SCORE Allocations for Bi-objective Ranking and Selection

GUY FELDMAN and SUSAN R. HUNTER, Purdue University

The bi-objective ranking and selection (R&S) problem is a special case of the multi-objective simulation optimization problem in which two conflicting objectives are known only through dependent Monte Carlo estimators, the decision space or number of systems is finite, and each system can be sampled to some extent. The solution to the bi-objective R&S problem is a set of systems with non-dominated objective vectors, called the set of Pareto systems. We exploit the special structure of the bi-objective problem to characterize the asymptotically optimal simulation budget allocation, which accounts for dependence between the objectives and balances the probabilities associated with two types of misclassification error. Like much of the R&S literature, our focus is on the case in which the simulation observations are bivariate normal. Assuming normality, we then use a certain asymptotic limit to derive an easily-implementable SCORE (Sampling Criteria for Optimization using Rate Estimators) sampling framework that approximates the optimal allocation and accounts for correlation between the objectives. Perhaps surprisingly, the limiting SCORE allocation exclusively controls for misclassification-by-inclusion events, in which non-Pareto systems are falsely estimated as Pareto. We also provide an iterative algorithm for implementation. Our numerical experience with the resulting SCORE framework indicates that it is fast and accurate for problems having up to ten thousand systems.

CCS Concepts: • **Mathematics of computing → Probability and statistics**; **Probabilistic algorithms**; • **Theory of computation → Discrete optimization**; • **Computing methodologies → Simulation theory**; • **Applied computing → Multi-criterion optimization and decision-making**;

Additional Key Words and Phrases: multi-objective simulation optimization, ranking and selection

ACM Reference Format:
Guy Feldman and Susan R. Hunter. 201X. SCORE Allocations for Bi-objective Ranking and Selection. *ACM Trans. Model. Comput. Simul.* 0, 0, Article 0 (201X), 50 pages. https://doi.org/0000001.0000001

1 INTRODUCTION

The simulation optimization (SO) problem is a nonlinear optimization problem in which the objective and constraint functions can only be observed with error as output from a Monte Carlo simulation. Such problems tend to arise when computer models are used to design complex systems under uncertainty — an increasingly popular practice [Powers et al. 2012]. Since the SO formulation is quite general, SO problems arise frequently in a variety of application areas, including agriculture [Hunter and McClosky 2016], energy [Marmidis et al. 2008; Subramanyan et al. 2011], and transportation [Osorio and Bierlaire 2013]. For additional examples and a library of SO problems, see the `simopt.org` website [Henderson and Pasupathy 2017].

Methods to solve the SO problem are often categorized by whether the feasible set contains categorical, integer-ordered, or continuous decision variables [Pasupathy and Henderson 2006]. Further, solution methods can be categorized by the number of performance measures posed as objectives and constraints. In the presence of a single objective and deterministic constraints, mature

Guy Feldman, Department of Statistics, Purdue University, 250 University St., West Lafayette, IN 47907; Susan R. Hunter, School of Industrial Engineering, Purdue University, 315 N. Grant St., West Lafayette, IN 47907, susanhunter@purdue.edu.

XXXX-XXXX/201X/0-ART0 $15.00
https://doi.org/0000001.0000001

solution methods are available for all types of feasible sets. For an overview of these methods and entry points into this literature, see Pasupathy and Ghosh [2013] and Fu [2015]. Recently, solution methods for a single objective with stochastic constraints have been proposed in the case of categorical variables [Andradóttir and Kim 2010; Lee et al. 2012; Pasupathy et al. 2015] and integer-ordered variables [Luo and Lim 2013; Nagaraj and Pasupathy 2016; Park and Kim 2015]. For methods with continuous variables, see, e.g., Ruszczyński and Shapiro [2003], Homem-de-Mello and Bayraksan [2015], and references therein. However, despite its mature development in the analogous deterministic context [Miettinen 1999, for example], few papers in the SO literature provide solution methods in the presence of multiple simultaneous objectives — a problem we call the multi-objective simulation optimization (MOSO) problem.

We formulate the MOSO problem as

$$\text{Problem } M: \quad \text{minimize}_{\mathbf{x}\in\mathcal{X}} \quad (\mathbf{E}[G_1(\mathbf{x},\xi)],\ldots,\mathbf{E}[G_d(\mathbf{x},\xi)]),$$

where $\mathcal{X} \subseteq \mathbb{R}^q$ is a known feasible set, ξ represents a random quantity, and each objective can be estimated as output from a Monte Carlo simulation. Since there may not exist a single point $\mathbf{x} \in \mathcal{X}$ that minimizes all objectives simultaneously, the solution to Problem M is called the efficient set or the Pareto set. We let the *efficient set* be the set of decision points $\mathbf{x} \in \mathcal{X}$ for which no other point $\mathbf{x}' \in \mathcal{X}, \mathbf{x} \neq \mathbf{x}'$ has objective values that are at least as good on all objectives, and strictly better on at least one objective. We refer to the image of the efficient set as the *Pareto set*.

We consider the context of solving Problem M when the goal is to identify the entire efficient set, the feasible set $\mathcal{X}$ is finite or comprised of categorical variables, and there are two objectives. Methods to solve SO problems in which $\mathcal{X}$ is finite are often called *ranking and selection* (R&S) methods (see Kim and Nelson 2006 for an overview). Such methods require the feasible set to be small enough to permit simulation of each decision point; the decision points are usually indexed by their objective values and called *systems*. Henceforth, we refer to systems with objective vectors in the Pareto set as *Pareto systems* (see §2.2 for terminology). R&S methods can be divided into two types: methods that provide a fixed-precision guarantee on the optimality gap of the returned systems, and methods that allocate a fixed simulation budget in a way that guarantees sampling efficiency [Hunter and Nelson 2017; Pasupathy and Ghosh 2013]. We fall in the latter category of fixed simulation budget methods; as such, we do not provide a fixed-precision guarantee on the optimality gap of the returned systems.

1.1 Questions Answered

To explore what we mean by allocating a fixed simulation budget in a way that guarantees sampling efficiency, consider a simple algorithm to solve Problem M: (a) allocate some non-zero proportion of a total sampling budget to each system, (b) sample and construct estimators of the objective vectors for each system, (c) return the indices of systems corresponding to the estimated Pareto set. Ideally, the estimated Pareto systems at the end of this procedure will correspond to the true Pareto systems; if not, a misclassification occurs. Under mild regularity conditions, as the total sampling budget tends to infinity, the probability of a misclassification decays to zero. Then we ask, *what proportion of the total sampling budget should be allocated to each system to maximize the rate of decay of the probability of misclassification, as the sampling budget tends to infinity?*

As may be expected given prior work in other SO contexts, notably, Glynn and Juneja [2004], Szechtman and Yücesan [2008], Hunter and Pasupathy [2013], and Pasupathy et al. [2015], we characterize the asymptotically optimal sampling allocation as the solution to a bi-level optimization problem where the "outer" problem is concave maximization, and the "inner" problems are convex minimization. Importantly, our allocation accounts for dependence between the objectives and balances the probabilities associated with two types of misclassification error: misclassification by exclusion (MCE), in which a Pareto system is falsely excluded from the set of estimated Pareto systems, and misclassification by inclusion (MCI), in which a non-Pareto system is falsely included in the set of estimated Pareto systems. Since solving for the optimal allocation may be computationally burdensome, then we ask, *what is the asymptotically optimal sampling allocation when the number of non-Pareto systems is large?*

As the number of non-Pareto systems tends to infinity in a certain rigorous sense, the Pareto systems receive a larger proportion of the sampling budget than the non-Pareto systems, and the optimal allocations for non-Pareto systems are inversely proportional to an intuitive measure called the score. When the random vectors corresponding to the objectives are bivariate normal, which is our focus, the *score* of a non-Pareto system is its squared standardized "distance" from the Pareto frontier in the objective function space. As in Pasupathy et al. [2015], we determine the relative allocations to the suboptimal systems by their scores. The sampling allocation we propose based on the scores is called the bi-objective Sampling Criteria for Optimization using Rate Estimators (SCORE) allocation.

We also highlight a key insight from this work that may be surprising: *when the number of non-Pareto systems is large relative to the number of Pareto systems, the optimal allocation exclusively controls for the probability of an MCI event.* To understand why this is true, for now, let MCE be the event that a Pareto system is falsely excluded by another Pareto system, while MCI is the event that a non-Pareto system is falsely included among the estimated Pareto systems, whether it is estimated as excluding a Pareto system or not. Further, assume we are in a theoretical framework in which we have access to all information about the systems, including their classifications as Pareto and non-Pareto. Then loosely speaking, as the number of non-Pareto systems tends to infinity, the Pareto systems each compete with more and more non-Pareto systems. Thus the Pareto systems receive a larger proportion of the total sampling budget than the non-Pareto systems. Indeed, they receive so many more samples that the probability of a Pareto system falsely excluding another Pareto system is small relative to the probability of a non-Pareto system being falsely included among the estimated Pareto systems. Thus the Pareto set appears "known" relative to the non-Pareto set, and the optimal allocation exclusively controls for MCI events.

Since the SCORE allocation requires knowing the true system performances on both objectives, which are unknown, we include a sequential sampling framework for implementation. We numerically compare the performance of the SCORE allocation and the sequential sampling framework with other popular allocations in the literature. We find that our implementation of the SCORE allocation performs well numerically. SCORE appears to be a fast and accurate heuristic allocation scheme for bi-objective R&S with 20 to $10,000$ systems, inspired by theoretical allocations that have limiting optimality guarantees on efficiency.

1.2 Other Relevant Work

When the goal of solving Problem M is to identify the entire efficient or Pareto set, few solution methods have been proposed in the SO literature. Arguably, the most well-known and popular method is the Multi-objective Optimal Computing Budget Allocation (MOCBA) [Lee et al. 2010], which is a multi-objective version of the popular Optimal Computing Budget Allocation [Chen et al. 2000] for a finite feasible set. Other recent work includes (a) M-MOBA [Branke and Zhang 2015; Branke et al. 2016], a multi-objective version of the small-sample expected value of information (EVI) procedures in Chick et al. [2010] for a finite feasible set; (b) MO-COMPASS [Li et al. 2015], which is a multi-objective version of COMPASS [Xu et al. 2010] for integer-ordered feasible sets; (c) Huang and Zabinsky [2014], who provide a branch-and-bound algorithm for integer-ordered or continuous feasible sets; and (d) Kim and Ryu [2011], Fliege and Xu [2011], and Bonnel and Collonge [2014], who provide methods for continuous feasible sets. We note that Butler et al. 2001 provide a utility function approach to multi-objective R&S and refer the reader to Hunter et al. [2017] for an overview of the existing MOSO literature.

