Submitted to *Operations Research*

Sequential Learning With a Similarity Selection Index

Yi Zhou
Department of Mathematics, University of Maryland, College Park, MD 20742, yzh@umd.edu

Michael C. Fu
Robert H. Smith School of Business & Institute for Systems Research, University of Maryland, College Park, MD 20742, mfu@umd.edu

Ilya O. Ryzhov
Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, iryzhov@umd.edu

We consider the ranking and selection problem of finding the best alternative in a setting where prior similarity information between the alternatives can be learned from data. Incorporating similarity information enables efficient budget allocation for faster identification of the best alternative in sequential selection. Using a new selection criterion, the similarity selection index, we develop two new allocation methods: one based on a mathematical programming characterization of the asymptotically optimal budget allocation, and the other based on a myopic expected improvement measure. For the former, we present a novel sequential implementation that provably learns the optimal allocation without tuning. For the latter, we also derive its asymptotic sampling ratios. We also propose a practical way to update the prior similarity information as new samples are collected. Numerical results illustrate the effectiveness of both methods.

Key words: ranking and selection, simulation optimization, spectral clustering

1. Introduction

In the ranking and selection problem (Chen et al. 2015), there is a finite set of feasible solutions ("systems" or "alternatives") with unknown performances that must be estimated from noisy observations, e.g., from expensive stochastic simulation experiments. A single experiment only provides information about a single alternative, creating a tradeoff: spending resources, such as simulation time, to learn about one alternative means that less information can be collected about the others. Much of the research in this area focuses on designing algorithms that sequentially allocate experiments to alternatives (adapting to the results of past assignments) for the purpose of identifying the highest-valued alternative as quickly as possible.

This problem can be approached using a wide variety of algorithmic concepts, including indifference-zone selection (Kim and Nelson 2001), value of information (Frazier et al. 2008, Chick

et al. 2010), optimal computing budget allocation (Chen et al. 2000, He et al. 2007), procedures based on stopping boundaries (Fan et al. 2016, Ma and Henderson 2017), and optimal sampling laws (Glynn and Juneja 2004, Hunter and Pasupathy 2013). These and many other papers assume that information is completely independent across alternatives, i.e., an experiment with one alternative provides no information about any others. Often, however, it is possible to exploit structural similarities or differences between alternatives to solve the problem more efficiently. For example, alternatives may have known common attributes that influence their performance. We may also have past performance data about other alternatives that come from the same context, giving us some prior knowledge about the problem at hand. Two examples of such settings are:

- In the simulation of automatic control systems, performance is affected by mechanical settings that are known to users before simulation. These settings may have been tested before for other systems. Thus, when optimizing a new system for which no test results are available, we may use these past data to help the simulation process (Gao and Lu 2021).
- In a clinical trial, an important concern is estimation of the dose-response relationship, which is influenced by factors such as age and gender. When determining the optimal dose for a new patient group (e.g., children), for which no data have yet been collected, we may use test results from other groups (e.g., adults) as prior information since the dose-response curves for both groups may follow similar patterns (Xue et al. 1999).

In these and other settings, past data provides prior information about new alternatives of interest – not only about their performance, but also about similarities between these performance values. As a consequence, when we conduct a new experiment on one alternative, the similarity information allows us to learn about other, similar alternatives. This can significantly reduce the expenditure of resources needed to identify the best alternative reliably, greatly improving the ability to solve large instances where exhaustive simulation may not be practical. The literature has recognized this advantage of similarity information and has generally handled it in two ways. In situations where alternatives are naturally described by certain attributes or features,

one can relate their values to these features using linear regression; see Han et al. (2016) or Shen et al. (2017) for examples of such parametric models. If no such features are available, one can instead quantify pairwise similarities in the form of a matrix. This approach is often adopted by Bayesian optimization methods, e.g., by Frazier et al. (2009) for ranking and selection or Scott et al. (2010) for continuous simulation optimization. In these methods, the matrix is included in the decision-maker's Bayesian belief about the unknown values and is updated after each experiment; furthermore, these "correlated beliefs" are also used inside the allocation procedure, leading to a significant increase in overall computational cost.

This paper investigates a different way to model similarities inside ranking and selection, which we call the *similarity selection index*, or *S-index* for short. This approach is motivated by spectral clustering techniques (Von Luxburg 2007), and was first proposed by Sun et al. (2019) in the *Proceedings of the Winter Simulation Conference* as a way to improve the performance of ranking and selection procedures.[1] One calculates the S-index by taking a certain linear transformation of the estimated values, which incorporates similarity information between them; one then uses these transformed values to select an alternative. Sun et al. (2019) provided some preliminary theoretical results (including several unproved conjectures) and numerical experiments in support of the S-index as a selection criterion, but did not develop any new algorithms.

In this approach, similarity information is represented by weights assigned to edges in a graph whose vertices are the alternatives. The linear transformation then adjusts the estimated values of these alternatives based on the weights. Thus, the weights play a similar role to that of correlated beliefs in Bayesian methods. However, the main distinguishing characteristic of this approach, relative to the simulation literature, is that similarity information is utilized only inside the selection criterion used to return the best alternative after the experiments have concluded, rather than inside the statistical model used to learn the unknown values. The learning process still assumes

[1] Sun et al. (2019) introduced the concept under the name "spectral index." In this paper, we use "similarity selection index" to avoid confusion with the spectrum of a matrix.

that the values of individual alternatives are estimated independently, as in traditional ranking and selection, but in the selection step, each alternative is evaluated using the S-index rather than the sample mean. In this way, we fully use the power of similarity information while reducing the computational cost of the allocation procedure.

Since the S-index is used for selection, rather than allocation, it can potentially be used in conjunction with any existing allocation method, which is how it was implemented by Sun et al. (2019). However, allocations should perform better if they are tailored to the new selection criterion. We propose two such allocation methods in this paper. The first method calculates a budget allocation by optimizing an approximation of the probability of correct selection (PCS) from a geometric perspective. This approximate problem, however, is shown to yield an asymptotically optimal solution, because it can be shown to be equivalent to the well-established problem of optimizing large deviations rates of PCS, first posed by Glynn and Juneja (2004).

The optimal budget allocation is a function of the unknown values, and cannot be directly implemented; one could approximate it by solving the optimization problem with plug-in estimators, but this is computationally cumbersome. We address this difficulty by developing a sequential implementation that provably converges to the optimal allocation, without tuning, and without exactly solving a convex optimization problem in every step. The design of this algorithm is quite distinct from existing sequential procedures, and leverages techniques from nonlinear optimization in a novel way. We also develop a second heuristic approach, based on value of information, which is simple to compute and performs remarkably well in numerical experiments.

The advantages of the S-index depend on the availability of accurate similarity information. To make this approach more robust against misspecification, we develop a procedure for learning prior similarities from existing data and sequentially updating them as new information is collected. It should be noted that, while correlated Bayesian beliefs have to deal with the same concern, in the simulation literature only Qu et al. (2015) and Zhang and Song (2017) have dealt with the problem of learning similarity structures to a significant degree. In general, identifying the most effective similarity structure is itself a subject for research (Malkomes and Garnett 2018).

In sum, our work makes the following contributions:

- We formulate an optimal budget allocation problem using a geometric approximation of PCS, and show that the solution to this problem optimizes the asymptotic convergence rate of PCS characterized using large deviations theory. There is a separate stream of literature focusing on large deviations-based allocations, so this result is of stand-alone interest.
- We propose a sequential implementation, leveraging the idea of reduced gradient methods in nonlinear optimization, and prove that this method asymptotically learns the optimal budget allocation with probability 1. We also propose an additional sequential heuristic that is computationally efficient and easy to implement.
- We develop a dynamic update procedure that learns a prior similarity graph from data and sequentially updates it as more samples are collected, significantly strengthening the applicability of our proposed algorithms.
- Numerical experiments comparing our proposed methods to each other and also to several existing benchmarks indicate that the new algorithms perform well. Furthermore, using the S-index for selection can also improve the performance of existing allocation methods.

The paper is organized as follows. Section 2 introduces the S-index and a class of similarity structures in which consistent selection can be guaranteed. Section 3 defines the notion of an optimal budget allocation under S-index selection, and presents a computationally efficient algorithm that can be guaranteed to learn that allocation asymptotically. Section 4 presents a second approach based on Bayesian value of information. Section 5 describes one way in which similarity structures can be updated over time as new information is acquired. Numerical experiments are given in Section 6. The Appendix contains all proofs and some additional results and discussions.

2. The S-index: Definition and Properties

We assume there are k alternatives with unknown values $\mu_1 > \mu_2 > \ldots > \mu_k$. For simplicity, we will develop the main concepts of the paper assuming that all k values are distinct. However, there is no loss of generality in this assumption, and Appendix D gives some additional discussion for the

case where multiple alternatives may have equal values. We model similarity information using a weighted graph representation $G = (V, S)$, where each alternative is viewed as a vertex $i \in V$. Let S be the similarity matrix, with $S_{i,j} = S_{j,i} \geq 0$ being the weight assigned to the edge between vertices $i, j \in V$.

Suppose now that the values μ_i are unknown, but we have access to estimates y_i of μ_i. In ranking and selection, y_i is usually the sample mean of all experiments conducted with alternative i (or the posterior mean, in a Bayesian framework). Now, consider the optimization problem

$$z = \arg\min_{u \in \mathbb{R}^k} \sum_{i=1}^{k} (u_i - y_i)^2 + \frac{\lambda}{2} \sum_{1 \leq i,j \leq k} S_{i,j}(u_i - u_j)^2. \tag{1}$$

The quantity z_i is called the *S-index* of alternative i. If $\lambda = 0$, we will have $z = y$, but when $\lambda > 0$ the solution of (1) will be regularized by the similarity information; consequently, two alternatives with larger similarity will also have similar S-indices. The choice of λ, which controls the relative importance of the similarity information, is left up to the user. We can rewrite (1) in matrix notation as

$$z = \arg\min_{u \in \mathbb{R}^k} (u - y)^T (u - y) + \lambda u^T L u, \tag{2}$$

where $L = L(S) = D - S$ and D is a diagonal matrix with $D_{i,i} = \sum_{j=1}^{k} S_{i,j}$. For simplicity, we let $S_{i,i} = 0$, since the diagonal entries of S do not affect L. The solution to (2) can be expressed in closed form as

$$z = (I + \lambda L)^{-1} y, \tag{3}$$

which allows us to compute the S-indices efficiently, given estimates y and a similarity matrix S. From Von Luxburg (2007), it can be seen that the matrix $I + \lambda L$ is symmetric positive definite (thus invertible) with minimum eigenvalue 1 and corresponding eigenvector $\mathbf{1}$. Sun et al. (2019) discusses alternate ways to compute (3) that do not require matrix inversion.

In much of what follows, we will make use of several properties of the inverse $Q = Q(S; \lambda) = (I + \lambda L(S))^{-1}$, stated below and proved in the Appendix. These properties hold for all similarity graphs.

THEOREM 1. *For a graph G with $S_{i,j} \geq 0$ and $S_{i,j} = S_{j,i}$, $Q = (I + \lambda(D - S))^{-1}$ has the following properties:*

(1) Q is symmetric positive definite, having largest eigenvalue 1 with eigenvector $\mathbf{1}$.

(2) $Q_{i,j} \geq 0$, $\forall i, j = 1, ..., k$, i.e., every entry of matrix Q is non-negative.

(3) $\forall i = 1, ..., k$, $Q_{i,i} > \max_{j \neq i} Q_{i,j}$, i.e., the diagonal entry is the largest entry in both its column and its row.

Classical ranking and selection uses $\arg\max_i y_i$ as the selection criterion, i.e., it predicts the alternative with the highest sample mean as being the best. Given enough samples, $y_i \to \mu_i$ and so the true best alternative will eventually be discovered under this criterion. We propose to use S-indices instead of sample means, so that $\arg\max_i z_i$ becomes the selection rule. Then, the first important question is whether the true best alternative can still be recovered using the linear transformation (3). To that end, Sun et al. (2019) introduced a particular class of similarity structures under which this can be guaranteed. The following is a slightly relaxed version of Definition 1 in Sun et al. (2019).

DEFINITION 1. A graph G is *aligned* if for every fixed $i = 1, ..., k$, the similarities $\{S_{i,j}\}$ satisfy $S_{i,j} \leq S_{i,m}$ for any $j < m < i$ or $j > m > i$.

We can then show that, if the original estimates are sufficiently accurate to recover the correct ordering of the alternatives, this ordering will also be preserved under S-indices. Consequently, selecting alternative $\arg\max_i z_i$ will still lead us to the true best alternative. The proof of Theorem 2, given in the Appendix, corrects a minor technical issue with a similar proof in Sun et al. (2019).

THEOREM 2. *For an aligned graph G and estimates $y_1 \geq y_2 \geq \cdots \geq y_k$, the S-index defined by $z = Qy = (I + \lambda L)^{-1} y$ preserves the ordering, i.e., $z_1 \geq z_2 \geq \cdots \geq z_k$.*

In the rest of the paper, y (and therefore z as well) will be a random vector, whose distribution depends on the number and outcomes of simulation experiments. The performance of a selection criterion based on such a vector can be measured in terms of the probability of correct selection

(PCS), i.e., the probability that the first component of the vector is also the highest-valued. For example, under S-index selection, the PCS can be written as

$$\mathrm{PCS}(z) = P\left(z_1 - z_i \geq 0, i \neq 1\right),$$

and analogously $\mathrm{PCS}\left(y\right)$ is the PCS under the original estimates y. If the similarity graph is very informative about the relationships between the values of alternatives, it is possible to show that $\mathrm{PCS}(z) \geq \mathrm{PCS}(y)$, though this cannot be guaranteed in general. Theorem 3 gives one example of a graph structure that has this property. Although most of our analysis, from Section 3 onwards, assumes that y and z are normally distributed, Theorem 3 does not require specific distributional assumptions, but rather arises from the structure of the similarity measure.[2] The result also does not require the graph to be aligned (and does not hold for all aligned graphs).

THEOREM 3. *If $S_{1,j} \leq S_{i,j}$ for any $i \neq j, i, j \neq 1$, for any allocation policy, $PCS(z) \geq PCS(y)$.*

Intuitively, if the best alternative is "separated" from the others by the similarity measure, PCS will be improved because our estimates of μ_i for suboptimal i will smooth out each other, while exerting little impact on our estimate of μ_1. This may give some intuition why, even if y_i is a minimum-variance unbiased estimator of μ_i, the S-indices z may nonetheless be more accurate than

[2] Theorem 3 establishes Conjecture 3 in Sun et al. (2019).

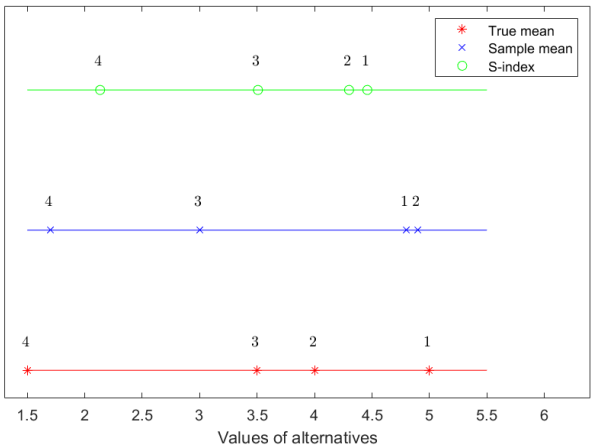

Figure 1 The numbers above the markers represent the indices of the alternatives. The S-indices smooth the sample means and correct the errors between the relative order of the largest two alternatives.

y in identifying the highest-valued alternative. A simple illustration is given in Figure 1, where the sample means suggest a wrong best alternative, while the S-indices give the correct best alternative by smoothing the sample means.

The preceding results treat the similarity graph as given. In practice, if there is reason to believe that alternatives have common attributes that strongly influence their values, one might construct the graph from past data on similar alternatives (see Section 6 for concrete examples). Such a graph generally will not be aligned, nor will it satisfy the conditions of Theorem 3. Nonetheless, it may still contain helpful information that can improve the performance of ranking and selection methods, especially in the early stages. Later, in Section 5, we will present a practical approach for updating the similarity graph over time, in a manner that eventually makes it aligned. For the moment, we will continue to assume that a suitable graph is available.

It is important to keep in mind that, in this methodology, the similarity measure is separate from the statistical mechanism used to estimate the unknown values. The computation of y itself need not use the similarity measure at all; the latter is only used to transform y in the final selection step. In this way, even though the concept of similarity between alternatives is used both in S-index selection and in correlated Bayesian belief models, they use it in different ways. In Bayesian methods, covariance matrices are built into the statistical model used to estimate values, and correlations decay over time (converge to zero) as more information is collected. Furthermore, prior covariances can be obtained from, e.g., spatial similarities that are not directly connected to the true performance values. In S-index selection, the similarity structure has to be derived directly from the true values if one uses an aligned graph. Thus, one can view S-indices as a purely frequentist framework with the ability to model correlated beliefs, but the similarity graph has to be constructed in a different way than in Bayesian methods; see Section 6 for practical examples.

3. Optimal Budget Allocation Under S-index Selection

We now develop methods for allocating the experimental budget under the assumption that the S-index is used for selection. Our main design principle is similar to that of optimal computing

budget allocation (Chen and Lee 2010): we first formulate a mathematical program whose solution is an optimal allocation in a certain theoretical sense. This "ideal" allocation will depend on the unknown values μ, and thus cannot be directly computed or implemented. However, we will design efficient sequential procedures that use only estimated values, but can be guaranteed to converge to the optimal allocation asymptotically.

In the following, we often denote the index of the best alternative by i^* instead of 1 to make the eventual implementation clearer. We suppose that the estimates y are constructed using simple frequentist statistics. Let $w_1, ..., w_k$ be a vector of i.i.d. standard normal random variables, and assume that

$$y = \mu + \Lambda w, \qquad \Lambda = \text{diag}\left(\frac{\sigma_1}{\sqrt{N_1}}, ..., \frac{\sigma_k}{\sqrt{N_k}}\right),$$

where $\sigma_1, ..., \sigma_k$ are fixed positive constants and $N_1, ..., N_k$ are positive integers, i.e., the simulation output is assumed to be normally distributed with standard deviation σ_i, and N_i being the number of experiments allocated to alternative i.

3.1. Geometric Approximation of the Probability of Correct Selection

From (3), we have $z = Q\mu + Q\Lambda w$, so $z \sim \mathcal{N}(Q\mu, Q\Lambda^2 Q^T)$. Let $P_{i,j} = Q_{i^*,j} - Q_{i,j}$, and define $m = P\mu$ and $\Sigma = P\Lambda^2 P^T$. Then, the joint distribution of the vector $\{z_{i^*} - z_i, i \neq i^*\}$ is $\mathcal{N}(m_{-i^*}, \Sigma_{-i^*})$, where the mean vector m_{-i^*} is obtained by deleting the i^*th element of m, and the covariance matrix Σ_{-i^*} is obtained by deleting the i^*th row and column of Σ. Suppose P_{-i^*}, defined similarly, has full rank. By the Cholesky decomposition, we can represent $\Sigma = U^T U$, where $U = [U_{i,j}]_{k\times k}$ is an upper triangular matrix. We are interested in the probability that $\arg\max_i z_i = \arg\max_i \mu_i$, which means that the correct alternative is identified by using the S-index for selection. The corresponding PCS is given by

$$\text{PCS}(z) = \left(\frac{1}{\sqrt{2\pi}}\right)^{k-1} \int \cdots \int_{\{\sum_{j\neq i^*}^{i} U_{j,i} w_j \geq -m_i, i\neq i^*\}} \prod_{i\neq i^*} e^{\frac{-w_i^2}{2}} dw_i \tag{4}$$

Since the terms in the integrand in (4) decay exponentially, the region near the origin dominates the value of the integral (Fu et al. 2007). Therefore, (4) can be approximately optimized by maximizing

the volume of the hypersphere contained inside the domain of integration and centered at the origin. This approximation methodology was also used in Peng et al. (2018) for a traditional ranking and selection problem.

The distance from the origin to the hyperplane $\sum_{j\neq i^*}^{i} U_{j,i}w_j = -m_i$, $i \neq i^*$ is given by

$$\frac{m_i}{\sqrt{\sum_{j\neq i^*}^{i} U_{j,i}^2}} = \frac{m_i}{\sqrt{\Sigma_{i,i}}}, \tag{5}$$

where $\Sigma_{i,i} = \sum_{j=1}^{k} \frac{\sigma_j^2}{N_j} P_{i,j}^2$. To maximize the volume of the hypersphere, it is sufficient to solve

$$\min_{\substack{\sum_{i=1}^{k} N_i = M \\ N_j \geq 0}} \max_{i\neq i^*} \frac{1}{({P_i}^T \mu)^2} \sum_{j=1}^{k} \frac{\sigma_j^2}{N_j} P_{i,j}^2, \tag{6}$$

which is equivalent to maximizing the smallest distance to any of the hyperplanes. The quantity M represents the total experimental budget; alternatively, one could set $M = 1$ and interpret N_j as the proportion of the budget to assign to alternative j. The objective in (6) can be linearized, giving rise to the equivalent problem

$$\begin{aligned} \text{minimize} \quad & \xi \\ \text{subject to} \quad & \sum_{j=1}^{k} \frac{a_{i,j}}{N_j} \leq \xi, \quad i \neq i^*, \\ & N_1 + N_2 + \cdots + N_k = M, \\ & N_j \geq 0 \quad j = 1, 2, \ldots, k, \end{aligned} \tag{7}$$

where $a_{i,j} = \frac{\sigma_j^2 P_{i,j}^2}{({P_i}^T \mu)^2}$, $i \neq i^*$, $j = 1, ..., k$.

From the properties of convex functions (Boyd and Vandenberghe 2004), it is easy to see that the mapping $x \mapsto \sum_{j=1}^{k} \frac{a_j}{x_j}$, where $a_j \geq 0$, is convex on $\mathbb{R}_+^k$. If we allow N_j to be continuous, (7) will be a convex program, guaranteed to have a unique optimal solution. In general, this solution cannot be computed analytically, with the exception of the special case $k = 3$, treated in the following result.

