \end{pmatrix} \text{ and } M_2 = \begin{pmatrix} 1 & 6 & 4 & 3 \\ 2 & 7 & 5 & 4 \\ 3 & 8 & 6 & 5 \\ 4 & 1 & 7 & 6 \\ 5 & 2 & 8 & 7 \\ 6 & 3 & 1 & 8 \\ 7 & 4 & 2 & 1 \\ 8 & 5 & 3 & 2 \end{pmatrix}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

under the cyclic addition procedure. It is easy to verify that $D = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$ is a 16×4 2-layer CCOP designs.

A larger example with $m = 10$ and $k = 2$ is provided in the supplementary material. We further extend the result of Theorem 1 to k-layer CCOP designs for any k when m is even in the following theorem.

Theorem 2. *For even m, consider the difference vectors of k base blocks $\Delta\vec{r_1}[i] = f(i, i, m-i)$ and $\Delta\vec{r_j} = \Delta\vec{r_1} + m(j-1)$, for $j = 2, \ldots, k$. Let 1 be the starting element of the base blocks. Each $\vec{r_j}$ leads to a different base block, generating k $km \times m$ submatrices M_j under cyclic addition. Then, D, obtained by stacking all M_j vertically for all j, is a $k^2m \times m$ k-layer CCOP design.*

Example 4. Consider a 3-layer CCOP design with ten components. Following Theorem 2, when $j = 1$ and 2, the difference vectors of the first two base blocks are the same as $\Delta\vec{r_1}$ and $\Delta\vec{r_2}$ from Example 3. For $j = 3$, the difference vector of the third base block is $\Delta\vec{r_3} = (9, 10, 11)$. The base blocks are $\vec{r_1} = (1, 2, 4, 7), \vec{r_2} = (1, 6, 12, 7)$, and $\vec{r_3} = (1, 10, 8, 7)$. These three base blocks generate three 12×4 submatrices M_1, M_2, and M_3 under the cyclic addition procedure. Consequently, D, obtained by stacking these nine submatrices, is a 36×4 3-layer CCOP design.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

A larger example for $m = 10$ and $k = 3$ is provided in the supplementary material.

4.2 Even-Level k-Layer CCOP Designs with Odd Numbers of Components

The following theorem extends Lemma 2 to 2-layer CCOP designs when m is odd.

Theorem 3. *When m is odd, the difference vector pairs of base blocks $\vec{r_1}$ and $\vec{r_2}$ are formulated as $\Delta\vec{r_1}[i] = f(i, i, m-i)$ and $\Delta\vec{r_1}[m-i] = f(i, i+m, 2m-i)$ and $\Delta\vec{r_2}[i] = f(i, m-i, i)$ and $\Delta\vec{r_2}[m-i] = f(i, 2m-i, i+m)$ for $i = 1, \ldots, \frac{m-1}{2}$. Let 1 be the starting element of the base blocks. A $2m \times m$ submatrix M_1 is obtained from $\vec{r_1}$ while another $2m \times m$ submatrix M_2 is obtained from $\vec{r_2}$ under the cyclic addition procedure. Then $D = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$ is a $4m \times m$ 2-layer CCOP design.*

An example with $m = 11$ and $k = 2$ is provided in the supplementary material.

We further extend the result of 1 to k-layer CCOP designs for any even number k when m is odd in the following theorem.

Theorem 4. *When m is odd and k is even, the difference vector pairs*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

are formulated as $\Delta\vec{r_j}[i] = f(i, i+2m(j-1), m-i+2m(j-1))$ *and* $\Delta\vec{r_j}[m-i] = f(i, i+m(2j-1), m-i_m(2j-1))$ *and* $\Delta\vec{r_{j+1}}[i] = f(i, m-i+2m(j-1)i+2m(j-1))$ *and* $\Delta\vec{r_{j+1}}[m-i] = f(i, m-i+(2j-1), i+m(2j-1))$ *for* $i = 1, \ldots, \frac{m-1}{2}$ *and odd* j *with* $1 \le j \le k-1$. *Let* 1 *be the starting element of the base blocks. Each* $\vec{r_j}$ *and* $\vec{r_{j+1}}$ *generates one base block, so they form* $\frac{k}{2}$ $km \times m$ *submatrices* M_j *and* $\frac{k}{2}$ $km \times m$ *submatrices* M_{j+1} *under cyclic addition. Then,* D, *obtained by vertically stacking all* M_j*'s and* M_{j+1}, *is a* $k^2m \times m$ k*-layer CCOP design.*