Given our context of finite feasible sets, the most appropriate competitors for our proposed sequential algorithm are MOCBA and M-MOBA. We compare the performance of our sequential algorithm with these methods in Section 8.

Remark 1. A preliminary version of this work appears in Hunter and Feldman [2015]. Also, Hunter and McClosky [2016] contains an asymptotically optimal allocation for the case of two independent objectives in the context of a plant breeding application. This paper is a significant outgrowth of Hunter and Feldman [2015] and subsumes the allocation provided in Hunter and McClosky [2016] for independent objectives. Neither Hunter and Feldman [2015] nor Hunter and McClosky [2016] provides a limiting SCORE framework. Feldman et al. [2015] provides analogous MOSO methods on finite sets for more than two objectives. Since the methods in Feldman et al. [2015] are more computationally burdensome than those we propose, we do not advocate using the methods of Feldman et al. [2015] in the bi-objective case. Thus we do not include these methods in numerical comparisons. Table 1 categorizes these papers according to some of their differences.

Table 1. The table provides a categorization of existing papers on multi-objective R&S by how many objectives they consider and whether they account for dependence between the objectives.

Dependence?	Exactly Two Objectives	Two or More Objectives
No	Hunter and McClosky [2016] and M-MOBA [Branke and Zhang 2015]	MOCBA [Lee et al. 2010]
Yes	This paper and its preliminary version, Hunter and Feldman [2015]	Feldman et al. [2015]

Remark 2. There is also work on the multi-objective multi-armed bandit problem, for example Yahyaa et al. [2014a], Yahyaa et al. [2014b], and Yahyaa et al. [2014c].

2 PROBLEM SETTING AND FORMULATION

We now provide a formal problem statement, describe terminology and notational conventions, and outline our assumptions. Due to space constraints, unless otherwise noted in the text, proofs for all results appear in the Online Appendix.

2.1 Problem Statement

We consider the MOSO problem with two objectives on a finite set. That is, we solve

$$\text{Problem } B: \quad \text{Find } \operatorname{argmin}_{k\in\mathcal{S}} \quad (\mathbf{E}[G(\mathbf{x}_k,\xi)], \mathbf{E}[H(\mathbf{x}_k,\xi)]),$$

where $\mathcal{S} := \{1, 2, \ldots, r\}$ is a finite set of system indices and ξ is a random quantity. Further, define $g_k := \mathbf{E}[G(\mathbf{x}_k,\xi)]$ and $h_k := \mathbf{E}[H(\mathbf{x}_k,\xi)]$ for all $k \in \mathcal{S}$. The objective vectors $(g_k, h_k) \in \mathbb{R}^2$ are unknown, but may be estimated by sample means. The solution to Problem B is the set of Pareto systems (see §2.2), $\mathcal{P} := \{\text{systems } i : \nexists \text{ system } k \in \mathcal{S} \text{ such that } (g_k, h_k) \leq (g_i, h_i)\}$, where a vector (g_k, h_k) dominates (g_i, h_i), written as $(g_k, h_k) \leq (g_i, h_i)$, if $g_k \leq g_i$ and $h_k < h_i$, or $g_k < g_i$ and $h_k \leq h_i$.

Now consider a method to solve Problem B in which we allocate a proportion $\alpha_k > 0$ of the total sampling budget to each system k, where $\sum_{k=1}^{r} \alpha_k = 1$. Once the total sampling budget has been expended, we return the set of estimated Pareto systems, $\hat{\mathcal{P}}$, constructed as follows. Let the vector of sample means after n samples be $(\bar{G}_k(n), \bar{H}_k(n)) := (\frac{1}{n}\sum_{j=1}^{n} G_{kj}, \frac{1}{n}\sum_{j=1}^{n} H_{kj})$ for all $k \in \mathcal{S}$, and define $(\hat{G}_k, \hat{H}_k) := (\bar{G}(\alpha_k n), \bar{H}(\alpha_k n))$ as the estimators of g_k and h_k after scaling the total sample size n by the proportional allocation to system k, $\alpha_k > 0$. Then the set of estimated Pareto systems is $\hat{\mathcal{P}} := \{\text{systems } i : \nexists \text{ system } k \in \mathcal{S} \text{ such that } (\hat{G}_k, \hat{H}_k) \leq (\hat{G}_i, \hat{H}_i)\}$.

If $\hat{\mathcal{P}} \neq \mathcal{P}$, then at least one system has been misclassified, that is, a Pareto system has been falsely estimated as non-Pareto, or a non-Pareto system has been falsely estimated as Pareto. As the sampling budget tends to infinity, the probability of misclassification tends to zero. Then we ask, what sampling budget $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_r)$ maximizes the rate of decay of the probability of misclassification?

Remark 3. While we focus on allocating the sample to maximize the rate of decay of the probability of misclassification, one could also allocate to minimize the expected number of misclassifications. Hunter and McClosky [2016] show that these two objectives result in identical asymptotic allocations when the objective estimators are independent. We anticipate that a similar result holds in the context of this paper.

2.2 Terminology and Notational Conventions

In general, we prefer to call the solution to Problem M "the efficient set" and call its image "the Pareto set" (see Ehrgott 2012 for a historical perspective on these terms). However, since R&S methods assume no structure in the decision space, we may work almost entirely in the objective space. Since we work in the objective space and index the systems by their objective function values in §3, we omit the term "efficient" in favor of the term "Pareto" throughout the paper.

When it is reasonable to do so, uppercase letters denote random variables or matrices, lowercase letters denote fixed quantities, script letters denote sets, and vectors are written in bold. For a set $\mathcal{C}$, the cardinality of $\mathcal{C}$ is denoted $|\mathcal{C}|$. For a function f, let $\nabla f(\mathbf{x})$ be the gradient of f with respect to $\mathbf{x} \in \mathbb{R}^q$, and $f'(x)$ the derivative of f with respect to $x \in \mathbb{R}$. For any 2-by-2 matrix A, let the eigenvalues of A be $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$. For any n-by-n matrices A and B, let $A \circ B$ denote their element-wise, or Hadamard, product. For a sequence of real numbers $\{a_n\}$, we say that $a_n = o(1)$ if $\lim_{n\to\infty}\{a_n\} = 0$ and $a_n = O(1)$ if $\{a_n\}$ is bounded, that is, if there exists $c > 0$ with $|a_n| < c$ for all n. Further, $a_n = \Theta(1)$ if $0 < \liminf a_n \leq \limsup a_n < \infty$. We use iff for "if and only if." Solutions to optimization problems are usually denoted with an asterisk, e.g., x^*. We use $\mathbb{I}_{[\cdot]}$ to denote the indicator function. Let $f^+ : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be a function such that $f^+(x) = x$ if $x \in \{x \in \mathbb{R} : x > 0\}$ and $f^+(x) = \infty$ if $x \leq 0$. For $a, b \in \mathbb{R}$, define $\min^+[a, b] := \min(f^+(a), f^+(b))$.

2.3 Assumptions

In what follows, we assume that the set of non-Pareto systems is nonempty. To estimate the unknown quantities g_k and h_k, we assume we obtain replicates of the random vector (G_k, H_k) from each system. We also assume the following.

ASSUMPTION 1. *Random vectors (G_k, H_k) are mutually independent for all $k \in \mathcal{S}$.*

That is, we develop a model to guide sampling that does not specifically account for correlation between systems, such as the correlation that would arise with the use of common random numbers (CRN). However, our model does not preclude the use of CRN during implementation. We also require the following technical assumption which is standard in optimal allocation literature, since it ensures all Pareto systems are distinguishable on each objective with a finite sample size.

ASSUMPTION 2. *We assume $g_i \neq g_k$ and $h_i \neq h_k$ for all $i \in \mathcal{P}$, $k \in \mathcal{S}$ such that $k \neq i$.*

Since we employ a large deviations analysis in Section 3, we require the following Assumptions 3 and 4, included here for completeness. We refer the reader to Dembo and Zeitouni [1998] for further explanation. Let $\langle \cdot, \cdot \rangle$ denote the dot product, and let $\Lambda^{(n)}_{(G_k,H_k)}(\boldsymbol{\theta}) = \log \mathbf{E}[e^{\langle \boldsymbol{\theta}, (\bar{G}_k(n), \bar{H}_k(n)) \rangle}]$, $\Lambda^{(n)}_{G_k}(\theta) = \log \mathbf{E}[e^{\theta \bar{G}_k(n)}]$, and $\Lambda^{(n)}_{H_k}(\theta) = \log \mathbf{E}[e^{\theta \bar{H}_k(n)}]$, be the cumulant generating functions of $(\bar{G}_k(n), \bar{H}_k(n))$, $\bar{G}_k(n)$, and $\bar{H}_k(n)$, respectively, where $\boldsymbol{\theta} \in \mathbb{R}^2$ and $\theta \in \mathbb{R}$. Let the effective domain of $f(\cdot)$ be $\mathcal{D}_f = \{x : f(x) < \infty\}$, and its interior $\mathcal{D}^\circ_f$. We make the following standard assumption.

Assumption 3. *For each system $k \in \mathcal{S}$,*

(1) *the limit $\Lambda_{(G_k,H_k)}(\boldsymbol{\theta}) = \lim_{n\to\infty} \frac{1}{n} \Lambda^{(n)}_{(G_k,H_k)}(n\boldsymbol{\theta})$ exists as an extended real number for all $\boldsymbol{\theta} \in \mathbb{R}^2$, where $\Lambda_{G_k}(\theta) = \lim_{n\to\infty} \frac{1}{n} \Lambda^{(n)}_{G_k}(n\theta)$ and $\Lambda_{H_k}(\theta) = \lim_{n\to\infty} \frac{1}{n} \Lambda^{(n)}_{H_k}(n\theta)$ for all $\theta \in \mathbb{R}$;*
(2) *the origin belongs to the interior of $\mathcal{D}_{\Lambda_{(G_k,H_k)}}$;*
(3) *$\Lambda_{(G_k,H_k)}(\boldsymbol{\theta})$ is strictly convex and C^∞ on $\mathcal{D}^\circ_{\Lambda_{(G_k,H_k)}}$;*
(4) *$\Lambda_{(G_k,H_k)}(\boldsymbol{\theta})$ is steep, that is, for any sequence $\{\boldsymbol{\theta}_n\} \in \mathcal{D}_{\Lambda_{(G_k,H_k)}}$ converging to a boundary point of $\mathcal{D}_{\Lambda_{(G_k,H_k)}}$, $\lim_{n\to\infty} |\nabla \Lambda_{(G_k,H_k)}(\theta_n)| = \infty$.*

Assumption 3 implies that by the Gärtner-Ellis theorem, the probability measure governing $(\bar{G}_k(n), \bar{H}_k(n))$ obeys the large deviations principle (LDP) with good, strictly convex rate function $I_k(x,y) = \sup_{\boldsymbol{\theta}\in\mathbb{R}^2}\{\langle \boldsymbol{\theta}, (x,y)\rangle - \Lambda_{(G_k,H_k)}(\boldsymbol{\theta})\}$ [Dembo and Zeitouni 1998, p. 44]. Further, $\bar{G}_k(n)$ and $\bar{H}_k(n)$ obey the LDP with good, strictly convex rate functions $J_k(x) = \sup_{\theta\in\mathbb{R}}\{\theta x - \Lambda_{G_k}(\theta)\}$, $K_k(y) = \sup_{\theta\in\mathbb{R}}\{\theta y - \Lambda_{H_k}(\theta)\}$. Let $(x,y) \in \mathcal{F}^\circ_{(G_k,H_k)} = \text{int}\{\nabla \Lambda_{(G_k,H_k)}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \mathcal{D}^\circ_{\Lambda_{(G_k,H_k)}}\}$, and let $\mathcal{F}^c_d$ denote the closure of the convex hull of the set $\{(g_k, h_k) : (g_k, h_k) \in \mathbb{R}^2, k \in \mathcal{S}\}$.

