This paper considers both approximate and exact designs for estimating the total effects under one crossover and two interference models. They are different from the traditional block designs in the sense that the assigned treatments also affect their neighboring plots, hence a design is understood as a collection of sequences of treatments. A notable result in literature is that the circular neighbor balanced design (CNBD) is optimal among designs which do not allow treatments to be neighbors of themselves. However, we find it necessary to allow self-neighboring, and further show that it is the best to allocate each treatment in a sub-block of adjacent plots with equal or almost equal numbers of replications. This explains why the efficiency of CNBD drops down to 50% as the sequence length, say $k$, increases. Unlike CNBD or the designs for direct effects, our proposed designs do not try to put as many treatments in a sequence as possible. The optimal number of distinct treatments in a sequence is around $\sqrt{2k}$ for crossover designs and $\sqrt{k}/0.96$ for interference models, whenever they are smaller than the total number of treatments under consideration. We systematically study necessary and sufficient conditions for any design to be universally optimal under the approximate design framework, based on which algorithms for deriving optimal or efficient exact designs are proposed. This hybrid nature of cohesively combining theories with algorithms makes our method more flexible than existing ones in the following aspects. (i) Not only symmetric designs are studied, general procedures for producing asymmetric designs are also provided. (ii) Our method applies to any form of within-block covariance matrix instead of specific forms. (iii) We cover all configurations of the numbers of treatments and sequence lengths, especially for large values of them when purely computational methods are not applicable. (iv) On top of the latter, we cover a continuous spectrum of the number of sequences instead of special numbers decided by combinatorial constraints.

1. Introduction. In the application of block designs, it is not uncommon that a treatment assigned to a plot may also have impact on the response of its neighboring plots [23], especially when the plots are small. Such neighbor effects naturally exist in serology, agriculture, horticulture, and forestry experiments, where the treatment could be a plant variety, fertilizer, irrigation, or pesticide [12, 29, 67]. In crop experiments, the neighbor effects have been recognized by [35, 44, 52, 68] due to competition for light, poaching of nutrients by plant roots as well as spread of disease pathogens. Particularly, the competition effect has also been discussed in other experiments by [17, 34, 45, 49, 51, 78] as associated with height difference, tillering ability, date to maturity, canopy size and root yield, respectively. For more applications, also see [6, 10, 14, 42, 48, 50, 61, 71, 77] on oil-seed rape, insect, kale, potatoes, cassava, disease screening, forest, fruit trees and sunflowers, respectively.

AMS 2000 subject classifications: Primary 62K05; secondary 62J05
Keywords and phrases: Approximate design theory, Interference model, Linear equations system, Symmetric designs, Universally optimal designs
Typically in those field experiments, the blocks are arranged in rows of plots and hence the design of such experiments boils down to the determination of the sequences of treatments at the presence of neighbor effects. See for instance [3, 5, 26, 31, 32, 87] on circular designs, for which each plot including the two end plots receives neighbor effects from both its left and right neighboring plots. This could be realized by setting up two guarding plots for each block, with treatments assigned but no response observed. The name circular comes from the particular arrangement that the left guarding plot is assigned the same treatment as the right end plot and similarly for the opposite end, so that the two end plots appear like having neighbor effects on each other. On the contrary, a noncircular design does not have guarding plots and hence each end plot receives only one neighbor effect from its inner neighbor. See [56, 57, 63, 86] among others.

A related topic is the crossover design, where each human or animal subject receives a sequence of treatments over multiple periods of time. The response observed at each period is attributed to the carryover effect of the treatment from the previous period besides the direct effect of the treatment from the current period. This design is very popular in clinical trials [16, 64], especially when the effect of treatments are short-lived and reversible [20, 30]. It is best suited to trials related to symptomatic but chronic conditions or diseases such as neurology, psychiatry, and pain treatment. According to [18], out of the 519 PubMed-indexed randomized trials published in December 2000, 22% of them (116) used crossover designs. In fact, the earliest conducted crossover design on record dates back to 1853 ([47], Sec. 1.4) on nutrition of crop plants. It has been used extensively in animal husbandry experiments since at least the 1930s. To name a few concrete examples, see [15, 21, 22, 33, 46, 69, 70, 72, 76, 80] for using crossover designs on cattle, animal feeding, psychology, biological assay, sensory testing, occupational stress, bioequivalence, industry, diets and milling of paper, respectively.

In view of the wide applications of the designs for these two types of related models, it is valuable to contribute to the toolbox of producing optimal or efficient designs for standard models as benchmarks for practical usage. There has been a rich body of literature for finding optimal or efficient crossover designs, including [9, 19, 37, 38, 39, 40, 55, 58, 59, 60, 84, 85] among others. Also see [13, 47, 66, 73, 74, 75, 79] for comprehensive reviews. In a natural crossover experiment, the first period does not receive a carryover effect from an earlier period treatment, hence noncircular designs are studied in most research. It is still possible to consider circular crossover designs, where the treatment at the last period is also applied to a pre-period so that its carryover effect contributes to the first period. Relevant results are rare. See [26, 54, 65] for examples. Throughout this paper, we call the model used in crossover designs as crossover models for conveniences.

For both models, most work have focused on optimal designs for estimating either the direct treatment effect or the neighbor/carryover effects. In reality, it is desirable to make the decision of selecting a single treatment to be applied over a larger spatial area or a range of time periods [11]. When the chosen treatment is in use, its only neighbor will be itself, and thus the parameter of interest shall be the sum of the direct and neighbor/carryover treatment effects. [7] proposed to adopt the name of total effect for this parameter after a historical remark. For both models, they have established the optimality of neighbor-balanced designs (CNBD) among the subclass of designs with no treatment as a neighbor of itself. Here we use the same acronym CNBD in referring to the two similar designs therein, namely CNBD and CNBD2 corresponding to the two models. For the interference model, the homogeneity and independence assumption therein is extended to two specific within-block covariance structures by [2, 43]. Unfortunately, as also noted by [7], the efficiency of CNBD for estimating the total effects drops down to 50% as the design size grows. This implies the necessity of including self-neighboring sequences in a design. [27] studied optimal designs allowing self-neighboring, however only considered symmetric designs generated by a single sequence
which is not flexible in number of sequences. They tabulated such designs with block lengths up to twelve with the derivation of larger size designs severely constrained by the fast growing computational burden. [63] studied optimal designs for a more parsimonious model where the neighbor effects are proportional to their own direct effects.

All the above mentioned research on total effects have focused on circular designs. Here, we shall explore the rarely studied problem of finding noncircular designs for estimating total effects and focus on three models, i.e., (1)–(3), that cover the most typical crossover and interference models. As far as we know, there is only one relevant work [28] which investigated total effects however under a different crossover model with self and mixed carryover effects [1]. While [28] studied a relatively more complicated model, we cover a much wider spectrum of situations for all the three frequently adopted models. Specifically, we provided the theoretical forms of supporting sequences instead of the computer programing, which could be quickly computationally infeasible as the design size enlarges. We allow all the three design size parameters, that is numbers of treatments, periods/plots, and subjects, to be any possible integers. On top of that, the within block covariance matrix is allowed to be any form. This great extent of flexibility is achieved by a hybrid approach of approximate design theory followed by efficient algorithms for deriving exact designs. In approximate design theory, we characterize all possible universally optimal designs in the whole design space for both models through linear equation systems regarding the sequence proportions. We also theoretically studies the precise form of sequences that are needed to construct optimal designs, which substantially reduces the size of these linear equation systems. These results are then converted into an integer quadratic programing problem as well as other straightforward construction methods for finding exact designs. It turns out that self-neighboring sequences indeed play important roles in our derived designs. In fact, the optimal designs consists of sequences which allocate each treatment in a sub-block of adjacent plots with equal or almost equal numbers of replications. This explains the low efficiency of CNBD outside the subclass. Unlike CNBD or the designs for direct effects, our proposed designs do not try to put as many treatments in a sequence as possible. The optimal number of distinct treatments in a sequence depends on the block length \( k \) in an interesting way. Specifically, it is around \( \sqrt{2k} \) for crossover designs and \( \sqrt{k}/0.96 \) for interference models, whenever these numbers are smaller than the total number of treatments under consideration.

Our approach is quite general in the sense that it produces optimal or nearly optimal designs for any within-block covariance matrix and any configuration of design sizes. Neither of these generality is shared by a combinatorial design such as CNBD. While [27] adopted the approximate design theory to achieve the flexibility on the number of treatments and the block length, their method is not flexible for the number of blocks due to the constraint to symmetric designs. Also their method quickly become infeasible as the block length grows, which is a common phenomena for a typical approximate design approach. Here, we shall provide theoretical forms of plausible block sequences so as to bypass the usual algorithmic search for it. As a result, any design size could be handled.

The rest of the paper is organized as follows. Section 2 formulates the design problems under one crossover model and two interference models into a unified optimization problem. Section 3 provides two systems of linear equations to characterize all possible universally optimal approximate designs, one for symmetric designs and one for general designs. Corresponding two general approaches for deriving optimal or efficient exact designs are provided in Section 4. To tackle the computational issues of these approaches especially for large designs, theoretical forms of the supporting sequences are obtained for the crossover model in Section 5 and for the two interference models in Section 6. Examples are provided in Section 7 to illustrate our theoretical results.
2. Problem formulation. Throughout the paper we consider designs in $\Omega_{n,k,t}$, the set of all possible block designs with $n$ blocks of size $k \geq 3$ ($k \geq 2$ for the crossover model) and $t \geq 2$ treatments. The response observed from the $j$th plot of block $i$, denoted as $Y_{dij}$, is modeled by the following equations

(1) $Y_{dij} = \mu + \beta_i + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + \epsilon_{ij},$
(2) $Y_{dij} = \mu + \beta_i + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + \lambda_{d(i,j+1)} + \epsilon_{ij},$
(3) $Y_{dij} = \mu + \beta_i + \tau_{d(i,j)} + \lambda_{d(i,j-1)} + \rho_{d(i,j+1)} + \epsilon_{ij},$

under crossover model (1), unidirectional interference model (2) and directional interference model (3) respectively. Here $\mu$ is the general mean, $\beta_i$ is the $i$th block effect, $\epsilon_{ij}$ is the error term with mean zero and $d(i,j)$ denotes the treatment assigned in the $j$th plot of block $i$ by the design $d$. For Model (3), $\tau_{d(i,j)}$ is the direct treatment effect of $d(i,j)$, $\lambda_{d(i,j-1)}$ is the neighbor effect of treatment $d(i,j-1)$ from the left neighbor, and $\rho_{d(i,j+1)}$ is that from the right neighbor. Model (2) assumes $\lambda_i = \rho_i$, $1 \leq i \leq t$, namely the neighbor effects are undirectional. For Model (1), $\lambda_{d(i,j-1)}$ is interpreted as the carryover effect. We focus on noncircular designs for all three models, that is to assume $\lambda_{d(i,0)} = \lambda_{d(i,k+1)} = \rho_{d(i,k+1)} = 0$.

If $Y_d$ is the vector of responses organised block by block, Models (1)–(3) can be written in matrix forms of

(4) $Y_d = U_1 n k \mu + T_d \beta + L_d \lambda + \epsilon,$
(5) $Y_d = U_1 n k \mu + T_d \beta + L_d \lambda + R_d \lambda + \epsilon,$
(6) $Y_d = U_1 n k \mu + T_d \beta + L_d \lambda + R_d \rho + \epsilon,$

where $\beta = (\beta_1, ..., \beta_n)'$, $\tau = (\tau_1, ..., \tau_t)'$, $\lambda = (\lambda_1, ..., \lambda_t)'$, $\rho = (\rho_1, ..., \rho_t)'$ with $'$ representing the transpose of a vector or a matrix. Also, $1_h$ represents a vector of $h$ ones, and $U = I_n \otimes 1_k$ with $\otimes$ being the Kronecker product and $I_k$ being the identity matrix of size $k$. Lastly, $T_d$, $L_d$ and $R_d$ represent the design matrices for the direct, left/carryover neighbor and right neighbor effects, respectively.