4.3 Odd-Level k-Layer CCOP Designs with Odd Numbers of Components

k-layer CCOP designs with an odd number of odd-level components are among the most challenging to construct across all dimensions. The following theorem outlines the construction method for k-layer CCOP designs when $m \equiv 1 \pmod 4$.

Theorem 5. *For* $m \equiv 1$ *(mod 4) and* $k = 3$, *construct the difference vector trios* $\Delta\vec{r_1}, \Delta\vec{r_1'}$ *and* $\Delta\vec{r_1''}$ *by* $\Delta\vec{r_1}[i] = f(i, i, m-i)$, $\Delta\vec{r_1}[m-i] = f(i, i+m, 2m-i)$, $\Delta\vec{r_1'}[\frac{m-1}{2} - i + 1] = f(i, i, m-i)$, *and* $\Delta\vec{r_1''}[\frac{m-1}{2} - i + 1] = i, i+m, 2m-i$ *for* $i = 1, \ldots, \frac{m-1}{2}$, *and* $\Delta\vec{r_1'}[\frac{m+1}{2}] = \frac{m+1}{2} + 2m$, $\Delta\vec{r_1''}[\frac{m+1}{2}] = \frac{m-1}{2} + 2m$, $\Delta\vec{r_1'}[\frac{m+1}{2} + i] = f(i, m-i+2m, i+2m)$, *and*

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

$\Delta\vec{r''_1}[\frac{m+1}{2}+i] = f(i, i+2m, m-i+2m)$ *for* $i = 1, \ldots, \frac{m-3}{2}$. *Let* 1 *be the starting element of the base blocks. Then,* $\vec{r_1}$, $\vec{r'_1}$, *and* $\vec{r''_1}$ *generate three* $3m \times m$ *submatrices* M_1, M'_1, *and* M''_1, *respectively, under cyclic addition. Then,* D, *obtained by stacking all three submatrices resulting from* $\vec{r_1}$, $\vec{r'_1}$, *and* $\vec{r''_1}$, *is a* $9m \times m$ *3-layer CCOP design.*

Theorem 6. *When* $m \equiv 1$ *(mod 4)* $k > 3$, *consider the difference vector trios for base blocks* $\Delta\vec{r_1}[i] = \Delta\vec{r'_1}[\frac{m-1}{2} - i + 1] = f(i, i, m-i)$ *and* $\Delta\vec{r_1}[m-i] = \Delta\vec{r''_1}[\frac{m-1}{2} - i + 1] = f(i, i+m, 2m-i)$ *for* $i = 1, \ldots, \frac{m-1}{2}$; $\Delta\vec{r'_1}[\frac{m+1}{2}] = \frac{m+1}{2}+2m$, $\Delta\vec{r''_1}[\frac{m+1}{2}] = \frac{m-1}{2}+2m$, $\Delta\vec{r'_1}[\frac{m+1}{2}+i] = f(i, m-i+2m, i+2m)$, *and* $\Delta\vec{r''_1}[\frac{m+1}{2}+i] = f(i, i+2m, m-i+2m)$ *for* $i = 1, \ldots, \frac{m-3}{2}$; *and* $\Delta\vec{r_{j+1}}[i] = f(i, i+jm, m-i+jm)$, $\Delta\vec{r_{j+1}}[m-i] = f(i, i+m(j+1), m-i+m(j+1))$, $\Delta\vec{r_{j+2}}[i] = f(i, m-i+jm, i+jm)$, *and* $\Delta\vec{r_{j+2}}[m-i] = f(i, m-i+jm+1, i+jm+1)$ *for* $i = 1, \ldots, \frac{m-1}{2}$ *and odd* j *with* $3 \leq j \leq k-2$. *Let* 1 *be the starting element of the base blocks. All the base blocks generate* k $km \times m$ *submatrices under the cyclic addition procedure. Then,* D, *obtained by stacking all the resulting submatrices, is a* $k^2m \times m$ k*-layer CCOP design.*