Assumption 4. *The closure of the convex hull of all points $(g_k, h_k) \in \mathbb{R}^2$ is a subset of the intersection of the interiors of the effective domains of the rate functions $I_k(x,y)$ for all $k \in \mathcal{S}$, that is, $\mathcal{F}^c_d \subset \cap_{k=1}^r \mathcal{F}^\circ_{(G_k,H_k)}$.*

3 CHARACTERIZATION OF THE OPTIMAL BUDGET ALLOCATION

Given that our goal is to determine the sample allocation $\boldsymbol{\alpha}$ that maximizes the rate of decay of the probability of misclassification, we first formulate the misclassification event in a way that facilitates analysis. We then analyze the rate of decay of the probability of misclassification as a function of $\boldsymbol{\alpha}$, and provide a characterization of the optimal budget allocation as the solution to a bi-level optimization problem. To avoid mathematical complications, we assume $n\alpha_k > 1$ for all $k \in \mathcal{S}$ in this section.

3.1 Formulation of the Misclassification Event

Recall that a misclassification event occurs if, after expending a total of n samples, the set of estimated Pareto systems, $\hat{\mathcal{P}}$, is not equal to the true set of Pareto systems, $\mathcal{P}$. If $\hat{\mathcal{P}} \neq \mathcal{P}$, then at least one of two events occurs: a Pareto system was falsely excluded from the set of estimated Pareto systems (MCE), or a non-Pareto system was falsely included in the set of estimated Pareto systems (MCI). Therefore we can formulate the misclassification event as MC := MCE ∪ MCI, where

$$\text{MCE} := \underbrace{\bigcup_{i\in\mathcal{P}} \bigcup_{k\in\mathcal{S}, k\neq i} (\hat{G}_k \le \hat{G}_i) \cap (\hat{H}_k \le \hat{H}_i)}_{\exists\, i \in \mathcal{P} \text{ dominated by some } k \in \mathcal{S}}; \quad \text{MCI} := \underbrace{\bigcup_{j\in\mathcal{P}^c} \bigcap_{k\in\mathcal{S}, k\neq j} (\hat{G}_j \le \hat{G}_k) \cup (\hat{H}_j \le \hat{H}_k)}_{\exists\, j \in \mathcal{P}^c \text{ not dominated by any } k \in \mathcal{S}}.$$

As the union of pairwise exclusion events, the MCE event is easy to analyze. However, the MCI event requires a non-Pareto system j to be falsely estimated as better than *every* system k on at least one objective. This event contains dependence that is difficult to analyze. In this section, we reformulate the MC event for easier analysis. First, we simplify the MCE event to consider only exclusion events between Pareto systems,

$$\text{MCE}_{\mathcal{P}} := \bigcup_{i\in\mathcal{P}} \bigcup_{i'\in\mathcal{P},\ i'\neq i} (\hat{G}_{i'} \leq \hat{G}_i) \cap (\hat{H}_{i'} \leq \hat{H}_i).$$

Then, we reformulate the MCI event so that it also appears as a union of pairwise exclusion events. We combine these results into the statement of Theorem 3.1, which states the equivalence of the MC and reformulated events.

To reformulate the MCI event as a union of pairwise exclusion events, first, we define new systems called phantom Pareto systems. To define these systems, reserve the indices $1, \ldots, p$ for the Pareto systems, such that $\mathcal{P} = \{1, \ldots, p\}, |\mathcal{P}| = p$. Then label the true Pareto systems by their ordered objective values, $g_1 < g_2 < \ldots < g_{p-1} < g_p < g_{p+1}$ and $h_0 > h_1 > h_2 > \ldots > h_{p-1} > h_p$, where $g_{p+1} := \infty$, and $h_0 := \infty$. Thus the objective values for the true Pareto systems are (g_i, h_i) for all $i \in \mathcal{P}$, where henceforth, we use i as an ordered index when we wish to refer to the ordered *Pareto systems*. An example of this labeling appears in Figure 1. Now construct the objective value corresponding to the ℓth *phantom Pareto system* as the coordinates (g_{i+1}, h_i) for $i = 0, 1, \ldots, p$, where we also place phantom Pareto systems at (g_1, ∞) and (∞, h_p) for a total of $p+1$ phantom Pareto systems. Henceforth, we use $\mathcal{P}^{\text{ph}} := \{0, 1, \ldots, p\}$ as the set of indices corresponding to the phantom Pareto systems, and we use ℓ as an ordered index when we wish to emphasize the ordered *phantom Pareto systems*; notice our labeling is such that $\ell = i$ (see Figure 1). *For the remainder of the paper, the indices ℓ and i are linked in this way.*

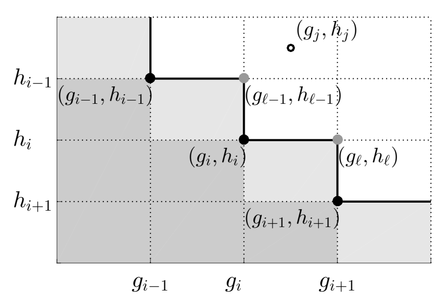

Fig. 1. Pareto systems $i-1$, i, $i+1$ are solid black dots, $i \in \{2, \ldots, p-1\}$. Phantom Pareto systems are solid gray dots, $\ell = i$. If the Pareto set were known, an MCE or MCI event would result from the non-Pareto system j being falsely estimated in the dark or light gray region, respectively.

To rewrite the MCI event using the phantom Pareto systems, we must estimate the objective values of the phantom Pareto systems. For the true Pareto systems, define $\hat{G}_{[i]}$ as the ith largest estimated first objective value and $\hat{H}_{[i]}$ as the ith smallest estimated second objective value. Thus $\hat{G}_{[1]} < \ldots < \hat{G}_{[p-1]} < \hat{G}_{[p]} < \hat{G}_{[p+1]}$ and $\hat{H}_{[0]} > \hat{H}_{[1]} > \hat{H}_{[2]} > \ldots > \hat{H}_{[p]}$, where $\hat{G}_{[p+1]} := \infty$ and $\hat{H}_{[0]} := \infty$ for all n. Now the estimated performances of the phantom Pareto systems are $(\hat{G}_{[i+1]}, \hat{H}_{[i]})$ for $i = 0, 1, \ldots, p$. Define misclassification by dominating an estimated phantom Pareto system as

$$\text{MCI}_{\text{ph}} := \bigcup_{j\in\mathcal{P}^c} \bigcup_{\ell\in\mathcal{P}^{\text{ph}},\ell=i} (\hat{G}_j \leq \hat{G}_{[i+1]}) \cap (\hat{H}_j \leq \hat{H}_{[i]}),$$

and rewrite the misclassification event as $\text{MC}_{\text{ph}} := \text{MCE}_{\mathcal{P}} \cup \text{MCI}_{\text{ph}}$. Theorem 3.1 states the equivalence of the probability of an MC event and the probability of an MC_{ph} event. A similar theorem was stated and proved in Hunter and McClosky [2016] under more restrictive assumptions.

Theorem 3.1. $\mathbf{P}\{\text{MC}\} = \mathbf{P}\{\text{MC}_{\text{ph}}\}$.

Henceforth, we use the notation $\mathbf{P}\{\text{MC}\}$ without loss of clarity.

3.2 The Rate of Decay of the Probability of a Misclassification Event

Now that we have formulated the MC event into the union of two MCE-like events, we are ready to analyze the rate of decay of $\mathbf{P}\{\text{MC}\}$ as a function of the sampling allocation vector $\boldsymbol{\alpha}$. Notice that for $b = \max(\mathbf{P}\{\text{MCE}_{\mathcal{P}}\}, \mathbf{P}\{\text{MCI}_{\text{ph}}\})$, we have $b \leq \mathbf{P}\{\text{MC}\} \leq 2b$, which, assuming the limits exist, implies

$$-\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MC}\} = \min\big(-\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}, -\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MCI}_{\mathrm{ph}}\}\big). \quad (1)$$

In what follows, we analyze the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$ and $\mathbf{P}\{\mathrm{MCI}_{\mathrm{ph}}\}$ separately.

First, consider the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$ in equation (1), since it is the most straightforward. For brevity, for all $i, i' \in \mathcal{P}$, $i' \neq i$, define the rate function

$$R_i(\alpha_i, \alpha_{i'}) := \inf_{x_{i'} \le x_i,\ y_{i'} \le y_i} \alpha_i I_i(x_i, y_i) + \alpha_{i'} I_k(x_{i'}, y_{i'}).$$

The following Lemma 3.2 states the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$ in terms of the pairwise rates $R_i(\alpha_i, \alpha_{i'})$ corresponding to a Pareto system i' dominating another Pareto system i. We do not provide a proof for Lemma 3.2; it follows from an analysis similar to those in Glynn and Juneja [2004], Hunter [2011], Li [2012], Feldman [2017].

Lemma 3.2. *The rate of decay of* $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$ *is*

$$-\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\} = \min_{i\in\mathcal{P}}\ \min_{i'\in\mathcal{P},\, i'\neq i} R_i(\alpha_i, \alpha_{i'}).$$

Lemma 3.2 states that the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$ is the slowest among the pairwise rates corresponding to one Pareto system falsely dominating another.