Our target here is to find the optimal design for the estimation of the total effect, namely $\phi = \gamma + \lambda$ for Model (1), $\phi = \gamma + 2\lambda$ for Model (2) and $\phi = \gamma + \lambda + \rho$ for Model (3). For this purpose, we shall re-parametrize those models as

(7) $Y_d = U_1 n k \mu + T_d \phi + \tilde{L}_d \lambda + \epsilon,$
(8) $Y_d = U_1 n k \mu + T_d \phi + (\tilde{L}_d + \tilde{R}_d) \lambda + \epsilon,$
(9) $Y_d = U_1 n k \mu + T_d \phi + \tilde{L}_d \lambda + \tilde{R}_d \rho + \epsilon,$

where $\tilde{L}_d = L_d - T_d$ and $\tilde{R}_d = R_d - T_d$. In other words, we have $\tilde{L}_d = (I_n \otimes H)T_d$ and $\tilde{R}_d = (I_n \otimes H')T_d$, where $H = (I_{i+j+1} - I_{i+j})1_{1 \leq i,j \leq k}$ with $I$ being the indicator function. Here we adopt a very mild condition for the covariance structure, i.e., $V(\epsilon) = I_n \otimes \Sigma$ with $\Sigma$ being an arbitrary positive definite $k \times k$ matrix. By similar arguments as in [55], the information matrix for $\phi$ under the three models are

(10) Model (1) : $C_d = C_{d00} - C_{d01} C_{d11} C_{d10},$
(11) Model (2) : $C_d = C_{d00} - (C_{d01} + C_{d02}) \left( \sum_{i,j=1}^{2} C_{dij} \right)^{-1} (C_{d10} + C_{d20}),$
(12) Model (3) : $C_d = C_{d00} - (C_{d01} C_{d02}) \left( \begin{array}{cc} C_{d11} & C_{d12} \\ C_{d21} & C_{d22} \end{array} \right)^{-1} \left( \begin{array}{c} C_{d10} \\ C_{d20} \end{array} \right),$
where $C_{dij} = G'(I_n \otimes \tilde{B})G_j$, $0 \leq i, j \leq 2$, with $G_0 = T_d$, $G_1 = \tilde{L}_d$, $G_2 = \tilde{R}_d$, and $\tilde{B} = \Sigma^{-1} - \Sigma^{-1}J_k\Sigma^{-1}/1_k\Sigma^{-1}1_k$ with $J_k = 1_k1_k'$. For technical conveniences, we shall define a projection matrix $B_k = I_k - k^{-1}J_k$. In fact, we have $\tilde{B} = B_k$ when $\Sigma = I_k$.

The block diagonal structure of the matrix $I_n \otimes \tilde{B}$ allows us to write each $C_{dij}$ in an additive form, which induces the approximate design framework. Let’s just examine

The presentation of theoretical results in this section would appear to be a re-organization of [59] with some implicit technical differences. One result that is of new form is part of an additive form, which induces the approximate design framework. Let’s just examine

3. Approximate design theory. We shall begin with the discussion on a subclass of measures which is guaranteed to contain some, albeit not all, universally optimal measures. Such a subclass is called a complete class in optimal design literature, see the series of papers [24, 25, 81, 82, 83] on the de la Garza phenomenon. While they identify the complete class through the Chebyshev system, here we shall use the symmetrization argument. The latter approach has been adopted by [59, 60], [56, 57], [58] for models (1), (3) and a crossover model with mixed carryover effects, respectively. While all of them are interested in the direct effects instead of the total effects, the re-parametrization (7)–(9) allows us to apply the symmetric argument to models (1)–(2) and (3) in the frameworks of [59] and [56], respectively.

For the traditional crossover model, [59] was also able to go beyond the complete class and derive an unusual result on the necessary and sufficient condition for any measure in $P$ (possibly outside the complete class) to be universally optimal. This is a substantial advance that does not normally exist in optimal design work for other models, and it is achieved by leveraging on the complete class results and exploring the special mathematical form of the information matrix like (10). Here we derive parallel results for models (1)–(3). This extension is arrived by tackling two difficulties, one is tied to the matrix transformations $L_d = L_d - T_d$ and $R_d = R_d - T_d$ induced by the re-parametrization, and the other is the substantial gap between structures of the informations matrices (10)–(11) and (12). As a result, the presentation of theoretical results in this section would appear to be a re-organization of that of [59] with some implicit technical differences. One result that is of new form is part
(iii) of Theorem 3.2, which provides the tool for us to improve on the computational complexity of the algorithm in [59] by magnitudes. See Step 1 of Algorithm 4.1 for details. With all these being said, we claim our main novelty lies in Sections 4 and 5 on theoretical forms of the supporting sequences, which is essential for assembling a machinery that could produce optimal or efficient exact designs for any $\Sigma, k, t, n$ and for any model. Particularly, the flexibility in $n$ and the capacity to deal with large values of $k$ and $t$ are the main challenges.

We now introduce some notations, whose meanings may vary across the three models. As will be shown in Theorem 3.2 part (i), a proper complete class here could be $\mathcal{P}_0 = \{ \xi : C_{\xi ij} \text{ is completely symmetric for all possible indices } i \text{ and } j \}$. We call an element in $\mathcal{P}_0$ to be a symmetric measure. One simple way to construct a symmetric measure is to partition $S$ into equivalence classes and assign equal proportion (could be zero) to all sequences within each equivalence class. Here, the equivalence class is defined as a class such that a new sequence generated by any treatment relabeling of any sequence in this class still belongs to this class. Specifically, for any sequence $s$, the equivalence class that it belongs to could be constructed as $\langle s \rangle = \{ \sigma(s) : \sigma \in \mathcal{G} \}$, where $\sigma(s)$ is a sequence generated by applying the treatment relabeling/permutation $\sigma$ on $s$, and $\mathcal{G}$ is the collection of all $t!$ possible permutations. For example, all binary sequences together form an equivalence class. By the group property of $\mathcal{G}$, we can see that two equivalence classes are either identical or mutually exclusive. One equivalence class could be represented by one way of partitioning $k$ balls into at most $t$ boxes. Let $m$ be the total number of distinct equivalence classes. When $k \leq t$, $m$ is the well known Bell number depending only on $k$, that is, $m = k^k L(k)$ with $L(k)$ decaying at the exponential rate in $k$. While this number is substantially smaller than $|S|$, it still grows faster than the exponential rate. See Table 1 for some examples. These equivalence classes can be generated by the Hindenburg’s algorithm (Andrews, [4]). When an equivalence class is large, we can also carefully select a subset of it to construct symmetric measures. See Example 7.4 for specific steps.

For any symmetric measure $\xi \in \mathcal{P}_0$, we have $C_{\xi ij} = c_{\xi ij} B_t/(t - 1) + 1^t_c C_{\xi ij} 1_t J_t/t^2$, where $c_{\xi ij} = tr(B_t C_{\xi ij} B_t)$, due to the orthogonality between $B_t$ and $J_t$. Applying them to (10)–(12) yields

\[
C_\xi = y_\xi B_t/(t - 1), \quad y_\xi = c_{\xi 00} - \ell_\xi Q_\xi^{-1} \ell_\xi,
\]

with $\ell_\xi$ and $Q_\xi$ defined in the following table under different models.

<table>
<thead>
<tr>
<th>$\ell_\xi$</th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{\xi 01}$</td>
<td>$\sum_{i,j=1}^t c_{\xi ij}$</td>
<td>$(c_{\xi ij})_{1 \leq i,j \leq 1}$</td>
<td>$(c_{\xi 01}, c_{\xi 02})$</td>
</tr>
<tr>
<td>$c_{\xi 11}$</td>
<td>$c_{\xi 11}$</td>
<td>$c_{\xi 01} + c_{\xi 02}$</td>
<td>$(c_{\xi 01}, c_{\xi 02})$</td>
</tr>
</tbody>
</table>

A similar representation also appeared in [56] for model (3), where the meanings of $\ell_\xi$ and $Q_\xi$ are different due to the interest in direct effects instead of total effects here. A substantial difference lies in the adoption of generalized inverse $Q_\xi^{-1}$ in [56] instead of direct inverse $Q_\xi^{-1}$ in (13). The latter is owing to Proposition 3.1 below, where we showed the non-singularity of $Q_\xi$ for all possible $\xi$. This subtle difference however become crucial in deriving general linear equation systems as optimality conditions for both symmetric and asymmetric measures in Theorems 3.2 and 3.6, and consequently the theoretical forms of supporting sequences in Sections 4 and 5.

**Proposition 3.1.** For any model in (1)–(3) and any sequence $s \in S$, we have $Q_s > 0$.  

To proceed, we shall examine (13) in more detail. Note that $y_\xi$ is a Schur complement of the matrix $F_\xi$ as defined in the left of the table. Observe the linearity $c_{\xi ij} = \sum_{s \in S} p_s c_{sij}$ with $c_{sij} = \text{tr}(B_i C_{sij} B_j)$ for $0 \leq i, j \leq 2$ due to the fact $C_{\xi ij} = \sum_{s \in S} p_s C_{sij}$. Propagating this linearity forward, we have $\ell_\xi = \sum_{s \in S} p_s \ell_s$, $Q_\xi = \sum_{s \in S} p_s Q_s$, and $F_\xi = \sum_{s \in S} p_s F_s$, with $\ell_s, Q_s, F_s$ equalling $\ell_\xi, Q_\xi, F_\xi$ when the measure $\xi$ is a degenerated measure with a single sequence $s$. These definitions allow us to express theoretical results in a unified form over the three models. Let $y^* = \max_{\xi \in \mathcal{P}} y_\xi$ and denote the support of a measure $\xi = \{p_s : s \in \mathcal{S}\}$ by $\mathcal{V}_\xi = \{s \in \mathcal{S} : p_s > 0\}$. Theorem 3.2 studies the optimality of symmetric measures.

**Theorem 3.2.** For any model in (1)–(3), (i) There exists a symmetric measure in $\mathcal{P}_0$, which is universally optimal among $\mathcal{P}$; (ii) A symmetric measure $\xi$ is universally optimal if and only if $y_\xi = y^*$; (iii) A symmetric measure $\xi \in \mathcal{P}_0$ is universally optimal among $\mathcal{P}$ if and only if $\det(F_\xi) > 0$ and

$$\max_{s \in \mathcal{S}} [\text{tr}(F_s F^{-1}_\xi) - \text{tr}(Q_s Q^{-1}_\xi)] = 1. \tag{14}$$

Besides, the maximum can be achieved by all $s \in \mathcal{V}_\xi$. (iv) A symmetric measure $\xi = \{p_s : s \in \mathcal{S}\} \in \mathcal{P}_0$ is universally optimal among $\mathcal{P}$ if and only if

$$\sum_{s \in \mathcal{S}} p_s (\ell_s + Q_s x^*) = 0, \tag{15}$$

$$\sum_{s \in \mathcal{S}} p_s = 1, \tag{16}$$

where $x^*$ is defined in (17).

In Theorem 3.2, part (i) indicates that we are able to find a universally optimal measure by restricting the search within $\mathcal{P}_0$. This together with (13) indicate that the optimization problem boils down to maximizing the scaler $y_\xi$, and thus part (ii) becomes self-evident. At this point, part (iii) naturally follows from a typical argument of the general equivalence theory (GET) for D-criterion by treating the scalar $y_\xi$ as the Schur complement of the matrix $F_\xi$. Equation (14) permits the usage of many GET based algorithms, such as Federov’s exchange algorithm, to derive a universally optimal symmetric measure. As a parallel result to part (iii), part (iv) allows us to find all universally optimal symmetric measures through a simple linear equation. Note the elements $c_{sij}, 0 \leq i, j \leq 2$, are the same for sequences from the same equivalence class, as a result the optimality conditions in parts (iii) and (iv) can be collapsed into conditions for the class proportion $p(o) = \sum_{s' \in (o)} p_{s'}$ instead of the sequence proportion $p_s$. In this regard, the dimensionality of the optimizing variable reduces from $|\mathcal{S}| = t^k$ to $m$.

Note that the implementation of part (iv) requires the knowledge of $x^*$, which could cost $O(m^2)$ computational operation by the pairwise-comparison algorithm based on Lemma 4.6 of [59]. Instead, one can get around of it by first deriving one optimal measure based on the GET, i.e. part (iii), which only requires $O(m)$ computational operations. This optimal measure, say $\xi_0$, does not has to be unique, but the $x^* = -Q^{-1}_{\xi_0} \ell_{\xi_0}$ value is unique regardless the choice of the optimal measure. Then we can use part (iv) to recover all possible universally optimal symmetric measures.