PROPOSITION 1. *Let $i_1, i_2 \neq i^*$ be denote the two distinct suboptimal alternatives in $\{1,2,3\}$, and define $r(\gamma) = \sum_{j=1}^{3} \frac{a_{i_1,j} - a_{i_2,j}}{\sqrt{\gamma a_{i_1,j} + (1-\gamma) a_{i_2,j}}}$. Let*

$$\gamma^* = \begin{cases} 1 & \text{if } r(0) > 0 \text{ and } r(1) > 0 \\ 0 & \text{if } r(0) < 0 \text{ and } r(1) < 0 \\ \gamma_0 & \text{otherwise,} \end{cases}$$

where γ_0 is the solution of $r(\gamma) = 0$, which can be efficiently obtained using a bisection algorithm. Then for $k = 3$, the optimal solution of (7) is given by

$$N_j^* = \frac{\sqrt{\gamma^* a_{i_1,j} + (1-\gamma^*) a_{i_2,j}}}{\sum_{j'=1}^{3} \sqrt{\gamma^* a_{i_1,j'} + (1-\gamma^*) a_{i_2,j'}}} M, \quad j = 1,2,3.$$

Whether or not we are in this special case, however, the solution depends on the unknown values μ (as well as on σ, which may also be unknown), and thus cannot be implemented directly. Instead, one might replace the unknown parameters in (7) with plug-in estimators computed from past experiments, and use the solution of the resulting approximate problem to allocate a portion of the budget. By iteratively resolving (7) with updated estimators, one will eventually arrive at the true optimal solution. If $k > 3$, the computational cost will be fairly high, due to the need to repeatedly solve convex programs, but it is possible to do this in principle using a solver such as CVX (Grant and Boyd 2014).

Algorithm 1 gives one possible implementation, which we call SIOCBA (S-index optimal computing budget allocation), where e_i denotes a k-vector of zeroes with the ith coordinate equal to 1. For simplicity, we treat σ as being known; in practice, the variances can be estimated using sample variances and updated as more data are collected. In the nth stage of sampling, M^n is the total budget spent thus far, y^n is the current vector of sample means, with $z^n = Qy^n$ being the corresponding vector of S-indices. We denote by $i^{*,n} = \arg\max_j (Qy^n)_j$ the index of the alternative currently believed to be the best (note that S-indices are used to identify this alternative), and use the quantities $P_{i,j}^n = Q_{i^{*,n},j} - Q_{i,j}$ and $a_{i,j}^n = \frac{\sigma_j^2 \left(P_{i,j}^n\right)^2}{\left(\left(P_i^n\right)^T y^n\right)^2}$, for all $i \neq i^{*,n}$, in the computation. The inputs i^*, $a_{i,j}$ and M in (7) are replaced by $i^{*,n}$, $a_{i,j}^n$ and M^n, respectively. The next experiment

Algorithm 1: S-index optimal computing budget allocation (SIOCBA).

Input: number of alternatives k, variances σ_i^2, total budget M, number of initial samples n_0, matrix Q.
Output: final selection $i^{*,M}$.
Initialization: $n=0$, $N_j^n=n_0$, $j=1,...,k$, $M^0=kn_0$.
Collect n_0 samples for each alternative and calculate sample means y^0.
while $M^0+n<M$ **do**
  $i^{*,n}\leftarrow \arg\max_j (Qy^n)_j$; `// find the best alternative using current S-index`
  $P^n\leftarrow \text{kron}(Q(i^{*,n},:),\text{ones}(k,1))$; `// update the optimization problem`
  $a_{i,j}^n\leftarrow \frac{\sigma_j^2\left(P_{i,j}^n\right)^2}{((P_i^n)^T y^n)^2}$, $i\neq i^{*,n}$, $j=1,..,k$;
  $M^n\leftarrow M^n+1$;
  $N^{n,*}\leftarrow$ optimal solution of the updated optimization problem (7);
  $j^n=\arg\max_j \frac{N_j^{*,n}}{N_j^n}$;
  $N^{n+1}\leftarrow N^n+e_{j^n}$;
  Sample alternative j^n and update the sample means y^{n+1};
  $n\leftarrow n+1$;
end
$i^{*,M}\leftarrow \arg\max_j (Qy^M)_j$

is then allocated to alternative $j^n=\arg\max_j \frac{N_j^{*,n}}{N_j^n}$, which favors those alternatives whose actual sample sizes N_j^n are small relative to the estimated optimal x_j^n (i.e., they are under-sampled).

Of course, even if the unknown values were known, the objective of (7) is based on an approximation of (4), not the exact PCS. However, it can be shown that the geometric approximation produces an asymptotically exact solution in the regime where $M\to\infty$. To show this, we turn to the asymptotic convergence rates of the probability of *incorrect* selection.

3.2. Large Deviations Analysis of Probability of Incorrect Selection

Let us now increase the experimental budget in a way that satisfies $\lim_{M\to\infty}\frac{N_j^M}{M}=x_j$ for each j, where each $x_j>0$ is a strictly positive constant. In other words, the budget becomes large, but each alternative receives a certain prespecified, nonzero proportion of it in the long term. For now, let us view the proportions x_j as fixed, but our ultimate goal will be to optimize them in a sense to be formalized later.

For $i\neq i^*$, define the "error set" $E_i=\{z: z_i\geq z_{i^*}\}$. We interpret $z\in E_i$ as a vector of possible S-indices under which the suboptimal alternative i appears to be better than the true best alternative

i^*. Let $z^M = Qy^M$ be the (random) vector of S-indices obtained after M experiments have been conducted. Then,

$$1 - PCS\left(z^M\right) = P\left(z^M \in \bigcup_i E_i\right),$$

that is, we select an incorrect alternative if z^M estimates at least one suboptimal alternative as being better than i^*. In the following, we will show that a certain $R_i > 0$ satisfies

$$\lim_{M\to\infty} \frac{1}{M} \log P\left(z^M \in E_i\right) = -R_i, \qquad i \neq i^*, \tag{8}$$

where, by the arguments in Glynn and Juneja (2004), it straightforwardly follows that

$$\lim_{M\to\infty} \frac{1}{M} \log P\left(z^M \in \bigcup_i E_i\right) = -\min_{i\neq i^*} R_i. \tag{9}$$

In words, for large M, $P\left(z^M \in E_i\right)$ behaves like $e^{-R_i \cdot M}$, and the overall probability of incorrect selection is governed by the smallest (slowest) of the rate exponents across $i \neq i^*$.

The derivation of (8) uses the Gärtner-Ellis theorem, as laid out in Ch. 1 of Dembo and Zeitouni (2009). Using this result, we can compute

$$R_i = \inf_{z\in E_i} I\left(z\right), \tag{10}$$

where I is the large deviations rate function of the sequence $\{z^M\}$. Since, for any M, $z^M = Qy^M$ is a linear transformation of a vector of independent normal random variables, it follows a multivariate normal distribution, and therefore I can be computed in closed form, as shown in the following result.

PROPOSITION 2. *The sequence $\{z^M\}$ obeys a large deviations law with rate function*

$$I\left(z\right) = \frac{1}{2}\left(z - Q\mu\right)^T Q\Gamma^{-1}Q^T\left(z - Q\mu\right),$$

where Γ is a diagonal matrix with $\Gamma_{j,j} = \frac{\sigma_j^2}{x_j}$.

Thus, the right-hand side of (10) is the optimal value of a convex optimization problem with a quadratic objective and a single linear constraint $v_i^T z \leq 0$, where v_i is a vector of zeros with

$v_{i,i} = -1$ and $v_{i,i^*} = 1$. Applying results in Section 2 of Zhou and Ryzhov (2022), we arrive at the closed-form solution

$$R_i = \frac{(v_i^T Q\mu)^2}{v_i^T Q\Gamma Q^T v_i}. \tag{11}$$

Note that R_i depends on the vector x of proportions through the diagonal matrix Γ.[3] One can then formulate the concave maximization problem

$$\begin{aligned} \text{maximize} \quad & \min_{i \neq i^*} R_i(x) \\ \text{subject to} \quad & \mathbf{1}^T x = 1, \quad x_j \geq 0, \quad j = 1, ..., k. \end{aligned} \tag{12}$$

Recalling (9), we can see that the objective function in (12) is the rate exponent for the probability of incorrect selection. By maximizing this quantity, this probability is made to converge to zero at the fastest possible rate.

At the same time, using the notation from Section 3.1, it is easy to see that $R_i(x) = 1/\sum_{j=1}^k \frac{a_{i,j}}{x_j}$. Therefore, (12) is equivalent to (6) with the experimental budget scaled to 1. This proves that, by using the geometric approximation of PCS, we obtain a budget allocation that optimizes the asymptotic convergence rate of the probability of incorrect selection.

Within the simulation community, there is an entire stream of literature studying large deviations-based problems similar to (12), beginning with the seminal work of Glynn and Juneja (2004), with some representative papers being Hunter and Pasupathy (2013), Pasupathy et al. (2014) and Gao et al. (2017). We find it is closely connected to the geometric approximation of Section 3.1, and the derivation of the optimal allocation under S-index selection is completely new.

3.3. A Provably Convergent Sequential Implementation Under S-index Selection (SIGD)

We have proposed SIOCBA to approximately maximize PCS when selecting with S-indices. However, when k is large, solving the nonlinear optimization problem (7) is expensive. Therefore, in this

[3] If $S = 0$, and consequently $Q = I$, we are back in the setting of classical ranking and selection, and it is easy to see that (11) reduces to the rate function derived in Example 1 of Glynn and Juneja (2004).

section, we propose a new sequential allocation policy that is more computationally efficient. We begin by deriving the convex dual of (7), since it will be more convenient to solve sequentially than the primal. We normalize $M = 1$, since we are interested in learning the asymptotically optimal solution as the budget becomes large.

THEOREM 4. *The primal problem*

$$\begin{aligned} \min_{x} \quad & \max_{i \neq i^*} \sum_{j=1}^{k} \frac{a_{i,j}}{x_j} \\ \textit{subject to} \quad & x_j \geq 0, \quad j = 1, ..., k, \\ & \mathbf{1}^T x = 1, \end{aligned} \tag{13}$$

has the convex dual problem

$$\begin{aligned} \max_{\gamma} \quad & \sum_{j=1}^{k} \sqrt{\sum_{i \neq i^*} \gamma_i a_{i,j}} \\ \textit{subject to} \quad & \gamma_i \geq 0, \quad i \neq i^*, \\ & \mathbf{1}^T \gamma = 1, \end{aligned} \tag{14}$$

and strong duality holds.

Using the KKT optimality conditions of these problems, one finds that the optimal primal-dual pair (x^*, γ^*) satisfies

$$x_j^* = \frac{\sqrt{\sum_{i \neq i^*} \gamma_i^* a_{i,j}}}{\sum_{j=1}^{k} \sqrt{\sum_{i \neq i^*} \gamma_i a_{i,j}}}, \quad j = 1, ..., k, \tag{15}$$

which allows us to easily convert a dual solution into a primal one. Of course, the index i^* and the inputs $a_{i,j}$ depend on the unknown values μ, so the dual solution still cannot be solved exactly. As in Section 3.1, one could potentially resolve it with the true values replaced by estimates, thus converging to the optimal solution over time, but we would like to avoid the cumbersome cost of solving a sequence of convex programs.

At a high level, our approach is based on the following concept. In nonlinear programming, one would typically solve (14) using an iterative gradient descent method, which repeats certain

computations on the same problem instance until the iterates appear to converge (and one knows from theory that their limit must be a stationary point of the optimization problem). Our sequential procedure resembles such a method, but we now run only one iteration of it in each stage of sampling. In other words, given estimates y^n, we perform a single iteration of the gradient algorithm on problem (14) with μ replaced by y^n. This yields a new iterate γ^n, which is likely not optimal for the current dual problem (if we wanted to solve it to optimality, we would need to run many iterations). We then use γ^n to choose an alternative for the next experiment, with the precise manner of this choice to be described further down. The results of the experiment produce a new vector y^{n+1} of estimates, and the process is repeated. In other words, each iteration of the gradient algorithm is performed on a slightly different *problem*, namely, the dual with a different set of estimates. Doing this is much faster than solving each individual problem to optimality, yet we prove that in this way the true solution of (14) is still recovered asymptotically.

The formal statement of the procedure is given in Algorithm 2. We first explain the notation used in the statement, and motivate the steps of the algorithm, before proceeding to the main convergence results. Denote by $g(\gamma; y) = -\sum_{j=1}^{k} \sqrt{\sum_{i \neq i^*(y)} \gamma_i a_{i,j}(y)}$ the objective function of the dual problem (14) with a generic vector y in place of μ, and the dependence of $i^*, a_{i,j}$ on y made explicit.

The initialization of the algorithm involves two small constants κ_0 and η, which are used for numerical stability and do not require any additional assumptions on the problem. In the nth iteration, Step 0 updates the inputs of the dual problem using the current estimates y^n. Step 0′ is included for numerical stability, due to the nondifferentiability of g on the boundary of the dual feasible region; as will be argued in our proof, this step will be invoked at most finitely many times and thus does not play a major role in convergence analysis.

Step 1 selects an index h^n for which $\gamma_{h^n}^{n-1} > 0$, using a small threshold η for numerical stability. This index is needed to determine a descent direction in which to move. It is not necessary to consider all possible directions; Lin et al. (2009) has shown that it is sufficient to consider elements of the set

$$D^h(\gamma) = \{e_i - e_h : i \neq h\} \bigcup \{e_h - e_i : i \neq h, \gamma_i > 0\}, \tag{16}$$

Algorithm 2: Sequential gradient descent algorithm for the dual problem (SIGD).

Input: A small constant κ_0 and $\eta < \frac{1}{k-1}$. Number of initial budgets N_0 for each alternative. Total Budget M.
Output:: Sample means y^{M-kN_0+1}.
Initialization: Compute y^1 using a first-stage sample (e.g., by running N_0 experiments for each alternative).
Repeat the following steps for $n = 1, 2, \ldots, M - kN_0$:
Step 0: Let $i^{*,n} = \arg\max_i [Qy^n]_i$, compute $a_{i,j}(y^n)$.
Step 0': Check whether $\min_j \sum_{i \neq i^{*,n}} \gamma_i^{n-1} a_{i,j}(y^n) > 0$. If this inequality does not hold, perturb γ^{n-1} to ensure the gradient is well defined.
Step 1: Choose some $h^n \in \left\{ i \neq i^{*,n} : \gamma_i^{n-1} \geq \eta \right\}$ with equal probability.
Step 2: Compute the descent direction d^n

$$\underset{d \in D^{h^n}(\gamma^{n-1})}{\arg\min} \; s^{\max}\left(d, \gamma^{n-1}\right) \nabla g\left(\gamma^{n-1}; y^n\right)^T d,$$

where $D^{h^n}(\gamma^{n-1})$ is as in (16), and $s^{\max}(d, \gamma^{n-1})$ is computed using (17). Let $V^n = \nabla g(\gamma^{n-1}; y^n)^T d^n$.
Step 3: If V^n does not satisfy (18) or (19), let $\gamma^n = \gamma^{n-1}$. Otherwise, choose $s^n = \text{LineSearch}(d^n, s^{\max}(d^n, \gamma^{n-1}), \gamma^{n-1}, y^n)$ and let $\gamma^n = \gamma^{n-1} + s^n d^n$.
Step 4: Compute x_j^n using (15) with inputs γ^n and $a_{i,j}(y^n)$.
Step 5: Choose $j^n = \arg\max_j \frac{x_j^n}{N_j^n}$, set $N_j^{n+1} = N_j^n + \mathbf{1}\{j = j^n\}$. Sample alternative j^n and compute updated sample means y^{n+1}.

for any h satisfying $\gamma_h > 0$. See Appendix A.2 for the details. This reduction of the set of feasible directions to the much smaller set D^h motivates a 2-coordinate descent algorithm. For any direction $d \in D^h(\gamma)$, we define $s^{\max}(d, \gamma)$ to be the largest stepsize s such that $\gamma + s \cdot d$ remains dual-feasible. This quantity can be explicitly computed as

$$s^{\max}(d, \gamma) = \begin{cases} \gamma_h, & \text{if } d = e_j - e_h, j \neq h \text{ and } \nabla g(\gamma; y)^T d < 0, \\ \gamma_j, & \text{if } d = e_h - e_j, j \neq h \text{ and } \nabla g(\gamma; y)^T d < 0, \\ 0, & \text{if } \nabla g(\gamma; y)^T d \geq 0. \end{cases} \tag{17}$$

Step 2 performs this computation for all $D^{h^n}(\gamma^{n-1})$ and selects the direction d^n of steepest descent. In the numerical optimization literature, techniques that minimize a first-order linear approximation of the objective are known as reduced gradient methods (Bazaraa et al. 2013).

Step 3 then determines how far to go in this direction. First, recall that y^n changes in every iteration, which can be seen as introducing noise into the gradient. For this reason, we prefer

not to change the iterate at all if the gradient appears to be small; we only update γ^{n-1} when $V^n = \nabla g(\gamma^{n-1}; y^n)^T d^n$ satisfies

$$V^n \geq \max\left\{-\kappa_0, -\left(\frac{\log n}{n}\right)^{1/4}\right\}, \tag{18}$$

$$s^{\max}\left(d^n, \gamma^{n-1}\right) V^n \geq \max\left\{-\kappa_0, -\left(\frac{\log n}{n}\right)^{1/2}\right\}, \tag{19}$$

where the rate $\frac{\log n}{n}$ is chosen to mitigate the effect of estimation error. When both conditions are satisfied, we update $\gamma^n = \gamma^{n-1} + s^n d^n$, where the stepsize s^n is chosen using Algorithm 3. Essentially, this procedure performs a line search to identify s satisfying conditions (20)-(22). Condition (20) ensures that the gradient is well-defined at the new dual solution, while (21)-(22) are based on the well-known Wolfe conditions (Nocedal and Wright 2006). It can be shown that Algorithm 3 will terminate in a finite number of iterations.

Step 4 uses (15) to convert the updated dual solution γ^n into a primal solution x^n, which represents an approximation of the optimal budget allocation. Finally, Step 5 allocates the next experiment to alternative $j^n = \arg\max_j \frac{x_j^n}{N_j^n}$, which again favors under-sampled alternatives.

Although Algorithm 2 appears to involve more technical complications than SIOCBA (Algorithm 1), in practice it is much more efficient, since it does not require the use of any convex programming

Algorithm 3: LineSearch$(d, s^{\max}, \gamma, y)$

Input: Descent direction d, maximum feasible stepsize $s^{\max}$, dual solution γ, estimated values y, parameters $\alpha, \alpha' > 0$ and $\tau \in (0,1)$.
Output: Stepsize s.
Let $s = s^{\max}$.
while *Any of the conditions*

$$\min_j \sum_{i \neq i^*} \gamma_i a_{i,j}(y) > 0, \tag{20}$$

$$g(\gamma + s \cdot d; y) \leq g(\gamma; y) + \alpha s \nabla g(\gamma; y)^T d, \tag{21}$$

$$\nabla g(\gamma + \lambda d; y)^T d \leq \alpha' \left|\nabla g(\gamma; y)^T d\right|, \quad \text{if } \nabla g(\gamma + \lambda d; y)^T d > 0. \tag{22}$$

not satisfied **do**
  $s \leftarrow \tau s$
end

solver. Instead of separating estimation and optimization, so that one has to solve a new problem for each set of sample means, we integrate these two aspects into a single iterative algorithm, which provably converges to the optimal solution (x^*, γ^*) of (13)-(14). The first major result establishes the consistency of the procedure: each alternative will receive a nonzero proportion of the budget asymptotically, ensuring that $y^n \to \mu$.

THEOREM 5. *Under Algorithm 2,* $\liminf_{n\to\infty} x_j^n > 0$ *for all* $j = 1, ..., k$.

Theorem 5 simplifies the subsequent analysis, since we will have $i^{*,n} = i^*$ for all sufficiently large n after some transient period of random length. We can then obtain the main convergence result, which is that Algorithm 2 learns the true optimal allocation of the budget in the limit. Note that Theorem 6 is much stronger than Theorem 5, as it implies through (15) and strong duality that, for any j, $x_j^n \to x_j^*$, where x_j^* is the optimal solution of the true primal problem (13). The highly technical proofs are detailed in the Appendix.

THEOREM 6. *Let* $\{\gamma^n\}$ *be the sequence of iterates generated by Algorithm 2. Every limit point of this sequence is a stationary point of the true dual problem (14).*

These results show that the SIGD algorithm learns the asymptotically optimal budget allocation, without tuning, and without the need for a convex programming solver. Recent work in ranking and selection has considered similar goals (e.g., the top-two method of Qin et al. 2017 or the tuning-free technique of Chen and Ryzhov 2019), but our setting here is much more complicated, because the objective in (14) is not separable across alternatives: each y_i affects multiple terms in the sum. The SIGD algorithm has little in common with the aforementioned methods, instead being based on numerical optimization techniques, and thus constitutes a novel contribution in its own right.

4. Myopic Allocation Under S-index Selection (SIMA)

In this section, we offer a different strategy for allocating the experimental budget, based on Bayesian value-of-information methodology. Since the S-index is itself a new selection criterion,

there are no existing allocation methods that can serve as natural benchmarks for the procedure proposed in Section 3.3. The myopic Bayesian procedure developed here provides such a benchmark, but is itself a very promising method.