Now let us turn our attention to the term corresponding to $\mathrm{MCI}_{\mathrm{ph}}$ in (1). The analysis for the rate of decay of the probability of an $\mathrm{MCI}_{\mathrm{ph}}$ event is a bit more involved: in addition to the possibility that a non-Pareto system j is estimated as dominating a phantom Pareto system, the Pareto systems themselves may be estimated as "out of order." In what follows, we do not directly state the rate of decay of $\mathbf{P}\{\mathrm{MCI}_{\mathrm{ph}}\}$. Instead, we show that the probability of events corresponding to $\mathrm{MCI}_{\mathrm{ph}}$ in which the Pareto systems are also estimated as out of order have rates of decay greater than or equal to the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$, and thus can never be the unique minimum rate in (1).

To explicitly denote the ordering of the Pareto systems, we require the following notation. First, recall that the Pareto systems are labeled "in order" from $1, 2 \ldots, p$. Then we define the ordered list $\mathcal{O} := \{(1,1), (2,2), \ldots, (p,p)\}$ as the positions of the true Pareto systems on each objective, where the first objective is labeled from smallest to largest, and the second objective is labeled from largest to smallest. Now define $\hat{\mathcal{O}}$ as the ordered list of estimated positions of the true Pareto systems. Thus the event that the Pareto systems are estimated in the correct order is $\hat{\mathcal{O}} = \mathcal{O}$. Define $\mathrm{MCI}_{\mathrm{ph}}$ *without* order statistics as $\mathrm{MCI}^*_{\mathrm{ph}} := \cup_{j\in\mathcal{P}^c} \cup_{\ell\in\mathcal{P}^{\mathrm{ph}},\, \ell=i} (\hat{G}_j \le \hat{G}_{i+1}) \cap (\hat{H}_j \le \hat{H}_i)$, where $\hat{G}_{p+1} := \infty, \hat{H}_0 := \infty$ for all n. Then the event $\mathrm{MCI}^*_{\mathrm{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}$ is the event that the Pareto systems are estimated in order, and a non-Pareto system is falsely included in the set of estimated Pareto systems. The following lemma states that only the rate of decay of the probability of this event can be a binding minimum in the overall rate of decay of $\mathbf{P}\{\mathrm{MC}\}$ in (1).

Lemma 3.3. *The rate of decay of* $\mathbf{P}\{\mathrm{MC}\}$ *is*

$$-\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MC}\} = \min\big(-\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}, -\lim_{n\to\infty}\tfrac{1}{n}\log\mathbf{P}\{\mathrm{MCI}^*_{\mathrm{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\}\big).$$

Again, because the Pareto systems being estimated out of order has a rate of decay that is greater than or equal to the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$, the rate of decay of $\mathbf{P}\{\mathrm{MC}\}$ can be simplified to a rate involving only pairs of non-Pareto systems and phantom Pareto systems. For all non-Pareto systems $j \in \mathcal{P}^c$ and all phantom Pareto systems $\ell \in \mathcal{P}^{\mathrm{ph}} = \{0, 1, \ldots, p\}$, recall that $\ell = i$ and define the rate function

$$
\begin{aligned}
&R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1}) \\
&:= \begin{cases} \inf\limits_{x_j \le x_1} & \alpha_j I_j(x_j, y_j) + \alpha_1 J_1(x_1) & \text{if } \ell = 0, \\ \inf\limits_{x_j \le x_{i+1},\ y_j \le y_i} & \alpha_j I_j(x_j, y_j) + \alpha_i K_i(y_i) + \alpha_{i+1} J_{i+1}(x_{i+1}) & \text{if } \ell \in \{1, \ldots, p-1\}, \\ \inf\limits_{y_j \le y_p} & \alpha_j I_j(x_j, y_j) + \alpha_p K_p(y_p) & \text{if } \ell = p, \end{cases} \\
&= \begin{cases} \inf\limits_{x} & \alpha_j J_j(x) + \alpha_1 J_1(x) & \text{if } \ell = 0, \\ \inf\limits_{x_j \le x_{i+1},\ y_j \le y_i} & \alpha_j I_j(x_j, y_j) + \alpha_i K_i(y_i) + \alpha_{i+1} J_{i+1}(x_{i+1}) & \text{if } \ell \in \{1, \ldots, p-1\}, \\ \inf\limits_{y} & \alpha_j K_j(y) + \alpha_p K_p(y) & \text{if } \ell = p, \end{cases}
\end{aligned}
$$

where $\alpha_0 := 1$, $\alpha_{p+1} := 1$, and equality of these two rates is explained in Online Appendix C. The following Theorem 3.4 states the rate of decay of the probability of misclassification, $\mathbf{P}\{\text{MC}\}$.

Theorem 3.4. *The rate of decay of the probability of misclassification is*

$$- \lim_{n \to \infty} \frac{1}{n} \log \mathbf{P}\{\text{MC}\} = \min \left(\min_{i \in \mathcal{P}} \min_{i' \in \mathcal{P}, i' \ne i} R_i(\alpha_i, \alpha_{i'}), \min_{j \in \mathcal{P}^c} \min_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1}) \right).$$

According to Theorem 3.4, we can retrieve the overall rate of decay of the probability of misclassification by calculating (a) the slowest among all pairwise false exclusion rates between the Pareto systems $i, i' \in \mathcal{P}$, $i \ne i'$, and (b) the slowest among the pairwise false inclusion rates between non-Pareto systems $j \in \mathcal{P}^c$ and phantom Pareto systems $\ell \in \mathcal{P}^{\text{ph}}$. Therefore the rate of decay of $\mathbf{P}\{\text{MC}\}$ is determined by the most likely misclassification event between two Pareto systems or between a non-Pareto system and a phantom Pareto system. We remind the reader that the rate in Theorem 3.4 accounts for dependence between the objectives.

3.3 Optimal Allocation Strategy

To determine the asymptotically optimal sampling allocation that maximizes the rate of decay of the probability of misclassification, $\mathbf{P}\{\text{MC}\}$, we consider the rate of decay of $\mathbf{P}\{\text{MC}\}$ in Theorem 3.4 as a function of the sampling allocation $\boldsymbol{\alpha}$. To determine the best value of $\boldsymbol{\alpha}$, we maximize the rate of decay of $\mathbf{P}\{\text{MC}\}$ by solving the following Problem Q, defined as

$$
\begin{aligned}
\text{Problem } Q: \quad & \text{maximize } z \text{ s.t.} \\
& R_i(\alpha_i, \alpha_{i'}) \ge z \text{ for all } i, i' \in \mathcal{P} \text{ such that } i' \ne i, \\
& R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1}) \ge z \text{ for all } j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i, \\
& \textstyle\sum_{k=1}^{r} \alpha_k = 1,\ \alpha_k \ge 0 \text{ for all } k \in \mathcal{S}.
\end{aligned}
$$

Thus at optimality in Problem Q, $\boldsymbol{\alpha}^*$ is the sampling allocation that maximizes the rate of decay of the probability of misclassification. The optimal rate is represented by z^*. Given a value of $(\alpha_i, \alpha_{i'})$, the value of $R_i(\alpha_i, \alpha_{i'})$ is obtained by solving

$$\text{Problem } R_{ii'}^{\text{MCE}}: \quad \text{minimize } \alpha_i I_i(x_i, y_i) + \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) \text{ s.t. } x_{i'} \le x_i,\quad y_{i'} \le y_i,$$

and given a value of $(\alpha_j, \alpha_i, \alpha_{i+1})$, the value of $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ is obtained by solving

$$
\begin{aligned}
\text{Problem } R_{j\ell}^{\text{MCI}}: \quad & \text{minimize } \alpha_j I_j(x_j, y_j) + \alpha_i K_i(y_i)\mathbb{I}_{[\ell \ne 0]} + \alpha_{i+1} J_{i+1}(x_{i+1})\mathbb{I}_{[\ell \ne p]} \\
& \text{s.t. } (x_j - x_{i+1})\mathbb{I}_{[\ell \ne p]} \le 0,\quad (y_j - y_i)\mathbb{I}_{[\ell \ne 0]} \le 0,
\end{aligned}
$$

where, for ease of notation, we write Problem $R^{\text{MCI}}_{j\ell}$ in its unsimplified form. We distinguish Problems $R^{\text{MCE}}_{ii'}$ and $R^{\text{MCI}}_{j\ell}$ as strictly convex optimization problems in $(x_i, y_i, x_{i'}, y_{i'})$ and (x_j, y_j, y_i, x_{i+1}), respectively, while $R_i(\alpha_i, \alpha_{i'})$ and $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ are their respective objective values at optimality. In the sections that follow, Problem $R^{\text{MCI}}_{j\ell}$ plays a prominent role. Thus we briefly discuss the properties of Problem Q. Then, we provide a more in-depth look at the properties of Problem $R^{\text{MCI}}_{j\ell}$.

3.3.1 Properties of Problem Q. Problem Q has $p \times (p-1)$ constraints corresponding to controlling the rate of decay of $\mathbf{P}\{\text{MCE}_{\mathcal{P}}\}$ and $(r-p)\times(p+1)$ constraints corresponding to controlling the rate of decay of $\mathbf{P}\{\text{MCI}_{\text{ph}}\}$. Also, each $R_i(\alpha_i, \alpha_{i'})$ and $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ are concave functions of $(\alpha_i, \alpha_{i'})$ and $(\alpha_j, \alpha_i, \alpha_{i+1})$, respectively [Boyd and Vandenberghe 2004, p. 81]. Thus Problem Q is a concave maximization problem. We emphasize the following important property of the rate z as a function of the allocation $\boldsymbol{\alpha}$: *If $\alpha_k = 0$ for any system k in Problem Q, then the rate $z = 0$.* Since equal allocation is feasible and results in a rate $z > 0$, at optimality in Problem Q, we have $z^* > 0$ and $\alpha^*_k > 0$ for all systems $k \in \mathcal{S}$.