In fact, it is also possible to derive the sufficient and necessary condition for a general measure $\xi \in \mathcal{P}$ to be universally optimal. Define the quadratic functions $q_s(x) = c_s + 2\ell_s x + x' Q_s x$ and $q_\xi(x) = c_\xi + 2\ell_\xi x + x' Q_\xi x$ so that $q_\xi(x) = \sum_{s \in \mathcal{S}} p_s q_s(x)$, with $x \in \mathbb{R}$ for Models (1)–(2) and $x \in \mathbb{R}^2$ for Model (3). One can verify that $y_\xi = \min_{x \in \mathbb{R}} q_\xi(x)$ (replace $\mathbb{R}$ by $\mathbb{R}^2$ for Model (3)) and the minimum is achieved at $x_\xi = -Q^{-1}_{\xi} \ell_\xi$. Define
\( r(x) = \max_{s \in S} q_s(x) \), which is strictly convex due to the strict convexity of \( q_s(x) \). Hence it has an attainable minimum value denoted by \( y_* = \min_x r(x) \) and a unique minimizer denoted by

\[(17) \quad x^* = \arg \min_x r(x) \]

so that \( r(x^*) = y_* \). Now we claim that it is sufficient to only consider sequences in the subset \( T = \{ s \in S : q_s(x^*) = y_* \} \subset S \) in the search of universally optimal measures. For symmetric measures, we have Corollary 3.4, which is a direct result of Theorem 3.2 and Proposition 3.3.

**Proposition 3.3.** For Models (1)–(3), we have (i) \( y_* = y^* \); (ii) \( y_\xi = y^* \) implies \( q_\xi(x^*) = y^* \) for any measure \( \xi \in \mathcal{P}_0 \); (iii) \( y_\xi = y^* \) implies \( V_\xi \subset T \) for any measure \( \xi \in \mathcal{P}_0 \).

**Corollary 3.4.** The optimality conditions in Theorem 3.2 parts (iii) and (iv) still hold if the set \( S \) in (14)–(16) is replaced by \( T \).

It is relatively more difficult to derive optimality conditions for a general measure in \( \mathcal{P} \). In general universally optimal measures are not unique, but Conditions (C.1)–(C.3) in Section 2 implicitly imply that the forms of the information matrix of them are unique. This conclusion together with Theorem 3.2 part (ii) naturally yields the follow condition of optimality of a general measure, not necessarily symmetric.

**Proposition 3.5.** A measure \( \xi \in \mathcal{P} \) is universally optimal if and only if

\[ C_\xi = y^* B_t / (t - 1). \]

This condition can serve to check the universal optimality of a measure, however does not help guide us to derive it. For the latter, Theorem 3.6 is very powerful for deriving both approximate and exact designs. Equations (18)–(20) indicate the necessity of \( V_\xi \subset T \) for any universally optimal measure in \( \mathcal{P} \). It is tempting to think of this as the result of Proposition 3.3 part (iii) and Proposition 3.5, which is only partially true. There is a need to show that \( C_\xi = y^* B_t / (t - 1) \) implies \( y_\xi = y^* \). While this is trivial for symmetric measures but not so for a general measure. Relevant arguments have to be done during the proof of the theorem.

**Theorem 3.6.** For Models (1)–(3), a measure \( \xi \in \mathcal{P} \) is universally optimal if and only if the following equations hold:

\[(18) \quad \sum_{s \in T} p_s [E_{s00} + E_{s01}(x^* \otimes B_t)] = y^* B_t / (t - 1), \]

\[(19) \quad \sum_{s \in T} p_s [E_{s10} + E_{s11}(x^* \otimes B_t)] = 0, \]

\[(20) \quad \sum_{s \in T} p_s = 1. \]

where the new notations are defined as follows with \( E_{s00} = E_{s01}^\prime \).

<table>
<thead>
<tr>
<th>( E_{s00} )</th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_{s00} )</td>
<td>( C_{s00} )</td>
<td>( C_{s00} )</td>
<td>( C_{s00} )</td>
</tr>
<tr>
<td>( E_{s01} )</td>
<td>( C_{s01} )</td>
<td>( C_{s01} + C_{s02} )</td>
<td>( (C_{s01}, C_{s02}) )</td>
</tr>
<tr>
<td>( E_{s11} )</td>
<td>( C_{s11} )</td>
<td>( \sum_{i,j=1}^{m} C_{sij} )</td>
<td>( (C_{sij})_{i,j=1,2} )</td>
</tr>
</tbody>
</table>
When $\xi$ is symmetric, it is easy to see that (19) reduces to (15), and also (18) becomes a direct result of (19) and Proposition 3.5. In other words, Theorem 3.6 is an extension of Theorem 3.2 part \((iv)\) in characterizing all universally optimal measures in $\mathcal{P}$ instead of only in $\mathcal{P}_0$. That means there is no constraint on $p_0$ to insure any particular structure like symmetry, except (20) and the non-negativity condition.

4. Computations of $\mathcal{T}$ and exact designs. As can be seen in Corollary 3.4 and Theorem 3.6, the subset of supporting sequences $\mathcal{T}$ plays a key role in reducing the computational complexity for finding universally optimal measures. This will be even more crucial for finding exact designs. Here, we provide a general algorithm of finding $\mathcal{T}$ based on the theoretical developments in Section 3.

ALGORITHM 4.1. (A general approach for finding $\mathcal{T}$)

Step 0. Specify representative sequences, $s_1, \ldots, s_m$, for each of the $m$ equivalence classes.

Step 1. Maximize $y_\xi$ over all $\xi$ supported on $\{s_1, \ldots, s_m\}$, and denote the maximizer by $\xi_0$.

Step 2. Calculate $x^* = -Q_{\xi_0}^{-1} \ell_{\xi_0}$ and $y^* = q_{\xi_0}(x^*)$.

Step 3. Identify the index set $A = \{i \in \mathbb{Z}_m : q_i(x^*) = y^*\}$.

Step 4. Recover $\mathcal{T} = \bigcup_{i \in A} \langle s_i \rangle$.

Step 0 could be executed by many off-the-shelf solvers such as restrictedparts($k, t$) in the software $\mathbb{R}$. The maximization in Step 1 could be carried out by an exchange algorithm or any other reasonable approaches based on the GET (14). The size of set $A$ in Step 3 is much smaller than $m$, and could often be only one or two. The computational complexity of Algorithm 4.1 is $O(m)$ with Step 1 being the bottleneck. As mentioned, Step 1 can also be achieved by the pair-wise comparison approach in [59], which however requires $O(m^2)$ computational operations. For fixed $t$, the value of $m$ is only a polynomial in $k$. However, when $t \geq k$, $m$ is a Bell number which increases in $k$ faster than the exponential rate. Table 3.6 shows that Algorithm 4.1 reduces the computational burden significantly keeping in mind that the numbers in the row of $|S|$ could be much larger when $t/k$ is larger than 1. Meanwhile, the value of $m$ could still be prohibitively large for the implementation of Algorithm 4.1 for large values of $k$ and $t$ such as $t \geq k \geq 13$. Hence there is a need of theoretical results on $\mathcal{T}$.

See Section 5 for Model (1) and Section 6 for Models (2) and (3).

<table>
<thead>
<tr>
<th>$k$</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>24</th>
<th>28</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>15</td>
<td>4140</td>
<td>4213597</td>
<td>1.05 \times 10^{19}</td>
<td>5.17 \times 10^{18}</td>
<td>4.46 \times 10^{17}</td>
<td>6.16 \times 10^{16}</td>
</tr>
<tr>
<td>$</td>
<td>S</td>
<td>$</td>
<td>256</td>
<td>16777216</td>
<td>8.92 \times 10^{15}</td>
<td>1.84 \times 10^{19}</td>
<td>1.05 \times 10^{26}</td>
</tr>
</tbody>
</table>

Now we discuss the approach for finding exact designs. Note that (18)–(20) essentially defines a linear subspace in the $|\mathcal{T}|$-dimensional space, where each point on it represents a universally optimal design. However, the chance is that none of these points corresponds to an exact design for a given $n$. To be specific, recall the sequence replications $n_s = np_s$, and (18)–(20) now translate into

$$\sum_{s \in \mathcal{T}} n_s \left[ E_{s00} + E_{s01}(x^* \otimes B_t) - y^*B_t/(t - 1) \right] = 0$$

and $\sum_{s \in \mathcal{T}} n_s = n$ with all $n_s$ being non-negative integers. When there is no such integer solution, a natural idea is to select the $n_s$ values so that each entry of the matrix in the left
of (21) is as close to zero as possible. When the closeness is measured by the square loss and each entry is considered equally important, we are basically choosing \( n_a \) to minimize the Frobenius norm of this matrix, which is essentially an integer quadratic programming (IQP) problem in \( n_a \). We use the Gurobi solver to implement the IQP in our examples. Since (21) stems from the condition of universal optimality (18)–(20), we expect the derived exact designs to be highly efficient or even optimal under various criteria.

Alternative to the IQP approach, we can construct symmetric designs based on Corollary 3.4 in the following two steps. Find the class weight \( p(\sigma_i) \) for \( i \in A \) as in Step 4 of Algorithm 4.1 through the linear equation

\[
\sum_{i \in A} p(\sigma_i) (\ell_{s_i} + Q_s x^*) = 0
\]

under the constraint of \( \sum_{i \in A} p(\sigma_i) = 1 \). Distribute the class weight \( p(\sigma_i) \) among the sequences in \( \langle s_i \rangle \) to ensure the resulting design is symmetric, that is \( C_{dij}, 0 \leq i, j \leq 2, \) are all completely symmetric. One simple way of achieving the latter is to assign equal weight \( p(\sigma_i)/|\langle s_i \rangle| \) to all sequences in \( \langle s_i \rangle \). A more sophisticated way of distributing the class weight to fewer number (typically \( t(t-1) \)) of sequences in each class is possible [8]. See [41] for recent account in detail. The number \( t(t-1) \) is attributed to the usage of type I orthogonal arrays, see Example 7.4 regarding how it works. Overall, this symmetric design approach improves on the computational efficiency against the IQP approach, yet at the price of inflexibility in \( n \).

5. Optimal designs for Model (1). The previous section proposed two approaches for deriving optimal designs, one based on (21) and the other (22). The key component of these two approaches is \( T \), which is derived by Algorithm 4.1. While \( T \) helps reduce the computational burden for optimal designs, the computation of \( T \) itself becomes more and more challenging as \( k \) and \( t \) continue to grow. Here we derive theoretical results on \( T \), which not only bring computational benefits but also provide deep understandings of the patterns of desired sequences and designs. This section focuses on Model (1) with the other two models deferred to the next section due to the intrinsic connection between them.

Here, we assume \( \Sigma \) to be of type-H, that is of the form \( \Sigma = aI_k + b\mathbf{1}_k' + 1_k b' \) for \( a \in \mathbb{R}^+ \) and \( b \in \mathbb{R}^k \). It covers the special case of completely symmetric matrix and the most often adopted case of identity matrix in relevant literature. Under such a condition, we have \( \bar{B} = B_k/a \) and hence the choice of optimal designs or measures will be independent of the values of \( a \) and \( b \). To introduce our result here, we fix the representative sequences of the first \( \min(k, t) \) equivalence classes as \( s_i = (1_{f_1}, 2 \cdot 1_{f_2}, \ldots, i \cdot 1_{f_i}) \) with the constraint of \( 1 + f_1 \geq f_2 \geq \cdots \geq f_2 \geq f_1 \) and \( \sum_{j=1}^i f_j = k \), for \( 1 \leq i \leq \min(k, t) \). When \( i \) divides \( k \), \( s_i \) allocates equal replications among the first \( i \) treatments with each treatment assigned in sub-blocks of adjacent plots. Otherwise, a proper number of treatments in the tail of the sequence receive one more replication. For each \( s_i \), \( 1 \leq i \leq \min(k, t) \), we notice that the swap of the sub-blocks \( j \cdot 1_{f_j} \) within the sequence among \( j < i \) does not change the values of \( c_{00}, c_{01}, \) or \( c_{11} \), and hence does not change the function \( q_n(x) \), which is essential in our search for \( T \). For this reason, we define \( [s_i] = U_k(\bar{s}) \) where the union is going over all the sequences \( \bar{s} \) resulting from such a swap on \( s_i \), including \( s_i \) itself. Theorem 5.1 shows that \( T \) simply consists of a single \( [s_i] \) with a proper choice of \( i \), and as a result an optimal design can be manually constructed.

**Theorem 5.1.** Suppose \( \Sigma \) is of type-H in Model (1). For \( k = 2 \), we have \( T = S \), and for \( k \geq 3 \) we have the following.
TABLE 2
The value of $i^\ast$ in Theorem 5.1

<table>
<thead>
<tr>
<th>$k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i^\ast$</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

(i) If $t < \sqrt{2k} - 1.5$, we have $T = [s_t]$.  

(ii) If $t > \sqrt{2k} - 1.5$, we have $T = [s_{i^\ast}]$, where $i^\ast = \arg\max_{i \in \mathbb{Z} \cap (\sqrt{2k} - 1.5, \sqrt{2k} + 1.5)} y_i^\ast$ with $y_i^\ast$ defined in (28).