Let $\{W_i^n\}$ be a sequence of i.i.d. samples from the distribution $\mathcal{N}(\mu_i, \sigma_i^2)$, representing the output of repeated experiments with alternative i. This time, however, we use the Bayesian model $\mu_i \sim \mathcal{N}\left(y_i^0, (\sigma_i^0)^2\right)$, with μ_i and μ_j independent for all $i \neq j$. Let $\{j^n\}_{n=0}^{\infty}$ be a sequence of alternatives chosen for experimentation, and denote by $\mathcal{F}^n$ the σ-algebra generated by $j^0, W_{j^0}^1, ..., j^{n-1}, W_{j^{n-1}}^n$. We allow $j^n \in \mathcal{F}^n$, meaning that decisions are made adaptively. It is well known (DeGroot 2005) that the conditional distribution of μ_i given $\mathcal{F}^n$ is $\mathcal{N}\left(y_i^n, (\sigma_i^n)^2\right)$, and the parameters can be updated recursively using

$$y_i^{n+1} = \begin{cases} \frac{(\sigma_i^n)^{-2} y^n + \sigma_i^{-2} W_i^{n+1}}{(\sigma_i^n)^{-2} + \sigma_i^{-2}} & i = j^n, \\ y_i^n & i \neq j^n. \end{cases}$$

$$\left(\sigma_i^{n+1}\right)^2 = \begin{cases} \left((\sigma_i^n)^{-2} + \sigma_i^{-2}\right)^{-1} & i = j^n, \\ (\sigma_i^n)^2 & i \neq j^n. \end{cases}$$

It can also be shown that the predictive distribution of y_i^{n+1}, conditional on $\mathcal{F}^n$, is $\mathcal{N}\left(\mu_i^n, (\tilde{\sigma}_i^n)^2\right)$, where $(\tilde{\sigma}_i^n)^2 = (\sigma_i^n)^2 - \left(\sigma_i^{n+1}\right)^2$.

As before, let $z^n = Qy^n$ be the S-indices computed from our point estimates at time n, with $i^{*,n} = \arg\max_i z_i^n$ being the alternative with the largest S-index. Define

$$\nu_j^n = \mathbb{E}\left[\mu_{i^{*,n+1}}^{n+1} - \mu_{i^{*,n+1}}^n \mid \mathcal{F}^n, j^n = j\right], \tag{23}$$

to be the expected improvement in the estimated value of alternative $i^{*,n+1}$ as a result of allocating one additional experiment to alternative j. If $S = 0$ and $Q = I$, the criterion in (23) becomes very similar to the knowledge gradient criterion of Frazier et al. (2008), with the difference that we use $i^{*,n+1}$ instead of $i^{*,n}$ in the second term to simplify the computation; as $i^{*,n}$ converges a.s., this difference does not have much practical significance. Since we are trying to identify $\arg\max_i \mu_i$, we use μ to express the value of the selected alternative, but the index of that alternative is now

determined using S-indices. Under the normality assumption, (23) can be computed in closed form as

$$\nu_j^n = \tilde{\sigma}_j^n \phi\left(\max_{i\neq j} \frac{z_i^n - z_j^n}{\tilde{\sigma}_j^n \left(Q_{j,j} - Q_{j,i}\right)}\right), \tag{24}$$

where ϕ is the standard normal density. This formula is obtained by direct computation combined with the fact, shown in Theorem 1, that $Q_{j,j} > Q_{j,i}$ for $j \neq i$. A myopic allocation policy under S-index selection policy can thus be proposed, which allocates the next experiment to alternative $j^n = \arg\max_j \nu_j^n$. We call this policy S-index Myopic Allocation (SIMA).

Specifically, when an aligned graph is available, it is possible to characterize the performance of SIMA in terms of the asymptotic sampling ratios $\lim_{n\to\infty} \frac{N_i^n}{N_j^n}$, which describe the limiting proportions $\lim_{n\to\infty} \frac{N_j^n}{n}$ of the budget allocated to each j. These proportions do not match the optimal proportions studied in Section 3, in keeping with previous work on ranking and selection (Ryzhov 2016), where methods based purely on expected improvement criteria also generally do not achieve optimal allocations without additional modifications (Chen and Ryzhov 2019, Peng and Fu 2017).

THEOREM 7. *Suppose that $k \geq 3$ with $\mu_1 > ... > \mu_k$ for simplicity. Given an aligned graph, the sampling ratios achieved by SIMA are given by*

$$\lim_{n\to\infty} \frac{N_i^n}{N_j^n} = \frac{\sigma_i\left(\frac{z_1 - z_j}{Q_{j,j} - Q_{j,1}}\right)}{\sigma_j\left(\frac{z_1 - z_i}{Q_{i,i} - Q_{i,1}}\right)}, \quad i,j \neq 1, \tag{25}$$

$$\lim_{n\to\infty} \frac{N_1^n}{N_2^n} = \frac{\sigma_1\left(Q_{1,1} - Q_{1,2}\right)}{\sigma_2\left(Q_{2,2} - Q_{2,1}\right)}, \tag{26}$$

where $z = Q\mu$ are the S-indices computed from the true values.

5. Dynamic Updating of the Similarity Graph

Up to this point, we have assumed the availability of a suitable similarity graph (e.g., an aligned graph). In practice, one might construct a graph S^0 from relevant past data before conducting any new experiments; see Section 6 for examples of how this might be done in specific applications. In any case, however, such a graph is unlikely to be aligned with respect to the true, unknown values μ. To avoid misdirecting the allocation procedure with inaccurate similarity information,

it is desirable to update S^0 over time as new information is acquired. At the same time, the computational cost of the update is a concern, because our algorithms rely heavily on the inverse matrix Q, and repeated computation of the inverse would compromise the computational efficiency of SIGD and SIMA.

We propose a simple updating scheme based on the observation that, given *any* S^0, it is possible to create an aligned graph S by simply permuting the rows and columns of S^0. This directly follows from Definition 1. Moreover, the matrix $Q = Q(S;\lambda)$ can also be computed by applying the same permutation operations on $Q^0 = Q(S^0;\lambda)$ instead of recomputing the inverse. In this way, we can update the similarity graph without high computational cost, while guaranteeing that we will eventually obtain an aligned graph with sufficiently many samples. Although we have mentioned before that an aligned graph does not guarantee $\mathrm{PCS}(z) \geq \mathrm{PCS}(y)$, it does ensure that S-index selection will reliably identify the best alternative asymptotically.

Formally, let r be an ordering of the alternatives, i.e., a permutation of the integers $1, ..., k$ with r_i indicating the position of alternative i in the ordering. Given two such orderings r, r', let $E(r, r')$ be a $k \times k$ matrix whose (j, j')th components are equal to 1 if there exists $i \in \{1, \ldots, k\}$ satisfying $r_j = r'_{j'} = i$, and zero otherwise. Such an E is a permutation matrix. Given the estimated values y^n at time n, let $\hat{r}^n$ be a ranking of the values y^n in descending order, i.e, a larger value of $\hat{r}^n_i$ means alternative i has a smaller estimated mean and is ranked lower. Suppose that we then sample alternative j^n. Then, its estimated value will be updated to $y^{n+1}_{j^n}$, and its position in the ranking will change to $\hat{r}^{n+1}_{j^n}$. If $\hat{r}^{n+1}_{j^n} \geq \hat{r}^n_{j^n}$, the other alternatives $i \neq j^n$ will be reranked according to the rule

$$\hat{r}^{n+1}_i = \begin{cases} \hat{r}^n_i, & \text{if } \hat{r}^n_i > \hat{r}^{n+1}_{j^n} \text{or } \hat{r}^n_i < \hat{r}^n_{j^n} \\ \hat{r}^n_i - 1, & \text{otherwise} \end{cases}. \tag{27}$$

In other words, the ranking can be updated without a sorting algorithm. A similar updating rule can be derived when $\hat{r}^{n+1}_{j^n} \leq \hat{r}^n_{j^n}$. Once $\hat{r}^{n+1}$ is computed, a permutation matrix $\hat{E}^{n+1} = E(\hat{r}^0, \hat{r}^{n+1})$ can be created. It is straightforward to show that the similarity matrix $\hat{S}^{n+1} = \left(\hat{E}^{n+1}\right)^\top S^0 \hat{E}^{n+1}$

is aligned with respect to y^{n+1}. Therefore, as n becomes large, $\hat{S}^n$ will converge to a graph that is aligned with respect to the true means μ, provided that each alternative is sampled infinitely often.

In practice, we may not wish to update the similarity matrix at every iteration, especially in the early stages where our estimated values are subject to high uncertainty. Furthermore, even if S^0 is not aligned, it may still contain some valuable information about the similarities between different values. Thus, we may wish to hold off on permuting S^0 until there is strong evidence in favor of doing so. To that end, we define another ranking $\tilde{r}^n$ which is only updated when there is a "significant" change in the ordering of the values y^n. Once $\tilde{r}^n$ is created, the similarity matrix is determined. We initialize $\tilde{r}^n$ as $\tilde{r}^0 = \hat{r}^0$. The change in $\hat{r}^n$ gives guidance on how we should change $\tilde{r}^n$, but as we discussed earlier, we only follow this guidance when we have high confidence. Therefore, we propose the following updating scheme of $\tilde{r}^n$. Again, suppose that we sample j^n at time n and find that $\hat{r}^{n+1}_{j^n} > \hat{r}^n_{j^n}$, suggesting that j^n should possibly be ranked lower. At this point we need to decide whether we would like to trust this information and increase $\tilde{r}^n_{j^n}$ accordingly. Define two sets

$$\begin{aligned}\tilde{\Omega}^{-,n} &= \{i = 1, \ldots, k : \tilde{r}^n_i > \tilde{r}^n_{j^n}\},\\ \hat{\Omega}^{+,n+1} &= \{i = 1, \ldots, k : y^{n+1}_i > y^{n+1}_{j^n}\} = \{i = 1, \ldots, k : \hat{r}^{n+1}_i < \hat{r}^{n+1}_{j^n}\}.\end{aligned} \tag{28}$$

The set $\tilde{\Omega}^{-,n}$ contains alternatives that were ranked lower than j^n according to $\tilde{r}^n$, and $\hat{\Omega}^{+,n+1}$ contains alternatives whose sample means are above j^n after the update. For $i \in \tilde{\Omega}^{-,n} \cap \hat{\Omega}^{+,n+1}$, the probability that $\mu_i < \mu_{j^n}$ can be approximated using a normal distribution, and if this probability is smaller than some small tolerance parameter $\kappa^{tol} < 0.5$, we have sufficient confidence to rank i more highly than j^n. We can reduce κ^{tol} if we wish to update the ranking more conservatively. Formally, we define

$$\tilde{\Omega}^{n+1} = \left\{ i \in \tilde{\Omega}^{-,n} \cap \hat{\Omega}^{+,n+1} : \Phi\left(\frac{y^{n+1}_{j^n} - y^{n+1}_i}{\sqrt{\frac{\sigma^2_i}{N^{n+1}_i} + \frac{\sigma^2_{j^n}}{N^{n+1}_{j^n}}}} \right) < \kappa^{tol} \right\}, \tag{29}$$

where $\Phi(\cdot)$ is the standard normal cdf. Then, the new rank of alternative j^n is set according to

$$\tilde{r}^{n+1}_{j^n} = \max_{i \in \hat{\Omega}^{n+1}} \hat{r}^{n+1}_i, \tag{30}$$

so alternative j^n is moved to the end of those alternatives that are sufficiently likely to be better based on both the updated sample means and the previous ranking. For alternatives $i \neq j^n$, their belief ranks can be updated following (27) by replacing $\hat{r}$ with $\tilde{r}$.

For the case where $\hat{r}^{n+1}_{j^n} < \hat{r}^n_{j^n}$, we can use symmetric computations based on the definitions

$$\begin{aligned}
\tilde{\Omega}^{+,n} &= \{i = 1, \ldots, k : \tilde{r}^n_i < \tilde{r}^n_{j^n}\}, \\
\hat{\Omega}^{-,n+1} &= \{i = 1, \ldots, k : y^{n+1}_i < y^{n+1}_{j^n}\} = \{i = 1, \ldots, k : \hat{r}^{n+1}_i > \hat{r}^{n+1}_{j^n}\}, \\
\tilde{\Omega}^{n+1} &= \left\{ i \in \tilde{\Omega}^{+,n} \cap \hat{\Omega}^{-,n+1} : \Phi\left(\frac{y^{n+1}_i - y^{n+1}_{j^n}}{\sqrt{\frac{\sigma^2_i}{N^{n+1}_i} + \frac{\sigma^2_{j^n}}{N^{n+1}_{j^n}}}} \right) < \kappa^{tol} \right\}.
\end{aligned} \tag{31}$$

For complete details, refer to Algorithm 4, which can be called after the allocation and sampling step in either of our proposed algorithms.

Our theoretical guarantees for both SIGD and SIMA are preserved when either procedure is combined with Algorithm 4. This is a consequence of the following result.

PROPOSITION 3. *Under any allocation method, Algorithm 4 will update the similarity graph finitely many times; in other words, there is a random but a.s. finite n_0 such that $\tilde{Q}^n = \tilde{Q}^{n_0}$ for $n \geq n_0$. Furthermore, if every alternative is measured infinitely often, then $\tilde{Q}^{n_0}$ must represent an aligned graph on that same sample path.*

By Proposition 3, there is always a finite time n_0 after which the allocation method runs under the same matrix $\tilde{Q}^{n_0}$. The behavior of SIGD and SIMA after time n_0 is identical to the situation where these algorithms are initialized with y^{n_0} at time 0 and run under the fixed matrix $\tilde{Q}^{n_0}$. Both methods are guaranteed to sample each alternative infinitely often under any fixed Q, so they will be consistent. Therefore, $\tilde{Q}^{n_0}$ must be aligned, whence the results of Theorems 6 and 7 are easily recovered. As a result, SIGD and SIMA can perform well even in settings where the initial similarity structure S^0 is uninformative or misleading, as will be demonstrated in our numerical experiments.

Algorithm 4: Sequential update of Q

Input: Q^0 associated with prior similarity matrix S^0, estimated means y^n, alternative j^n sampled at step n and its new estimated mean $y^{n+1}_{j^n}$, previous orders $\hat{r}^n$ and $\tilde{r}^n$, tolerance parameter κ^{tol}.
Output: updated orders $\hat{r}^{n+1}$ and $\tilde{r}^{n+1}$, updated matrix $\tilde{Q}^{n+1}$ at step $n+1$.
Find $\hat{r}^{n+1}_{j^n}$ by a binary search of y^{n+1};
Update $\hat{r}^{n+1}$ by (27);
$\tilde{r}^{n+1}_{j^n} \leftarrow \tilde{r}^n_{j^n}$;
if $\hat{r}^{n+1}_{j^n} > \hat{r}^n_{j^n}$ *or* $(\hat{r}^{n+1}_{j^n} = \hat{r}^n_{j^n}$ *and* $\hat{r}^{n+1}_{j^n} > \tilde{r}^n_{j^n})$ **then**
  Find $\tilde{\Omega}^{n+1}$ by (28) and (29);
  $\tilde{r}^{n+1}_{j^n} \leftarrow \max_{i\in\tilde{\Omega}^{n+1}} \hat{r}^{n+1}_i$ if $\tilde{\Omega}^{n+1}$ is not empty;
else
  if $\hat{r}^{n+1}_{j^n} < \hat{r}^n_{j^n}$ *or* $(\hat{r}^{n+1}_{j^n} = \hat{r}^n_{j^n}$ *and* $\hat{r}^{n+1}_{j^n} < \tilde{r}^n_{j^n})$ **then**
    Find $\tilde{\Omega}^{n+1}$ by (31);
    $\tilde{r}^{n+1}_{j^n} \leftarrow \min_{i\in\tilde{\Omega}^{n+1}} \hat{r}^{n+1}_i$ if $\tilde{\Omega}^{n+1}$ is not empty;
  end
end
Update $\tilde{r}^{n+1}$ similar to (27);
$\tilde{E}^{n+1} \leftarrow E\,(\tilde{r}^0, \tilde{r}^{n+1})$;
$\tilde{Q}^{n+1} \leftarrow \left(\tilde{E}^{n+1}\right)^\top Q^0 \tilde{E}^{n+1}$.

6. Numerical Experiments

In this section, we present numerical comparisons on four test problems. The following allocation methods were implemented:

- *SIGD.* We use Algorithm 2 to learn the optimal allocation. At the beginning of the sequential allocation process, we run additional iterations of the coordinate descent algorithm to obtain a good starting point.

- *SIMA.* The nth experiment is allocated to alternative $\arg\max_j \nu^n_j$, where ν^n_j is given by (24).

- *Correlated knowledge gradient (CKG)* of Frazier et al. (2009). This method, a standard benchmark in the literature, models similarity between alternatives using correlated Bayesian beliefs. For a fair comparison, we set the prior covariance matrix to be $a(S+bI)$, where b is the smallest positive number such that $S+bI$ is symmetric positive definite and a is a prior variance parameter.

- *Optimal computing budget allocation (OCBA).* The fully sequential OCBA method of Chen et al. (2000) is another standard benchmark from the ranking and selection literature.

Each of these methods is tested using two selection criteria, namely sample (or posterior) means and S-indices. This allows us to assess the value of using the S-index for selection separately from any one particular choice of allocation method. Performance is measured in two ways, namely PCS and the expected opportunity cost (EOC), defined as $\text{EOC} = \max_j \mu_j - \mathbb{E}\mu_{i^M}$, where i^M is the selected alternative when the total budget is M. In general, allocation procedures use one of the two metrics in their derivation: in the above list, SIGD and OCBA are derived to optimize PCS, while SIMA and CKG are specialized for EOC. Empirically, however, we find that each policy performs similarly (relative to the others) under either criterion. Thus, though there is a version of OCBA specialized for EOC (He et al. 2007), we did not implement it in these experiments.

In all the experiments, we use $\lambda = 2$ as the regularization parameter; in the Appendix, we present additional numerical results showing that performance is insensitive to this value.

6.1. Experiment 1: A Servo System Selection Problem

A servo system is an automatic control system where the controller provides commands to activate motions (Zhang 2008). A critical characteristic of a servo system is the rise time, which is the time required for the signal to change to a higher position. This time depends on the mechanical configuration of the system, and may be estimated from simulation. Thus, we may consider a ranking and selection problem where different configurations are alternatives, and performance is represented by rise time.

Our test problem is adapted from a real-world dataset originating from the UCI Machine Learning Repository (Dheeru and Casey 2017). The dataset contains 167 instances, each of which has a distinct configuration of 4 settings (two gain settings and two choices of mechanical linkages). We randomly choose 35% of these instances to serve as the alternatives in our test problem; their recorded rise times are treated as their true μ_i values, and normally distributed noise with $\sigma_i = 0.1\mu_i + 0.2$ is added during experiments.

We construct an initial similarity graph S^0 by using the remaining 65% of instances as past data. We run a random forest algorithm on these data, with the mechanical settings as features,

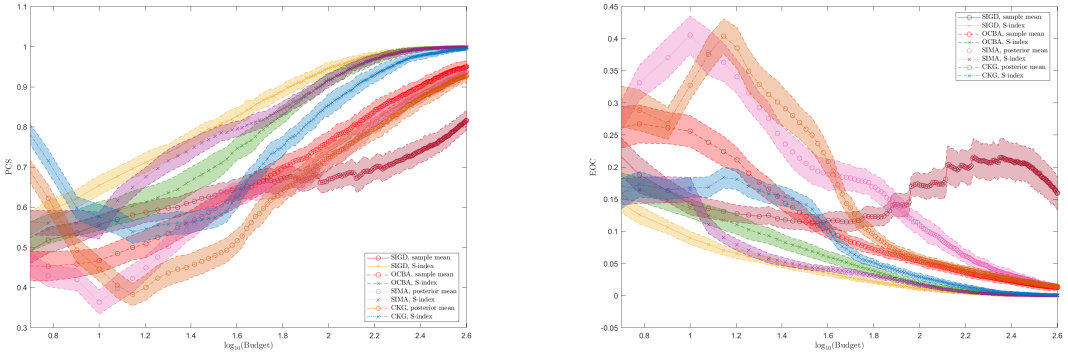

Figure 2 Case 1: PCS and EOC results with standard error bounds obtained by 10000 independent runs (40 macro-replications, 250 micro-replications) with a **fixed** *good* prior similarity matrix.

and use the predictions made by this model as initial estimated values y^0 of the alternatives. Prior similarities are computed as $S^0_{i,j} = e^{-|y^0_i - y^0_j|}$.

Comparisons are conducted on two versions of the problem, which use different randomly chosen sets of alternatives. Both versions construct S^0 in the same way as described above, but the quality of this similarity structure varies. In Case 1, S^0 is "good," meaning that it is generally accurate in describing the relationships between the values μ_i, although the similarity graph is not aligned with respect to μ. In Case 2, S^0 is "bad," in that it provides little useful guidance about μ. Figures 2 and 3 show empirical results for both cases (averaged over 40 macro-replications) when S^0 is *fixed* (no dynamic updating) for the duration of the experiment. In Case 1, when S^0 is good (though not aligned), we find that S-index selection always yields better performance than selection based on

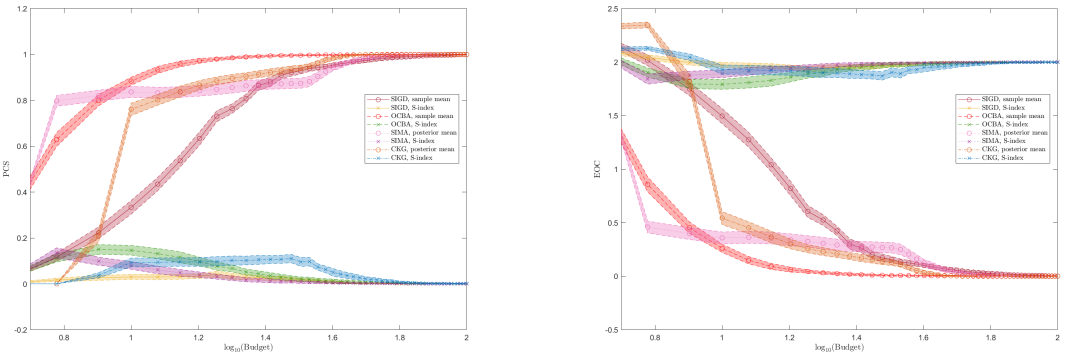

Figure 3 Case 2: PCS and EOC results with standard error bounds obtained by 10000 independent runs (40 macro-replications, 250 micro-replications) with a **fixed** *bad* prior similarity matrix.