3.3.2 Properties of Problem $R^{\text{MCI}}_{j\ell}$. Along with primal feasibility, the following KKT conditions are necessary and sufficient for global optimality in the strictly convex Problem $R^{\text{MCI}}_{j\ell}$. Let $(x^*_j, y^*_j, y^*_i, x^*_{i+1})$ be the solution to Problem $R^{\text{MCI}}_{j\ell}$, where $y^*_0 := 0, x^*_{p+1} := 0$. Letting $\lambda_x \geq 0$ and $\lambda_y \geq 0$ be dual variables, we have complementary slackness conditions $\lambda_x(x^*_j - x^*_{i+1}) = 0$ if $\ell \neq p$, $\lambda_y(y^*_j - y^*_i) = 0$ if $\ell \neq 0$, and stationarity conditions

$$\alpha_j \frac{\partial I_j(x^*_j, y^*_j)}{\partial x_j} + \lambda_x \mathbb{I}_{[\ell \neq p]} = 0, \qquad \alpha_j \frac{\partial I_j(x^*_j, y^*_j)}{\partial y_j} + \lambda_y \mathbb{I}_{[\ell \neq 0]} = 0, \tag{2}$$

$$\alpha_{i+1} \frac{\partial J_{i+1}(x^*_{i+1})}{\partial x_{i+1}} - \lambda_x = 0 \quad \text{if } \ell \neq p, \qquad \alpha_i \frac{\partial K_i(y^*_i)}{\partial y_i} - \lambda_y = 0 \quad \text{if } \ell \neq 0. \tag{3}$$

In the solution to Problem $R^{\text{MCI}}_{j\ell}$, the variables x^*_j, y^*_j, y^*_i, and x^*_{i+1} are each functions of the proportional allocations to non-Pareto system j and Pareto systems i and $i+1$, $(\alpha_j, \alpha_i, \alpha_{i+1})$. When this dependence must explicitly be denoted, for brevity, define

$$\mathfrak{z}^*_j(\alpha_j, \alpha_i, \alpha_{i+1}) := (x^*_j(\alpha_j, \alpha_i, \alpha_{i+1}), y^*_j(\alpha_j, \alpha_i, \alpha_{i+1})).$$

Now notice that under Assumptions 2–4, from the KKT conditions for Problem $R^{\text{MCI}}_{j\ell}$, the value of the rate function $I_j(\mathfrak{z}^*_j(\alpha_j, \alpha_i, \alpha_{i+1})) > 0$ at optimality in Problem $R^{\text{MCI}}_{j\ell}$. This result is stated formally in Lemma 3.5; we omit the proof.

Lemma 3.5. *If $\alpha_j > 0, \alpha_i > 0, \alpha_{i+1} > 0$, then $I_j(\mathfrak{z}^*_j(\alpha_j, \alpha_i, \alpha_{i+1})) > 0$ in Problem $R^{\text{MCI}}_{j\ell}$ for all non-Pareto systems $j \in \mathcal{P}^c$ and all phantom Pareto systems $\ell \in \mathcal{P}^{\text{ph}}, \ell = i$.*

A lemma regarding the locations of the solutions to Problem $R^{\text{MCI}}_{j\ell}$ appears in Online Appendix D.

4 LIMITING APPROXIMATION TO THE OPTIMAL ALLOCATION

Since Problem Q is a bi-level optimization problem, it may take some time to solve for the optimal allocation when the number of systems is large. While the computational time could be reduced by solving the inner problems in parallel, we believe it is useful to see if the optimal allocation can be simplified for large problem instances. In this section, we send the number of non-Pareto systems to infinity while keeping the number of Pareto systems finite and equal to p. This limiting regime enables us to write the relative allocations between the non-Pareto systems in closed form.

Before proceeding, we emphasize two key points about our limiting regime. First, we do not intend that the SCORE framework be implemented as-written when the number of systems is infinite. We are, after all, providing optimal allocations for R&S problems in which the number of systems is finite. We seek only a simplifying framework that would be a good approximation to the asymptotically optimal allocation when the number of non-Pareto systems is large relative to the number of Pareto systems. Further, it may seem natural that if the number of non-Pareto systems tends to infinity, the number of Pareto systems should also tend to infinity, but perhaps at a slower rate. While such a regime may be intuitively appealing, it is not clear how the Pareto systems should be added to achieve a meaningful limiting allocation framework in a large deviations regime. Thus in what follows, we keep the number of Pareto systems finite and equal to p. We also remind the reader that, like much of the R&S literature, our emphasis is on the case in which the underlying distributions are normal. Thus we make a normality assumption in §4.1. This assumption simplifies the proofs and assists our intuition regarding dependence by allowing us to model the dependence between the objectives as correlation.

4.1 Preliminaries for the Limiting Allocation

Recall that $r = |\mathcal{S}| = p + |\mathcal{P}^c|$ is the total number of systems, and in what follows, only $|\mathcal{P}^c|$ will tend to infinity while p remains constant.

4.1.1 Assumptions. We make four assumptions on the way non-Pareto systems are added to ensure a meaningful limiting allocation.

Assumption 5. *There exists a compact set $\mathcal{C}_1 \subset \mathbb{R}^2$ such that $(g_k, h_k) \in \mathcal{C}_1$ for all $k \in \mathcal{S}$, and such that $\mathcal{C}_1 \subset \mathcal{F}_d^c$. (See Assumption 4 for notation.)*

Since all rate functions are strictly convex with a unique minimum at the location of the mean, there exists another compact set $\mathcal{C} \supseteq \mathcal{C}_1$ that contains the locations of the solutions to all Problems $R_{j\ell}^{\text{MCI}}$. Let β be the diameter of a circle that covers $\mathcal{C}$; β appears in the Online Appendix.

Assumption 6. *For all $k \in \mathcal{S}$, the rate functions $I_k(x,y)$ have the quadratic form $I_k(x,y) = \frac{1}{2}\begin{bmatrix} g_k-x \\ h_k-y \end{bmatrix}^\intercal \Sigma_k^{-1} \begin{bmatrix} g_k-x \\ h_k-y \end{bmatrix}$ for all $(x,y) \in \mathbb{R}^2$, where $\Sigma_k := \begin{bmatrix} \sigma_{g_k}^2 & \rho_k\sigma_{g_k}\sigma_{h_k} \\ \rho_k\sigma_{g_k}\sigma_{h_k} & \sigma_{h_k}^2 \end{bmatrix}$. Further, there exist constants $c_a < 1$ and $c_b > 1$ such that the eigenvalues of Σ_k are bounded as $0 < c_a \le \lambda_{\min}(\Sigma_k) \le \lambda_{\max}(\Sigma_k) \le c_b < \infty$ for all $k \in \mathcal{S}$.*

Assumption 6 implies $J_k(x) = (x-g_k)^2/(2\sigma_{g_k}^2)$, $K_k(y) = (y-h_k)^2/(2\sigma_{h_k}^2)$ for all $k \in \mathcal{S}$.

Assumption 7. *There exists $\epsilon > 0$ such that (g_j, h_j) satisfies (a) $\inf\{|h_j - h_i| : i \in \mathcal{P}\} > \epsilon$, $\inf\{|g_j - g_i| : i \in \mathcal{P}\} > \epsilon$, and (b) $\inf\{|(h_j-h_i)/\sigma_{h_j} - \rho_j(g_j - g_{i+1})/\sigma_{g_j}| : i \in \mathcal{P}\} > \epsilon$, $\inf\{|(g_j - g_{i+1})/\sigma_{g_j} - \rho_j(h_j - h_i)/\sigma_{h_j}| : i \in \mathcal{P}\} > \epsilon$ for all $j \in \mathcal{P}^c$.*

Assumption 8. *For all Pareto systems $i \in \mathcal{P}$, there exists a non-Pareto system $j \in \mathcal{P}^c$ such that $h_j \le h_{i-1}$ or $g_j \le g_{i+1}$.*

Assumption 5 ensures that the systems that are added continue to compete with the Pareto systems and do not become irrelevant in the limit. Assumption 7 ensures the non-Pareto systems $j \in \mathcal{P}^c$ are added to $\mathcal{C}_1$ so that they do not systematically approach the Pareto front, and so that they do not approach the lines $y = h_i + \rho_j(\sigma_{h_j}/\sigma_{g_j})(x - g_{i+1})$ and $y = h_i + (1/\rho_j)(\sigma_{h_j}/\sigma_{g_j})(x - g_{i+1})$ for all $i \in \mathcal{P}$. Notice that Assumption 7(b) follows from Assumption 7(a) and Assumption 6 when the correlation $\rho_j = 0$. Assumptions 5 and 7 are analogous to assumptions in Pasupathy et al. [2015].

We differ from Pasupathy et al. [2015] in Assumptions 6 and 8. While Pasupathy et al. [2015] assume the rate functions have upper and lower bounding quadratics on a compact set (a mild assumption), we simplify the analysis by assuming the rate functions are quadratic. Sufficient conditions to ensure we have appropriate quadratic rate functions are (a) we obtain i.i.d. replicates of the bivariate normal random vector (G_k, H_k) with parameters $(g_k, h_k, \sigma_{g_k}^2, \sigma_{h_k}^2, \rho_k)$ for all $k \in \mathcal{S}$ where $\sigma_{g_k}^2, \sigma_{h_k}^2$ denote variance, ρ_k denotes correlation between the objectives, and (b) the variance and correlation values are uniformly bounded as $0 < \sigma_a^2 \le \sigma_{g_k}^2 \le \sigma_b^2 < \infty$, $0 < \sigma_a^2 \le \sigma_{h_k}^2 \le \sigma_b^2 < \infty$ for $\sigma_a^2 < 1$, $\sigma_b^2 > 1$, and $|\rho_k| \le \rho_b$ for $\rho_b \in (0.5, 1)$. (Note that the existence of such bounds follows from the condition on the eigenvalues of Σ_k. The independence of the replicates can be relaxed under the conditions in Assumption 3.) Thus we require that the systems be added to $\mathcal{C}_1$ in such a way that their corresponding rate functions cannot become too shallow (less than σ_a^2) or too steep (larger than σ_b^2), and so that they cannot degenerate to a single dimension ($|\rho_k|$ approaching 1). We conjecture that our analysis holds in the case of bounding quadratics, but we do not show it.

Finally, Assumption 8 implies that there does not exist a Pareto system $i \in \{2, \ldots, p-1\}$ such that Pareto systems $i-1, i$, and $i+1$ dominate *all* of the non-Pareto systems (see Figure 2); note that the assumption always holds for $i \in \{1, p\}$. Assumption 8 ensures that when we relax Problem Q to contain only constraints corresponding to MCI in §4.2, Pareto system i receives a positive sample allocation at optimality for all $i \in \mathcal{P}$. To see why, suppose there are three Pareto systems $i-1, i$, and $i+1$, and suppose *all* non-Pareto systems are in the shaded region of Figure 2. Further, suppose all variances associated with Pareto systems $i-1, i+1$ and the non-Pareto systems are relatively small, while the variances associated with Pareto system i are relatively large. Then it is expedient for Pareto systems $i-1$ and $i+1$ to do all the work of excluding the non-Pareto systems. If there exists a non-Pareto system j outside the shaded region, then Pareto systems $i-1$ and $i+1$ can no longer do all the work of excluding the non-Pareto systems; hence Pareto system i receives positive sample allocation. We view Assumption 8 as mild for two reasons. First, in §4.2, we send the cardinality of non-Pareto systems to infinity in an "even" way under Assumption 9. Thus we view Assumption 8 as requiring an initial level of evenness among the non-Pareto systems. Second, at optimality in Problem Q, all Pareto systems receive positive allocation due to the constraints corresponding to MCE, regardless of the system configuration. Since the SCORE framework in §5 includes the MCE constraints, this assumption does not impact implementation.