In fact, the interval $(\sqrt{2k} - 1.5, \sqrt{2k} + 1.5)$ contains three integers, $\text{round}(\sqrt{2k})$ and $\text{round}(\sqrt{2k}) \pm 1$. This is evidenced by Table 2, which lists many values of $i^\ast$. To see the significance of Theorem 5.1, consider $k = 10$ with $i^\ast = 4$. As a result, we have $T = [s_t]$ for $t = 2$ or 3, and $T = [s_4]$ with $s_4 = (1122333444)$ for $t \geq 4$. For the latter case, swapping treatment 3 with 1 and 2 in $s_4$, we obtain two new sequences $(3332211444)$ and $(1133322444)$ that also belong to $[s_4]$. Applying treatment relabeling on each of these three sequences results in three equivalence classes, which together assemble $[s_4]$. For a sequence in $[s_4]$, the treatment at its right end has to be replicated 3 times even though the treatments in other locations could be replicated either 2 or 3 times. Since all sequences in $[s_4]$ have the same $t_e$ and $Q_n$, (22) becomes redundant for the case of $T = [s_4]$ and it is sufficient to take any one of the three equivalence classes to construct a symmetric universally optimal design. Similar arguments apply to $s_t$ for $t = 2$ and $t = 3$. To derive an exact design for an arbitrarily given $n$, one can apply the IQP algorithm to $[s_{\min(4,t)}]$. See Example 7.1.

6. Optimal and efficient designs for Models (2)–(3). We investigate Models (2)–(3) jointly here due to the intrinsic connections between them. To illustrate the connections, we first need to distinguish the notations $(x^*, y^*, T)$ for Models (2) and (3) by $(x_2^*, y_2^*, T_2)$ and $(x_3^*, y_3^*, T_3)$, respectively. Proposition 6.1 indicates that we can find $T_3$ by working on the simpler task of finding $T_2$ when $\Sigma$ is persymmetric. Theorem 6.2 shows that an efficient or optimal design under Model (3) is automatically efficient or optimal under Model (2). To introduce the theorem, we define the dual of a sequence $s = (t_1, t_2, ..., t_k)$ as by reserving the positions of the treatments, that is $s' = (t_k, ..., t_1)$. A measure is said to be self dual if $p(s) = p(s')$ for any $s \in S$.

**Proposition 6.1.** If $\Sigma$ is persymmetric, we have: $x_3^* = (x_2^*, x_2^*)$, $y_3^* = y_2^*$ and $T_3 = T_2$.

**Theorem 6.2.** If $\Sigma$ is persymmetric, we have the following conclusions.

(i) The universal optimality of a measure under Model (3) implies its universal optimality under Model (2).

(ii) If the measure is symmetric and self-dual, then the reverse of (i) holds.

(iii) For any criterion function satisfying (C.1)–(C.3), the efficiency of any measure under Model (2) is greater than or equal to that under Model (3).

(iv) The efficiency of a symmetric and self-dual measure is same under Models (2) and (3).

To derive an exact design for an arbitrarily given configuration of $k, t, n$, we can start with deriving $T_2$ along with $x_2^*$ and $y_2^*$ by Algorithm 4.1 and recover $(x_3^*, y_3^*, T_3)$ by Proposition 6.1. Then the IQP approach based on (21) can be applied to $(x_3^*, y_3^*, T_3)$ in deriving exact
### Table 3
Representative sequences of $T_2$ and $x^*_2$ for Theorem 6.3. The integers $\lambda \geq 2$, $\mu \geq 3$, and the number $b = (12\mu^2 - 18\mu - 1 + \sqrt{144\mu^4 - 432\mu^3 + 396\mu^2 - 84\mu + 1})/(12(4\mu - 5)\mu)$ are used where applicable.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$k$</th>
<th>representative sequences of equivalence classes in $T_2$</th>
<th>$x^*_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 2$</td>
<td>$3$</td>
<td>$s_1 = (1122), s_2 = (121)$</td>
<td>$5/14$</td>
</tr>
<tr>
<td></td>
<td>$2\lambda$</td>
<td>$s_1 = (1'<em>\lambda[2 \cdot 1'</em>\lambda], s_2 = (1'_\lambda \otimes (12))$</td>
<td>$(2\lambda - 1)/(4\lambda - 3)$</td>
</tr>
<tr>
<td></td>
<td>$2\lambda + 1$</td>
<td>$s_1 = (1'<em>\lambda[2 \cdot 1'</em>\lambda], s_2 = (1'_\lambda \otimes (12)[1], s'_1$</td>
<td>$(8\lambda^2 - 3)/(16\lambda^2 + 4\lambda - 3)$</td>
</tr>
<tr>
<td>$t = 3$</td>
<td>$3$</td>
<td>$s_1 = (123)$</td>
<td>$8t/(23t - 1)$</td>
</tr>
<tr>
<td></td>
<td>$4$</td>
<td>$s_1 = (1122), s_2 = (1212)$</td>
<td>$2/5$</td>
</tr>
<tr>
<td></td>
<td>$5$</td>
<td>$s_1 = (11122), s_2 = (11232), s_3 = (11233), s'_1, s'_2$</td>
<td>$2/5$</td>
</tr>
<tr>
<td></td>
<td>$6$</td>
<td>$s_1 = (111222), s_2 = (111232), s'_2$</td>
<td>$(25 + \sqrt{241})/96$</td>
</tr>
<tr>
<td></td>
<td>$7$</td>
<td>$s_1 = (1112333), s_2 = (1112323), s'_2$</td>
<td>$(29 + \sqrt{589})/126$</td>
</tr>
<tr>
<td></td>
<td>$8$</td>
<td>$s_1 = (11122333), s_2 = (12121333), s'_2$</td>
<td>$3/7$</td>
</tr>
<tr>
<td></td>
<td>$3\mu$</td>
<td>$s_1 = (1'<em>\mu[2 \cdot 1'</em>\mu][3 \cdot 1'<em>\mu], s_2 = (1'</em>\mu[1])_1 \otimes (23)[2], s'_2$</td>
<td>$b$</td>
</tr>
<tr>
<td></td>
<td>$3\mu + 1$</td>
<td>$s_1 = (1'<em>\mu[2 \cdot 1'</em>\mu][3 \cdot 1'<em>\mu], s_2 = (1'</em>\mu[1])_1 \otimes (32), s'_1, s'_2$</td>
<td>$(2\mu - 1)/(4\mu - 3)$</td>
</tr>
<tr>
<td></td>
<td>$3\mu + 2$</td>
<td>$s_1 = (1'<em>\mu[2 \cdot 1'</em>\mu][3 \cdot 1'<em>\mu], s_2 = (1'</em>\mu[1])_1 \otimes (32)[3], s'_2$</td>
<td>$(2\mu - 1)/(4\mu - 1)$</td>
</tr>
<tr>
<td>$t \geq 4$</td>
<td>$3, 4, 5, 7$</td>
<td>same as $t = 3$</td>
<td>same as $t = 3$</td>
</tr>
<tr>
<td></td>
<td>$6$</td>
<td>$s_1 = (111222), s_2 = (111232), s_3 = (121343), s'_2$</td>
<td>$(25 + \sqrt{241})/96$</td>
</tr>
<tr>
<td></td>
<td>$8$</td>
<td>$s_1 = (11122333), s_2 = (12123434)$</td>
<td>$(41 + \sqrt{1153})/176$</td>
</tr>
</tbody>
</table>

Designs for Model (3). The same design can be used under Model (2). If only $k$ and $t$ are specified by a practitioner with $n$ being relatively flexible, we can potentially replace the IQP approach by the symmetric design approach based on (22). In addition, if a practitioner is quite certain about the equality of the left and right neighboring effects, i.e., Model (2), we can apply one of these two approaches directly to $(x^*_2, y^*_2, T_2)$ to driven an exact design.

The algorithmic approach described above comprehensively covers various practical situations. However, as pointed out in the previous section, it is essential to derive theoretical forms of $\mathcal{T}$ for two purposes. One is to bypass the computational bottleneck in the implementation of Algorithm 4.1, and the other is to have insight on preferred block sequences and designs. Theorem 6.3 provides the exact forms of $\mathcal{T}$ for all values of $t$ and $k$, except when $t \geq 4$ and $k \geq 9$. Theorem 6.4 proposes plausible sequences for producing efficient designs for $k \geq 5$ and any values of $t$. These results together cover all combinations of $k$ and $t$, and are stronger than existing results for circular designs. For the latter, [26] focused on symmetric designs with $k \leq 12$. The only existing design criterion is [7]’s CNBD. [26] used a tighter bound than [7] to verify that the efficiency of CNBD is 1, 0.712 and 0.592 for the $k$ values of 4, 8 and 12, respectively. For the crossover model, [7] showed that the efficiency of CNBD drops down to the limit of 50% as $k$ goes to infinity. This seems to be also true for the interference model. We are the first to study noncircular designs for the estimation of total effects under the interference model, and Theorem 6.4 shows that the efficiency of our proposed design converges to 1 at the rate of $(0.04/\sqrt{0.96})k^{-1/2}$.

The total effects are not estimable for any $\Sigma$ under Models (2) and (3) when $k = 2$. Hence, we shall only consider designs with $k \geq 3$ in this section. Proposition 6.1 assumes $\Sigma$ to be persymmetric and Theorems 6.3 and 6.4 assume $\Sigma$ to be of type-H. It is not uncommon to see covariance structures in literature satisfying both conditions such as complete symmetric matrices and any matrix proportional to the identity matrix. The latter is frequently adopted in studies of circular designs.

**Theorem 6.3.** Suppose $\Sigma$ is of type H in Model (2), $T_2$ and $x^*_2$ are derived and displayed in Table 3 for cases of (i) $t = 2$ and $k \geq 3$, (ii) $t = 3$ and $k \geq 3$, and (iii) $t \geq 4$ and $8 \geq k \geq 3$. If $\Sigma$ is also persymmetric, we have: $x^*_3 = (x^*_2, x^*_2)$ and $T_3 = T_2$ for Model (3).
TABLE 4
The lower bound of $e_R$ in Theorem 6.4 for $15 \geq k \geq 8$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_R$ lower bound</td>
<td>0.9320</td>
<td>0.9481</td>
<td>0.9561</td>
<td>0.9686</td>
<td>0.9690</td>
<td>0.9748</td>
<td>0.9801</td>
<td>0.9814</td>
</tr>
</tbody>
</table>

Under Model (2), we have $c_{sij} = c_{s'ij}$ for $0 \leq i, j \leq 2$. Hence we are still able to build optimal designs when the dual sequences $s'_1$ and $s'_2$ in Table 6.3 are excluded. To tackle the uncovered case $(iv) t \geq 4$ and $k \geq 9$, we propose to build efficient designs based on relatively simpler supporting set, say $R$, rather than $T$. The quality of $R$ can be quantified by its efficiency $e_R = y_R / y^*$, where $y_R = \max \xi_i \in \mathcal{R} y_i$. In other words, $e_R$ calculates the efficiency of the best measure we could potentially construct using sequences in $R$. For $1 \leq i \leq \min(k, t)$, let $s_i = (1'_{f_1}, 3 \cdot 1'_{f_2}, \ldots, i \cdot 1'_{f_i}, \ldots, 4 \cdot 1'_{f_1}, 2 \cdot 1'_{f_2})$ with the constraint of $1 + f_1 \geq f_1 \geq f_2 \cdots \geq f_{i-1} \geq f_i$ and $\sum_{j=1}^i f_j = k$. These sequences are similar to the ones defined in Section 5 in the sense that each treatment $1 \leq j \leq i$ therein appears in a sub-block of adjacent plots with almost equal replications. When $i$ does not divide $k$, the sequence $s_i$, $1 \leq i \leq \min(k, t)$, in Section 5 assigns one more replication to treatments in the right end of the sequence whereas here we assign one more replication to treatments in both the left and right ends of the sequence.

**Theorem 6.4.** Suppose $\Sigma$ is of type $H$. Let $R = \langle s_i \rangle$ under Model (2) and $R = \langle s_i \rangle \cup \langle s'_i \rangle$ under Model (3) with $i^* = \arg \max_{i \in \mathcal{Z}_i \cap (\sqrt{k/0.96 - 1.5}, \sqrt{k/0.96 + 1.5})} q_{s_{i^*}}(0.4)$. For $k \geq 16$, we have $e_R \geq 1 - v(k, t)$, where

$$v(k, t) = \frac{0.04 \sqrt{k/0.96} + 0.26 + 0.18 k + t - 2}{k - 1.962 \sqrt{k} - 1.36} - \frac{(0.4 \sqrt{1/0.96} - 0.25) k^{-1/2}}$$

decreases in both $k$ and $t$. This bound is asymptotically tight in the sense that $(1 - e_R) \sqrt{k} \to 0.04 / \sqrt{0.96}$ and $v(k, t) \sqrt{k} \to 0.04 / \sqrt{0.96}$ uniformly for any $t$ as $k \to \infty$. For each $15 \geq k \geq 8$, the uniform lower bound of $e_R$ for all $t$ is listed in Table 4.