The remaining part of this subsection examines several unique structures of k-layer CCOP designs when there is an odd number of level and $m \equiv 3 \pmod 4$.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

4.3 Odd-Level k-Layer CCOP Designs with Odd Numbers of Components

Theorem 7. *When $m = 3$, a k-layer CCOP design does not exist for any odd k.*

Theorem 8. *For $m = 7$ and $k = 3$, consider the difference vector trios for base blocks $\vec{r_1}$, $\vec{r_1'}$, and $\vec{r_1''}$ as follows: $\Delta\vec{r_1} = (6, 2, 4, 11, 9, 13)$, $\Delta\vec{r_1'} = (3, 5, 1, 10, 20, 16)$, and $\Delta\vec{r_1''} = (17, 19, 15, 18, 12, 8)$. Assuming 1 is the starting element of the base blocks, and for $1 \leq a \leq 21 - 1$, the vectors $\vec{r_1}$, $\vec{r_1'}$, and $\vec{r_1''}$ generate three 21×7 submatrices M_1, M_1', and M_1'', respectively, under cyclic addition. The final design matrix D is obtained by stacking all three submatrices, resulting in a 63×7 3-layer CCOP design. Similar to Theorem 6, this approach can be extended to obtain k-layer CCOP designs for any odd $k > 3$.*

In general, cyclic addition-structured solutions are not available for $m \equiv 3 \pmod 4$. In graph theory, finding such solutions is equivalent to partitioning k directed complete graphs K_m^* into $k^2 m$ Hamiltonian paths. The method proposed in Tillson (1980) offers a possible approach to construct a matrix without relying on cyclic addition structure. In practice, for an OofA experiment, it is relatively straightforward to incorporate a placebo component (dummy variable) or an unused ground level (zero dose) to create k-layer CCOP designs using Theorem 2 and 4.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

5. Applications

This section is divided into two parts: the first addresses multi-level simulations, and the second covers 1-level simulations along with data analysis. In the first part, we demonstrate that our results for $k > 1$ simulations outperform those of the DOA and QS designs. In the second part, we show through simulations and data analysis that even for $k = 1$, our design and analysis surpass existing methods in terms of selecting the optimal setting. Our primary goal is to identify the optimal order for response optimization using a small design, rather than focusing on model fitting. We emphasize the performance of the true rank identified as the optimal order by both the design and the model.

The procedure for applying our design and analysis method with any number of levels is outlined below and is supported by numerical studies:

(i) Select the initial design with n points from our design.

(ii) Fit our distance-based model and, if necessary, use the Dantzig selector to select variables..

(iii) Conduct a confirming experiment for points with the top or bottom n fitted values, depending on whether maximization or minimization is being pursued.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

(iv) Select the setting with the optimal observation.

5.1 Simulation with $k > 1$

We compare the 2-layer CCOP design with DBLM, the DOA design with the PWO model, and the QS design with the kriging model. For this multi-level simulation, we use a numerical setting similar to that in Mee (2020). Starting with $Y_0 = 1$, the response follows the pattern below.

1. If the next component is $\begin{cases} 1, \text{ then add } c_{11} \\ 5, \text{ then add } c_{12} \end{cases}$.
2. If the next component is $\begin{cases} 2, \text{ then subtract } c_{21} \\ 6, \text{ then subtract } c_{22} \end{cases}$.
3. If the next component is $\begin{cases} 3, \text{ then multiply by } c_{31} \\ 7, \text{ then multiply by } c_{32} \end{cases}$.
4. If the next component is $\begin{cases} 4, \text{ then divide by } c_{41} \\ 8, \text{ then divide by } c_{42} \end{cases}$.

In this case, $m = 4$, $k = 2$, and components are considered the same if they have the same value modulo 4. To generate the level differences, we

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

create a half-normal distribution c_i with $\sigma = 2\sqrt{\pi}$ and a Bernoulli trial b_i with probability $\frac{1}{2}$, such that $c_{i1} = c_i + b_i + 3(1-b_i)$ and $c_{i2} = c_i + 3b_i + (1-b_i)$ for all i. We repeat the data generation process 100 times. Figure 1 displays the final true rank identified by our method with 16 initial and 16 follow-up runs, compared to a DOA with 36 runs and a QS design with 16 initial runs and 16 follow-up runs. Since the run size is too small for active learning, the follow-up experiment for the QS design replicates the third step of our analysis procedure.