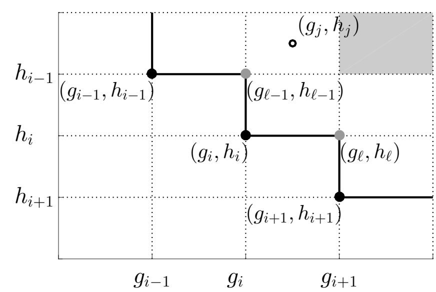

Fig. 2. Under Assumption 8, for each Pareto system $i \in \{2, \ldots, p-1\}$, there exists at least one non-Pareto system j outside the shaded region.

4.1.2 Rate Functions Under the Normality Assumption. We write $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ under Assumption 6 in the following Proposition 4.1. For brevity, define the indicators $\mathbb{I}^g_{j\ell} := \mathbb{I}_{[\lambda_x > 0,\ \ell \ne p]}$ and $\mathbb{I}^h_{j\ell} := \mathbb{I}_{[\lambda_y > 0,\ \ell \ne 0]}$ at optimality in Problem $R^{\text{MCI}}_{j\ell}$. Recalling that $\ell = i$, intuitively, $\mathbb{I}^g_{j\ell} > 0$ means that non-Pareto system j "competes" with Pareto system $i+1$ on the g objective via the phantom Pareto system ℓ, and $\mathbb{I}^h_{j\ell} > 0$ means that non-Pareto system j "competes" with Pareto system i on the h objective via phantom Pareto system ℓ (see Figure 2). To further simplify the rate function, when $\alpha_j > 0, \alpha_i > 0, \alpha_{i+1} > 0$, for all $j \in \mathcal{P}^c$ and phantom Pareto systems $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$, define

$$w_g(\alpha_j, \alpha_{i+1}) := \frac{\sigma^2_{g_j}/\alpha_j}{\sigma^2_{g_j}/\alpha_j + \sigma^2_{g_{i+1}}/\alpha_{i+1}} \text{ if } \ell \ne p \quad \text{and} \quad w_h(\alpha_j, \alpha_i) := \frac{\sigma^2_{h_j}/\alpha_j}{\sigma^2_{h_j}/\alpha_j + \sigma^2_{h_i}/\alpha_i} \text{ if } \ell \ne 0,$$

where $0 < w_g(\alpha_j, \alpha_{i+1}) < 1$ and $0 < w_h(\alpha_j, \alpha_i) < 1$ can be interpreted as weights. For readability and compactness, we often denote these weights as w_g and w_h, respectively, where the appropriate dependencies can be deduced from context.

Notice that the expression for $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ in Proposition 4.1 simplifies to one of three cases: the one-dimensional rate corresponding to system j being estimated as better than Pareto system $i+1$ on objective g, the one-dimensional rate corresponding to system j being estimated as better than

Pareto system i on objective h, or a bivariate rate of j dominating the phantom Pareto system ℓ. Since $\hat{G}_{i+1}$ and $\hat{H}_i$ are independent, only ρ_j appears in the rate. Further, in the expressions that follow, only w_g and w_h are functions of $(\alpha_j, \alpha_i, \alpha_{i+1})$.

Proposition 4.1. *Under Assumption 6, for each non-Pareto system $j \in \mathcal{P}^c$ and phantom Pareto system $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$, if $\alpha_j > 0, \alpha_i > 0, \alpha_{i+1} > 0$, then*

(1) *the rate function $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ is*

$$R_{j\ell}(\alpha_j,\alpha_i,\alpha_{i+1}) = \frac{\alpha_j}{2}\begin{bmatrix}(g_j-g_{i+1})\mathbb{I}^g_{j\ell}\\(h_j-h_i)\mathbb{I}^h_{j\ell}\end{bmatrix}^\intercal \begin{bmatrix}\sigma^2_{g_j}/w_g & \rho_j\sigma_{g_j}\sigma_{h_j}\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell}\\ \rho_j\sigma_{g_j}\sigma_{h_j}\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell} & \sigma^2_{h_j}/w_h\end{bmatrix}^{-1}\begin{bmatrix}(g_j-g_{i+1})\mathbb{I}^g_{j\ell}\\(h_j-h_i)\mathbb{I}^h_{j\ell}\end{bmatrix}$$

$$= \frac{\alpha_j}{2(1-\rho_j^2 w_g w_h \mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell})}\Big(\frac{(g_j-g_{i+1})^2}{\sigma^2_{g_j}}w_g\mathbb{I}^g_{j\ell} - \frac{2\rho_j(g_j-g_{i+1})(h_j-h_i)}{\sigma_{g_j}\sigma_{h_j}}w_g w_h \mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell} + \frac{(h_j-h_i)^2}{\sigma^2_{h_j}}w_h\mathbb{I}^h_{j\ell}\Big),$$

$$\text{where } \mathbb{I}^g_{j\ell} > 0, \mathbb{I}^h_{j\ell} = 0 \text{ iff } \ell \neq p, g_j > g_{i+1}, \frac{(h_j-h_i)}{\sigma_{h_j}} \le \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}w_g,$$

$$\mathbb{I}^g_{j\ell} = 0, \mathbb{I}^h_{j\ell} > 0 \text{ iff } \ell \neq 0, h_j > h_i, \frac{(g_j-g_{i+1})}{\sigma_{g_j}} \le \rho_j\frac{(h_j-h_i)}{\sigma_{h_j}}w_h, \text{ and}$$

$$\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell} > 0 \text{ iff } \ell \notin \{0,p\}, \frac{(g_j-g_{i+1})}{\sigma_{g_j}} > \rho_j\frac{(h_j-h_i)}{\sigma_{h_j}}w_h, \frac{(h_j-h_i)}{\sigma_{h_j}} > \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}w_g;$$

(2) *the rate functions in $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ corresponding to systems j, i, and $i+1$ are*

$$I_j(\mathfrak{z}^*_j(\alpha_j,\alpha_i,\alpha_{i+1})) = \frac{(1-\rho_j^2 w_h(2-w_h)\mathbb{I}^h_{j\ell})}{2(1-\rho_j^2 w_g w_h\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell})^2}\frac{(g_j-g_{i+1})^2}{\sigma^2_{g_j}}w_g^2\mathbb{I}^g_{j\ell} + \frac{(1-\rho_j^2 w_g(2-w_g)\mathbb{I}^g_{j\ell})}{2(1-\rho_j^2 w_g w_h\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell})^2}\frac{(h_j-h_i)^2}{\sigma^2_{h_j}}w_h^2\mathbb{I}^h_{j\ell}$$
$$- \rho_j\frac{[(1-\rho_j^2)w_g w_h-(1-w_g)(1-w_h)]}{(1-\rho_j^2 w_g w_h\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell})^2}\frac{(g_j-g_{i+1})(h_j-h_i)}{\sigma_{g_j}\sigma_{h_j}}w_g w_h\mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell},$$

$$K_i(y^*_i(\alpha_j,\alpha_i,\alpha_{i+1})) = \mathbb{I}^h_{j\ell}\frac{(1-w_h)^2}{2[1-\rho_j^2 w_g w_h\mathbb{I}^g_{j\ell}]^2}\frac{\sigma^2_{h_j}}{\sigma^2_{h_i}}\Big[\frac{(h_j-h_i)}{\sigma_{h_j}} - \mathbb{I}^g_{j\ell}\rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}w_g\Big]^2,$$

$$J_{i+1}(x^*_{i+1}(\alpha_j,\alpha_i,\alpha_{i+1})) = \mathbb{I}^g_{j\ell}\frac{(1-w_g)^2}{2[1-\rho_j^2 w_g w_h\mathbb{I}^h_{j\ell}]^2}\frac{\sigma^2_{g_j}}{\sigma^2_{g_{i+1}}}\Big[\frac{(g_j-g_{i+1})}{\sigma_{g_j}} - \mathbb{I}^h_{j\ell}\rho_j\frac{(h_j-h_i)}{\sigma_{h_j}}w_h\Big]^2.$$

In what follows, we work with $\boldsymbol{\alpha}$ directly, instead of working with w_g and w_h. However, we preview the result of our limiting regime here: by sending the number of non-Pareto systems to infinity, $w_g \to 1$ and $w_h \to 1$ in $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ for all $j \in \mathcal{P}^c$, $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$.

While the value of $I_j(\mathfrak{z}^*_j(\alpha_j, \alpha_i, \alpha_{i+1}))$ is always strictly positive at optimality by Lemma 3.5, it may be that $K_i(y^*_i(\alpha_j, \alpha_i, \alpha_{i+1})) = 0$ or $J_{i+1}(x^*_{i+1}(\alpha_j, \alpha_i, \alpha_{i+1})) = 0$, in which case either Pareto system i or Pareto system $i+1$ does not appear in the rate function in Problem $R^{\text{MCI}}_{j\ell}$, respectively. This fact raises the possibility that a particular Pareto system i does not appear in the rate function for Problem $R^{\text{MCI}}_{j\ell-1}$ *or* Problem $R^{\text{MCI}}_{j\ell}$, in which case the non-Pareto system j does not compete with the Pareto system i at all (see Figure 2). The following Lemma 4.2 states that such a case is impossible.

Lemma 4.2. *Under Assumption 6, if the allocations $\alpha_j > 0$, $\alpha_{i-1} > 0$, $\alpha_i > 0$, and $\alpha_{i+1} > 0$, then $\max\big(J_i(x^*_i(\alpha_j, \alpha_{i-1}, \alpha_i)), K_i(y^*_i(\alpha_j, \alpha_i, \alpha_{i+1}))\big) > 0$ for all $j \in \mathcal{P}^c, i \in \mathcal{P}$.*

4.2 Allocation to Non-Pareto Systems

Since we send the number of non-Pareto systems to infinity, we relax the constraints in Problem Q that pertain only to Pareto systems and MCE events. Thus in this section, we concern ourselves not with Problem Q, but with its relaxation:

$$\text{Problem } \widetilde{Q}: \qquad \text{maximize } \tilde{z} \text{ s.t.}$$
$$R_{j\ell}(\tilde{\alpha}_j, \tilde{\alpha}_i, \tilde{\alpha}_{i+1}) \geq \tilde{z} \text{ for all } j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i,$$
$$\textstyle\sum_{k=1}^{r} \tilde{\alpha}_k = 1, \; \tilde{\alpha}_k \geq 0 \; \text{ for all } k \in \mathcal{S}.$$

Under our assumptions, the KKT conditions are necessary and sufficient for global optimality in Problem $\widetilde{Q}$. We first use Problem $\widetilde{Q}$ to derive insights on the optimal allocation as the number of non-Pareto systems tends to infinity. In §4.4, we show that under mild conditions, for a large enough set of non-Pareto systems, the solutions to Problems Q and $\widetilde{Q}$ are equal. Since they play a prominent role in the results that follow, we present the KKT conditions for Problem $\widetilde{Q}$ in Theorem 4.3.