Theorem 6.4 derives a lower bound of $e_R$ for our proposed supporting set $R$ without knowing the true $T$ or optimal designs. It further shows that this bound is asymptotically
tight. But there is some gap for \( k \) slightly larger than 16. A tighter yet more complicated bound could be found in (42). Figure 1 displays these two bounds for the particular case of \( t = 4 \), while the change of \( t \) value does not affect the comparison of the two curves very much. Theorem 6.4 also provides numerical bounds of the efficiencies of \( e_R \) for \( 15 \geq k \geq 8 \). Smaller values of \( k \) can be similarly dealt with except we shall adopt different sequences. For \( k = 5, 6, 7 \), we have \( e_R \) as 0.9571, 0.9259 and 0.9948 if we take \( R \) as \( \{(11233), (111222)\} \), and \( \{(1112333)\} \), respectively. Note that the proposed sequences in Theorem 6.4 belong to \( T \) in most cases in Table 3. On the other hand, we observe that \( T \) also contains another type of sequence as labelled as \( s_2 \) in Table 3, which mixes together the two treatments on one end of the sequence. For \( k \geq 9 \) and \( t \geq 4 \), blending such sequences with the proposed sequences in Theorem 6.4 indeed further improve the efficiency. But the marginal gain of efficiency is not significant compared to the additional complexity involved.

Note that CNBD is optimal among designs not allowing a treatment to be a neighbor of itself. This strongly indicates the necessity of including self-neighboring sequences in a design. Such observations is validated by Theorems 6.3 and 6.4 for noncircular designs. Particularly for large \( k \), Theorem 6.4 suggests treatments to appear as adhesively as possible, which is the exact opposite of not allowing self-neighboring. The optimal number of distinct treatments to be included in a single sequence is one of the three integers \( \text{round}(\sqrt{k/0.96}) \) and \( \text{round}(\sqrt{k/0.96}) \pm 1 \). This number echoes with the number \( \sqrt{2k} \) adopted for crossover designs in alignment with the numbers (1:2) of neighbor effects in the crossover and interference models.

7. Examples. This section illustrates the implementations of the algorithms in Section 4 and the applications of Theorems 5.1, 6.3, 6.4 through five examples. We consider two forms of \( \Sigma \). One is the identity matrix which is of type-H, so that Theorems 5.1, 6.3, 6.4 are directly applicable to derive \( T \) or reasonable subsets of sequences. The other is in the AR(1) form \( \Sigma = (0.2^{i−j})_{1 \leq i,j \leq k} \) which is not of type-H, hence Algorithm 4.1 is needed instead for producing \( T \). Examples 7.1 and 7.2 consider these two types of \( \Sigma \) for flexible choices of \( k \) and \( t \). Example 7.3 further shows the flexibility of our method regarding \( n \). These examples work on relatively small \( k \) and \( t \) to save the space. We devote Examples 7.4 and 7.5 to large \( k \) and \( t \), where highly efficient designs are manually constructed according to Theorems 5.1 and 6.4 along with a combinatorial technique. For an exact design \( d \), let \( 0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{t−1} \) be all the eigenvalues of \( C_d \). Its A-, D-, E-, T-efficiencies are defined as follows.

\[
E_A(d) = \frac{(t−1)^2}{ny^*\Sigma^{i=1}_{i=1}\lambda^i_{-1}}, \quad E_D(d) = \frac{t−1}{ny^*}(\Pi^{t−1}_{i=1}\lambda^i_{i−1})^{\frac{1}{r−1}}, \\
E_E(d) = \frac{(t−1)\lambda_1}{ny^*}, \quad E_T(d) = \frac{\Sigma^{i=1}_{i=1}\lambda^i_{i}}{ny^*}.
\]

Example 7.1. Consider \( k = 7, t \in \{2, 4\} \) and \( \Sigma = I \). For Model (1), the interval \((\sqrt{2k}−1.5, \sqrt{2k}+1.5)\) in Theorem 5.1 contains the integers 3, 4, and 5 with the corresponding \( y^* \) value in (28) as 3.50, 3.35, and 3.14, respectively. We can see that \( y^*_3 \) is the largest among them. That means we have \( T = [s_2] \sim (1112333) \) for \( t = 2 \) and \( T = [s_3] \sim (1112222) \) for \( t = 3 \), including the case of \( t = 4 \). The first two rows of Table 5 display exact designs with \( n = 5 \) derived by applying the IQP algorithm in Section 4 to these sequences. The A- and D- efficiencies of both designs are all above 98.5%. Model (3) can be similarly dealt with by applying IQP to sequences based on Theorem 6.3 and highly efficient exact designs with \( n = 7 \) are displayed in the last two rows of Table 5 for \( t = 2 \) and \( t = 4 \).
Table 5
Some exact designs and the lower bounds of efficiencies, $\Sigma = I$.

<table>
<thead>
<tr>
<th>$n, k, t, model$</th>
<th>exact design</th>
<th>$A$-efficiency</th>
<th>$D$-efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5, 7, 2, (1)</td>
<td>(1112222)$\times 2$, (2221111)$\times 3$</td>
<td>0.9980</td>
<td>0.9980</td>
</tr>
<tr>
<td>5, 7, 4, (1)</td>
<td>(1133444), (2211444), (2233111), (3311222), (3322444)</td>
<td>0.9858</td>
<td>0.9877</td>
</tr>
<tr>
<td>7, 7, 2, (3)</td>
<td>(1112222)$\times 2$, (2112222), (2121212), (2221111)$\times 2$, (2222111)</td>
<td>0.9774</td>
<td>0.9774</td>
</tr>
<tr>
<td>7, 7, 4, (3)</td>
<td>(1112444), (1114333), (2221444), (2224111), (3131222), (3334111), (4442333)</td>
<td>0.9495</td>
<td>0.9667</td>
</tr>
</tbody>
</table>

Table 6
Some exact designs and the lower bounds of efficiencies, $\Sigma = \Sigma_2$.

<table>
<thead>
<tr>
<th>$n, k, t, model$</th>
<th>exact design</th>
<th>$A$-efficiency</th>
<th>$D$-efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>5, 7, 2, (1)</td>
<td>(1112222)$\times 2$, (2221111)$\times 3$</td>
<td>0.9989</td>
<td>0.9989</td>
</tr>
<tr>
<td>5, 7, 4, (1)</td>
<td>(1332244), (2331144), (3442211), (4112233), (4332211)</td>
<td>0.9890</td>
<td>0.9919</td>
</tr>
<tr>
<td>7, 7, 2, (3)</td>
<td>(1111222)$\times 2$, (1112222)$\times 2$, (2221111)$\times 2$, (2222111)</td>
<td>0.9722</td>
<td>0.9722</td>
</tr>
<tr>
<td>7, 7, 4, (3)</td>
<td>(1112333), (1113222), (2221444), (2224111), (3331222), (3331444), (4442111)</td>
<td>0.9456</td>
<td>0.9669</td>
</tr>
</tbody>
</table>

Fig 2. $A$, $D$, $E$, $T$- efficiencies of exact designs for different $n$ with $k = 4$ and $t = 3$ under Model (1).

Example 7.2. Consider $\Sigma = \Sigma_1$ and the same values of $k, t, n$ as in Example 7.1. Theorems 5.1, 6.3 and 6.4 are not applicable and we shall apply Algorithm 4.1 to obtain $T$ directly. Exact designs are again derived by applying IQP to $T$ and displayed in Table 6.

The first two examples have shown that our method is flexible for arbitrary choices of $\Sigma$, the number of treatments $t$ and the block length $k$. The next example shows that our method is even flexible for the number of blocks $n$.

Example 7.3. Consider $\Sigma = \Sigma_1$ and a wide spectrum of $n$ values. For this purpose, we carry out a series of calculation like Example 7.2 for all $n$ between 10 and 52 when $k = 4$ and $t = 3$. The $A$, $D$, $E$ and $T$- efficiencies of these designs are plotted against $n$ in Figures 2 and 3 for Models (1) and (3), respectively. Also see Figure 4 for $k = 10$ under Model (3).
The first three examples have all derived exact designs by applying IQP to $T$. One potential issue is that $|T|$ increases in $t$ at the rate of factorial scale. There is a need of restricting to a subset of $T$ when $t$ is large. One such subset can be constructed based on an orthogonal array of type I ($OA_I$) of strength 2. We demonstrate the specific steps in Example 7.4 for a moderate $t = 5$ and extend it to larger $t$ in Example 7.5. Efficient designs with smaller $n$ for relatively large $t$ is still possible if we further apply IQP to these sequences generated by $OA_I$.

**Example 7.4.** Consider $k = 11$, $t = 5$ and $\Sigma = I$. Theorem 6.4 suggests the set of sequences $\langle s_3 \rangle$ with $s_3 = (11112223333)$. To avoid listing all elements in $\langle s_3 \rangle$, one can generate a symmetric exact design based on $s_3$ as follows. Find an $OA_I$ with 3 rows and 5 levels, denoted by $M$, as follows.
DESIGNS FOR TOTAL EFFECTS

\[ \mathcal{M} = \begin{pmatrix}
1 & 5 & 5 & 2 & 5 & 1 & 2 & 4 & 1 & 4 & 4 & 3 & 1 & 4 & 3 & 3 & 3 & 5 & 2 & 2 \\
5 & 2 & 1 & 4 & 4 & 2 & 5 & 2 & 4 & 5 & 3 & 2 & 3 & 1 & 4 & 1 & 5 & 3 & 3 & 1 \\
4 & 1 & 3 & 5 & 2 & 5 & 3 & 3 & 3 & 1 & 5 & 4 & 2 & 2 & 1 & 5 & 2 & 4 & 1 & 4
\end{pmatrix}. \]

By definition of \( OA_1 \), each pair of distinct symbols from \( \mathbb{Z}_5 \) appears equally often, exactly once here, for all three \( 2 \times 20 \) subarrays of \( \mathcal{M} \). The first column \( (1,5,4) \) means that we shall replace the symbols \( 1,2,3 \) in \( s_3 \) by \( 1,5,4 \) respectively, and hence produce a new sequence from \( s_3 \) as \((1115554444)\). Applying all columns of \( \mathcal{M} \) to \( s_3 \) results in a new array

\[ \mathcal{M}_{s_3} = \begin{pmatrix}
2 & 1 & 1 & 3 & 1 & 2 & 3 & 5 & 2 & 5 & 5 & 4 & 2 & 5 & 4 & 4 & 1 & 3 & 3 \\
2 & 1 & 1 & 3 & 1 & 2 & 3 & 5 & 2 & 5 & 5 & 4 & 2 & 5 & 4 & 4 & 1 & 3 & 3 \\
2 & 1 & 1 & 3 & 1 & 2 & 3 & 5 & 2 & 5 & 5 & 4 & 2 & 5 & 4 & 4 & 1 & 3 & 3 \\
5 & 2 & 4 & 1 & 3 & 1 & 4 & 4 & 4 & 2 & 1 & 5 & 3 & 3 & 2 & 1 & 3 & 5 & 2 & 5 \\
5 & 2 & 4 & 1 & 3 & 1 & 4 & 4 & 4 & 2 & 1 & 5 & 3 & 3 & 2 & 1 & 3 & 5 & 2 & 5 \\
1 & 3 & 2 & 5 & 5 & 3 & 1 & 3 & 5 & 1 & 4 & 3 & 4 & 2 & 5 & 2 & 1 & 4 & 4 & 2 \\
1 & 3 & 2 & 5 & 5 & 3 & 1 & 3 & 5 & 1 & 4 & 3 & 4 & 2 & 5 & 2 & 1 & 4 & 4 & 2 \\
1 & 3 & 2 & 5 & 5 & 3 & 1 & 3 & 5 & 1 & 4 & 3 & 4 & 2 & 5 & 2 & 1 & 4 & 4 & 2
\end{pmatrix}. \]

The design \( \mathcal{M}_{s_3} \) is symmetric and has the same values of \( A-, D-, E-, \) and \( T\)-efficiencies \([8, 41]\). These alphabetical efficiencies all equal to the efficiency of \( \langle s_3 \rangle \), particularly we have \( e_{\langle s_3 \rangle} = 0.9709 \) for both Models (2) and (3).