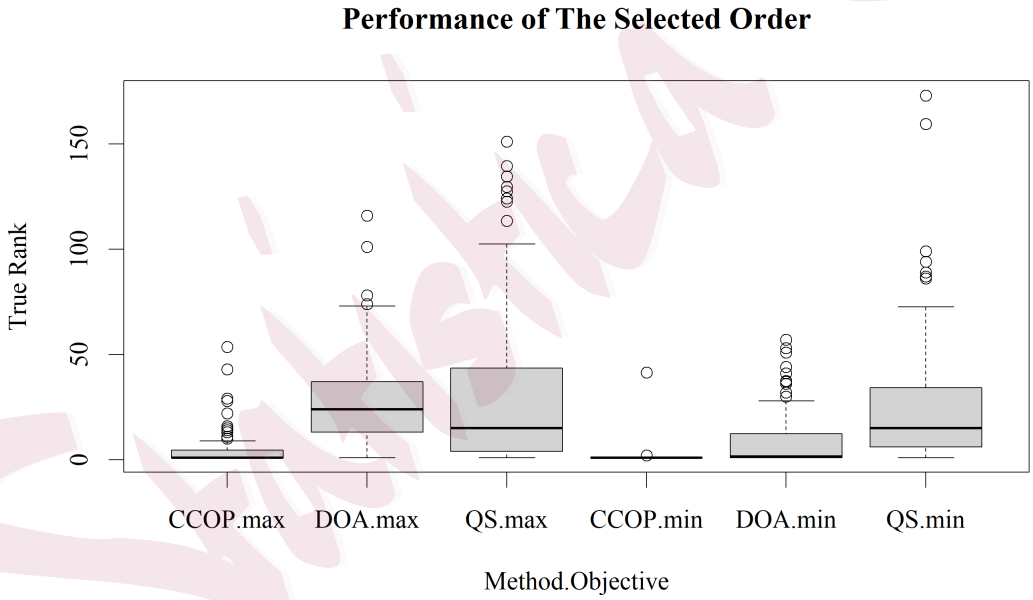

Figure 1: Side-by-Side Boxplots of the True Rank of the Order Selected.

In Figure 1, the x-axis shows the design name and the objective (e.g.,

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

DOA.max indicates a DOA design with maximization as the objective). The y-axis represents the true rank of the order selected by the design and its corresponding model across 100 replications. The side-by-side boxplot demonstrates that, in most cases, our method selects sequences with a better rank than those of the DOA and QS designs, both for maximization and minimization tasks. In addition, our run size is at most 32, which is equal to or smaller than that of the other methods. While our multi-level design and basic modeling method show promise in handling multi-level situations, further development is needed. For instance, optimizing the follow-up experiment for small primary designs and constructing more precise models remain open challenges.

5.2 Simulation and Real Data Analysis with $k = 1$

Table 1 presents the results of a full OofA experiment based on Mee (2020) and Yang, Sun and Xu (2021). It includes two cases: Y_1, a numerical example, and Y_2, which uses real data. The true function for Y_1 is detailed in Mee (2020). The maximizer for Y_1 is 4132 with a value of 37, and the minimizer is 4231 with a value of -8. The response Y_2 pertains to a drug combination experiment (a larger-the-better problem), where sequences of four anti-tumor drugs are evaluated based on their inhibition of tumor cell

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

proliferation. The naive best order for maximizing Y_2 is 1342, with a value of 56.5. Since the true mean is unknown, this result is provided for reference only. Table 2 summarizes the overall results of the analysis. Further details on CCOP design with DBLM, MP design with PWO model, and Bayesian D-optimal design with PWO model are provided below. Details on the two designs from Stokes and Xu (2022) with CP-related models are provided in the appendix.