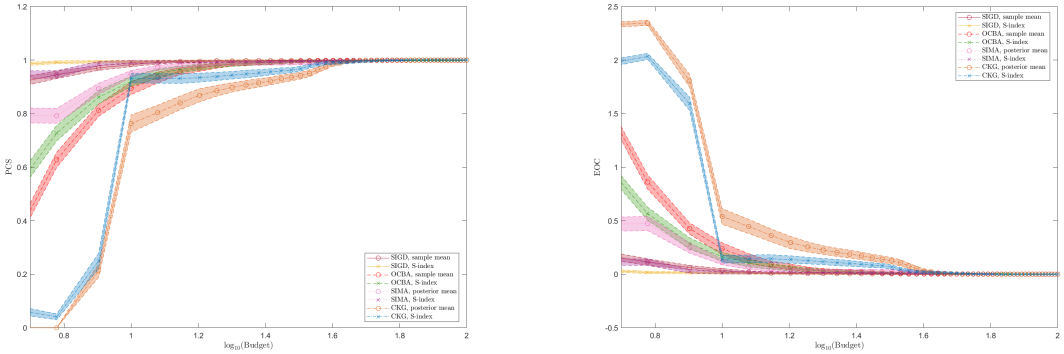

Figure 4 Case 2: PCS and EOC results with standard error bounds obtained by 10000 independent runs (40 macro-replications, 250 micro-replications) with **dynamic updating** of a *bad* prior similarity matrix.

sample means, regardless of which method is used to allocate the simulation budget; furthermore, SIGD achieves the best performance overall, with SIMA coming in second.

However, in Case 2, S-index selection performs poorly for all of the methods, because S^0 is providing little useful information and may even be misleading. To solve this problem, we apply dynamic updating using Algorithm 4 with $\kappa^{tol} = 0.1$. Figure 4 shows that the updating strategy greatly improves the performance of S-index selection in Case 2, so that SIGD and SIMA are once more outperforming the other benchmarks. In this way, we see that S-index selection can offer significant advantages over sample mean-based selection, even when a good similarity structure is not available ahead of time.

6.2. Experiment 2: A Dose Selection Problem

The objective of a dose-finding clinical trial is to identify the maximal response to a drug (Nasrollahzadeh and Khademi 2020). The literature has shown that different factors, such as gender and age, often exert great impact on the dose-response curve. Often, two such curves for different patient categories may be accurately represented using different parameter values within the same model (Xue et al. 1999).

For the purposes of this example, suppose that there are two age groups. For Age Group 1, clinical trials have been conducted and the dose-response curve $\chi_1(v)$, for doses $v \in \{v_1, \ldots, v_k\}$, is known. We now wish to find the dose v_{i^*} that achieves the maximal response for Age Group 2.

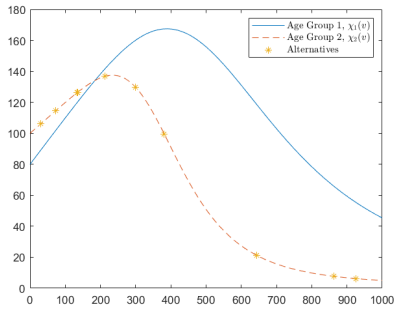

Figure 5 Settings for the dose selection problem. $\chi_1(v) = \chi(v;[2,80,0.3,600,4])$, $\chi_2(v) = \chi(v;[2,100,0.2,400,5])$.

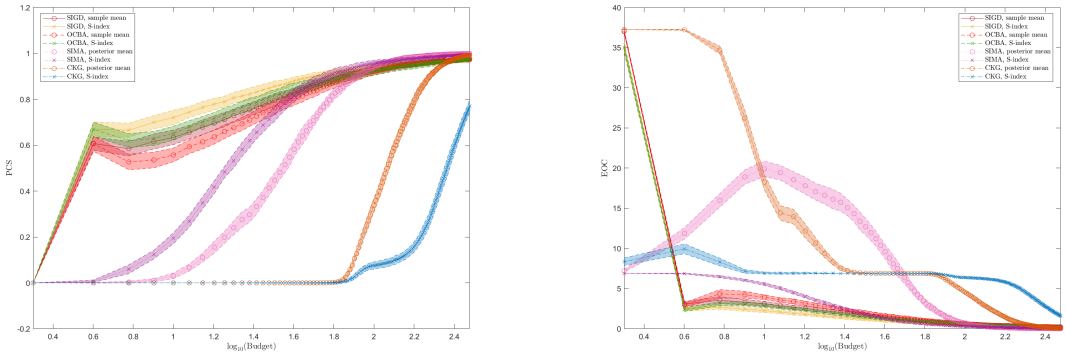

Figure 6 PCS and EOC results with standard error bounds for Experiment 2 obtained by 10000 independent runs (40 macro-replications, 250 micro-replications)

Then, we can use $y_i^0 = \chi_1(v_i)$ as the prior estimates of the dose response values $\mu_i = \chi_2(v_i)$ for Age Group 2, and construct S^0 from these values in the same way as in Experiment 1. This similarity graph is unlikely to be aligned with respect to the true dose-response values for Age Group 2, so we will use dynamic updating in this experiment.

We assume that the underlying dose-response curves for both groups follow the Brain-Cousens model $\chi(v;c) = c_1 + \frac{c_2 - c_1 + c_3 v}{\exp(1+c_5(\log(v)-\log(c_4)))}$ with different parameters (Ritz et al. 2015), as shown in Figure 5. We randomly choose $k = 10$ doses from $\mathcal{U}[0,1000]$ as alternatives. Each of them has mean $\mu_i = \chi_2(v_i)$ and standard deviation $\sigma_i = 0.1\mu_i + 2$, $i = 1,\ldots,k$. Figure 6 shows that SIGD with S-index selection achieves the best performance.

6.3. Experiment 3: M/G/1 Queues

This experiment is taken from Qu et al. (2015) with a few modifications. There are 20 alternatives, each of which corresponds to a first-come-first-serve (FCFS) M/G/1 queue. Arrivals to each queue

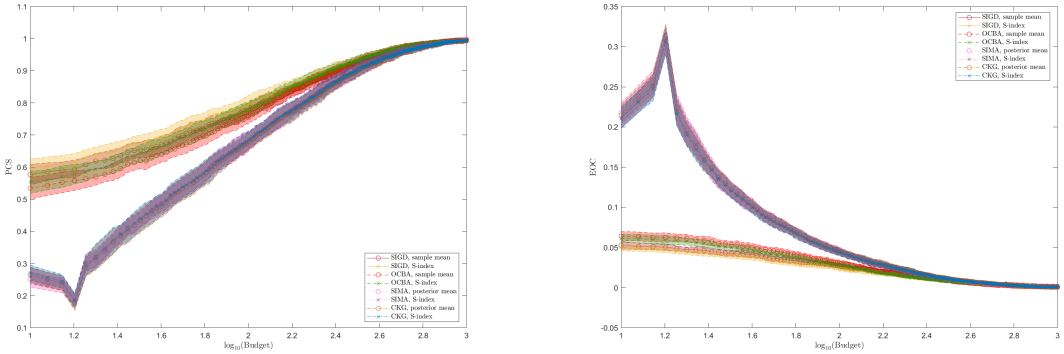

Figure 7 PCS and EOC results with standard error bounds for Experiment 3 based on 10000 independent replications (20 macro-replications, 250 micro-replications.

follow a Poisson process with rate 0.05, and the service times of the ith queue are i.i.d. according to a generalized Pareto distribution with scale parameter ς_i, $i = 1, ..., 20$, set as $\frac{1}{0.15+0.005i}$, location parameter $l_i = \frac{\varsigma_i(0.1i+1)}{10}$ and shape parameter $\frac{1}{4}$. We assume that we know the true steady-state waiting time of the last 5 queues, and our objective is to find the slowest queue among the first 15. The true mean steady-state waiting time (used as the value μ_i) can be calculated using the Pollaczek-Khinchin formula.

For each queue, a single experiment consists of simulating the queue until 500 customers have been served and averaging the waiting time of the last 200 customers. Note that this value may not be exactly normally distributed. We run 200000 macro-replications in advance to estimate the true variance. Prior values y_i^0 for the 15 queues with unknown waiting times are obtained by applying simple linear regression to the 5 queues with known waiting times, and S^0 is constructed from these values in the same way as in Experiments 1-2. We apply dynamic updating of the similarity structure with $\kappa^{tol} = 0.05$. The results are shown in Figure 7. The best performance is achieved by SIGD with either selection criterion.

6.4. Experiment 4: Rosenbrock Test Function

The Rosenbrock function is a well-known benchmark in the continuous optimization literature. The function is defined on $\mathbb{R}^2$ as $f(u) = (1-u_1)^2 + 100(u_2 - u_1^2)^2$. For the purposes of this experiment, suppose that our alternatives correspond to pairs of values taken from the set $\{0, 0.25, 0.5, 0.75, 1\}$.

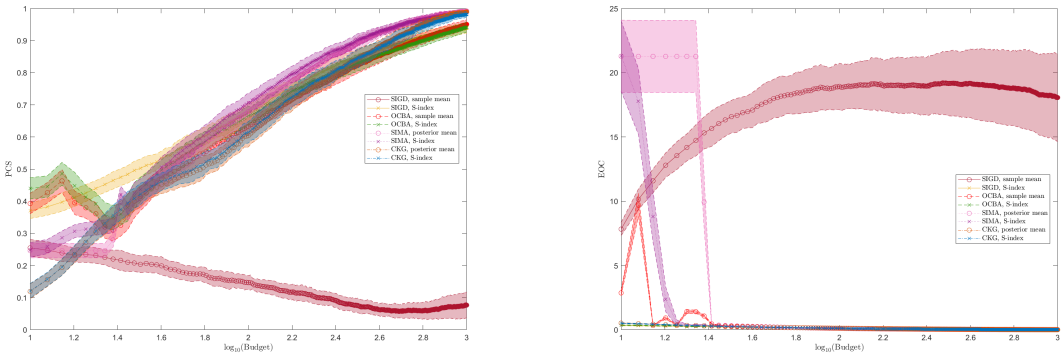

Figure 8 PCS and EOC results with standard error bounds for Experiment 4 based on 10000 independent replications (20 macro-replications, 250 micro-replications.

Thus, there are 25 alternatives; the mean of alternative i is the negative of the corresponding Rosenbrock function value, and the standard deviation of the simulation noise is given by $\sigma_i = 0.2\mu_i + 2$.

We further suppose that we already know the true function values for 9 other points whose coordinates are chosen randomly from $\mathcal{U}[0,1]$. We perform 2-degree polynomial regression on these values to obtain predictions y_i^0 for the 25 alternatives of interest. The initial similarity structure S^0 is constructed from these values in the same way as in the other experiments. The results, shown in Figure 8, show that SIGD and SIMA with S-index selection achieve the best performance, with SIGD performing better in the early stages of sampling, and SIMA pulling ahead later.

6.5. Summary of Results and Computation Time

In Experiment 1, we saw that our proposed algorithms can outperform the benchmark methods when the prior similarity matrix S^0 is sufficiently informative. In practice, it is unlikely that prior information is sufficient to provide accurate information about the relationships between the unknown performance values. We have seen, however, that the advantages of SIGD and SIMA can

Table 1 Average computation time (ms) (including simulation) over 100 independent macro-replications.

	k	SIGD	SIMA	OCBA	CKG
Experiment 1	31	0.22	0.13	0.09	2.09
Experiment 2	10	0.25	0.09	0.10	0.64
Experiment 3	15	0.32	0.18	0.21	1.18
Experiment 4	25	0.21	0.08	0.12	1.24

be recovered when the similarity matrix is updated using the permutation strategy of Section 5. Furthermore, both methods are more computationally efficient than CKG, the only benchmark method with the ability to consider similarity information, because CKG uses a time-consuming sorting step when calculating the expected value of information. The average computation time for each experiment, conducted on a 3.6 GHz Intel Core i7 PC with 16 GB of RAM, is reported in Table 1.

7. Conclusion

We have presented a new framework for using similarity between alternatives in ranking and selection to enable more efficient identification of the best alternative. Because the S-index similarity measure is built into the selection step, we retain the computational efficiency of traditional methods that model each alternative independently of the others.

In addition, we have developed two allocation methods adapted to the new selection criterion. The first is based on a mathematical program that optimizes the probability of correct selection. To avoid having to repeatedly solve the mathematical program to optimality, we have developed a novel sequential algorithm based on reduced gradient methods from numerical optimization. The second approach uses a simple myopic calculation based on the Bayesian value of information. Both approaches can be integrated with a scheme for dynamically updating the similarity structure based on new information, thus avoiding the need to rely on inaccurate or misleading prior information. Numerical experiments show that both procedures perform well with this additional subroutine.

In practice, the quality of the similarity information used to calculate S-indices is critical. Our dynamic updating scheme was effective in utilizing this information to learn the similarity (or Q) matrix, but developing ways to tailor the similarity structure to specific problem settings is key to practical implementation of the algorithms developed here and thus an important avenue for future research.

Acknowledgments

This work was supported in part by AFOSR under Grant FA95502010211.

References

Bazaraa MS, Sherali HD, Shetty CM (2013) *Nonlinear Programming: Theory and Algorithms* (John Wiley & Sons).

Boyd SP, Vandenberghe L (2004) *Convex Optimization* (Cambridge University Press).

Chen CH, Chick SE, Lee LH, Pujowidianto NA (2015) Ranking and selection: efficient simulation budget allocation. Fu MC, ed., *Handbook of Simulation Optimization*, 45–80 (Springer).

Chen CH, Lee LH (2010) *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation* (World Scientific).

Chen CH, Lin J, Yücesan E, Chick SE (2000) Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems* 10(3):251–270.

Chen Y, Ryzhov IO (2019) Complete expected improvement converges to an optimal budget allocation. *Advances in Applied Probability* 51(1):209–235.

Chick SE, Branke J, Schmidt C (2010) Sequential sampling to myopically maximize the expected value of information. *INFORMS Journal on Computing* 22(1):71–80.

DeGroot MH (2005) *Optimal Statistical Decisions* (John Wiley & Sons).

Dembo A, Zeitouni O (2009) *Large Deviations Techniques and Applications (2nd ed.)* (Springer Berlin Heidelberg).

Dheeru D, Casey G (2017) UCI machine learning repository. `http://archive.ics.uci.edu/ml`.

Ding J, Zhou A (2007) Eigenvalues of rank-one updated matrices with some applications. *Applied Mathematics Letters* 20(12):1223–1226.

Fan W, Hong LJ, Nelson BL (2016) Indifference-zone-free selection of the best. *Operations Research* 64(6):1499–1514.

Frazier PI, Powell WB, Dayanik S (2008) A knowledge-gradient policy for sequential information collection. *SIAM Journal on Control and Optimization* 47(5):2410–2439.

Frazier PI, Powell WB, Dayanik S (2009) The knowledge-gradient policy for correlated normal beliefs. *INFORMS Journal on Computing* 21(4):599–613.

Fu MC, Hu JQ, Chen CH, Xiong X (2007) Simulation allocation for determining the best design in the presence of correlated sampling. *INFORMS Journal on Computing* 19(1):101–111.

Gao S, Chen W, Shi L (2017) A new budget allocation framework for the expected opportunity cost. *Operations Research* 65(3):787–803.

Gao T, Lu W (2021) Machine learning toward advanced energy storage devices and systems. *iScience* 24(1):101936:1–101936:33.

Glynn P, Juneja S (2004) A large deviations perspective on ordinal optimization. *Proceedings of the 2004 Winter Simulation Conference*, 577–586.

Grant M, Boyd S (2014) CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`.

Han B, Ryzhov IO, Defourny B (2016) Optimal learning in linear regression with combinatorial feature selection. *INFORMS Journal on Computing* 28(4):721–735.

He D, Chick SE, Chen CH (2007) Opportunity cost and OCBA selection procedures in ordinal optimization for a fixed number of alternative systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37(5):951–961.

Hunter SR, Pasupathy R (2013) Optimal sampling laws for stochastically constrained simulation optimization on finite sets. *INFORMS Journal on Computing* 25(3):527–542.

Kim SH, Nelson BL (2001) A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation* 11(3):251–273.

Lin CJ, Lucidi S, Palagi L, Risi A, Sciandrone M (2009) Decomposition algorithm model for singly linearly-constrained problems subject to lower and upper bounds. *Journal of Optimization Theory and Applications* 141(1):107–126.

Ma S, Henderson SG (2017) An efficient fully sequential selection procedure guaranteeing probably approximately correct selection. Chan WKV, D'Ambrogio A, Zacharewicz G, Mustafee N, Wainer G, Page E, eds., *Proceedings of the 2017 Winter Simulation Conference*, 2225–2236.

Malkomes G, Garnett R (2018) Automating Bayesian optimization with Bayesian optimization. *Advances in Neural Information Processing Systems* 31:5984–5994.

Nasrollahzadeh AA, Khademi A (2020) Optimal stopping of adaptive dose-finding trials. *Service Science* 12(2-3):80–99.

Nocedal J, Wright S (2006) *Numerical Optimization* (Springer Science & Business Media).

Pasupathy R, Hunter SR, Pujowidianto NA, Lee LH, Chen CH (2014) Stochastically constrained ranking and selection via score. *ACM Transactions on Modeling and Computer Simulation* 25(1):1–26.

Peng Y, Chong EK, Chen CH, Fu MC (2018) Ranking and selection as stochastic control. *IEEE Transactions on Automatic Control* 63(8):2359–2373.

Peng Y, Fu MC (2017) Myopic allocation policy with asymptotically optimal sampling rate. *IEEE Transactions on Automatic Control* 62(4):2041–2047.

Qin C, Klabjan D, Russo D (2017) Improving the expected improvement algorithm. Guyon I, Von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, eds., *Advances in Neural Information Processing Systems*, volume 30, 5381–5391.

Qu H, Ryzhov IO, Fu MC, Ding Z (2015) Sequential selection with unknown correlation structures. *Operations Research* 63(4):931–948.

Ritz C, Baty F, Streibig C, Gerhard D (2015) Dose-response analysis using R. *PloS ONE* 10(12):e0146021.

Ryzhov IO (2016) On the convergence rates of expected improvement methods. *Operations Research* 64(6):1515–1528.

Scott WR, Powell WB, Simão HP (2010) Calibrating simulation models using the knowledge gradient with continuous parameters. Johansson B, Jain S, Montoya-Torres J, Hugan J, Yücesan E, eds., *Proceedings of the 2010 Winter Simulation Conference*, 1099–1109.

Shen H, Hong LJ, Zhang X (2017) Ranking and selection with covariates. Chan WKV, D'Ambrogio A, Zacharewicz G, Mustafee N, Wainer G, Page E, eds., *Proceedings of the 2017 Winter Simulation Conference*, 2137–2148.

Sun G, Li Y, Fu MC (2019) A spectral index for selecting the best alternative. Mustafee N, Bae KHG, Lazarova-Molnar S, Rabe M, Szabo C, Haas P, Son YJ, eds., *Proceedings of the 2019 Winter Simulation Conference*, 3404–3415.

Von Luxburg U (2007) A tutorial on spectral clustering. *Statistics and Computing* 17(4):395–416.

Xue F, Zhang YM, Liao X, Liu JH, An G (1999) Influences of age and gender on dose response and time course of effect of atracurium in anesthetized adult patients. *Journal of Clinical Anesthesia* 11(5):397–405.

Zhang P (2008) *Industrial Control Technology: A Handbook for Engineers And Researchers* (William Andrew).

Zhang Q, Song Y (2017) Moment-matching-based conjugacy approximation for Bayesian ranking and selection. *ACM Transactions on Modeling and Computer Simulation* 27(4):26:1–26:23.

Zhou J, Ryzhov IO (2022) A new rate-optimal design for linear regression. *Submitted for publication* .

Appendix A: Convergence Analysis of Algorithm 2

Theorems 5 and 6 are proved below, in Sections A.3 and A.4. Before proceeding to these proofs, we first give some necessary background and preliminary technical results in Sections A.1 and A.2. The proofs of other results stated in the paper are deferred to Section B.

A.1. Properties of the Dual Objective Function

We now show some useful properties of the dual objective $g(\cdot; y^n)$. First, note that $\bar{y} = \lim_{n\to\infty} y^n$ is always well-defined: if $N_j^n \to \infty$, then $y_j^n \to \mu_j$, whereas if $\sup_n N_j^n < \infty$, the sample mean is only updated finitely many times. Theorem 5 will show that $\bar{y} = \mu$, but at the moment this has not yet been established.

Let $i' = \arg\max_j [Q\bar{y}]_j$ be the alternative selected under the limiting estimates $\bar{y}$. Similarly, let $\bar{a}_{i,j} = a_{i,j}(\bar{y})$. From the definition of $a_{i,j}$, it is easy to see that

$$|a_{i,j}(y^n) - \bar{a}_{i,j}| \le K_a \|y^n - \bar{y}\|, \qquad j = 1, ..., k,\ i \ne i'. \tag{32}$$

Consequently, we have $a_{i,j}(y^n) \to \bar{a}_{i,j}$, so the sequence $\{a_{i,j}(y^n)\}$ is bounded. We can therefore let

$$a_{i,j}^{\max} = \sup_n a_{i,j}(y^n), \qquad a_{i,j}^{\min} = \inf_n a_{i,j}(y^n),$$

and we will have $a_{i,j}^{\min} > 0$ for any $i \ne i'$ satisfying $Q_{i',j} \ne Q_{i,j}$. This fact leads to the following technical lemma establishing a kind of uniform convergence of g.

LEMMA 1. *There exists a constant $K_g > 0$ such that, for all n,*

$$|g(\gamma; y^n) - g(\gamma; \bar{y})| \le K_g \|y^n - \bar{y}\|, \qquad \forall \gamma \in \mathcal{G}.$$

Consequently, $g(\gamma; y^n)$ converges uniformly to $g(\gamma; \bar{y})$ on $\mathcal{G}$.