**Example 7.5.** Consider \( k = 17 \), \( t \in \{7, 11\} \) and \( \Sigma = I \). The representative sequences could be found by Theorem 5.1 for Model (1) and by Theorem 6.4 for Models (2) and (3). For \( t = 7 \), symmetric designs with \( n = 42 \) runs can be constructed based on an \( OA_1 \) by using similar steps in Example 7.4. Their alphabetical efficiencies are 1 under Models (1) and 0.9864 under Models (2) and (3). For \( t = 11 \), symmetric designs with \( n = 110 \) runs can be similarly constructed with their efficiencies as 1 under Model (1) and 0.9874 under both Models (2) and (3).

**Appendix: A**

Corollary 3.4 is a direct result of Theorem 3.2 and Proposition 3.3. Proposition 3.5 can be proved following the process in Section 2.2 of [60]. Thus, we only need to give the detailed proofs of Propositions 3.1 and 3.3, Theorems 3.2, 3.6, 5.1, 6.1, 6.2 and 6.4. The proof of Theorem 6.3 is more tedious and is so deferred to Appendix B. Before diving into the proofs, we shall give some notations which will be frequently used in both Appendix A and B. For sequence \( s = (t_1 \ldots t_k) \), let \( \gamma_s = \sum_{i=1}^{k-1} \delta_{t_i, t_{i+1}} \), \( \psi_s = \sum_{i=2}^{k-1} \delta_{t_{i-1}, t_{i+1}} \), \( f_{s,j} = \sum_{i=1}^{k} \delta_{t_i, j} \), \( \chi_s = \sum_{i=1}^{k} f_{s,i}^2 \). Here \( \delta_{ij} \) is the Kronecker delta.

**Proof of Proposition 3.1.** We prove the result for Model (3). For other two models, the analysis is similar but only easier. Let \( x = (x_1, x_2) \in \mathbb{R}^2 \). Due to \( Q_s \geq 0 \), it will be sufficient to show that \( x'Q_s x = 0 \) implies \( x = (0,0) \). Firstly, we have

\[ 0 = x'Q_s x = \sum_{i,j=1}^{2} x_i x_j tr(B_t C_{sij} B_t) = tr \left( \sum_{i,j=1}^{2} x_i x_j B_t C_{sij} B_t \right). \]
Since \(\sum_{i,j=1}^{2} x_i x_j B_t C_{sij} B_t\) is nonnegative definite, we have
\[
0 = \sum_{i,j=1}^{2} x_i x_j B_t C_{sij} B_t = B_t(x_1 \tilde{L}_s + x_2 \tilde{R}_s)\tilde{B} B_t(x_1 \tilde{L}_s + x_2 \tilde{R}_s) B_t,
\]
from which we conclude
\[
(24) \quad \tilde{B}(x_1 \tilde{L}_s + x_2 \tilde{R}_s) B_t = 0.
\]
From (24), it follows that each column of \((x_1 \tilde{L}_s + x_2 \tilde{R}_s) B_t\) has identical entries, hence columns of \(B_t(x_1 \tilde{L}_s + x_2 \tilde{R}_s)\)' are equal. For \(s = (t_1, t_2, \ldots, t_k)\), let \(e_{t_i}(1 \leq i \leq k)\) be a vector of length \(t\) with \(t_i\)-th component as 1 and the rest components as 0. By convention we have \(e_{0} = e_{k+1} = 0\). The \(i\)th column of \((x_1 \tilde{L}_s + x_2 \tilde{R}_s)\)' is given by \(\alpha_i = x_1 e_{t_{i-1}} - (x_1 + x_2) e_{t_i} + x_2 e_{t_{i+1}}, 1 \leq i \leq k\). Thus we have \(B_t \alpha_i = B_t \alpha_j, 1 \leq i, j \leq k\), and hence \(\alpha_i - \alpha_j = \lambda_i j 1_t\) for some \(\lambda_{ij} \in \mathbb{R}\). Particularly, for \(2 \leq i \leq k\), we have
\[
\lambda_{i,i-1} 1_t = -x_1 e_{t_{i-2}} + (2x_1 + x_2) e_{t_{i-1}} - (x_1 + 2x_2) e_{t_i} + x_2 e_{t_{i+1}}.
\]
Next we discuss (25) in the following four cases A–D.

Case A: \(t \geq 3\) and \(\lambda_{21} \neq 0\) or \(\lambda_{k,k-1} \neq 0\). In fact, only \(t = 3\) is possible in this case. Suppose \(\lambda_{21} \neq 0\), then \(t_1, t_2, t_3\) shall be different from each other. Then (25) with \(i \neq 2\) implies \(2x_1 + x_2 = -(x_1 + 2x_2) = x_2\), hence \(x_1 = x_2 = 0\). The argument for \(\lambda_{k,k-1} \neq 0\) is the same by symmetry.

Case B: \(t \geq 3\) and \(\lambda_{21} = \lambda_{k,k-1} = 0\). (a) When \(t_1, t_2, t_3\) are different from each other, we have \(2x_1 + x_2 = -(x_1 + 2x_2) = x_2 = 0\) by (25) with \(i \neq 2\). (b) \(t_1 = t_2 \neq t_3\), we have \(2x_1 + x_2 = -(x_1 + 2x_2) = x_2 = 0\). (c) \(t_1 = t_3 \neq t_2\), we have \(2x_1 + x_2 = x_2 = 0\) and \(x_1 = 2x_2 = 0\). (d) \(t_2 = t_3 \neq t_1\), we have \(2x_1 + x_2 = x_2 = 0\). Each of these equation systems leads to \(x_1 = x_2 = 0\). (e) When \(t_1 = t_2 = t_3\), we have \(2x_1 + x_2 = x_1 = x_2 = 0\). Suppose \(x_2 \neq 0\), then by (25) with \(i \neq 3\), one gets \(t_3 = t_4\). By iterating the process, we have \(t_1 = t_2 = \cdots = t_k\). But in this case \(-x_2 e_{t_k} = \lambda_{k,k-1} 1_t = 0\), which implies \(x_2 = 0\).

Case C: \(t = 2\) and \(\lambda_{21} \neq 0\). The (a) and (e) of Case B are not impossible here. When \(t_1 = t_2 \neq t_3\), we have \(x_1 = x_2 = 0\). When \(t_1 = t_3 \neq t_2\), we have \(x_1 = x_2 = 0\). When \(t_2 = t_3 \neq t_1\), then \(x_1 = x_2 = 0\). When \(t_2 = t_3 = t_1\), then \(x_1 = x_2 = 0\). In this case, \(x_2(\frac{5}{3} e_{t_1} - \frac{5}{3} e_{t_2} + e_{t_3})\) = \(\lambda_{32} 1_t\), which implies \(x_1 = x_2 = 0\). (d) \(t \geq 3\) and \(\lambda_{21} = 0\). It is not possible to have (a) for Case B. The cases of (b), (c), (d) and (e) of Case C can be argued in the same way as in Case B. Thus, \(Q_s\) is positive.

**Proof of Theorem 3.2.** Theorem 3.2(i) and (ii) can be prove by similar arguments in Section 4 of [59]. Past (iii) naturally follows from a typical argument of the general equivalence theory (GET) for D-criterion by treating the scalar \(y_{\xi}\) as the Schur complement of the matrix \(F_{\xi}\). The detailed proof is given in Appendix B. For (15) and (16) in (iv), we give a brief analysis as follows.

Let \(\nabla q_s(x)\) (resp. \(\nabla q_{\xi}(x)\)) be the gradient of the bivariate function \(q_s(\cdot)\) (resp. \(q_{\xi}(\cdot)\)) evaluated at point \(x\). Note that (15) is equivalent to \(\sum_{s \in T} p_s \nabla q_s(x) = 0\). Suppose (15) and (16) hold, then \(q_s(x)\) reaches its minimum at \(x^*\) and hence \(y_s = q_s(x^*) = \sum_{s \in T} p_s q_s(x^*) = \sum_{s \in T} p_s y_s = y_s = y^*\). So \(\xi\) is universally optimal due to the former conclusions and hence the sufficiency of the theorem. The necessity follows from the three former conclusions in view of \(\nabla q_s(x) = \sum_{s \in S} p_s \nabla q_s(x)\).

**Proof of Proposition 3.3.** Since \(\max_{\xi \in \mathbb{P}} q_{\xi}(x) = r(x)\), we have \(q_{\xi}^* = \min_{x \in \mathbb{R}^2} q_{\xi}(x) \leq \min_{x \in \mathbb{R}^2} \max_{\xi \in \mathbb{P}} q_{\xi}(x) = y_s\), which implies \(y^* = \max_{\xi \in \mathbb{P}} q_{\xi}^* \leq y_s\). Define \(\mathcal{T}_0 = \{s :\)
where \( \tilde{\xi} \) is the \((i,j)\)-th element of \( \tilde{B} \). Note that \( \tilde{B} \) is nonnegative definite with rank \( k - 1 \) and \( B1_k = 0 \). Hence for \( x = (x_1, x_2)' \in \mathbb{R}^2 - \{(0,0)\} \), one has \( x'\tilde{Q}_\xi x = \tilde{x}'\tilde{B}\tilde{x} > 0 \), where \( \tilde{x} = (x_1, \ldots, x_2) \in \mathbb{R}^k \). Thus, \( \tilde{Q}_\xi \) is positive. Since \( Q_\xi \) is also positive according to Proposition 3.1, we conclude \( E_{\xi 11} \) is positive.

To proceed, we shall define valid sub-index sets in the following Lemma A.6, which will be frequently used in the proof of Theorem 5.1. The definition of \( q_s(x) \) is given before (17). We shall emphasize here that, to distinguish with the \( x_1^*, x_2^*, x_3^* \) for Models (1)–(3), we use \( q_s(z) \) instead of \( q_s(x) \) in all analysis hereinafter.

**LEMMA A.6.** Let \( I \) denote an index set such that \( \{(s_i)\}_{i \in I} \) forms all equivalence classes in \( S \). Let \( y^*_i \) denote the minimum value of \( q_{s_i}(z) \) achieved at \( z_i^* \). For an arbitrary \( I^* \subset I \), if \( y^*_i > q_{s_i}(z_i^*) \) hold for all \( i \in I^* \) and \( j \in I \setminus I^* \), then \( T \subset \bigcup_{i \in I^*} \langle s_i \rangle \). We call such \( I^* \) as valid sub-index sets.

**PROOF.** Obviously \( y^* \geq \max \{ y^*_i : i \in I^* \} \). Suppose \( x^* < \min \{ z_i^* : i \in I^* \} \), then we can find \( i_0 \) such that \( q_{s_{i_0}}(x^*) \geq 0 \) and so \( q_{s_{i_0}}(z) > q_{s_{i_0}}(z^*) \) for any \( i \in I^* \). Thus, we must have \( i_0 \in I^* \) or it leads to a contradiction against the condition that \( y^*_i > q_{s_i}(z_i^*) \) hold for all \( i \in I^* \) and \( j \in I \setminus I^* \). But we also have \( z_i^* \leq x^* \) since \( q_{s_{i_0}}(x^*) \geq 0 \), which leads to contradiction.
Proof of Theorem 5.1. The proof can be broken down into three steps: (i) \( T \subset \bigcup_{i=1}^{\min(t,k)} [s_i] \).
(ii) The union in the earlier step can be further narrowed to three indices. (iii) The quadratic function \( q_s(z) \) does not cross each other within a short interval for sequences identified in the earlier step. As a result, \( T \) consists of only one \([s_i]\).

We shall begin with step (i). A lengthy computation yields
\[
\tilde{q}_s(z) = q_{s,0} + q_{s,1}z + q_{s,2}z^2,
\]
where \( q_{s,0} = k - \frac{x}{2}, q_{s,1} = 2\{\gamma_s - k + \frac{2f_{s,t_k}}{k} - \frac{1}{k}\}, q_{s,2} = 2k - 2\gamma_s - \frac{2f_{s,t_k}}{k} + \frac{1}{k} - 1 + \frac{1-k}{kt} \).

Consider \( q'_s(z) \) which is the derivation of \( q_s(z) \). Some direct analysis reveals that \( q'_s(z) \) is increased or unchanged and \( q_{s,2} \) is decreased or unchanged. Note that \( z^* \in (0,1) \). First, we need to show that only \([s_i]\) can be contained in \( T \).

This sequence is the ruling sequence. No other sequence shall occur since, by sorting an arbitrary sequence into the form of \( s_0 \), we have (i) \( q_{s,1} \) is increased or unchanged and (ii) \( q_{s,2} \) is decreased or unchanged. Note that \( z^* \in (0,1) \), the increase in \( q_{s,1}z \) is larger than the decrease in \( q_{s,2}z^2 \).