	Z_{12}	Z_{13}	Z_{23}	Z_{14}	Z_{24}	Z_{34}	Y_1	Y_2
1234	1.00	1.00	1.00	1.00	1.00	1.00	12	41.1
1243	1.00	1.00	1.00	1.00	1.00	-1.00	12	37.5
1324	1.00	1.00	-1.00	1.00	1.00	1.00	19.5	55.4
1342	1.00	1.00	-1.00	1.00	-1.00	1.00	17	56.5
1423	1.00	1.00	1.00	1.00	-1.00	-1.00	2	43.3
1432	1.00	1.00	-1.00	1.00	-1.00	-1.00	17	51.2
2134	-1.00	1.00	1.00	1.00	1.00	1.00	12	46.1
2143	-1.00	1.00	1.00	1.00	1.00	-1.00	12	27.8
2314	-1.00	-1.00	1.00	1.00	1.00	1.00	-3	39.5
2341	-1.00	-1.00	1.00	-1.00	1.00	1.00	2	46.4
2413	-1.00	1.00	1.00	-1.00	1.00	-1.00	32	34.4
2431	-1.00	-1.00	1.00	-1.00	1.00	-1.00	2	39.4
3124	1.00	-1.00	-1.00	1.00	1.00	1.00	4.5	53.5
3142	1.00	-1.00	-1.00	1.00	-1.00	1.00	2	51.2
3214	-1.00	-1.00	-1.00	1.00	1.00	1.00	4.5	50.8
3241	-1.00	-1.00	-1.00	-1.00	1.00	1.00	9.5	51.4
3412	1.00	-1.00	-1.00	-1.00	-1.00	1.00	7	52.9
3421	-1.00	-1.00	-1.00	-1.00	-1.00	1.00	7	53.4
4123	1.00	1.00	1.00	-1.00	-1.00	-1.00	22	39.1
4132	1.00	1.00	-1.00	-1.00	-1.00	-1.00	37	46.4
4213	-1.00	1.00	1.00	-1.00	-1.00	-1.00	22	37.2
4231	-1.00	-1.00	1.00	-1.00	-1.00	-1.00	-8	42.1
4312	1.00	-1.00	-1.00	-1.00	-1.00	-1.00	7	46.8
4321	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	7	41.8

Table 1: The Dataset for the Illustrative Example

We evaluate six designs and six models in total. Our 1-layer CCOP

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Response	Design Name	Run Size	Model	Rank(Max)	Rank(Min)
Y_1	1-layer CCOP	4+3	DBLM, Eq (5.1)	1	1
Y_1	1-layer CCOP*	4+3	DBLM, Eq (5.2)	1	1
Y_1	MP	7	PWO, Eq (5.3)	1	2
Y_1	Bayesian D	7	PWO, Eq (5.4)	1	2
Y_1	$F_{8,4}$	8	1st-order, Eq (S3.1)	19	13
Y_1	$F_{8,4}$	8	Quadratic, Eq (S3.2)	3	3
Y_1	$F_{12,4}$	12	2nd-order, Eq (S3.3)	3	13
Y_1	$F_{12,4}$	12	Nominal CP, Eq (S3.4)	13	13
Y_2	1-layer CCOP	4+3	DBLM, Eq (5.5)	1	-
Y_2	1-layer CCOP*	4+3	DBLM, Eq (5.6)	2	-
Y_2	MP	7	PWO, Eq (5.7)	5	-
Y_2	Bayesian D	7	PWO, Eq (5.8)	5	-
Y_2	$F_{8,4}$	8	1st-order, Eq (S3.5)	19	-
Y_2	$F_{8,4}$	8	Quadratic, Eq (S3.6)	19	-
Y_2	$F_{12,4}$	12	2nd-order, Eq (S3.7)	19	-
Y_2	$F_{12,4}$	12	Nominal CP, Eq (S3.8)	13	-

Table 2: Data Analysis Summary Table

design suggests four runs 1243, 2314, 3421, and 4132. To validate that our results are not due to chance, we provide another 1-layer CCOP, denoted as CCOP* in Table 2, which proposes four different runs 1423, 234, 3241, and 4312. Note that this CCOP is not derived from our construction. Both CCOP designs are fitted using the DBLM. The MP design selects seven runs 1234, 2314, 2143, 1342, 3241, 4213, and 4312. The Bayesian D-optimal design selects seven runs 1243, 1324, 2341, 3412, 3214, 4213, and 4312. Both the MP design and the Bayesian D-optimal design are fitted using the PWO model.