Proof. For any $\gamma \in \mathcal{G}$,

$$\begin{aligned}
|g(\gamma; y^n) - g(\gamma; \bar{y})| &= \left| \sum_{j=1}^k \sqrt{\sum_{i\ne i'} \gamma_i a_{i,j}(y^n)} - \sum_{j=1}^k \sqrt{\sum_{i\ne i'} \gamma_i \bar{a}_{i,j}} \right| \\
&\le \sum_{j=1}^k \frac{\sum_{i\ne i'} \gamma_i |a_{i,j}(y^n) - \bar{a}_{i,j}|}{\sqrt{\sum_{i\ne i'} \gamma_i a_{i,j}(y^n)} + \sqrt{\sum_{i\ne i'} \gamma_i \bar{a}_{i,j}}} \\
&\le \sum_{j=1}^k \sum_{i: Q_{i',j} \ne Q_{i,j}} \frac{\gamma_i |a_{i,j}(y^n) - \bar{a}_{i,j}|}{\sqrt{\gamma_i a_{i,j}(y^n)} + \sqrt{\gamma_i \bar{a}_{i,j}}} \\
&= \sum_{j=1}^k \sum_{i: Q_{i',j} \ne Q_{i,j}} \frac{\sqrt{\gamma_i} |a_{i,j}(y^n) - \bar{a}_{i,j}|}{\sqrt{a_{i,j}(y^n)} + \sqrt{\bar{a}_{i,j}}} \\
&\le \sum_{j=1}^k \sum_{i: Q_{i',j} \ne Q_{i,j}} \frac{|a_{i,j}(y^n) - \bar{a}_{i,j}|}{\sqrt{\bar{a}_{i,j}}} \qquad (33) \\
&\le \frac{k^2 K_a}{\sqrt{\min_{i: Q_{i',j} \ne Q_{i,j}} a_{i,j}^{\min}}} \|y - \bar{y}\|, \qquad (34)
\end{aligned}$$

where (33) is obtained by $\gamma_i \in [0,1]$ and $a^n_{i,j} \geq 0$, and (34) holds from (32) and the boundedness of $\{a_{i,j}(y^n)\}$. $\square$

LEMMA 2. *There exist constants $0 < b_1 < b_2$ such that $0 < b_1 \leq g(\gamma; y^n) \leq b_2$ for all n and all $\gamma \in \mathcal{G}$.*

Proof: By Lemma 1, it suffices to show that $g(\gamma; \bar{y})$ is bounded between two negative numbers. It is easy to see that for any $\gamma \in \mathcal{G}$,

$$g(\gamma; \bar{y}) = -\sum_{j=1}^{k} \sqrt{\sum_{i \neq i'} \gamma_i \bar{a}_{i,j}} \geq -\sum_{j=1}^{k} \sqrt{\sum_{i \neq i'} \bar{a}_{i,j}},$$

meaning that $g(\gamma; \bar{y})$ is bounded below by a strictly negative number. For the upper bound, we have

$$g(\gamma; \bar{y}) = -\sum_{j=1}^{k} \sqrt{\sum_{i \neq i'} \gamma_i \bar{a}_{i,j}} \leq -\sqrt{\sum_{i \neq i'} \gamma_i \bar{a}_{i,i'}} - \sum_{j \neq i'} \sqrt{\gamma_j \bar{a}_{j,j}} < 0.$$

The proof is thus completed. $\square$

LEMMA 3. *Let $\{\gamma^n\}$ be the sequence generated by Algorithm 2, and suppose that*

$$\liminf_{n \to \infty} \sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(y^n) > 0 \tag{35}$$

for all j. Then, there exists a constant $K > 0$ such that, for all n,

$$\|\nabla g(\gamma^{n-1}; y^n) - \nabla g(\gamma^{n-1}; \mu)\| \leq K \|y^n - \mu\|.$$

Proof. From (32), for any j and $\gamma^{n-1} \in \mathcal{G}$, we have

$$\left| \sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(y^n) - \sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(\mu) \right| = \left| \sum_{i \neq i'} \gamma_i^{n-1} \left(a_{i,j}(y^n) - a_{i,j}(\mu) \right) \right| \leq K_a \|y^n - \mu\|. \tag{36}$$

From (35), it also follows that

$$\liminf_{n \to \infty} \sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(\mu) > 0.$$

Therefore, there exists a constant $b > 0$ such that

$$\sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(y^n) > b, \qquad \sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(\mu) > b$$

for all n and for all j. Then, for $i \neq i'$,

$$\begin{aligned} & \left| \left[\nabla g(\gamma^{n-1}; y^n) \right]_i - \left[\nabla g(\gamma^{n-1}; \mu) \right]_i \right| \\ = & \left| \sum_{j=1}^{k} \frac{a_{i,j}(y^n)}{\sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(y^n)} - \sum_{j=1}^{k} \frac{a_{i,j}(\mu)}{\sum_{i \neq i'} \gamma_i^{n-1} a_{i,j}(\mu)} \right| \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{j=1}^{k} \frac{\left|a_{i,j}\left(y^n\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(y^n\right)\right)-a_{i,j}\left(\mu\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)\right|}{\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(y^n\right)\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)} \\
&\leq \frac{1}{b^2}\sum_{j=1}^{k}\left|a_{i,j}\left(y^n\right)\left(\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(y^n\right)\right)\right)-a_{i,j}\left(\mu\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)\right| \\
&\leq \frac{1}{b^2}\sum_{j=1}^{k}\left(\left|a_{i,j}\left(y^n\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(y^n\right)\right)-a_{i,j}\left(y^n\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)\right|+\right. \\
&\quad \left.\left|a_{i,j}\left(y^n\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)-a_{i,j}\left(\mu\right)\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)\right|\right) \\
&\leq \frac{1}{b^2}\sum_{j=1}^{k}\max_{Q_{i',j}\neq Q_{i,j}} a_{i,j}^{\max}\left(\left|\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(y^n\right)\right)-\left(\sum_{i\neq i'}\gamma_i^{n-1}a_{i,j}\left(\mu\right)\right)\right|+\left|a_{i,j}\left(y^n\right)-a_{i,j}\left(\mu\right)\right|\right) \\
&\leq \frac{2kK_a\max_j\max_{Q_{i',j}\neq Q_{i,j}} a_{i,j}^{\max}}{b^2}\|y^n-\mu\|,
\end{aligned}
$$

which completes the proof. □

A.2. Properties of Algorithm 2

Below, we discuss several properties of the sequential algorithm that explain various design choices used in its construction. We also prove properties that will be important later for showing the main results.

Nondifferentiability of g. Step 0′ of Algorithm 2 is designed to avoid instability resulting from γ^{n-1} being too close to the boundary of the feasible region

$$\mathcal{G}=\left\{\gamma\in\mathbb{R}^{k-1}:\gamma_i\geq 0,\,\mathbf{1}^T\gamma=1\right\}.$$

This is because $g\left(\gamma;y\right)$ is nondifferentiable only on the boundary, as we will now show. First, from the definition of g it follows that ∇g is not well-defined if and only if there exists $j\in\{1,...,k\}$ such that

$$\sum_{i\neq i^*(y)}\gamma_i a_{i,j}\left(y\right)=0, \tag{37}$$

where $i^*\left(y\right)=\arg\max_i y_i$.

We first observe that $a_{i,j}>0$ for all j and $i\neq i^*$. Since $Q_{i^*,i^*}-Q_{i,i^*}>0$ by Theorem 1, we have $a_{i,i^*}>0$. On the other hand, for $j\neq i^*$, we have $Q_{i^*,j}-Q_{j,j}<0$ by the same result, so $a_{j,j}>0$. Since all $\gamma_i,a_{i,j}\geq 0$, it follows that (37) can hold if and only if $\gamma_j=0$. This motivates Step 0′, in which we perturb γ^{n-1} to avoid this situation. Nonetheless, the issue of nondifferentiability is not critical to the long-term performance of the algorithm, as shown in the following result.

LEMMA 4. *Let $\gamma \in \mathcal{G}$ and suppose that $\sum_{i \neq i^*} \gamma_i a_{i,j}(y) > 0$ for all j. Then, $\sum_{i \neq i^*} \gamma_i a_{i,j}(y') > 0$ for all j and y'.*

Proof. From the premise, we know that for each j, there exists at least one i with both $\gamma_i, a_{i,j}(y) > 0$. Note that $a_{i,j}(y) = \frac{\sigma_j^2 P_{i,j}^2}{\left(P_i^T y\right)^2}$, so changing y to y' will not change whether or not $a_{i,j}$ is zero or nonzero. The conclusion then follows. □

Since $i^{*,n}$ converges to i', for all sufficiently large n, the sum $\sum_{i \neq i^*} \gamma_i a_{i,j}(y)$ will be computed for a fixed i'. Once this happens, Lemma 4 ensures that Step 0′ will no longer be invoked after some finite number of iterations.

Sufficiency of reduced direction set. Because the dual problem (14) has a fairly simple feasible region, verifying the optimality of a dual solution can also be simplified. Given $\gamma \in \mathcal{G}$, the set of all feasible directions is given by

$$\mathcal{D}(\gamma) = \left\{ d \in \mathbb{R}^{k-1} : \mathbf{1}^T d = 0, d_i \geq 0 \text{ if } \gamma_i = 0 \right\}.$$

However, Lin et al. (2009) showed that $\mathcal{D}(\gamma)$ can be generated by a much smaller set

$$D^h(\gamma) = \{e_i - e_h : i \neq h\} \bigcup \{e_h - e_i : i \neq h, \gamma_i > 0\}$$

for any h with $\gamma_h > 0$. In other words, $\text{Cone}\left(D^h(\gamma)\right) = \mathcal{D}(\gamma)$. Consequently, we obtain the following simplified optimality condition.

LEMMA 5. *(Lin et al. 2009) Fix any h with $\gamma_h > 0$. A dual solution $\gamma \in \mathcal{G}$ is a stationary point of (14) if and only if $\nabla g(\gamma; y)$ is well-defined and*

$$\nabla g(\gamma; y)^T d \geq 0, \qquad \forall d \in D^h(\gamma).$$

Lemma 5 motivates the use of a reduced direction set in Step 2 of Algorithm 2, as well as the arbitrary selection of h in Step 1. Note that, for η small enough, there will always be at least one h with $\gamma_h \geq \eta$ due to the equality constraint in $\mathcal{G}$.

Modified line search. Step 3 in Algorithm 2 uses the modified line search procedure in Algorithm 3. We now prove that this procedure terminates in finite time. The proof is based on Proposition 4.1 in Lin et al. (2009) and Lemma 3.1 in Nocedal and Wright (2006), and also uses Lemmas 5 and 3 from Section A.1.

Note that, in Step 3 of Algorithm 2, we may not call the line search procedure in every time stage n. Thus, technically, the following result should be applied to the subsequence of time stages at which line search is called; however, to avoid overcomplicating the notation, we work with the sequence of *distinct* iterates obtained from line search and relabel their indices as $\{1, 2, ...\}$.

LEMMA 6. *Let γ^n be the solution obtained from the nth time that line search is called (thus, γ^{n-1} denotes the dual solution obtained from the previous call, y^n is the vector of most recent sample means, and d^n is the most recent descent direction). Then:*

(i) At step n, a stepsize s^n can always be found in a finite number of iterations such that conditions (20), (21) and (22) are satisfied with $\gamma = \gamma^{n-1}$ and $y = y^n$.

(ii) Suppose that $\gamma^n \to \bar{\gamma}$ and

$$\lim_{n\to\infty} g\left(\gamma^{n-1};\mu\right) - g\left(\gamma^{n-1} + s^n d^n;\mu\right) = 0.$$

Then,

$$\lim_{n\to\infty} s^{\max}\left(d^n,\gamma^{n-1}\right)\nabla g\left(\gamma^{n-1};\mu\right)^T d^n = 0 \tag{38}$$

$$\lim_{n\to\infty} s^{\max}\left(d^n,\gamma^{n-1}\right)\nabla g\left(\gamma^{n-1};y^n\right)^T d^n = 0. \tag{39}$$

Proof. We know that $\nabla g\left(\gamma^{n-1};y^n\right)^T d^n < 0$ by Lemma 5. Since $\min_j \sum_{i\neq i^{*,n}} \gamma_i a_{i,j}\left(y^n\right) = 0$ can only occur on the boundary of $\mathcal{G}$, condition (20) must be satisfied for any $s \in (0, s^{\max}(d^n,\gamma^{n-1}))$. Since $g\left(\cdot\,;y^n\right)$ is smooth on $\mathcal{G}$, there must exist some $s_0 > 0$ such that $\nabla g(\gamma^{n-1} + sd^n\,;y^n)^T d^n < 0$ for any $s \in (0,s_0)$. In other words, when the stepsize is small enough, condition (22) can be neglected. Since $g\left(\cdot;y^n\right)$ is bounded below by Lemma 2, while the mapping $\ell(s) = g\left(\gamma^{n-1};y^n\right) + \alpha s\nabla g\left(\gamma^{n-1};y^n\right)^T d^n$ is unbounded below, it follows that $\ell(s)$ must intersect $g\left(\gamma^{n-1}+sd^n;y^n\right)$. Let s_1 be the smallest such intersection point; then, for any $s \in (0,s_1)$, condition (21) must be satisfied. It is easy to see that there exists some positive integer m such that

$$\tau^m s^{\max}\left(d^n,\gamma^{n-1}\right) < \min\left\{s^{\max}\left(d^n,\gamma^{n-1}\right), s_0, s_1\right\},$$

which is sufficient to establish (i).

With regard to (ii), (38) follows directly from Proposition 4.1(ii) in Lin et al. (2009), while (39) follows from Lemma 3. □

A.3. Proof of Theorem 5

Proof. To establish the main result, it is sufficient to show

$$\liminf_{n\to\infty} \sum_{i\neq i'} \gamma_i^n a_{i,j}\left(y^n\right) > 0, \tag{40}$$

and apply (15) together with Lemma 2. Note that we have replaced $i^{*,n}$ in (40) by $i' = \lim_{n\to\infty} i^{*,n}$ as in Section A.1. In the following, we assume without loss of generality that $i^{*,n} = i'$, as this will be true for sufficiently large n and we are interested in the asymptotic behaviour of the algorithm.

When $j = i'$, (40) follows from the boundedness of $\{a_{i,j}(y^n)\}$, established previously. Thus, it is sufficient to handle the case $j \neq i'$. We proceed by contradiction: suppose that there exists $j \neq i'$ with

$$\liminf_{n\to\infty} \sum_{i\neq i'} \gamma_i^n a_{i,j}(y^n) = 0.$$

Then, there exists a subsequence $\{n_t\}$ such that $\sum_{i\neq i'} \gamma_i^{n_t} a_{i,j}(y^{n_t}) \to 0$. Because $\{a_{i,j}(y^n)\}$ is uniformly bounded, it must be the case that $\gamma_i^{n_t} \to 0$ for all i satisfying $Q_{i',j} \neq Q_{i,j}$.

Define

$$C = \max_{i\neq i'} \sum_{j'=1}^{k} \sqrt{a_{i,j'}^{\max}},$$

$$C_1 = \frac{\min_{i:Q_{i',j}\neq Q_{i,j}} \left(a_{i,j}^{\min}\right)^2}{\sum_{i:Q_{i',j}\neq Q_{i,j}} a_{i,j}^{\max}},$$

$$c_0 = \frac{C_1}{C^2}\eta.$$

Note that $\frac{C_1}{C^2} \leq 1$ and $c_0 \leq \eta$. Letting b_1, b_2 be the values in Lemma 2, choose ε satisfying

$$0 < \varepsilon < \min\left\{ \frac{C_1}{C^2\left(\frac{1+\alpha'}{\sqrt{\eta}} + \frac{\alpha'}{\sqrt{c_0}}\right)^2}, \frac{C_1}{\left(\sqrt{\frac{3}{2}}\frac{C}{\eta} + \frac{3\alpha'(b_2-b_1)}{\alpha\eta}\right)^2}, c_0, \frac{\eta}{3} \right\}. \tag{41}$$

There exists t_0 such that, for any $t > t_0$, $\gamma_i^{n_t} \leq \varepsilon$. Since the direction is in the form of $e_h - e_{h'}$, $h \neq h'$, $h, h' \neq i'$, the value of γ^{n_t-1} must be described by one and only one of the following three cases:

1) For all i satisfying $Q_{i',j} \neq Q_{i,j}$, $\gamma_i^{n_t-1} \leq c_0$.

2) There exists i satisfying $Q_{i',j} \neq Q_{i,j}$ such that $\gamma_i^{n_t-1} \in (c_0, \eta)$ and $\gamma_{j'}^{n_t-1} \leq \varepsilon$ for $j' \neq i$ satisfying $Q_{i',j} \neq Q_{j',j}$.

3) There exists i satisfying $Q_{i',j} \neq Q_{i,j}$ such that $\gamma_i^{n_t-1} \geq \eta$, and $\gamma_{j'}^{n_t-1} \leq \varepsilon$ for $j' \neq i$ satisfying $Q_{i',j} \neq Q_{j',j}$.

We consider each case separately.

In case 1),

$$\sum_{i\neq i'} \gamma_i^{n_t-1} a_{i,j}^{n_t} \leq c_0 \sum_{Q_{i',j}\neq Q_{i,j}} a_{i,j}^{n_t} \leq c_0 \sum_{Q_{i',j}\neq Q_{i,j}} a_{i,j}^{\max}.$$

By assumption, $\gamma_i^{n_t-1} \leq c_0 \leq \eta$ for all i satisfying $Q_{i',j} \neq Q_{i,j}$, and therefore the index h^{n_t} selected in Step 1 of Algorithm 2 cannot be such an i. The corresponding component of the gradient $\nabla g(\gamma^{n_t-1}; y^{n_t})$ must then be bounded below, as can be seen from

$$\left[\nabla g(\gamma^{n_t-1}; y^{n_t})\right]_{h^{n_t}} = -\frac{1}{2}\sum_{i=1}^{k} \frac{a_{h^{n_t},i}(y^{n_t})}{\sqrt{\sum_{j'\neq i'} \gamma_{j'}^{n_t-1} a_{j',i}(y^{n_t})}}$$

$$\geq -\frac{1}{2}\sum_{i=1}^{k}\frac{a_{h^{n_t},i}\left(y^{n_t}\right)}{\sqrt{\gamma_{h^{n_t}}^{n_t-1}a_{h^{n_t},i}\left(y^{n_t}\right)}}$$
$$= -\frac{1}{2}\sum_{i=1}^{k}\sqrt{\frac{a_{h^{n_t},i}\left(y^{n_t}\right)}{\gamma_{h^{n_t}}^{n_t-1}}}$$
$$\geq -\frac{1}{2\sqrt{\eta}}\sum_{i=1}^{k}\sqrt{a_{h^{n_t},i}\left(y^{n_t}\right)}$$
$$\geq -\frac{C}{2\sqrt{\eta}}. \tag{42}$$

On the other hand, for i satisfying $Q_{i',j}\neq Q_{i,j}$, the corresponding component of the gradient satisfies

$$\begin{aligned}\left[\nabla g(\gamma^{n_t-1};y^{n_t})\right]_i &= -\frac{1}{2}\sum_{j'=1}^{k}\frac{a_{i,j'}\left(y^{n_t}\right)}{\sqrt{\sum_{j''\neq i'}\gamma_{j''}^{n_t-1}a_{j'',j'}\left(y^{n_t}\right)}}\\ &\leq -\frac{a_{i,j}\left(y^{n_t}\right)}{2\sqrt{\sum_{j''\neq i'}\gamma_{j''}^{n_t-1}a_{j'',j}\left(y^{n_t}\right)}}\\ &\leq -\frac{a_{i,j}^{\min}}{2\sqrt{c_0\sum_{j'':Q_{i',j}\neq Q_{j'',j}}a_{j'',j}^{\max}}}\\ &\leq -\frac{Ca_{i,j}^{\min}}{2\sqrt{\eta\min_{j'':Q_{i',j}\neq Q_{j'',j}}a_{j'',j}^{\min}}}\\ &\leq -\frac{C}{2\sqrt{\eta}}\\ &\leq \left[\nabla g(\gamma^{n_t-1};y^{n_t})\right]_{h^{n_t}}.\end{aligned}$$

Therefore, the direction d^{n_t} chosen in Step 2 cannot be $e_{h^{n_t}}-e_i$. Consequently, $\gamma_i^{n_t}\geq\gamma_i^{n_t-1}$, i.e., if $\gamma_i^{n_t-1}\leq c_0$ and $Q_{i',j}\neq Q_{i,j}$, then the ith component of γ^{n_t-1} cannot decrease at step n_t. Therefore, the only possible way to have $\sum_{i\neq i'}\gamma_i^{n_t}a_{i,j}\left(y^{n_t}\right)\to 0$ is $\gamma_i^{n_t}\equiv 0$ for all i satisfying $Q_{i',j}\neq Q_{i,j}$, but this is prevented by Step $0'$ of Algorithm 2 and condition (20).