Moreover, for any given \( t' \), simple analysis reveals that we can balance the frequencies of each symbol in \( s_0 \) and use the longest subsequence for \( a_{t'} \) to minimize \( q_{s,0} \) without decreasing \( \tilde{q}_s(z) \) for any given \( z \), i.e., \( f_{s_0,a_{t'}} \geq f_{s_0,a_i} \geq (f_{s_0,a_i} - 1) \) for any \( 1 \leq i \neq j \leq t' \). Note that the quadratic function \( \tilde{q}_s(z) \) is same for all equivalence classes in \([s_i]\), we can restrict our following analysis to \((s_i)\).

Now we will show the results in (i) and (ii). First, we give no restriction to \( t \). Consider the the \( \tilde{q}_s(z) \) function. Direct calculation reveals that it is symmetric about
\[
x = 0.5 + \frac{1 + \frac{1}{k} - \frac{1-k}{kt} - \frac{2f_{s,t_k}}{k}}{4k - 4\gamma_s - 4f_{s,t_k} + 4 + \frac{2}{k} - 4 + \frac{1(1-k)}{k}} = 0.5 + \Delta_s.
\]

The \( \Delta_s \) used in (27) is only for the convenience in the following analysis. We will show that the choices of \( i \leq \sqrt{2k} - 1.5 \) or \( i \geq \sqrt{2k} + 1.5 \) do not appear in \( T \), thus \( T \) contains at most 4 sequences. Let \( I^* = \{l : \sqrt{2k} - 1.5 < i < \sqrt{2k} + 1.5\} \). It is sufficient to show that \( I^* \) is a valid sub-index set, whose definition can be found in Lemma A.6. Obviously, \( I^* \) usually contains 3 sequences and rarely contains 2 sequences only when \( \sqrt{2k} = p + 0.5 \) for some integer \( p \). We give the following discussion for the case of 3 sequences. The case of 2 sequences can be analyzed analogously but more easily. Let \( I^* = \{i_1, i_2, i_3 : i_3 = i_1 + 1 = i_1 + 2, \sqrt{2k} - 1.5 < i_1 < i_3 < \sqrt{2k} + 1.5\} \).

Let \( x \) denote the number of different treatments. Indexed by \( i_1, i_2 \) and \( i_3 \), we have.
\[
y^*_i = k - \frac{x_{s_i}}{k} + 2\{\gamma_{s_i} - k + \frac{2f_{s_i,t_k}}{k} - \frac{1}{k}\}(0.5 + \Delta_{s_i})
+ \left\{2k - 2\gamma_{s_i} - \frac{2f_{s_i,t_k}}{k} + \frac{1}{k} - 1 + \frac{1-k}{kt}\right\}(0.5 + \Delta_{s_i})^2
\]
and similarly \( y^*_i \) and \( y^*_i \), where \( C \) is a constant independent of \( s \) but only determined by \( k \) and \( t \). Immediately, we can define
\[
\tilde{q}_{x^*_i}(s_i) = \tilde{q}_s(x^*_i) = k - \frac{x_{s_i}}{k} + 2\{\gamma_{s_i} - k + \frac{2f_{s_i,t_k}}{k} - \frac{1}{k}\}(0.5 + \Delta_{s_i})
+ \left\{2k - 2\gamma_{s_i} - \frac{2f_{s_i,t_k}}{k} + \frac{1}{k} - 1 + \frac{1-k}{kt}\right\}(0.5 + \Delta_{s_i})^2
\]
and similarly \( \tilde{q}_{x^*_i}(s_i) \), where \( C \) is a constant independent of \( s \) but only determined by \( k \) and \( t \). Immediately, we can define
\[
\tilde{q}_{x^*_i}(s_i) = \tilde{q}_s(x^*_i) = k - \frac{x_{s_i}}{k} + 2\{\gamma_{s_i} - k + \frac{2f_{s_i,t_k}}{k} - \frac{1}{k}\}(0.5 + \Delta_{s_i})
+ \left\{2k - 2\gamma_{s_i} - \frac{2f_{s_i,t_k}}{k} + \frac{1}{k} - 1 + \frac{1-k}{kt}\right\}(0.5 + \Delta_{s_i})^2
\]
and similarly \( \tilde{q}_{x^*_i}(s_i) \), where \( C \) is a constant independent of \( s \) but only determined by \( k \) and \( t \). Immediately, we can define
\[
\tilde{q}_{x^*_i}(s_i) = \tilde{q}_s(x^*_i) = k - \frac{x_{s_i}}{k} + 2\{\gamma_{s_i} - k + \frac{2f_{s_i,t_k}}{k} - \frac{1}{k}\}(0.5 + \Delta_{s_i})
+ \left\{2k - 2\gamma_{s_i} - \frac{2f_{s_i,t_k}}{k} + \frac{1}{k} - 1 + \frac{1-k}{kt}\right\}(0.5 + \Delta_{s_i})^2
\]
and similarly \( q_{x_{i_2}^*}(s_i) \) and \( q_{x_{i_3}^*}(s_i) \). From Lemma A.6, we only need to prove the conditions therein.

Suppose \( 0 \leq \sqrt{2k} - \lfloor \sqrt{2k} \rfloor < 0.5 \), we are going to prove that \( i_1 = \lfloor \sqrt{2k} \rfloor - 1 \), \( i_2 = \lfloor \sqrt{2k} \rfloor \), \( i_3 = \lfloor \sqrt{2k} \rfloor + 1 \). Consider (29). By the same notation \( i \) which represent the number of different treatments, we have \((k/i_1) \leq \chi_{s_{i_1}}/k \leq (k/i_1) + \sqrt{2k}/4\). Thus

\[
q_{x_{i_1}^*}(s_i) \geq -\frac{k}{i_1} - \frac{i_1}{2} - \frac{\sqrt{2k}}{4} + (1.5 + 2\Delta_{s_{i_1}} - 2\Delta_{x_{i_1}^*}) \cdot \frac{f_{s_{i_1}, t_k}}{k} - 2\Delta_{s_{i_1}}\gamma_{s_{i_1}} + C_1,
\]

where \( C_1 \) only depends on \( k \) and \( t \).

\[
q_{x_{i_2}^*}(s_i) - q_{x_{i_3}^*}(s_i) \leq \frac{k}{i_1} + \frac{i_1}{2} - \frac{i}{2} + \frac{\sqrt{2k}}{4} + \frac{C_2}{k},
\]

where \( C_2 \) only depends on \( k \) and \( t \). We should do similar arguments for \( i_2 \) and \( i_3 \). For \( i_2 \) and \( i_3 \), we can show that \( k/i_1 + i_1/2 - k/i - i/2 \leq -\sqrt{2k}/2 \) for any \( i \in I \setminus I^* \). Thus, generally, for any \( i \in I \setminus I^* \), we have \( q_{x_{i_2}^*}(s_{i_2}) - q_{x_{i_3}^*}(s_{i_3}) \geq \sqrt{k}/4 \) and \( q_{x_{i_3}^*}(s_{i_3}) - q_{x_{i_2}^*}(s_{i_2}) \geq \sqrt{k}/4 \) for any \( k \geq 8 \). The following arguments are also based on \( k \geq 8 \). Now we can focus on \( s_{i_1} \).

In order to prove the conditions in Lemma A.6, we now only need to show that \( q_{x_{i_2}^*}(s_{i_2}) - q_{x_{i_3}^*}(s_{i_3}) > 0 \) for any \( i \in I \setminus I^* \). Consider the derivative of \( q_{s_i}(x) \) over the interval \([M_L, M_U]\)

where \( M_L = \min\{x_{i_1}^*, x_{i_2}^*, x_{i_3}^*\} = 0.5 + \min(\Delta_{s_{i_1}}, \Delta_{s_{i_2}}, \Delta_{s_{i_3}}) \) and \( M_U = \max\{x_{i_1}^*, x_{i_2}^*, x_{i_3}^*\} = 0.5 + \max(\Delta_{s_{i_1}}, \Delta_{s_{i_2}}, \Delta_{s_{i_3}}) \).

It is obvious from (27) that \( -4/k < \min(\Delta_{s_{i_1}}, \Delta_{s_{i_2}}, \Delta_{s_{i_3}}) < 4/k \). For any \( i \in I \), let \( d_i = \max_{x \in [M_L, M_U]} |q_{s_i}(x)| \). Note that, for any \( x \in [M_L, M_U] \), we have \(|x - 0.5| < 1/k\).

\[
q_{s_i}^*(x) = 2\left\{ \gamma_{s_i} - k + \frac{2f_{s_i, t_k}}{k} \right\} + \left\{ 2k - 2\gamma_{s_i} - \frac{2f_{s_i, t_k}}{k} + \frac{1}{k} \right\} \cdot 2x
\]

\[
= \frac{2f_{s_i, t_k}}{k} - \frac{1}{k} - 1 + \frac{1}{kt} x + 2k - 2\gamma_{s_i} - \frac{2f_{s_i, t_k}}{k} + \frac{1}{k} - 1 + \frac{1}{kt} x \cdot 2(x - 0.5),
\]

We immediately have \( q_{s_i}^*(x) \) is no less than \(-4\) and no greater than \(4\). Thus, we know that, for any sequence \( s_i, \), \( q_{s_i}^*(x_1) - q_{s_i}^*(x_2) \leq 4(M_U - M_L) \leq 8/k \). Now we consider two cases. Now we assume \( k \geq 16 \).

Case 1. \( q_{s_{i_2}}(x_{i_2}^*) - q_{s_{i_3}}(x_{i_3}^*) > 0 \) and \( q_{s_{i_3}}(x_{i_2}^*) - q_{s_{i_1}}(x_{i_3}^*) > 0 \). We can simply remove \( i_1 \) and \( I^* = \{i_2, i_3\} \) forms the corresponding set in Lemma A.6.

Case 2. \( q_{s_{i_3}}(x_{i_2}^*) - q_{s_{i_1}}(x_{i_3}^*) < 0 \) or \( q_{s_{i_1}}(x_{i_2}^*) - q_{s_{i_3}}(x_{i_3}^*) < 0 \). Without loss of generality, suppose \( q_{s_{i_3}}(x_{i_2}^*) - q_{s_{i_1}}(x_{i_3}^*) < 0 \). For any \( i \in I \setminus I^* \), we have \( q_{s_{i_1}}(x_{i_2}^*) \geq q_{s_{i_1}}(x_{i_3}^*) \geq q_{s_{i_1}}(x_{i_1}^*) - 8/k \geq q_{s_{i_1}}(x_{i_1}^*) - 8/k + \sqrt{k}/4 \geq q_{s_{i_1}}(x_{i_1}^*) - 16/k + \sqrt{k}/4 \geq q_{s_{i_1}}(x_{i_1}^*) \) since \( k \geq 16 \). Thus we can take \( I^* = \{i_1, i_2, i_3\} \).

Similarly, we can show that (i) \( I^* = \{ \lfloor \sqrt{2k} \rfloor, \lceil \sqrt{2k} \rceil + 1 \} \) for \( \sqrt{2k} - \lfloor \sqrt{2k} \rfloor = 0.5 \), and (ii) \( I^* = \{ \lceil \sqrt{2k} \rceil - 1, \lfloor \sqrt{2k} \rfloor, \lceil \sqrt{2k} \rceil + 1 \} \) for \( 0 \leq \sqrt{2k} - \lfloor \sqrt{2k} \rfloor < 0.5 \).

Since \( \sqrt{2k} \) is fixed for all \( k \), we can give \( I^* \) an index \( k \) as \( I^*_k \). Let \( I^*_k \) denote the smallest set for which we can prove it contains \( T \). Use the same argument, we have \( (i) I^*_{k,t} = Z_{t_1} \cap I^*_k \) when \( Z_{k,t} \) is not empty, \( (ii) I^*_k = t \) when \( Z_{k,t} \) is empty. With this result, we immediately have the results in Theorem 5.1. We should mention here that the former analysis uses the assumption that \( k \geq 16 \). For the case of \( k < 16 \), we have used computer codes to search for the \( T \) and it is verified that the results in Theorem 5.1 still hold. The computational complexity is \( k \) since we only need to consider \( s_i \) with \( i \leq \min(t, k) \).