The DBLM fitted model for the 1-layer CCOP design, $\{1243, 2314, 3421, 4132\}$,

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

is given by:

$$\hat{Y} = 1.889d(1243, X) + 6.889d(2314, X) + 3.556d(3421, X) - 6.444d(4132, X). \tag{5.1}$$

According to this model, the top four fitted values are selected from 4132, 1324, 2413, and 3241. After a follow-up experiment for these top values, the maximizer 4132 is identified among the 7 runs. Similarly, the bottom four fitted values are 2314, 1423, 3142, and 4231. The minimizer 4231 is identified among the 7 runs.

To illustrate the effectiveness of our process even when the true maximizer is not included in our design, we consider another 1-layer CCOP design, $\{1423, 2134, 3241, 4312\}$. The DBLM fitted model for this design is

$$\hat{Y} = 2.722d(1423, X) - 0.611d(2134, X) + 0.222d(3241, X) + 1.056d(4312, X). \tag{5.2}$$

After the follow-up experiment for 2134, 1324, 2413, and 4132 with top fitted values, and 1423, 2314, 3142, and 4231 with bottom fitted values, we again identify the maximizer 4132 and the minimizer 4231 in 7 runs.

In contrast, the fitted PWO model for the MP design $\{1234, 2314, 2143, 1342, 3241, 4213, 4312\}$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

is

$$\hat{y} = 10.438 - 0.938X_{12} + 8.438X_{13} - 1.875X_{23} - 4.375X_{14} - 0.625X_{24} + 0.938X_{34}. \tag{5.3}$$

This model identifies the maximizer 4132 with 7 runs, matching our total run size. However, it fails to identify the minimizer 4231, instead identifying 2314 with a value of -3, while 4231 has a value of -8.

The Bayesian D-optimal design $\{1243, 1324, 2341, 3412, 3214, 4213, 4312\}$ results in the model

$$\hat{y} = 11.375 - 1.875X_{12} + 9.375X_{13} - 3.750X_{23} - 2.500X_{14} - 0.625X_{24} + 1.496 \cdot 10^{-15}X_{34}, \tag{5.4}$$

This model is very similar to that of the MP design. It identifies the maximizer 4132 but fails to identify the minimizer 4231, instead identifying the order 2314 with the value -3 as the smallest fitted value.

For the real data Y_2, similar procedures are followed. The naive optimizer 1342 is not included in the two 1-layer CCOP designs used. The

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

fitted model for the first 1-layer CCOP design, $\{1243, 2314, 3421, 4132\}$, is:

$$\hat{Y} = 7.144d(1243, X) + 6.478d(2314, X) + 1.844d(3421, X) + 4.178d(4132, X). \tag{5.5}$$

After the follow-up experiment for 3421, 1342, 2134, and 4213 with top fitted values, we identify the naive optimizer 1342 with $Y_2 = 56.5$ in 7 runs. The second 1-layer CCOP design, $\{1423, 2134, 3241, 4312\}$, has the fitted model:

$$\hat{Y} = 6.411d(1423, X) + 5.478d(2134, X) + 3.711d(3241, X) + 5.244d(4312, X). \tag{5.6}$$

After the follow-up experiment for 3241, 1324, 2413, and 4132 with top fitted values, we identify 1324 with $Y_2 = 55.4$ in 7 runs, which ranks second among all observations and was not included in the initial design.

In contrast, the MP design $\{1234, 2314, 2143, 1342, 3241, 4213, 4312\}$ results in:

$$\hat{y} = 43.388 + 0.563X_{12} + 0.238X_{13} - 4.475X_{23} - 1.475X_{14} - 3.225X_{24} + 6.088X_{34}. \tag{5.7}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Similarly, the Bayesian D-optimal design $\{1243, 1324, 2341, 3412, 3214, 4213, 4312\}$ gives:

$$\hat{y} = 45.312 + 1.312X_{12} + 0.988X_{13} - 4.475X_{23} - 2.275X_{14} + 1.112X_{24} + 4.475X_{34}. \tag{5.8}$$

Both MP and Bayesian D-optimal designs identify the sequence 3412 with $Y_2 = 52.9$.