Theorem 4.3. *Let $\lambda_{j\ell} \geq 0$ for all $j \in \mathcal{P}^c$ and all $\ell \in \{0, 1, \ldots, p\}$ be dual variables associated with Problem $\widetilde{Q}$, and recall that the phantom Pareto labels are $\ell = i$ for all $i \in \mathcal{P}$. Under Assumptions 6 and 8, at optimality in Problem $\widetilde{Q}$, $\tilde{\alpha}_k^* > 0$ for all $k \in \mathcal{S}$ and*

1. *for each non-Pareto system $j \in \mathcal{P}^c$, there exists a phantom Pareto system $\ell^* \in \mathcal{P}^{\text{ph}}$, $\ell^* = i^*$ such that $\lambda_{j\ell^*} > 0$, which implies that the rate $\tilde{z}^* = R_{j\ell^*}(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*) > 0$;*
2. *for each Pareto system $i \in \mathcal{P}$, there exists a non-Pareto system $j^* \in \mathcal{P}^c$ such that the quantities $\lambda_{j^*\ell-1} J_i(x_i^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) > 0$ or $\lambda_{j^*\ell} K_i(y_i^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0$, which implies that the rate $\tilde{z}^* = \min\big(R_{j^*\ell-1}(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*), R_{j^*\ell}(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)\big) > 0$;*
3. *for all non-Pareto systems $j, j' \in \mathcal{P}^c$,*
$$\frac{\sum_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} \lambda_{j\ell} I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))}{\sum_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} \lambda_{j'\ell} I_{j'}(\mathfrak{z}_{j'}^*(\tilde{\alpha}_{j'}^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))} = 1; \tag{4}$$
4. *for all Pareto systems $i, i' \in \mathcal{P}$, letting the phantom Pareto label for i' be ℓ',*
$$\frac{\sum_{j \in \mathcal{P}^c} \lambda_{j\ell-1} J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell} K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))}{\sum_{j \in \mathcal{P}^c} \lambda_{j\ell'-1} J_{i'}(x_{i'}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'-1}^*, \tilde{\alpha}_{i'}^*)) + \lambda_{j\ell'} K_{i'}(y_{i'}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'}^*, \tilde{\alpha}_{i'+1}^*))} = 1; \tag{5}$$
5. *for all Pareto systems $i \in \mathcal{P}$,*
$$\sum_{j \in \mathcal{P}^c} \frac{\lambda_{j\ell-1} J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell} K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))}{\sum_{\ell' \in \mathcal{P}^{\text{ph}}, \ell' = i'} \lambda_{j\ell'} I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'}^*, \tilde{\alpha}_{i'+1}^*))} = 1. \tag{6}$$

Proof. Let ν and $\lambda_{j\ell} \geq 0$ for all $j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}$ be dual variables. Then we have the complementary slackness conditions $\lambda_{j\ell}(R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*) - \tilde{z}^*) = 0$ for all $j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i$, and the (simplified) stationarity conditions

$$\textstyle\sum_{j \in \mathcal{P}^c} [\lambda_{j\ell-1} J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell} K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))] = \nu \quad \forall i \in \mathcal{P}; \tag{7}$$
$$\textstyle\sum_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} \lambda_{j\ell} I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) = \nu \quad \forall j \in \mathcal{P}^c; \tag{8}$$
$$\textstyle\sum_{j \in \mathcal{P}^c} \sum_{\ell \in \mathcal{P}^{\text{ph}}} \lambda_{j\ell} = 1. \tag{9}$$

See Online Appendix G for a complete proof. □

In Theorem 4.3, Parts (1) and (2) ensure the existence of a binding constraint in Problem $\widetilde{Q}$ for each non-Pareto system j and Pareto system i, respectively. Parts (3) and (4) determine the relative allocations between the non-Pareto systems and between the Pareto systems, respectively. Part (5) determines the relative allocation between a Pareto system i and the non-Pareto systems j that compete with it.

Observe that as the number of non-Pareto systems added according to Assumptions 5–8 grows, the overall rate of decay of $\mathbf{P}\{\text{MCI}_{\text{ph}}\}$ in Problem $\widetilde{Q}$ will decrease. If this fact is not intuitive, it can be seen by noticing that adding non-Pareto systems that compete with the Pareto systems in a non-trivial way implies that we are adding binding constraints to Problem $\widetilde{Q}$ that decrease its optimal value. Thus under our assumptions, as $|\mathcal{P}^c| \to \infty$, $\tilde{z}^* \to 0$. (Notice that now, we consider a sequence of Problems $\widetilde{Q}(r)$ that are indexed by r, and quantities such as $\tilde{z}^*$, $\lambda_{j\ell}$ for all $j \in \mathcal{P}^c$ and $\ell \in \mathcal{P}^{\text{ph}}$, and $\tilde{\alpha}_k^*$ for all $k \in \mathcal{S}$ are functions of r and could be denoted as $\tilde{z}^*(r)$, $\lambda_{j\ell}(r)$ and $\tilde{\alpha}_k^*(r)$, respectively. We often suppress this notation unless it is helpful for clarity.) Proposition 4.5 states the rate at which $\tilde{z}^* \to 0$. Before we state the proposition, we present Lemma 4.4, which provides bounds on $R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)$ and the rate functions that comprise it; these bounds are useful in the proofs of several subsequent results. The constants used in Lemma 4.4 are defined in Online Appendix H.

Lemma 4.4. *Let $\kappa_R^{\text{L}}, \kappa_R^{\text{U}}$ be positive, finite constants that do not depend on the system indices, and let* $\min^+$ *be an operator that returns the smallest positive element in a list (see §2.2 for a definition). Under Assumptions 5–8, for each $j \in \mathcal{P}^c$ and $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$,*

$$\tilde{\alpha}_j^* \kappa_R^{\text{L}} \min{}^+\Big[\frac{\mathbb{I}_{j\ell}^g}{1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*}, \frac{\mathbb{I}_{j\ell}^h}{1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*}\Big] \le R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*) \le \tilde{\alpha}_j^* \kappa_R^{\text{U}}\Big[\frac{\mathbb{I}_{j\ell}^h}{1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*} + \frac{\mathbb{I}_{j\ell}^g}{1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*}\Big] \tag{10}$$

$$\kappa_R^{\text{L}} \min{}^+\Big[\frac{\mathbb{I}_{j\ell}^g}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*)^2}, \frac{\mathbb{I}_{j\ell}^h}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*)^2}\Big] \le I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) \le \kappa_R^{\text{U}}\Big[\frac{\mathbb{I}_{j\ell}^h}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*)^2} + \frac{\mathbb{I}_{j\ell}^g}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*)^2}\Big] \tag{11}$$

$$\kappa_R^{\text{L}} \frac{\mathbb{I}_{j\ell}^h}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2} \le K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) \le \kappa_R^{\text{U}} \frac{\mathbb{I}_{j\ell}^h}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2} \tag{12}$$

$$\kappa_R^{\text{L}} \frac{\mathbb{I}_{j\ell}^g}{(1+\tilde{\alpha}_{i+1}^*/\tilde{\alpha}_j^*)^2} \le J_{i+1}(x_{i+1}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) \le \kappa_R^{\text{U}} \frac{\mathbb{I}_{j\ell}^g}{(1+\tilde{\alpha}_{i+1}^*/\tilde{\alpha}_j^*)^2}. \tag{13}$$

Proposition 4.5. *Under Assumptions 5–8, as $|\mathcal{P}^c| \to \infty$, $\tilde{z}^* = O(1/|\mathcal{P}^c|)$.*

Thus by Proposition 4.5, as the number of non-Pareto systems grows, the asymptotically optimal rate of decay of the probability of MC decreases to zero as $O(1/|\mathcal{P}^c|)$.

For each Pareto system $i \in \mathcal{P}$ and for each value of the total number of systems r, let $\mathcal{P}^c(i, r)$ denote the set of non-Pareto systems j that have a binding constraint with Pareto system i in Problem $\widetilde{Q}(r)$; here, we explicitly denote the dependence of Problem $\widetilde{Q}$ on r. That is, for all Pareto systems $i \in \mathcal{P}$ and all r, define

$$\mathcal{P}^c(i,r) := \{j \in \mathcal{P}^c :\ \lambda_{j\ell-1}(r) J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell}(r) K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0\}.$$

In Problem $\widetilde{Q}(r)$, we say that non-Pareto system j and Pareto system i *bind* with each other if $j \in \mathcal{P}^c(i, r)$, and we say that non-Pareto system j and phantom Pareto system ℓ *bind* with each other if $\lambda_{j\ell}(r) > 0$. Notice that the interior expression in the definition of $\mathcal{P}^c(i, r)$ equals the expression in the numerator of equation (6). Equation (6) determines the relative allocation between a Pareto system i and the non-Pareto systems j that bind with it via phantom Pareto system $\ell - 1$ or ℓ. By Theorem 4.3, at least one non-Pareto system binds with each Pareto system, so that $|\mathcal{P}^c(i, r)| \ge 1$ for all $i \in \mathcal{P}$ and all r. Also by Theorem 4.3, each non-Pareto system binds with at least one Pareto system, so that each $j \in \mathcal{P}^c$ belongs to at least one set $\mathcal{P}^c(i, r)$ for all r.

We make the following additional observations about the set $\mathcal{P}^c(i, r)$. First, since $\tilde{\alpha}_k^* > 0$ for all $k \in \mathcal{S}$ by Theorem 4.3, then for all $j \in \mathcal{P}^c, i \in \mathcal{P}$, a rate function term corresponding to Pareto system i must appear in one of $R_{j\ell-1}(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)$ or $R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)$ by Lemma 4.2. Thus $J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0$, which implies $\mathbb{I}_{j\ell-1}^g(r) + \mathbb{I}_{j\ell}^h(r) > 0$, for all $i \in \mathcal{P}, j \in \mathcal{P}^c$, and all r. If we also have non-Pareto system $j \in \mathcal{P}^c(i, r)$, then $\lambda_{j\ell-1}(r)\mathbb{I}_{j\ell-1}^g(r) + \lambda_{j\ell}(r)\mathbb{I}_{j\ell}^h(r) > 0$.

This result follows because for a non-Pareto system j to bind with a Pareto system i, we require a rate function corresponding to system i to appear in the binding constraint in Problem $\widetilde{Q}(r)$.

To make statements about the limiting relative allocations as $|\mathcal{P}^c| \to \infty$, we require notions of the *limiting* allegiances between non-Pareto systems and Pareto systems. (In what follows, the reader may find it helpful to consult Figure 1 or 2.) First, notice that when a new non-Pareto system j enters Problem $\widetilde{Q}(r)$, it will bind with one or more Pareto systems i via its phantom Pareto systems $\ell - 1$ or ℓ. Among these phantom Pareto systems, we require that there exists one primary phantom Pareto system for each non-Pareto system j, defined as follows.