In case 2), similarly, we have $h^{n_t}\neq i$ for any i satisfying $Q_{i',j}\neq Q_{i,j}$. To have $\gamma_i^{n_t}\leq\varepsilon$ for all such i, the direction d^{n_t} chosen in Step 2 must be $e_{h^{n_t}}-e_i$ with maximum feasible stepsize $s^{\max}\left(d^{n_t},\gamma^{n_t-1}\right)=\gamma_i^{n_t-1}$ along this direction, so that the Algorithm can enter into the line search step at n_t. Furthermore, the selected stepsize s^{n_t} must satisfy $s^{n_t}\geq c_0-\varepsilon$ and the modified Wolfe condition

$$\nabla g\left(\gamma^{n_t};y^{n_t}\right)^T d^{n_t}>0 \qquad \Rightarrow \qquad \nabla g\left(\gamma^{n_t};y^{n_t}\right)^T d^{n_t}\leq\alpha'\left|\nabla g\left(\gamma^{n_t-1};y^{n_t}\right)^T d^{n_t}\right|, \tag{43}$$

where $\gamma^{n_t}=\gamma^{n_t-1}+s^{n_t}d^{n_t}$. Since $d^{n_t}=e_{h^{n_t}}-e_i$, we have $\gamma_{h^{n_t}}^{n_t}\geq\gamma_{h^{n_t}}^{n_t-1}\geq\eta$. Similarly to (42), we can obtain

$$\left[\nabla g\left(\gamma^{n_t-1};y^{n_t}\right)\right]_{h^{n_t}}\in\left[-\frac{C}{2\sqrt{\eta}},0\right], \qquad \left[\nabla g\left(\gamma^{n_t};y^{n_t}\right)\right]_{h^{n_t}}\in\left[-\frac{C}{2\sqrt{\eta}},0\right].$$

Similarly, since $\gamma_i^{n_t-1} > c_0$, we also have

$$\left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_i \in \left[-\frac{C}{2\sqrt{c_0}}, 0\right].$$

Then (43) yields

$$\begin{aligned}
\left[\nabla g\left(\gamma^{n_t}; y^{n_t}\right)\right]_i &\geq \left[\nabla g\left(\gamma^{n_t}; y^{n_t}\right)\right]_{h^{n_t}} - \alpha' \left|\left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_{h^{n_t}} - \left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_i\right| \\
&\geq \left[\nabla g\left(\gamma^{n_t}; y^{n_t}\right)\right]_{h^{n_t}} - \alpha' \left(\left|\left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_{h^{n_t}}\right| + \left|\left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_i\right|\right) \\
&\geq -\frac{C}{2\sqrt{\eta}} - \alpha'\left(\frac{C}{2\sqrt{\eta}} + \frac{C}{2\sqrt{c_0}}\right) \\
&= -\frac{C}{2}\left(\frac{1+\alpha'}{\sqrt{\eta}} + \frac{\alpha'}{\sqrt{c_0}}\right). \qquad (44)
\end{aligned}$$

Since, by assumption, we have $\gamma_i^{n_t} \leq \varepsilon$ for any i satisfying $Q_{i',j} \neq Q_{i,j}$, it follows that

$$\sum_{i \neq i'} \gamma_i^{n_t} a_{i,j}\left(y^{n_t}\right) \leq \varepsilon \sum_{i: Q_{i',j} \neq Q_{i,j}} a_{i,j}^{\max},$$

and, consequently,

$$\left[\nabla g\left(\gamma^{n_t}; y^{n_t}\right)\right]_i \leq -\frac{a_{i,j}^{\min}}{2\sqrt{\varepsilon \sum_{i: Q_{i',j} \neq Q_{i,j}} a_{i,j}^{\max}}} < -\frac{C}{2}\left(\frac{1+\alpha'}{\sqrt{\eta}} + \frac{\alpha'}{\sqrt{c_0}}\right), \qquad (45)$$

which contradicts (44). Therefore, for case 2), it is not possible to have $\gamma_i^{n_t} \leq \varepsilon$ for all i satisfying $Q_{i',j} \neq Q_{i,j}$.

In case 3), there are two subcases: either (i) $Q_{i',j} = Q_{h^{n_t},j}$, or (ii) $i = h^{n_t}$. Subcase (i) is handled similarly to case 2, resulting in the desired contradiction. Therefore, we can focus on the second subcase $i = h^{n_t}$, and we have

$$\left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_i \geq -\frac{C}{2\sqrt{\eta}}$$

since $\gamma_i^{n_t-1} > \eta$. It is necessary for the value of $\gamma_i^{n_t-1}$ to decrease such that $\gamma_i^{n_t} \leq \varepsilon$, which implies that $d^{n_t} = e_{i''} - e_i$ for some $i'' \neq i, i'$ and

$$s^{n_t} \geq \eta - \varepsilon \geq \frac{2\eta}{3} > \varepsilon.$$

First of all, notice that i'' cannot satisfy $Q_{i',j} \neq Q_{i'',j}$, otherwise we would have $\gamma_{i''}^{n_t} = \gamma_{i''}^{n_t-1} + s^{n_t} > \varepsilon$. Then, condition (21) yields

$$\begin{aligned}
g\left(\gamma^{n_t}; y^{n_t}\right) &\leq g\left(\gamma^{n_t-1}; y^{n_t}\right) + \alpha s^{n_t} \nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)^T d^{n_t} \\
&= g\left(\gamma^{n_t-1}; y^{n_t}\right) + \alpha s^{n_t}\left(\nabla\left[g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_{i''} - \left[\nabla g\left(\gamma^{n_t-1}; y^{n_t}\right)\right]_i\right).
\end{aligned}$$

Then,

$$\begin{aligned}0 &> \nabla\left[g(\gamma^{n_t-1};y^{n_t})\right]_{i''} - \left[\nabla g(\gamma^{n_t-1};y^{n_t})\right]_i \\ &\geq \frac{1}{\alpha s^{n_t}}\left(g\left(\gamma^{n_t};y^{n_t}\right) - g\left(\gamma^{n_t-1};y^{n_t}\right)\right) \\ &\geq -\frac{b_2-b_1}{\alpha s^{n_t}} \\ &\geq -\frac{3\left(b_2-b_1\right)}{2\alpha\eta},\end{aligned}$$

where the third line is obtained by Lemma 2, and the fourth line is due to the fact that $s^{n_t} \geq \frac{2\eta}{3}$. Since $\gamma_{i''}^{n_t} = \gamma_{i''}^{n_t-1} + s^{n_t} \geq \frac{2\eta}{3}$, we have

$$\left[\nabla g\left(\gamma^{n_t};y^{n_t}\right)\right]_{i''} \geq -\frac{\sqrt{3}C}{2\sqrt{2\eta}}.$$

Then, repeating the arguments of (43) and (44), we have

$$\begin{aligned}\left[\nabla g\left(\gamma^{n_t};y^{n_t}\right)\right]_i &\geq \left[\nabla g\left(\gamma^{n_t};y^{n_t}\right)\right]_{i''} - \alpha'\left|\left[\nabla g\left(\gamma^{n_t-1};y^{n_t}\right)\right]_{i''} - \left[\nabla g\left(\gamma^{n_t-1};y^{n_t}\right)\right]_i\right| \\ &\geq -\frac{\sqrt{3}C}{2\sqrt{2\eta}} - \alpha'\left(\frac{3\left(b_2-b_1\right)}{2\alpha\eta}\right). \qquad (46)\end{aligned}$$

By the definition of ε, similar to (45), we find that (46) contradicts the assumption that $\gamma_i^{n_t} \leq \varepsilon$ for any i satisfying $Q_{i',j} \neq Q_{i,j}$. This completes the proof. □

A.4. Proof of Theorem 6

With Theorem 5, we can now replace $i^{*,n}$ by i^* in all calculations made by the algorithm, since we now know that $\lim_{n\to\infty} y^n = \mu$. We first prove a technical lemma showing that ∇g is Lipschitz continuous in γ (previously Lemma 3 considered Lipschitz continuity in the estimated value y).

LEMMA 7. *Let $\{\gamma^n\}$ be the sequence generated by Algorithm 2, and define $\mathcal{B}^n = Conv\left(\{\gamma^{n-1},\gamma^n\}\right)$, $n=1,2,....$ There exists a constant $K_\gamma > 0$, such that for all n,*

$$\|\nabla g\left(\gamma;y^n\right) - \nabla g\left(\gamma';y^n\right)\| \leq K_\gamma\|\gamma-\gamma'\|, \quad \forall\gamma,\gamma' \in \mathcal{B}^n.$$

Proof. From the boundedness of $\{a_{i,j}\left(y^n\right)\}$, we can see that, for any j and any $\gamma,\gamma' \in \mathcal{G}$,

$$\left|\sum_{i\neq i^*}\left(\gamma_i-\gamma_i'\right)a_{i,j}^n\right| \leq \max_{i,j\,:\,Q_{i^*,j}\neq Q_{i,j}} a_{i,j}^{\max}\|\gamma-\gamma'\|_\infty \leq K_r\|\gamma-\gamma'\|, \qquad (47)$$

for some constant $K_r > 0$. Combining (40) with (47), we have

$$\liminf_{n\to\infty}\sum_{i\neq i^*}\gamma_i^n a_{i,j}\left(y^n\right) > 0.$$

Therefore, there exists a constant $b_r > 0$ such that

$$\sum_{i \neq i^*} \gamma_i^{n-1} a_{i,j}\left(y^n\right) \geq b_r, \qquad \sum_{i \neq i^*} \gamma_i^{n} a_{i,j}\left(y^n\right) \geq b_r$$

for all j and all n. Since $\gamma \mapsto \sum_{i \neq i^*} \gamma_i a_{i,j}\left(y^n\right)$ is an affine function, we have

$$\sum_{i \neq i^*} \gamma_i a_{i,j}\left(y^n\right) \geq b_r$$

for any $\gamma \in \mathcal{B}^n$. Now, for any $i \neq i^*$ and any $\gamma, \gamma' \in \mathcal{B}^n$, we obtain

$$\begin{aligned}
\left|\left[\nabla g\left(\gamma; y^n\right)\right]_i - \left[\nabla g\left(\gamma'; y^n\right)\right]_i\right| &= \left|\sum_{j=1}^{k} \frac{a_{i,j}\left(y^n\right)}{\sum_{i \neq i^*} \gamma_i a_{i,j}\left(y^n\right)} - \sum_{j=1}^{k} \frac{a_{i,j}\left(y^n\right)}{\sum_{i \neq i^*} \gamma_i' a_{i,j}\left(y^n\right)}\right| \\
&\leq \sum_{j=1}^{k} a_{i,j}\left(y^n\right) \frac{\left|\sum_{i \neq i^*} \gamma_i' a_{i,j}\left(y^n\right) - \sum_{i \neq i^*} \gamma_i a_{i,j}\left(y^n\right)\right|}{\left(\sum_{i \neq i^*} \gamma_i a_{i,j}\left(y^n\right)\right)\left(\sum_{i \neq i^*} \gamma_i' a_{i,j}\left(y^n\right)\right)} \\
&\leq \frac{k K_r \max_{i,j: Q_{i^*,j} \neq Q_{i,j}} a_{i,j}^{\max}}{b_r^2} \|\gamma - \gamma'\|,
\end{aligned}$$

which completes the proof. □

We now prove Theorem 6. Let $\bar{\gamma} \in \mathcal{G}$ be a limit point of $\{\gamma^{n-1}\}$, so that there exists a subsequence $\gamma^{n_t - 1} \to \bar{\gamma}$. We can pick h such that $\bar{\gamma}_h \geq \eta$ and $\mathcal{D}\left(\bar{\gamma}\right) = \text{Cone}\left(D^h\left(\bar{\gamma}\right)\right)$. Furthermore, there exists t_0 such that, for all $t > t_0$, this same h satisfies $\gamma^{n_t - 1} \geq \eta$. Since h^{n_t} is chosen from among all such h with equal probability, we can further extract a subsequence $\{n_u\} \subseteq \{n_t\}$ such that $h^{n_u} \equiv h$ for all sufficiently large u.

Suppose that $\bar{\gamma}$ is not a stationary point of (14) under the true values μ. Then, there is a feasible $\bar{d} \in D^h\left(\bar{\gamma}\right)$ such that

$$\nabla g\left(\bar{\gamma}; \mu\right)^T \bar{d} < 0. \tag{48}$$

By the convergence of $\{\gamma^{n_u - 1}\}$, we have $\bar{d} \in D^h\left(\gamma^{n_u - 1}\right)$ for all sufficiently large u. Since $y^n \to \mu$, we can find $c_1 > 0$ such that

$$\nabla g\left(\gamma^{n_u - 1}; y^{n_u}\right)^T \bar{d} \leq -c_1 < 0 \tag{49}$$

for all sufficiently large u. Combining the convergence of $\{\gamma^{n_u - 1}\}$ with Proposition A.1 in Lin et al. (2009), we obtain a constant $c_2 > 0$ such that $s^{\max}\left(\bar{d}, \gamma^{n_u - 1}\right) \geq c_2$ for all sufficiently large u. Then, at iteration n_u, Step 2 of Algorithm 2 yields

$$s^{\max}\left(d^{n_u}, \gamma^{n_u - 1}\right) \nabla g\left(\gamma^{n_u - 1}; y^{n_u}\right)^T d^{n_u} \leq s^{\max}\left(\bar{d}, \gamma^{n_u - 1}\right) \nabla g\left(\gamma^{n_u - 1}; y^{n_u}\right)^T \bar{d} \leq -c_1 c_2 < 0. \tag{50}$$

Define three subsequences

$$E_1 = \left\{ n : V^n \geq \max\left\{ -\kappa_0, -\left(\frac{\log n}{n}\right)^{1/4} \right\} \right\},$$
$$E_2 = \left\{ n : s^{\max}\left(d^n, \gamma^{n-1}\right) V^n \geq \max\left\{ -\kappa_0, -\left(\frac{\log n}{n}\right)^{1/2} \right\} \right\},$$
$$E_3 = \{1, 2, ...\} \setminus (E_1 \cup E_2).$$

Notice that the line search method is invoked, and thus γ^n is updated, only along E_3. We also observe that

$$\liminf_{n \in E_1} s^{\max}\left(d^n, \gamma^{n-1}\right) V^n \geq 0, \qquad \liminf_{n \in E_2} s^{\max}\left(d^n, \gamma^{n-1}\right) V^n \geq 0. \tag{51}$$

Along E_3, we do line search, which terminates when s^n satisfies (21)-(22).

Using Lemma 7 and the convexity of $g(\cdot; y^n)$, we obtain

$$g\left(\gamma^{n-1} + sd^n; y^n\right) \leq g\left(\gamma^{n-1}; y^n\right) + s\nabla g\left(\gamma^{n-1}; y^n\right)^T d^n + \frac{s^2 K_\gamma}{2}\|d^n\|_2^2.$$

Then, (21) will be satisfied if we choose s such that

$$g\left(\gamma^{n-1}; y^n\right) + s\nabla g\left(\gamma^{n-1}; y^n\right)^T d^n + \frac{\lambda^2 K_\gamma}{2}\|d^n\|_2^2 \leq g\left(\gamma^{n-1}; y^n\right) + \alpha s \nabla g\left(\gamma^{n-1}; y^n\right)^T d^n,$$

so

$$s \leq \frac{(\alpha - 1)\nabla g\left(\gamma^{n-1}; y^n\right)^T d^n}{K_\gamma} = \frac{(\alpha - 1) V^n}{K_\gamma}.$$

Define $s^{K,n} = \frac{(\alpha-1)V^n}{K_\gamma}$ and $s^{*,n} = \arg\min_s g\left(\gamma^{n-1} + sd^n\right)$. Let $s^{W,n}$ be the largest s that satisfies (22). By the convexity of $g(\cdot; y^n)$, we have $s^{W,n} \geq s^{*,n}$.

Now suppose that $s^{\max}\left(d^n, \gamma^{n-1}\right) \leq \min\left\{s^{K,n}, s^{W,n}\right\}$. It follows that $s^n = s^{\max}\left(d^n, \gamma^{n-1}\right)$ and

$$g\left(\gamma^{n-1}; y^n\right) - g\left(\gamma^n; y^n\right) \geq -\alpha s^{\max}\left(d^n\right) V^n. \tag{52}$$

On the other hand, if $s^{\max}\left(d^n, \gamma^{n-1}\right) > \min\left\{s^{K,n}, s^{W,n}\right\}$, we consider two cases. In the first case, $s^{W,n} \geq s^{K,n}$, so $s^n \geq \tau s^{K,n}$, and

$$g\left(\gamma^{n-1}; y^n\right) - g\left(\gamma^n; y^n\right) \geq \frac{\alpha\tau(1-\alpha)\left(V^n\right)^2}{K_\gamma}. \tag{53}$$

In the second case, $s^{K,n} \geq s^{W,n}$. Then $s^n \in [\tau s^{*,n}, s^{W,n}]$. Since $g(\cdot; y^n)$ is convex, we have

$$g\left(\gamma^{n-1} + s^n d^n; y^n\right) \leq \max\left\{g\left(\gamma^{n-1} + \tau s^{*,n} d^n; y^n\right), g\left(\gamma^{n-1} + s^{W,n} d^n; y^n\right)\right\}.$$

Since

$$\begin{aligned} g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^{n-1}+\tau s^{*,n}d^n;y^n\right) &\geq g\left(\gamma^{n-1};y^n\right)-\left((1-\tau)\,g\left(\gamma^{n-1};y^n\right)+\tau g\left(\gamma^{n-1}+s^{*,n}d^n;y^n\right)\right) \\ &= \tau\left(g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^{n-1}+s^{*,n}d^n;y^n\right)\right) \\ &\geq \tau\left(g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^{n-1}+s^{K,n}d^n;y^n\right)\right) \\ &\geq \frac{\alpha\tau\left(1-\alpha\right)\left(V^n\right)^2}{K_\gamma}, \end{aligned}$$

and, similarly,

$$\begin{aligned} g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^{n-1}+s^{W,n}d^n;y^n\right) &\geq g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^{n-1}+s^{K,n}d^n;y^n\right) \\ &\geq \frac{\alpha\left(1-\alpha\right)\left(V^n\right)^2}{K_\gamma}, \end{aligned}$$

we again obtain (53). Combining (52) and (53), we arrive at

$$g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^n;y^n\right)\geq\min\left\{\frac{\alpha\tau\left(1-\alpha\right)\left(V^n\right)^2}{K_\gamma},-\alpha\tau s^{\max}\left(d^n,\gamma^{n-1}\right)V^n\right\}.$$

Combining Lemma 1 with the law of the iterated logarithm, we have

$$\begin{aligned} &g\left(\gamma^{n-1};\mu\right)-g\left(\gamma^n;\mu\right) \\ &= g\left(\gamma^{n-1};\mu\right)-g\left(\gamma^{n-1};y^n\right)+g\left(\gamma^{n-1};y^n\right)-g\left(\gamma^n;y^n\right)+g\left(\gamma^n;y^n\right)-g\left(\gamma^n;\mu\right) \\ &\geq \min\left\{\frac{\alpha\tau\left(1-\alpha\right)\left(V^n\right)^2}{K_\gamma},-\alpha\tau s^{\max}\left(d^n,\gamma^{n-1}\right)V^n\right\}+\left(g\left(\gamma^{n-1};\mu\right)-g\left(\gamma^{n-1};y^n\right)\right)+\left(g\left(\gamma^n;y^n\right)-g\left(\gamma^n;\mu\right)\right) \\ &= \min\left\{\frac{\alpha\tau\left(1-\alpha\right)\left(V^n\right)^2}{K_\gamma},-\alpha\tau s^{\max}\left(d^n,\gamma^{n-1}\right)V^n\right\}+O\left(\sqrt{\frac{\log\log n}{n}}\right). \end{aligned}$$

By the definition of E_3, we have

$$g\left(\gamma^{n-1};\mu\right)-g\left(\gamma^n;\mu\right)\geq\Omega\left(\sqrt{\frac{\log n}{n}}\right)+O\left(\sqrt{\frac{\log\log n}{n}}\right),\qquad n\in E_3,$$

which implies that, for sufficiently large $n\in E_3$, we have $g\left(\gamma^{n-1};\mu\right)\geq g\left(\gamma^n;\mu\right)$. In other words, the effect of estimation error is removed along E_3, and the objective value becomes monotonically decreasing.

Let $w\left(n\right)=\max\left\{w_0\in E_3:w_0<n\right\}$. Since γ^n is updated only along E_3, we have

$$g\left(\gamma^{w(n)};\mu\right)=g\left(\gamma^{w(n)+1};\mu\right)=\cdots=g\left(\gamma^{n-1};\mu\right)\geq g\left(\gamma^n;\mu\right). \tag{54}$$

Since $g(\cdot;\mu)$ is bounded below, the sequence $\left\{g\left(\gamma^n;\mu\right):n\in E_3\right\}$ converges. Since $g\left(\gamma;\mu\right)$ is continuous in $\gamma\in\mathcal{G}$, it follows that $\left\{\gamma^n:n\in E_3\right\}$ also converges to some limit $\hat{\gamma}\in\mathcal{G}$. Taking limits of both sides of (54) yields

$$\lim_{n\in E_3}g\left(\gamma^{n-1};\mu\right)-g\left(\gamma^{n-1}+s^nd^n;\mu\right)=0.$$

Applying Lemma 6(ii), we have

$$\lim_{n\in E_3} s^{\max}\left(d^n,\gamma^{n-1}\right)V^n=0. \tag{55}$$

From (51) and (55), we have $\liminf_{n\to\infty} s^{\max}(d^n,\gamma^{n-1})V^n\geq 0$, which contradicts (50). Therefore, $\bar{\gamma}$ must be a stationary point.

Appendix B: Other Proofs

Below, we give the full proofs for results that were stated in the text.

B.1. Proof of Theorem 1

Proof. (1) can be directly concluded from the fact that the inverse of Q, $I+\lambda L$, is positive semidefinite and has largest eigenvalue 1 with eigenvector $\mathbf{1}$.

To show (2), we recall that Sun et al. (2019) showed that the linear system $(I+\lambda L)z=y$ can be solved using an iterative algorithm

$$z^{(0)}=y,\quad z^{(t)}=z^{(t-1)}+a\left(y-(I+\lambda L)z^{(t-1)}\right),t\geq 1, \tag{56}$$

which is guaranteed to converge to the true solution if $a>0$ is sufficiently small. For any i, (56) can be rewritten as

$$z_i^{(t)}=ay_i+\left(1-a-a\sum_{j\neq i}S_{i,j}\right)z_i^{(t-1)}+a\lambda\sum_{j\neq i}S_{i,j}z_j^{(t-1)}. \tag{57}$$

Consequently, when a is sufficiently small, $y_i\geq 0$ implies $z_i^{(t)}\geq 0$, $i=1,...,k$. This implies that every entry of Q is non-negative.