In conclusion, our designs are constructed systematically and can be easily formulated using general formulas. In many cases, our method outperforms other designs in identifying the optimal order. While our model fitting may not be perfect, our approach is effective in ranking response sequences. Although it is not ideal for estimating the entire response surface, it excels at evaluating the relative ranking between responses.

6. Summary and Discussion

In this work, we propose a systematic construction method for k-layer CCOP designs under the following scenarios: (1) an even number of components; (2) an odd number of k-level components where k is even; and (3) an odd number of k-level components where k is odd. Our k-layer CCOP designs follow a similar structure to those in Williams (1949). This new

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

class of designs has the following properties:

1. They can be systematically constructed using our theorems.

2. They are highly cost-efficient, with a run size of at most $2mk^2$.

3. All ordered pairs occur exactly once in k-layer CCOP designs.

4. They can accommodate components with any level of user interest.

We introduce a new procedure for data analysis by first defining important information as the smallest units that can describe the ordering. These units are used to construct the CCOP design and its corresponding model, DBLM. Our design is the smallest to encompass all necessary information. Both for $k = 1$ and $k > 1$, our design and model demonstrate cost-efficiency in identifying the optimal order, even though the model is defined in a simple and naive manner.

It is important to note that our systematic construction is incomplete for the case where $m = 4p + 3$ and $p > 1$, which is expected to be highly complex, if not infeasible. This issue will be addressed in a future extension of this paper. In additional, all components of k-layer CCOP designs constructed here are assumed to be at the same level, which may not be realistic in some scenarios. For instance, different medicines may have varying

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

dosages available for patients. Therefore, constructing mixed-level k-layer (nearly-)CCOP designs represents a natural extension of this work.

In practice, OofA designs have numerous applications. In clinical trials, they can be used to identify the optimal orders and dosages of medicines that result in the best patient outcomes. For supercomputing scheduling problems, a one-level OofA design can effectively schedule jobs to minimize waiting time. Moreover, OofA designs are applicable in various industrial experiments where experimental components are not restricted to a single level. Moreover, our designs can serve as a "smart initialization" method Hsu and Phoa (2018) for metaheuristic algorithms in ordering problems, such as the traveling salesman problem Yen and Phoa (2021).

Supplementary Materials

The supplementary material consists of two appendix sections: (1) Examples, (2) Proof of Theorems, and (3) Simulation and Real Data Analysis with $k = 1$ for CP-related Models.

Acknowledgments

The authors would like to thank Dr. Yuan-Lung Lin for his suggestions on the structure of CCOP designs. This work was supported by Academia

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

Sinica (Taiwan) Thematic Project grant numbers AS-TP-109-M07 and AS-IA-112-M03, and the National Science and Technology Council of Taiwan grant numbers 111-2118-M-001-007-MY2 and 113-2628-M-001-010-MY3. Huang is supported by the doctoral student scholarship provided by the Institute of Statistical Science, Academia Sinica. Part of this work was included in a chapter of the Ph.D. dissertation of Ms. Jing-Wen Huang. The authors would also like to thank the reviewers and the editors for their excellent comments and suggestions to improve the quality of this paper during the revision.

References

Bashkirtseva, I., Ryashko, L., López, Á. G., Seoane, J. M., and Sanjuán, M. A. (2021). The effect of time ordering and concurrency in a mathematical model of chemoradiotherapy. *Communications in Nonlinear Science and Numerical Simulation*, **96**, 105693.

Cuadras, C., and Arenas, C. (1990). A distance-based regression model for prediction with mixed data. *Communications in Statistics A - Theory and Methods*, **19**, 2261-2279.

Dean, A. M., Morris, M., Stufken, J., and Bingham, D. (Eds.). (2015). *Handbook of design and analysis of experiments (Vol. 7)*. Boca Raton: CRC Press.

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

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

treating oral cancer. *Anti-cancer drugs*, **26(5)**, 547-554.