Now we are going to show that there is only one sequence. Note that we have already show that there is at most three sequences. Consider the set \( \{ \lfloor \sqrt{2k} \rfloor - 1, \lfloor \sqrt{2k} \rfloor, \lceil \sqrt{2k} \rceil, \lceil \sqrt{2k} \rceil + 1 \} \).
Obviously, this set always contains $I^*$. There are at most four different numbers in this set and we can write them as $i_1, i_2, i_3, i_4$. Similar to (28) we can write $y_{i_1}^*, y_{i_2}^*, y_{i_3}^*$, $y_{i_4}^*$ achieved at $x_{i_1}^* = 0.5 + \Delta s_{i_1}, x_{i_2}^* = 0.5 + \Delta s_{i_2}$, $x_{i_3}^* = 0.5 + \Delta s_{i_3}, x_{i_4}^* = 0.5 + \Delta s_{i_4}$. Let $UB = \max(x_{i_1}^*, x_{i_2}^*, x_{i_3}^*, x_{i_4}^*)$ and $LB = \min(x_{i_1}^*, x_{i_2}^*, x_{i_3}^*, x_{i_4}^*)$. It is easy to show that the derivative of all these four sequences has the same order of $O(1/k)$ in $[LB, UB]$ and the scale of $UB - LB$ is of order $O(1/k^2)$. Thus, the variation of the corresponding four $\tilde{q}$ among $[LB, UB]$, denoted by $v$, is of order $O(1/k^3)$. From (28), we write $y_{i_1}^* - y_{i_2}^*$ as follows.

$$y_{i_1}^* - y_{i_2}^* = \begin{cases} \left(-\frac{\chi_{s_{i_1}}}{k} + \frac{\gamma_{s_{i_1}}}{2} + 1.5 \frac{f_{s_{i_1},t_k}}{k} \right) - \left(-\frac{\chi_{s_{i_2}}}{k} + \frac{\gamma_{s_{i_2}}}{2} + 1.5 \frac{f_{s_{i_2},t_k}}{k} \right) \\
\quad + \left(2\Delta s_{i_1} - 2\Delta^2 s_{i_1} \right) \cdot \frac{f_{s_{i_1},t_k}}{k} - \left(2\Delta s_{i_2} - 2\Delta^2 s_{i_2} \right) \cdot \frac{f_{s_{i_2},t_k}}{k} \\
\quad + \left(-2\Delta^2 s_{i_1} \gamma_{s_{i_1}} - (-2\Delta^2 s_{i_2} \gamma_{s_{i_2}}) \right). \end{cases}$$

(33)

(34)

(35)

Consider (33). There are two cases: (i) If $-2\chi_{s_{i_1}} + k\gamma_{s_{i_1}} + 3f_{s_{i_1},t_k} \neq -2\chi_{s_{i_2}} + k\gamma_{s_{i_2}} + 3f_{s_{i_2},t_k}$, then $|y_{i_1}^* - y_{i_2}^*| \geq 1/2k + o(1/k)$. Note that $v = O(1/k^3)$. Thus, $\tilde{q}_{s_{i_1}}(x)$ and $\tilde{q}_{s_{i_2}}(x)$ do not intersect in $[LB, UB]$. (ii) If $-2\chi_{s_{i_1}} + k\gamma_{s_{i_1}} + 3f_{s_{i_1},t_k} = -2\chi_{s_{i_2}} + k\gamma_{s_{i_2}} + 3f_{s_{i_2},t_k}$, (33) equals zero and we can move on to (34) and (35). There are also two cases: (i) If $f_{s_{i_1},t_k} \neq f_{s_{i_2},t_k}$, then $|y_{i_1}^* - y_{i_2}^*| \geq 1/4k^2 + o(1/k^2)$. Note that $v = O(1/k^3)$. Thus, $\tilde{q}_{s_{i_1}}(x)$ and $\tilde{q}_{s_{i_2}}(x)$ does not intersect in $[LB, UB]$. (i) If $f_{s_{i_1},t_k} = f_{s_{i_2},t_k}$, then $|y_{i_1}^* - y_{i_2}^*| \geq 1/16k^2 + o(1/k^2)$ since $\gamma_{s_{i_1}} \neq \gamma_{s_{i_2}}$. Note that $v = O(1/k^3)$. Thus, $\tilde{q}_{s_{i_1}}(x)$ and $\tilde{q}_{s_{i_2}}(x)$ does not intersect in $[LB, UB]$. This analysis reveals that, among these four alternative curves, no two curves intersect in $[LB, UB]$. Thus, there exist one curve, among all these four curves, whose $q_s(x)$ value is uniformly higher than the rest three curves over this interval.

Let $Set_{alter} = \{[\sqrt{2k}] - 1, [\sqrt{2k}], [\sqrt{2k}], [\sqrt{2k}] + 1\}$. After all this analysis, we have $i^* = \arg \max_{i \in Set_{alter}} y_{i}^*$, where $y_{i}^*$ is defined in (28). And this $i^*$ is the one used in the main part of Theorem 5.1. The analysis uses the order of $k$, which can be shown valid for $k \geq 64$. For smaller $k$, we have used computer codes to verify the results.

In the rest of this appendix and also throughout Appendix B, the discussions are all based on Models (2) and (3). For the convenience of later analysis, we introduce the following notation under Model (2).

$$g_s(z) = k(1 - 4z) + 2(3k - 1 - (k + t - 2)/(kt))z^2 + B_s - A_s/k,$$

where $B_s = 4\chi_s(1 - 2z)z + 2\psi_s z^2$ and $A_s = \chi_s - 2z(f_{s,t_1} + f_{s,t_k}) + 2\delta_{t_1,t_k} z^2$. Also, we define $g_s(z) = B_s - A_s/k$.

**Proof of Theorem 6.1.** When a sequence is reversed from $s$ to $s'$, we can easily show that the $q_s(x) = q_{s'}(x)$ and $q_s(x_1, x_2) = q_{s'}(x_1, x_2) = q_s(x_2, x_1)$. That is to say $q_s(x)$ is symmetric about $x_1 = x_2$. Thus we have $x^* = \{z^*, z^*\}$ for some $z^*$. Based on this result, (i)–(iii) are trivial to prove.

**Proof of Theorem 6.2.** Part (i) is trivial by comparing the $q_s(x)$ function for two models. When $\xi$ is dual, $q_\xi(x)$ is symmetric about $x_1 = x_2$. It indicates $\min_{x \in [0, 1]} q_\xi(x) = \min_{x \in [0, 1]} q_\xi(x)$ and part (ii) follows. Part (iii) can be proved from the fact that $pr^\perp(L_d|R_d) \leq$
DESIGNS FOR TOTAL EFFECTS

Here we only prove Theorem 6.2(iv). Note that we have $c_{ξ01} = c_{ξ02}$, $c_{ξ11} = c_{ξ12}$ and $c_{ξ12} = c_{ξ21}$ for any symmetric and self-dual measure $ξ$. The $y_k$ in (13) can be calculated as $c_{ξ00} - 2c_{ξ01}^2/(c_{ξ11} + c_{ξ12})$ for both Models (2) and (3). Thus, the efficiency of $ξ$ under these two models are the same.

**Proof of Theorem 6.4.** We shall begin with the proof of $x_s^* \in [0.4, 0.5]$. Let us focus on $A_s$. It is easy to show that, when $i$ different treatments are included in the sequence, the $χ_s$ term in $A_s$ is minimized by having each treatment for either $|k/i|$ or $|k/i| + 1$ times. Note that, when $z = 0.4$, $B_s = 0.32 γ_s + 0.32 ψ_s$. Thus, for any sequence $s$ we can reformulate it into the form of $L = (1...1, 2...2, ..., i...i)$ if $t_1 \neq t_k$ or $L' = (1...1, 2...2, ..., i...i, 1...1)$ if $t_1 = t_k$ without changing $A_s$, and $B_s$ is increased or unchanged since the increase in $γ_s$ is larger than or equal to the decrease of $ψ_s$. For such sequences, we have $q_s'(0.4) \leq 0$ when $k \geq 4$ and $q_s'(0.4) < 0$ when $k \geq 6$. Thus, we can always find a sequence of pattern $L$ such that it maximizes $q_s(z)$ among all $s \in S$. Note that, when $z = 0.5$, $B_s = 0.5 ψ_s$. Thus, for any sequence $s$ we can reformulate it into the form of $(1212...1212, 3...3, 4...4, ...)$. Without changing $A_s$, and $B_s$ is increased or unchanged since $ψ_s$ is increased or unchanged. For such sequences, we have $q_s'(0.5) > 0$. We have proved that $q_s$ is strictly convex of any $s$ and is $\max_{s \in S} q_s$. As a result $x_s^* \in (0.4, 0.5).

Now we shall move on to find the index $i \leq \min(k, t)$ which maximizes $q_s(0.4)$, and then evaluate the efficiency of $(s_i)$. By direct calculations we have

\begin{align}
q_s(0.4) &= 0.36k - k^{-1}χ_s + 0.32γ_s + 0.32ψ_s - U_{0.4}(k, t, s), \\
q_s'(0.4) &= 0.8k + 1.6ψ_s - 2.4γ_s - H_{0.4}(k, t, s), \\
q_s(0.5) &= 0.5k - k^{-1}χ_s + 0.5ψ_s - U_{0.5}(k, t, s), \\
q_s'(0.5) &= 2k + 2ψ_s - 4γ_s - H_{0.5}(k, t, s),
\end{align}

where (i) $H_{0.4}(k, t, s) = 1.6 + \frac{1.6(k+t-2)}{k} f_{s,t} + \frac{1.6}{k} f_{s,t} + 2 f_{s,t}$ with $U_{0.4}(k, t, s) = 0.32 - 0.8 f_{s,t} + 2 f_{s,t}$, and (ii) $H_{0.5}(k, t, s) = 2 + \frac{2(k+t-2)}{k} + 2 f_{s,t} - 2 f_{s,t}$ with $U_{0.5}(k, t, s) = 0.5 - 2 f_{s,t} + f_{s,t} + 0.5 k+t-2 + 0.5 f_{s,t}$. For sequences of two types $L$ and $L'$, it is not hard to show that no such sequence can have any treatment repeating more than $0.8k$ times. For type $L'$ sequence, we can move the first several $1$ to the tail and balance the sequence into the form of the sequence $s_i$ defined before Theorem 6.4 such that $q_s(0.4)$ is increased or unchanged. For type $L$ sequence, we can simply balance the sequence such that $q_s(0.4)$ is increased or unchanged. Similarly, we can also prove that, one can move the last several $1$'s in $L'$ sequence to the front such that it becomes $L$ and $q_s(0.4)$ is not decreased. Thus, there must be a sequence $s_i$ for some $1 \leq i \leq \min(t, k)$ such that it belongs to $T$. For all $s_i$, $1 \leq i \leq \min(t, k)$, we can calculate $q_{s_i}(0.4)$ and choose the largest one, indexed by $i_0$. Similar to the analysis in the proof of Theorem 5.1, we can show that the optimal $k$ has at most three choices that $i_0 \in \{\text{round}(\sqrt{k}/0.96) - 1, \text{round}(\sqrt{k}/0.96), \text{round}(\sqrt{k}/0.96) + 1\}.

\begin{align}
q_s(z) &= q_{s,0} + q_{s,1}z + q_{s,2}z^2,
\end{align}

where $q_{s,0} = k - \frac{χ_s}{2} - q_{s,1} = 2(2γ_s - 2k + f_{s,t} + f_{s,t})$, $q_{s,2} = 6k - 8γ_s + 2ψ_s - 2k+t-2 - 2 f_{s,t}/k - 2$. For $1 \leq i \leq \min(t, k)$, since the $y^*$ for Model (2) satisfies $\max_i y_i^* \leq y^* \leq \max_s q_s(0.4)$, we can see that

\begin{align}
e_R \geq \frac{\max_i y_i^*}{\max_s q_s(0.4)} \geq \frac{y_{i_0}^*}{q_{s_{i_0}}(0.4)}
\end{align}
\[ y^*_k = k - \frac{\chi_{s_i0}}{k} - \frac{4(2\gamma_{s_i0} - 2k + f_{s_i0,t_1} + f_{s_i0,t_k})^2}{4(6k - 8\gamma_{s_i0} + 2\psi_{s_i0} - 2k + t - 2kt_s + 2t_k)} \]
\[ = \frac{0.36k - k^{-1}\chi_{s_i0} + 0.32\gamma_{s_i0} + 0.32\psi_{s_i0} - U_{0.4}(k, t, s_i0)}{0.36k - k^{-1}\chi_{s_i0} + 0.32\gamma_{s_i0} + 0.32\psi_{s_i0} - U_{0.4}(k, t, s_i0)} \]

where \( i_0 \in \{ \text{round}(\sqrt{k/0.96}) - 1, \text{round}(\sqrt{k/0.96}), \text{round}(\sqrt{k/0.96}) + 1 \} \). A more tedious but straightforward analysis derives the \( v(k, t) \) in Theorem 6.4.

**Acknowledgements.** Dr. Kong’s research is partially supported by NSFC grant 11801033 and Beijing Institute of Technology Research Fund Program for Young Scholars. Dr. Zheng’s research is partially supported by National Science Foundation, DMS-1830864.

**REFERENCES**


[78] STADLER (1921). University of Missouri Agricultural Experimental Station Research Bulletin 49.