For (3), let $Q=[Q_1,Q_2,...,Q_k]$, we have $(I+\lambda L)Q_l=e_l$, where $e_l=[0,...,0,1,0,...,0]^T$, 1 is the lth element. From (3), we know Q_l is the solution to the following optimization problem:

$$\arg\min_{x\in\mathbb{R}^k}\sum_{i\neq l}^{k}x_i^2+(x_l-1)^2+\frac{\lambda}{2}\sum_{1\leq i,j\leq k}S_{i,j}(x_i-x_j)^2. \tag{58}$$

Then we prove the results by contradiction. Assume the optimal solution is x^* and define $\Omega_1=\{i: x_i^*>x_l^*\}$. Suppose $x_l^*\neq\max_i x_i^*$, i.e., $\Omega_1\neq\emptyset$. Now we construct $\bar{x}$ by

$$\bar{x}_i=\begin{cases}x_i^*, & i\notin\Omega_1\\ x_l^*, & i\in\Omega_1\end{cases}.$$

Denote the value of the objective in (58) by $\psi(x)$ for a given x. Then,

$$\psi(\bar{x})=\sum_{i\neq l}^{k}\bar{x}_i^2+(\bar{x}_l-1)^2+\frac{\lambda}{2}\sum_{1\leq i,j\leq k}S_{i,j}(\bar{x}_i-\bar{x}_j)^2 \tag{59}$$

$$<\sum_{i\neq l}^{k}(x_i^*)^2+(x_l^*-1)^2+\frac{\lambda}{2}\sum_{1\leq i,j\leq k}S_{i,j}(\bar{x}_i-\bar{x}_j)^2. \tag{60}$$

Now we discuss the term $S_{i,j}(\bar{x}_i-\bar{x}_j)^2$ for four cases.

1) for $i \notin \Omega_1$ and $j \in \Omega_1$, we have $x_i^* \le x_l^* < x_j^*$, therefore, $(\bar{x}_i - \bar{x}_j)^2 = (x_i^* - x_l^*)^2 < (x_i^* - x_j^*)^2$;
2) for $i \in \Omega_1$ and $j \notin \Omega_1$, same as case (1);
3) for $i \notin \Omega_1$ and $j \notin \Omega_1$, $(\bar{x}_i - \bar{x}_j)^2 = (x_i^* - x_j^*)^2$;
4) for $i \in \Omega_1$ and $j \in \Omega_1$, $(\bar{x}_i - \bar{x}_j)^2 = (x_l^* - x_l^*)^2 = 0 \le (x_i^* - x_j^*)^2$.

In summary, we have

$$\psi(\bar{x}) < \sum_{i \neq l}^{k} (x_i^*)^2 + (x_l^* - 1)^2 + \frac{\lambda}{2} \sum_{1 \le i,j \le k} S_{i,j}(x_i^* - x_j^*)^2 = \psi(x^*),$$

which contradicts the assumption that x^* is optimal. Therefore, $x_l^* = \max_j x_j^*$. We further prove that $x_l^* > \max_{i \neq l} x_i^*$. From (1) and (2) in this theorem, we know that $x_i^* \in [0,1]$ and x^* can't be a multiple of $\mathbf{1}$. Therefore, there must be a positive gap between x_l^* and some other component of x^*. Let $\Omega_2 = \{i \neq l : x_i^* = x_l^*\}$ and suppose $\Omega_2 \neq \emptyset$. Let ε be a positive number less than $\min\{x_l^* - \max_{j \in \Omega_2^c \setminus \{l\}} x_j^*, \min_{i \in \Omega_2}\{\frac{2x_l^*}{1+\lambda S_{l,i}}\}\}$. Construct $\bar{x}$ by

$$\bar{x}_i = \begin{cases} x_i^*, & i \notin \Omega_2 \\ x_l^* - \varepsilon, & i \in \Omega_2 \end{cases}.$$

Notice that for $i \in \Omega_2$ and $j \in \Omega_2^c \setminus \{l\}$,

$$(\bar{x}_i - \bar{x}_j)^2 < (x_i^* - x_j^*)^2.$$

Then we have

$$\begin{aligned} \psi(\bar{x}) - \psi(x^*) &< \sum_{i \in \Omega_2} \bar{x}_i^2 - \sum_{i \in \Omega_2} {x_i^*}^2 + \lambda \sum_{i \in \Omega_2} S_{l,i}(x_l^* - \bar{x}_i)^2 \\ &= \sum_{i \in \Omega_2} \left((x_l^* - \varepsilon)^2 - (x_l^*)^2 + \lambda S_{l,i}(x_l^* - (x_l^* - \varepsilon))^2 \right) \\ &= \sum_{i \in \Omega_2} \varepsilon \left((1 + \lambda S_{l,i})\varepsilon - 2x_l^* \right) < 0, \end{aligned}$$

which leads to a contradiction. □

B.2. Proof of Theorem 2

Here, we give a corrected version of the proof of Theorem 2 (Sun et al. 2019), which had $D_{i,i} = \sum_{j \neq i} S_{i,j}$, where $S_{i,i}$ is not included in the sum. In their proof, they manipulated $S_{i,i}$ such that $D_{i,i} = D_{j,j}$, $i \neq j$, which is not consistent with their setting. We follow the same general arguments, but correct the technical issues.

Proof. At iteration 0, let $z^{(0)} = y$, we have $z_1^{(0)} \ge z_2^{(0)} \ge \cdots \ge z_k^{(0)}$. At iteration $t > 0$, assume $z_1^{(t-1)} \ge z_2^{(t-1)} \ge \cdots \ge z_k^{(t-1)}$ holds, using the iterative algorithm in (56), we need to show that $z_i^{(t)} \ge z_{i+1}^{(t)}$ holds for any $1 \le i \le k-1$. Since

$$z_i^{(t)} = a y_i + (1-a) z_i^{(t-1)} + a\lambda \sum_{j \neq i} S_{i,j} \left(z_j^{(t-1)} - z_i^{(t-1)} \right), \tag{61}$$

$$z_{i+1}^{(t)} = a y_{i+1} + (1-a) z_{i+1}^{(t-1)} + a\lambda \sum_{j \neq i+1} S_{i+1,j} \left(z_j^{(t-1)} - z_{i+1}^{(t-1)} \right), \tag{62}$$

taking the difference, we have:

$$\begin{aligned} z_i^{(t)} - z_{i+1}^{(t)} =& a(y_i - y_{i+1}) + (1-a)\left(z_i^{(t-1)} - z_{i+1}^{(t-1)}\right) \\ &+ a\lambda\left(\sum_{j\neq i} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right) - \sum_{j\neq i+1} S_{i+1,j}\left(z_j^{(t-1)} - z_{i+1}^{(t-1)}\right)\right). \end{aligned}$$

Since

$$\begin{aligned} &\sum_{j\neq i} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right) - \sum_{j\neq i+1} S_{i+1,j}\left(z_j^{(t-1)} - z_{i+1}^{(t-1)}\right) \\ =&\sum_{j<i} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right) - \sum_{j<i} S_{i+1,j}\left(z_j^{(t-1)} - z_{i+1}^{(t-1)}\right) + \sum_{j>i+1} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right) \\ &- \sum_{j>i+1} S_{i+1,j}\left(z_j^{(t-1)} - z_{i+1}^{(t-1)}\right) - 2S_{i,i+1}\left(z_i^{(t-1)} - z_{i+1}^{(t-1)}\right) \\ =&\sum_{j<i} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right) - \sum_{j<i} S_{i+1,j}\left(z_j^{(t-1)} - z_{i+1}^{(t-1)}\right) + \sum_{j>i+1} S_{i+1,j}\left(z_{i+1}^{(t-1)} - z_j^{(t-1)}\right) \\ &- \sum_{j>i+1} S_{i,j}\left(z_i^{(t-1)} - z_j^{(t-1)}\right) - 2S_{i,i+1}\left(z_i^{(t-1)} - z_{i+1}^{(t-1)}\right) \\ =&\sum_{j<i} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right) - \sum_{j<i} S_{i+1,j}\left(z_j^{(t-1)} - z_i^{(t-1)} + z_i^{(t-1)} - z_{i+1}^{(t-1)}\right) + \sum_{j>i+1} S_{i+1,j}\left(z_{i+1}^{(t-1)} - z_j^{(t-1)}\right) \\ &- \sum_{j>i+1} S_{i,j}\left(z_i^{(t-1)} - z_{i+1}^{(t-1)} + z_{i+1}^{(t-1)} - z_j^{(t-1)}\right) - 2S_{i,i+1}\left(z_i^{(t-1)} - z_{i+1}^{(t-1)}\right) \\ =&\sum_{j<i} (S_{i,j} - S_{i+1,j})\left(z_j^{(t-1)} - z_i^{(t-1)}\right) + \sum_{j>i+1} (S_{i+1,j} - S_{i,j})\left(z_{i+1}^{(t-1)} - z_j^{(t-1)}\right) \\ &- \left(\sum_{j<i} S_{i+1,j} + \sum_{j>i+1} S_{i,j} + 2S_{i,i+1}\right)\left(z_i^{(t-1)} - z_{i+1}^{(t-1)}\right), \end{aligned}$$

we have,

$$\begin{aligned} z_i^{(t)} - z_{i+1}^{(t)} =& a(y_i - y_{i+1}) + \left[1 - a - a\lambda\left(\sum_{j<i} S_{i+1,j} + \sum_{j>i+1} S_{i,j} + 2S_{i,i+1}\right)\right]\left(z_i^{(t-1)} - z_{i+1}^{(t-1)}\right) \\ &+ a\lambda\left[\sum_{j<i} (S_{i,j} - S_{i+1,j})\left(z_j^{(t-1)} - z_i^{(t-1)}\right) + \sum_{j>i+1} (S_{i+1,j} - S_{i,j})\left(z_{i+1}^{(t-1)} - z_j^{(t-1)}\right)\right]. \end{aligned} \tag{63}$$

Since $S_{i,j} \geq S_{i+1,j}$, $z_j^{(t-1)} \geq z_i^{(t-1)}$ when $j < i$, and $S_{i+1,j} \geq S_{i,j}$, $z_{i+1}^{(t-1)} \geq z_j^{(t-1)}$, when $j > i+1$,

$$\sum_{j<i} (S_{i,j} - S_{i+1,j})\left(z_j^{(t-1)} - z_i^{(t-1)}\right) + \sum_{j>i+1} (S_{i+1,j} - S_{i,j})\left(z_{i+1}^{(t-1)} - z_j^{(t-1)}\right) \geq 0.$$

Also, $y_i \geq y_{i+1}$ and $z_i^{(t-1)} \geq z_{i+1}^{(t-1)}$, if a is sufficiently small, we have

$$z_i^{(t)} \geq z_{i+1}^{(t)},$$

which completes the proof. □

B.3. Proof of Theorem 3

We use the same induction approach as in the proof of Theorem 2 by assuming $z^{(0)} = y$. Suppose that $z_1^{(0)} \geq z_i^{(0)}$ for any $i = 1, \ldots, k$. Assume that at iteration $t-1$, $z_1^{(t-1)} \geq z_i^{(t-1)}$, then at iteration t,

$$\begin{aligned}
z_1^{(t)} - z_i^{(t)} =& a(y_1 - y_i) + (1-a)\left(z_1^{(t-1)} - z_i^{(t-1)}\right) \\
&+ a\lambda\left(\sum_{j\neq 1} S_{1,j}\left(z_j^{(t-1)} - z_1^{(t-1)}\right) - \sum_{j\neq i} S_{i,j}\left(z_j^{(t-1)} - z_i^{(t-1)}\right)\right) \\
=& a(y_1 - y_i) + (1-a)\left(z_1^{(t-1)} - z_i^{(t-1)}\right) \\
&+ a\lambda\left(\sum_{j\neq 1} S_{1,j}\left(z_j^{(t-1)} - z_1^{(t-1)}\right) - \sum_{j\neq i} S_{i,j}\left(\left(z_1^{(t-1)} - z_i^{(t-1)}\right) + \left(z_j^{(t-1)} - z_1^{(t-1)}\right)\right)\right) \\
=& a(y_1 - y_i) + (1 - a - a\lambda\sum_{j\neq i} S_{i,j})\left(z_1^{(t-1)} - z_i^{(t-1)}\right) + \lambda\left(\sum_{j\neq 1,i}(S_{i,j} - S_{1,j})\left(z_1^{(t-1)} - z_j^{(t-1)}\right)\right)
\end{aligned}$$

Since $S_{i,j} \geq S_{1,j}$, $z_1^{(t-1)} \geq z_j^{(t-1)}$, if a is sufficiently small, we have

$$z_1^{(t)} \geq z_i^{(t)}, \quad i = 1, \ldots, k.$$

Thus, we have shown

$$\{y_1 \geq y_i,\, i = 1, \ldots, k\} \subseteq \{z_1 \geq z_i,\, i = 1, \ldots, k\},$$

which completes the proof. □

B.4. Proof of Proposition 1

Here, we let $M = 1$. For a general M, we only need to scale the optimal proportion by M to get the allocation vector. Then, for $k = 3$, the optimization problem (6) can be written in the form

$$\min_{\Omega} \max\left\{\frac{a_1}{x_1} + \frac{a_2}{x_2} + \frac{a_3}{x_3}, \frac{b_1}{x_1} + \frac{b_2}{x_2} + \frac{b_3}{x_3}\right\}, \tag{64}$$

where $\Omega = \{x : 0 < x_j < 1, \mathbf{1}^T x = 1\}$, $a_j \geq 0$, $b_j \geq 0$, $j = 1, 2, 3$. One can derive, either directly or by applying Theorem 4, the dual problem

$$\begin{aligned}
\max_{\gamma} \quad & \left(\sum_{j=1}^{k} \sqrt{\gamma_1 a_j + \gamma_2 b_j}\right)^2 \\
\text{subject to} \quad & \gamma_1 \geq 0, \gamma_2 \geq 0, \\
& \gamma_1 + \gamma_2 = 1,
\end{aligned} \tag{65}$$

which can be further simplified to

$$\begin{aligned} \max_{\gamma} \quad & \left(\sum_{j=1}^{k} \sqrt{\gamma_1 a_j + (1-\gamma_1) b_j}\right)^2 \\ \text{subject to} \quad & 0 \le \gamma_1 \le 1. \end{aligned} \tag{66}$$

Denote by $g(\gamma_1)$ the objective function of (66). It is easy to see that g is concave. Taking the derivative, we have

$$g'(\gamma_1) = \left(\sum_{j=1}^{k} \sqrt{\gamma_1 a_j + (1-\gamma_1) b_j}\right) \sum_{j=1}^{k} \left(\frac{a_j - b_j}{\sqrt{\gamma_1 a_j + (1-\gamma_1) b_j}}\right). \tag{67}$$

The primal (64) is convex, and it is easy to see that Slater's condition (and thus, strong duality) holds. Neglecting the factor $\sum_{j=1}^{k} \sqrt{\gamma_1 a_j + (1-\gamma_1) b_j} > 0$ in (67), we obtain the function $r(\cdot)$ in the statement of the theorem. The results then follow from the properties of one-variable concave optimization problems on $[0,1]$. $\square$

B.5. Proof of Proposition 2

Proof. Let

$$\Psi_M(\theta) = \mathbb{E}\left[e^{\theta z^M}\right] = e^{\theta^T Q\mu + \frac{1}{2}\theta^T Q \Lambda^2 Q^T \theta}$$

be the moment generating function of z^M, where Λ is a diagonal matrix with $\Lambda_{jj} = \frac{\sigma_j}{\sqrt{x_j \cdot M}}$. Here the sample size $N_j \approx x_j \cdot M$ is allowed to be fractional; however, as we will be passing to an asymptotic regime, this is not a major issue. Next, we calculate the scaled limit of the log-mgf, given by

$$\begin{aligned} \Psi(\theta) &= \lim_{M\to\infty} \frac{1}{M} \log \Psi_M(M\theta) \\ &= \lim_{M\to\infty} \frac{1}{M} \left(M\theta^T Q\mu + \frac{1}{2} M^2 \theta^T Q \Lambda^2 Q^T \theta\right) \\ &= \theta^T Q\mu + \lim_{M\to\infty} \frac{1}{2} \theta^T Q \left(M\Lambda^2\right) Q^T \theta \\ &= \theta^T Q\mu + \frac{1}{2} \theta^T Q \Gamma Q^T \theta. \end{aligned}$$

By the Gärtner-Ellis theorem, the rate function is the Fenchel-Legendre transform of Ψ, defined as

$$\begin{aligned} I(z) &= \sup_{\theta} \theta^T z - \Psi(\theta) \\ &= \sup_{\theta} \theta^T (z - Q\mu) - \frac{1}{2} \theta^T Q \Gamma Q^T \theta. \end{aligned} \tag{68}$$

The supremum is achieved at

$$\theta^* = Q\Gamma^{-1} Q^T (z - Q\mu),$$

and plugging this back into (68) yields the desired result. $\square$

B.6. Proof of Theorem 4

Proof. As in (7), we can linearize the objective of (13) by adding a scalar variable ξ. We then obtain the Lagrangian

$$L(x,\xi,\gamma,\beta) = \xi + \sum_{i\neq i^*}\gamma_i\left(\sum_{j=1}^k \frac{a_{i,j}}{x_j} - \xi\right) + \beta\left(\sum_{j=1}^k x_j - 1\right).$$

The dual function is obtained from the Lagrangian (Ch. 5, Boyd and Vandenberghe 2004) by computing

$$\begin{aligned} g(\gamma,\beta) &= \inf_{x,\xi} L(x,\gamma,\beta) \\ &= \inf_{x,\xi}\left(1-\sum_{i\neq i^*}\gamma_i\right)\xi + \sum_{i\neq i^*}\gamma_i\sum_{j=1}^k\frac{a_{i,j}}{x_j} + \beta\left(\sum_{j=1}^k x_j - 1\right) \\ &= \inf_{x,\xi}\left(1-\sum_{i\neq i^*}\gamma_i\right)\xi + \sum_{j=1}^k\left(\frac{\sum_{i\neq i^*}\gamma_i a_{i,j}}{x_j} + \beta x_j\right) - \beta. \end{aligned} \tag{69}$$

From (69), we can see that

$$g(\gamma,\beta) = \begin{cases} 2\sqrt{\beta}\sum_{j=1}^k\sqrt{\sum_{i\neq i^*}\gamma_i a_{i,j}} - \beta, & \sum_{i\neq i^*}\gamma_i = 1 \text{ and } \beta\geq 0, \\ -\infty, & \text{otherwise.}, \end{cases}$$

Therefore, the dual problem of (13) is given by

$$\begin{aligned} \max_{\gamma,\beta} \quad & 2\sqrt{\beta}\sum_{j=1}^k\sqrt{\sum_{i\neq i^*}\gamma_i a_{i,j}} - \beta \\ \text{subject to} \quad & \gamma_i \geq 0, \quad i\neq i^*, \\ & \sum_{i\neq i^*}\gamma_i = 1, \\ & \beta \geq 0. \end{aligned} \tag{70}$$

It is possible to simplify (70) by removing β entirely. Taking the derivative of the dual objective with respect to β, and setting it equal to zero, we obtain

$$\frac{\sum_{j=1}^k\sqrt{\sum_{i\neq i^*}\gamma_i a_{i,j}}}{\sqrt{\beta}} - 1 = 0$$

so

$$\beta = \left(\sum_{j=1}^k\sqrt{\sum_{i\neq i^*}\gamma_i a_{i,j}}\right)^2.$$

Substituting this expression back into (70), we obtain (14), as required. It can be easily seen that the primal problem is strictly feasible, so strong duality holds. □

B.7. Proof of Theorem 7

We first state three technical lemmas, which will be proved in separate subsections of the Appendix.

LEMMA 8. *Suppose that $k \geq 3$ with $\mu_1 > \ldots > \mu_k$. Given any fixed similarity matrix S with $S_{i,j} \geq 0$ and $S_{i,j} = S_{j,i}$, $i, j = 1, \ldots, k$, the sampling ratios achieved by SIMA are given by*

$$\lim_{n\to\infty} \frac{N_i^n}{N_j^n} = \frac{c_j}{c_i},$$

where

$$c_j = \left| \max_{u \neq i} \frac{z_u - z_j}{\sigma_j(Q_{j,j} - Q_{j,u})} \right|, \tag{71}$$

and $z = (I + \lambda L(S))^{-1}\mu = Q(S)\mu$.

LEMMA 9. *Let $\mathbb{S}^k$ be the set of all real possible similarity matrices $S(= [S_{i,j}]_{k\times k})$ for aligned graphs, $\mathbb{Y}^k = \{y \in \mathbb{R}^k | y_1 \geq y_2 \geq ... \geq y_k\}$ and $\Omega^k = \{2, ..., k\}$. Given $\lambda > 0$, $\mathcal{D}^k(S, y, \zeta, j) : \mathbb{S}^k \times \mathbb{Y}^k \times (\mathbb{R}^+)^k \times \Omega^k \mapsto \mathbb{R}$ is defined by*

$$\mathcal{D}^k(S, y, \zeta, j) = \det\left([y, W_2, ..., W_{j-1}, \mathbf{1}, W_{j+1}, ..., W_k]\right), \tag{72}$$

where $W_i = [-\lambda S_{i,1}, ..., -\lambda S_{i,i-1}, 1 + \lambda D_{i,i} + \lambda\zeta_i, -\lambda S_{i,i+1}, ..., -\lambda S_{i,k}]^T$ $(i \in \Omega^k \setminus \{j\})$ with $D_{i,i} = \sum_{l\neq i} S_{i,l}$, and $\zeta_1 = \zeta_j = 0$. Then for finite integer $k \geq 3$, $\mathcal{D}^k(S, y, \zeta, j) \geq 0$ for any S, y, ζ, j in its domain.

REMARK 1. To clarify the slight abuse of notation, we specifically discuss the following two cases: If $j = 2$, $\mathbf{1}$ is right behind y, i.e., there is no W_2 in the matrix in (72). If $j = k$, $\mathbf{1}$ is the last column and there is no W_k.

REMARK 2. Let $B = I + \lambda L$, then $W_i = B_i + \lambda\zeta_i e_i$, where B_i is the ith column of B, meaning that W_i is obtained by simply adding some non-negative number to the ith component of B_i.