Chen, J., Han, X., Lin D. K., Yang, L. and Zhou, Y. (2023). On Ordering Problems: A Statistical Approach. *Statistica Sinica*, **33(3)**, 1-20.

Chen, J., Mukerjee, R., Lin, D. K. (2020). Construction of optimal fractional order-of-addition designs via block designs. *Statistics and Probability Letters*, **161**, 108728.

Hsu, T.C., and Phoa, F.K.H. (2018). A smart initialization on the swarm intelligence based method for efficient search of optimal minimum energy design. *Advances in Swarm Intelligence*, Springer Volume **10941**, 78-87.

Lin, D. K., Peng, J. (2019). Order-of-addition experiments: A review and some new thoughts. *Quality Engineering*, **31(1)**, 49-59.

Mee, R. W. (2020). Order-of-addition modeling. *Statistica Sinica*, **30**, 1543-1559.

Peng, J., Mukerjee, R., Lin, D.K.J. (2019). Design of order-of-addition experiment. *Biometrika* **106(3)**, 683–694.

Rios, N., Winker, P., Lin, D.K.J. (2022). TA algorithms for D-optimal OofA Mixture designs. *Computational Statistics and Data Analysis* **168**, 107411.

Stokes, Z. (2021). *Advancements in the Design and Analysis of Order-of-Addition Experiments.* University of California, Los Angeles.

Stokes, Z., Xu, H. (2022). A position-based approach for design and analysis of order-of-addition experiments. *Statistica Sinica*, **32**, 1467-1488.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

REFERENCES

Tillson, T.W. (1980). A Hamiltonian decomposition of $K_{2^{m*}}$, $2m > 8$. *Journal of Combinatorial Theory, Series B*, **29**, 68-74.

Tsai, S. (2023). Dual-Orthogonal Arrays for order-of-addition two-level factorial experiments. *Technometrics*, **65(3)**, 388-395.

Van Nostrand, R.C. (1995). Design of experiments where the order of addition is important. In *ASA Proceedings in Section of Physical and Engineering Science*, 155-160. American Statistical Association, Alexandria, Virginia.

Voelkel, J.G. (2019). The design of order-of-addition experiment. *Journal of Quality Technology*, **51(3)**, 230-241.

Wang, C., Mee, R. W. (2022). Saturated and supersaturated order-of-addition designs. *Journal of Statistical Planning and Inference*, *219*, 204-215.

Wang, A., Xu, H. and Ding, X. (2020). Simultaneous Optimization of Drug Combination Dose-Ratio Sequence with Innovative Design and Active Learning. *Advanced Therapeutics*, **3**, 1900135.

Winker, P., Chen, J., Lin, D. K. (2020). The construction of optimal design for order-of-addition experiment via threshold accepting. In *Contemporary experimental design, multivariate analysis and data mining* (pp. 93-109). Springer, Cham.

Williams, E.J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research, Series A*, **2**, 149-168.

Statistica Sinica: Preprint
doi:10.5705/ss.202023.0357

REFERENCES

Xiao, Q., Xu, H. (2021). A mapping-based universal Kriging model for order-of-addition experiments in drug combination studies. *Computational Statistics and Data Analysis*, **157**, 107155.

Xiao, Q., Wang, Y., Mandal, A. and Deng, X. (2022). Modeling and active learning for experiments with quantitative-sequence factors. *Journal of the American Statistical Association*, **119(545)**, 407-421.

Yang, J.F., Sun, F. and Xu, H. (2021). A Component-Position model, analysis and design for order-of-Addition experiments. *Technometrics*, **63**, 212-224.

Yen, P.C., and Phoa, F.K.H. (2021). Traveling salesman problem via swarm intelligence. *Advances in Swarm Intelligence*, Springer Volume **12689**, 106-115.

Zhao, Y.N., Lin, D.K.J., Liu, M.Q. (2021). Designs for order-of-addition experiments. *Journal of Applied Statistics*, **48(8)**, 1475-1495.

National Tsing Hua University, Hsinchu, Taiwan

E-mail: (miffy8485@gmail.com)

Institute of Statistical Science, Academia Sinica, Taiwan

E-mail: (fredphoa@stat.sinica.edu.tw)