LEMMA 10. *Given an aligned graph, suppose $k = 3, 4, 5, ...$, then for $j = 2, ..., k$, $\arg\max_{u\neq j} \frac{z_u - z_j}{Q_{j,j} - Q_{j,u}} = 1$, and for $j = 1$, $\arg\max_{u\neq 1} \frac{z_u - z_1}{Q_{1,1} - Q_{1,u}} = 2$.*

Lemma 8 gives the asymptotic sampling ratios using SIMA for general similarity graph. Specifically, when an aligned graph is used, with Lemmas 9 and 10, c_j defined in (71) can be computed explicitly by

$$c_j = \begin{cases} \frac{z_1 - z_2}{\sigma_1(Q_{1,1} - Q_{1,2})}, & j = 1 \\ \frac{z_1 - z_j}{\sigma_j(Q_{j,j} - Q_{j,1})}, & j \neq 1 \end{cases}.$$

Then the results in Theorem 7 follow easily.

B.8. Proof of Proposition 3

We first argue that, under any allocation method, both $\hat{r}^n$ and $\tilde{r}^n$ will be updated finitely many times. First, we note that, under any policy, the estimated means y^n converge to a limit $\hat{y}$: if alternative i is measured infinitely often, then $\hat{y}_i = \mu_i$, otherwise y_i^n will be updated finitely many times and thus converges trivially. Then, $\hat{r}^n$ will converge to a limit $\hat{r}$ that corresponds to a ranking of the limiting values of y^n. We have $\hat{y}_i \neq \hat{y}_j$ for $i \neq j$ because the true values are distinct, and for alternatives i that are measured finitely many times, the limiting values $\hat{y}_i$ depend on normally distributed noise. Then, there exists $n_0 > 0$ such that $\hat{r}^n = \hat{r}$ for all $n > n_0$.

Next, we argue that $\tilde{r}^n$ converges. Let Ω be the set of alternatives that are sampled infinitely often, and take $i \in \Omega$ (we know that Ω must be non-empty). Let $\{n_m\}$ be the subsequence of time stages at which i is sampled.

We first consider the case $\hat{r}_i^{n_m+1} > \tilde{r}_i^{n_m}$, which is equivalent to the inequality $\hat{r}_i > \tilde{r}_i^{n_m}$ for large enough m. Because we can further take m to be large enough that no $j \in \Omega^c$ is sampled at any time $n > n_m$, the set

$$\bar{\Omega}^{n_m+1} := \left\{ j \in \hat{\Omega}^{+,n_m+1} : \Phi\left(\frac{y_j^{n_m+1} - y_i^{n_m+1}}{\sqrt{\frac{\sigma_j^2}{N_j^{n_m+1}} + \frac{\sigma_i^2}{N_i^{n_m+1}}}} \right) < \kappa^{tol} \right\}$$

will cease to be updated (i.e., will contain the same elements) after some sufficiently large m. Moreover, if $\tilde{\Omega}^{n_m+1} \neq \emptyset$, $\tilde{\Omega}^{-,n_m+1} \subsetneq \tilde{\Omega}^{-,n_m}$ by the rank updating step (30). Therefore, eventually, $\tilde{\Omega}^{n_m+1}$ will become empty, meaning that $\tilde{r}_i$ will not be updated. The case where $\hat{r}_i < \tilde{r}_i^{n_m}$ can be handled symmetrically. Since the rank of alternatives in Ω^c can only change when there is a change in the rank of alternatives in Ω, it follows that $\tilde{r}_i^n$ will also be updated finitely many times for $i \in \Omega^c$. Therefore, for all i, the ranking $\tilde{r}_i^n$ will be updated only finitely many times.

Now suppose that all of the alternatives are sampled infinitely often. Clearly, $y^n \to \mu$ and $\hat{r}^n \to r$, where r is the true ranking of the alternatives based on the values μ. We also know that there exists $\tilde{r}$ and some n_1 such that, for any $n > n_1$, $\tilde{r}^n = \tilde{r}$. Consider an arbitrary alternative j and suppose that $\tilde{r}_j < \hat{r}_j$. Then there must exist an alternative i, such that $\tilde{r}_j < \tilde{r}_i$ and $\hat{r}_j > \hat{r}_i$. The latter inequality implies $\mu_j < \mu_i$. For all large enough $n > n_1$, we then have

$$\Phi\left(\frac{y_j^{n+1} - y_i^{n+1}}{\sqrt{\frac{\sigma_i^2}{N_i^{n+1}} + \frac{\sigma_j^2}{N_j^{n+1}}}} \right) < \kappa^{tol}.$$

By the updating step (30), we should then change the ranking so that $\tilde{r}_j^n > \tilde{r}_i^n$, contradicting the fact that $\tilde{r}^n$ is no longer updated for $n > n_1$. Therefore, the assumption that $\tilde{r}_j < \hat{r}_j$ cannot hold. Similar arguments can be made for the case where $\tilde{r}_j > \hat{r}_j$. Therefore, $\tilde{r} = r$, which means that the limit of $\tilde{S}^n$ must be aligned with respect to the true means μ.

B.9. Proof of Lemma 8

To prove Lemma 8, we basically follow Ryzhov (2016). First consider a modified version of SIMA policy, which is given by

$$\bar{\nu}_j^n = \tilde{\sigma}_j^n \phi\left(-\max_{u\neq j} \frac{z_u - z_j}{\tilde{\sigma}_j^n (Q_{j,j} - Q_{j,u})}\right), \tag{73}$$

where $z = Q(S)\mu$. Suppose N_j^n is the number of measurements of alternative j up to step n, and let $O_j^n = N_j^n(N_j^n+1)$, we have $\tilde{\sigma}_j^n = \frac{\sigma_j}{\sqrt{O_j^n}}$. By the definition of c_j in (71), $\bar{\nu}_j^n$ can be written as:

$$\bar{\nu}_j^n = \frac{\sigma_j}{\sqrt{O_j^n}} \phi\left(-c_j\sqrt{O_j^n}\right). \tag{74}$$

Construct a continuous function

$$\bar{\nu}_j(w) = \frac{\sigma_j}{\sqrt{w}} \phi\left(-c_j\sqrt{w}\right), \quad w \geq 0. \tag{75}$$

When w takes value O_j^n, (75) becomes (74). Notice that the standard normal p.d.f has the following property:

PROPOSITION 4. *Fix $c_1, c_2 \geq 0$. Then*

$$\lim_{x\to\infty} \frac{\phi(-c_1\sqrt{x})}{\phi(-c_2\sqrt{x})} = \begin{cases} \infty, & c_1 < c_2 \\ 1, & c_1 = c_2 \\ 0, & c_1 > c_2 \end{cases}. \tag{76}$$

Proof. Since

$$\lim_{x\to\infty} \frac{\phi(-c_1\sqrt{x})}{\phi(-c_2\sqrt{x})} = \lim_{x\to\infty} \frac{e^{-c_1^2 x}}{e^{-c_2^2 x}} = \lim_{x\to\infty} e^{-(c_1^2-c_2^2)x},$$

the results follow easily. □

The remainder of the proof is very similar to Ryzhov (2016), as we only need to replace $f(x) = x\Phi(x) + \phi(x)$ in that paper by $\phi(x)$.

B.10. Proof of Lemma 9

Proof. Without loss of generality, we may assume $\lambda = 1$, since for $\lambda \neq 1$, we can scale S and ζ by λ to get the results. We prove the lemma by induction. For $k = 3$, $j = 2$,

$$\mathcal{D}^3(S, y, \zeta, 2) = \begin{vmatrix} y_1 & 1 & -S_{3,1} \\ y_2 & 1 & -S_{3,2} \\ y_3 & 1 & 1 + D_{3,3} + \zeta_3 \end{vmatrix} = (y_1 - y_3)(1 + D_{3,3} + \zeta_3 + S_{3,2}) - (y_2 - y_3)(1 + D_{3,3} + \zeta_3 + S_{3,1}) \geq 0.$$

The last inequality holds because $S_{3,2} \geq S_{3,1}$ and $y_1 \geq y_2$. For $k = 3$, $j = 3$,

$$\mathcal{D}^3(S, y, \zeta, 3) = \begin{vmatrix} y_1 & -S_{2,1} & 1 \\ y_2 & 1 + D_{2,2} + \zeta_2 & 1 \\ y_3 & -S_{2,3} & 1 \end{vmatrix} = (y_1 - y_3)(1 + D_{2,2} + \zeta_2 + S_{2,3}) - (y_2 - y_3)(-S_{2,1} + S_{2,3}) \geq 0.$$

Suppose for $k=n-1$, the statement is true. Denote the matrix $[y, W_2, ..., W_{j-1}, \mathbf{1}, W_{j+1}, ..., W_n]$ by $A^{(n)}$. We have:

$$\begin{aligned} A^{(1)} =& [y, B_2, ..., B_{j-1}, \mathbf{1}, B_{j+1}, ..., B_n], \\ A^{(t)} =& A^{(t-1)} + \zeta_t e_t e_t^T, \quad t=2,3,...,n, \end{aligned}$$

where B_i, $i=2,..,j-1,j+1,...,n$, is the ith column of $B=I+L$ associated with similarity matrix S. From Theorem 2, we know there exists $z \in \mathbb{Y}^n$ such that $A^{(1)} = B[z, e_2, ..., e_{j-1}, \mathbf{1}, e_{j+1}, ..., e_n]$. Therefore $\det(A^{(1)}) = \det(B)(z_1 - z_j) \geq 0$. By the matrix determinant lemma (Ding and Zhou 2007), for $t=2,...,j-1,j+1,...,n$,

$$\det\left(A^{(t)}\right) = \det\left(A^{(t-1)}\right) + \zeta_t e_t^T \left(\text{adj}\left(A^{(t-1)}\right)\right) e_t = \det\left(A^{(t-1)}\right) + \zeta_t \left(\text{adj}\left(A^{(t-1)}\right)\right)_{t,t}, \tag{77}$$

where $\left(\text{adj}\left(A^{(t-1)}\right)\right)_{t,t}$ is the (t,t)th component of the adjoint matrix of $A^{(t-1)}$. Notice that the (t,t)th component is on the diagonal of $\text{adj}\left(A^{(t-1)}\right)$, therefore $\left(\text{adj}\left(A^{(t-1)}\right)\right)_{t,t} = \det\left(C^{(t)}\right)$, where

$$\begin{aligned} C^{(t)} &= \left[y^{(t)}, W_2^{(t)}, ..., W_{t-1}^{(t)}, B_{t+1}^{(t)}, ..., B_{j-1}^{(t)}, \mathbf{1}, B_{j+1}^{(t)}, ..., B_k^{(t)}\right] \quad \text{for } t \leq j-2, \\ C^{(t)} &= \left[y^{(t)}, W_2^{(t)}, ..., W_{t-1}^{(t)}, \mathbf{1}, B_{j+1}^{(t)}, ..., B_k^{(t)}\right] \quad \text{for } t = j-1, \\ C^{(t)} &= \left[y^{(t)}, W_2^{(t)}, ..., W_{j-1}^{(t)}, \mathbf{1}, B_{t+1}^{(t)}, ..., B_k^{(t)}\right] \quad \text{for } t = j+1, \\ C^{(t)} &= \left[y^{(t)}, W_2^{(t)}, ..., W_{j-1}^{(t)}, \mathbf{1}, W_{j+1}^{(t)}, ..., W_{t-1}^{(t)}, B_{t+1}^{(t)}, ..., B_k^{(t)}\right] \quad \text{for } t \geq j+2, \end{aligned}$$

and $y^{(t)}$, $W_i^{(t)}$, $i \leq t-1$, and $B_i^{(t)}$, $i \geq t+1$, are obtained by deleting the tth component of y, W_i and B_i, respectively. Then,

$$\begin{aligned} W_{i,i}^{(t)} =& W_{i,i} = 1 + \sum_{l \neq i}^{n} |W_{i,l}| + \zeta_i = 1 + \sum_{l \neq i}^{n-1} |W_{i,l}^{(t)}| + S_{i,t} + \zeta_i, \quad i = 2, ..., t-1, \\ B_{i,i-1}^{(t)} =& B_{i,i} = 1 + \sum_{l \neq i}^{n} |B_{i,l}| = 1 + \sum_{l \neq i-1}^{n-1} |B_{i,l}^{(t)}| + S_{i,t}, \quad i = t+1, ..., k. \end{aligned} \tag{78}$$

Here $W_{i,i}^{(t)}$ and $B_{i,i-1}^{(t)}$ are on the diagonal of $C^{(t)}$. Denote the matrix obtained by deleting the tth column and row of S by $S_{(n-1)\times(n-1)}^{(t)}$. And let $\zeta^{(t)}$ be a $n-1$ dimensional vector with $\zeta_1^{(t)} = 0$. If $t \leq j-1$, let $\zeta_{j-1}^{(t)} = 0$; otherwise, let $\zeta_j^{(t)} = 0$. All the other components of $\zeta^{(t)}$ are set to be $C_{i,i}^{(t)} - \left(1 + \sum_{l \neq i}^{n-1} C_{i,l}^{(t)}\right)$, which are always non-negative from (78). Define $j^{(t)} = j-1$ if $t \leq j-1$; $j^{(t)} = j$, otherwise. It's easy to see $(S_{(n-1)\times(n-1)}^{(t)}, y^{(t)}, \zeta^{(t)}, j^{(t)}) \in dom(\mathcal{D}^{n-1})$ and $\det(C^{(t)}) = \mathcal{D}^{n-1}\left(S_{(n-1)\times(n-1)}^{(t)}, y^{(t)}, \zeta^{(t)}, j^{(t)}\right)$. By our assumption that the statement is true for $k=n-1$, we have

$$\left(\text{adj}\left(A^{(t-1)}\right)\right)_{t,t} = \det(C^{(t)}) \geq 0.$$

From (77), we know that if $\left(\text{adj}\left(A^{(t-1)}\right)\right)_{t,t} \geq 0$ and $\det\left(A^{(t-1)}\right) \geq 0$, then $\det\left(A^{(t)}\right) \geq 0$. Since $\det\left(A^{(1)}\right) \geq 0$, by induction, we can get that for $k = n$ and any $(S, y, \zeta, j) \in dom(\mathcal{D}^n)$,

$$\mathcal{D}^n(S, y, \zeta, j) = \det\left(A^{(n)}\right) \geq 0.$$

The proof is completed. □

B.11. Proof of Lemma 10

Proof. For $j \neq 1$, if $u > j$, since $z_1 \geq z_j \geq z_u$ and $Q_{j,j} > \max\{Q_{j,1}, Q_{j,u}\}$, we can easily obtain $\frac{z_1 - z_j}{Q_{j,j} - Q_{j,1}} \geq \frac{z_u - z_j}{Q_{j,j} - Q_{j,u}}$. For $1 < u < j$, let $A = [z, e_2, ..., e_{u-1}, \mathbf{1}, e_{u+1}, ...e_{j-1}, Q_j, e_{j+1}, ..., e_k]$. Then $\frac{z_1 - z_j}{Q_{j,j} - Q_{j,1}} \geq \frac{z_u - z_j}{Q_{j,j} - Q_{j,u}}$ is equivalent to $\det(A) \geq 0$. Let $B = I + \lambda L$ and construct Υ by

$$\Upsilon = [\mu, B_2, ..., B_{u-1}, \mathbf{1}, B_{u+1}, ..., B_{j-1}, e_j, B_{j+1}, ..., B_k].$$

Since $z = Q\mu$, $\mathbf{1} = Q\mathbf{1}$ and $e_i = QB_i$, we have $A = (I + \lambda L)^{-1}\Upsilon = Q\Upsilon$. Suppose $\Upsilon^{(j)}_{(k-1)\times(k-1)} = [\mu^{(j)}, B_2^{(j)}, ..., B_{u-1}^{(j)}, \mathbf{1}, B_{u+1}^{(j)}, ..., B_{j-1}^{(j)}, B_{j+1}^{(j)}, ..., B_k^{(j)}]$ is the matrix obtained by deleting the jth column and row of Υ, then $\det(\Upsilon) = \det(\Upsilon^{(j)})$. By Lemma 9, we have $\det(\Upsilon^{(j)}) \geq 0$. Furthermore, $\det(A) = \det(Q)\det(\Upsilon) = \det(Q)\det(\Upsilon^{(j)}) \geq 0$.

For $j = 1$, let $A = [Q_1, z, e_3, ..., e_{u-1}, \mathbf{1}, e_{u+1}, ..., e_k]$. Then $\frac{z_2 - z_1}{Q_{1,1} - Q_{1,2}} \geq \frac{z_u - z_1}{Q_{1,1} - Q_{1,u}}$ is equivalent to $\det(A) \geq 0$. Now construct Υ by

$$\Upsilon = [e_1, \mu, B_3, ..., B_{u-1}, \mathbf{1}, B_{u+1}, ..., B_k],$$

Delete the first row and column of Υ, we obtain $\Upsilon^{(1)}_{(k-1)\times(k-1)} = [\mu^{(1)}, B_3^{(1)}, ..., B_{u-1}^{(1)}, \mathbf{1}, B_{u+1}^{(1)}, ..., B_k^{(1)}]$, then $\det(\Upsilon) = \det(\Upsilon^{(1)}) \geq 0$ by Lemma 9 again. Finally, we have

$$\det(A) = \det(Q)\det(\Upsilon) \geq 0.$$

The proof is completed. □

Appendix C: Experiments on Sensitivity to λ

We conduct experiments to test the sensitivity to λ of our proposed allocation algorithms combined with dynamic updating of the similarity structure for the servo system selection problem (Case 2 of Experiment 1). The results are shown in Figures 9 and 10. From those results, we can see that performances are not very sensitive to λ, especially for large budgets.

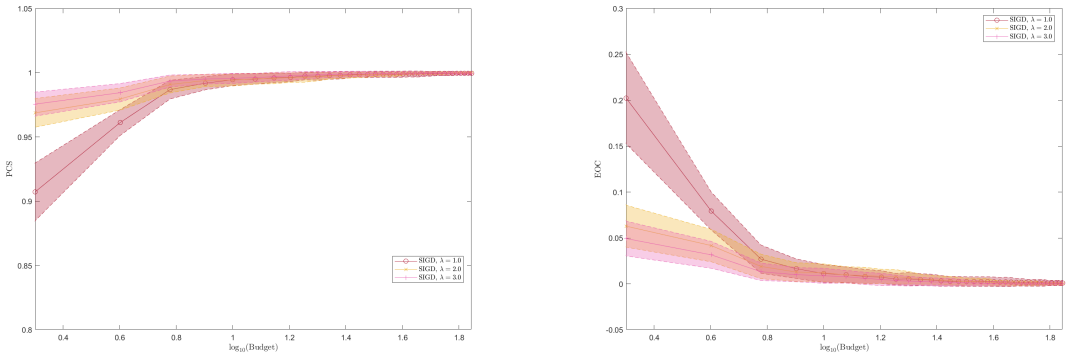

Figure 9 Case 2: Average PCS and EOC results for different values of λ using SIGD, obtained by 10000 independent runs (40 macro-replications, 250 micro-replications). Dotted curves represent standard errors.

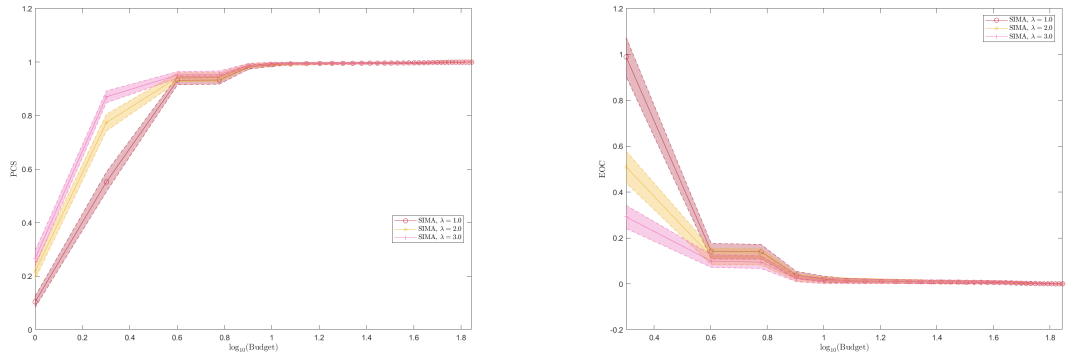

Figure 10 Case 2: Average PCS and EOC results for different values of λ using SIMA, obtained by 10000 independent runs (40 macro-replications, 250 micro-replications). Dotted curves represent standard errors.

Appendix D: Handling Alternatives with Equal Values

In this section, we consider the case where there exist alternatives sharing the same value, i.e., $\mu_i = \mu_j$ for some $i \neq j$. For this special case, the definition of an aligned graph is adapted as follows.

Definition 2. A similarity graph S is aligned if, for any $i = 1, \ldots, k$, we have $S_{i,j} \leq S_{i,m}$ for $j < m < \min\{l \leq i : \mu_l = \mu_i\}$ as well as all $j > m > \max\{l \geq i : \mu_l = \mu_i\}$.

Essentially, Definition 2 ensures that the similarity relationships satisfy the monotonicity requirement of Definition 1 for all alternatives whose values are not tied. When there are multiple alternatives with the same value, their similarity relationships can have any arbitrary ordering. The result of Theorem 2 is preserved under Definition 2 if we neglect the relative order of S-indices between alternatives that have the same true mean. Thus, when y^n converges to μ, we can always select the best alternative i^* by the S-indices , such that $\mu_{i^*} = \max_{i=1,\ldots,k} \mu_i$.

Moreover, in the dynamic updating strategy of the similarity matrix S^n, to prevent permuting S infinitely many times, we can set $\hat{r}_i^n < \hat{r}_j^n$ only when $y_i^n - y_j^n > C_d\sqrt{\frac{\log n}{n}}$ for some fixed $C_d > 0$.

By the law of the iterated logarithm, for any sample path, the similarity matrix will be updated finitely many times. Consequently, we will obtain an aligned graph in the sense of Definition 2 after finitely many time steps, on the condition that each alternative is sampled infinitely often on that same sample path.