Definition 4.6. Let non-Pareto system $j \in \mathcal{P}^c$ enter Problem $\widetilde{Q}(r)$ at $r = r_{j0} < \infty$. The phantom Pareto system $\ell^* \in \mathcal{P}^{\text{ph}}$ is the *primary phantom Pareto system* for j if $\lambda_{j\ell^*}(r) > 0$ for all $r \geq r_{j0}$, and $\lambda_{j\ell}(r) = o(\lambda_{j\ell^*}(r))$ for all other phantom Pareto systems $\ell \in \mathcal{P}^{\text{ph}}, \ell \neq \ell^*$.

The first condition ensures that the non-Pareto system j binds with the phantom Pareto system ℓ^* for all $r \geq r_{j0}$, so that $j \in \mathcal{P}^c(i^*, r)$ or $j \in \mathcal{P}^c(i^* + 1, r)$ for all $r \geq r_{j0}$. Using the shadow price interpretation of $\lambda_{j\ell}$'s in Problem $\widetilde{Q}$, the second condition implies that the greatest gain to the rate $\tilde{z}^*$ will be achieved by perturbing the MCI rate constraint associated with system j and phantom ℓ^*, by more than a constant. Since the primary phantom Pareto system ℓ^* is a function of the non-Pareto system $j \in \mathcal{P}^c$, we denote it as $\ell^*(j)$ whenever this notation is helpful for clarity. Otherwise, the dependency on system j is implied.

While the assumption that each non-Pareto j has a primary phantom Pareto ℓ^* may feel somewhat artificial, we believe that it will arise naturally, for example, when non-Pareto systems are added according to a uniform distribution (provided our assumptions are maintained). To understand why, consider what it means for one non-Pareto system j to bind with more than one phantom Pareto system ℓ. In a scenario with multiple Pareto systems ($p \geq 2$) and only one non-Pareto system j_1, the non-Pareto system j_1 will bind with *all* of the Pareto systems $i \in \mathcal{P}$ via at least one of their phantoms $\ell - 1, \ell$, due to Theorem 4.3. However, as new non-Pareto systems are added uniformly across the set $\mathcal{C}_1$, the new non-Pareto systems bind with the Pareto systems "closest" to them, and j_1 will cease to bind with Pareto systems that are "far away" from it — those Pareto systems will bind with other, closer, non-Pareto systems. Therefore intuitively, non-Pareto systems binding with multiple Pareto systems, and thus multiple phantom Pareto systems, may arise when (a) there are not very many non-Pareto systems, or (b) when the non-Pareto systems are not "evenly distributed," as might arise when all non-Pareto systems are uniquely dominated by the same Pareto system. Therefore we anticipate that the number of non-Pareto systems binding with multiple phantom Pareto systems ℓ decreases as non-Pareto systems are added "evenly."

In addition to assuming each non-Pareto system has a primary phantom Pareto system, we also require that the number of non-Pareto systems binding with each Pareto system increase to infinity. Specifically, for all Pareto systems $i \in \mathcal{P}$, define $\mathcal{P}^c(i)$ as

$$\mathcal{P}^c(i) := \{j \in \mathcal{P}^c : j \in \mathcal{P}^c(i, r) \text{ for all } r \geq r_{j0} \text{ and } \ell^*(j) \in \{\ell - 1, \ell\}\}.$$

The set $\mathcal{P}^c(i)$ contains the non-Pareto systems $j \in \mathcal{P}^c$ that bind with Pareto system $i \in \mathcal{P}$, via phantom Pareto $\ell^*(j) \in \{\ell - 1, \ell\}$, in every Problem $\widetilde{Q}(r)$ after r_{j0}. For each Pareto system $i \in \{2, \ldots, p - 1\}$, we further require that there exists at least one $j^* \in \mathcal{P}^c(i)$ such that (g_{j^*}, h_{j^*}) is non-dominated by Pareto systems $i - 1$ and $i + 1$ for all r; that is, j^* is outside the shaded region in Figure 2.

In what follows, we send $|\mathcal{P}^c(i)| \to \infty$ for all $i \in \mathcal{P}$. This condition ensures that the number of non-Pareto systems binding with each Pareto system i in Problem $\widetilde{Q}(r)$ goes to infinity with r. To ensure evenness of the non-Pareto systems in $\mathcal{C}_1$, the cardinality of each set $\mathcal{P}^c(i)$ must remain within a constant of the total number of non-Pareto systems. We formalize these assumptions in

Assumption 9, where $\{r\}$ denotes the sequence of the total number of competing systems. We numerically evaluate such a regime in §4.5.

Assumption 9. *We assume that (a) for each non-Pareto system $j \in \mathcal{P}^c$, there exists a primary phantom Pareto system $\ell^*(j) \in \mathcal{P}^{\text{ph}}$, (b) for each Pareto system $i \in \mathcal{P}$, there exists a non-Pareto system $j^* \in \mathcal{P}^c(i)$ such that $h_{j^*} \le h_{i-1}$ or $g_{j^*} \le g_{i+1}$ for all r, and (c) there exists $\kappa \in (0, \infty)$ such that $|\mathcal{P}^c(i)| \ge \kappa |\mathcal{P}^c|$ for all Pareto systems $i \in \mathcal{P}$ and all r.*

Under the regularity conditions in Assumption 9, we ensure that each system receives nonzero sample in the limit. Theorem 4.7 provides results on the limiting allocations.

Theorem 4.7. *Under Assumptions 5–9, as $|\mathcal{P}^c(i)| \to \infty$ for all Pareto systems $i \in \mathcal{P}$:*

1. *There exists $\kappa_1 > 0$ such that $I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*)) \ge \kappa_1$ for all $j \in \mathcal{P}^c$ with primary phantom Pareto system $\ell^*(j) \in \mathcal{P}^{\text{ph}}$, $\ell^*(j) = i^*(j)$, and all $r \ge r_{j0}$.*
2. *The allocation $\tilde{\alpha}_j^* = \Theta(\tilde{z}^*)$ for all $j \in \mathcal{P}^c$.*
3. *The dual variables $\nu = \Theta(1/|\mathcal{P}^c|)$ and $\lambda_{j\ell^*(j)} = \Theta(1/|\mathcal{P}^c|)$ for all $j \in \mathcal{P}^c$ with primary phantom Pareto system $\ell^*(j) \in \mathcal{P}^{\text{ph}}$.*
4. *There exists $\kappa_4 \in (0, \infty)$ such that $\tilde{\alpha}_i^* / \tilde{\alpha}_{j^*}^* > \kappa_4$ for all $i \in \mathcal{P}, j^* \in \mathcal{P}^c(i)$ such that $h_{j^*} \le h_{i-1}$ or $g_{j^*} \le g_{i+1}$, and all $r \ge r_{j^*0}$.*
5. *There exists $\kappa_2 < \infty$ such that $\kappa_1/\tilde{z}^* \le 1/\tilde{\alpha}_j^* + 1/\tilde{\alpha}_i^* \le \kappa_2/\tilde{z}^*$ for all $i \in \mathcal{P}, j \in \mathcal{P}^c$, $r \ge r_{j0}$.*
6. *There exists $\kappa_6 \in (0, \infty)$ such that $\tilde{\alpha}_i^* / \tilde{\alpha}_{i'}^* < \kappa_6$ for all $i, i' \in \mathcal{P}$, and all r.*
7. *For all $i \in \mathcal{P}, j \in \mathcal{P}^c$, the ratio of squared allocations $\tilde{\alpha}_j^{*2} / \tilde{\alpha}_i^{*2} = \Theta(1/|\mathcal{P}^c|)$.*
8. *The rate $\tilde{z}^* = \Theta(1/|\mathcal{P}^c|)$ and the allocations $\tilde{\alpha}_j^* = \Theta(1/|\mathcal{P}^c|)$ for all $j \in \mathcal{P}^c$.*
9. *In $R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)$, the rate function $K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) = \Theta(1/|\mathcal{P}^c|)$ and the rate function $J_{i+1}(x_{i+1}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) = \Theta(1/|\mathcal{P}^c|)$ for all $j \in \mathcal{P}^c$, $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$.*
10. *In Problem $R_{j\ell}^{\text{MCI}}$, $y_i^* \to h_i$ if $\ell \ne 0$ and $x_{i+1}^* \to g_{i+1}$ if $\ell \ne p$ for all $j \in \mathcal{P}^c$, $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$.*

The primary results in Theorem 4.7 appear in Parts (7)–(10). Because each Pareto system is competing with an increasingly large number of non-Pareto systems, Parts (7) and (8) state that in the limit, each Pareto system i will receive many more samples than the non-Pareto systems j. Parts (9) and (10) state that in the limit in Problem $R_{j\ell}^{\text{MCI}}$, the rate functions corresponding to both Pareto systems i and $i+1$ tend to zero, while we know the rate function corresponding to system j remains positive by Part (1). Thus in the limit, the rate function corresponding to j is evaluated over the region in which non-Pareto system j would dominate the phantom Pareto system ℓ, $x_j \le g_{i+1}, y_j \le h_i$. Loosely speaking, in this asymptotic regime, the Pareto systems receive so many samples that, relative to the non-Pareto systems, the Pareto systems appear known.

This last result leads us directly to the main result of the paper, presented in Theorem 4.8. We do not provide a proof; notice that it follows by applying Theorem 4.7.

Theorem 4.8. *Under Assumptions 5–9, as $|\mathcal{P}^c(i)| \to \infty$ for all $i \in \mathcal{P}$,*

$$R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)/\tilde{\alpha}_j^* = \inf_{x_j \le g_{i+1},\ y_j \le h_i} I_j(x_j, y_j) \quad \text{for all } j \in \mathcal{P}^c,\ \ell \in \mathcal{P}^{\text{ph}}, \ell = i.$$

To see the implications of Theorem 4.8, define the *score* $\mathbb{S}_j$ as

$$\mathbb{S}_j := \min_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} \big(\inf_{x_j \le g_{i+1},\ y_j \le h_i} I_j(x_j, y_j) \big) \quad \text{for all } j \in \mathcal{P}^c.$$

Then it follows that in the limit, $\tilde{z}^* = \min_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)$ for all $j \in \mathcal{P}^c$, and we have $\tilde{z}^*/\tilde{\alpha}_j^* = \min_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i}(\inf_{x_j \le g_{i+1},\ y_j \le h_i} I_j(x_j, y_j))$ for all $j \in \mathcal{P}^c$. Therefore the relative allocations between the non-Pareto systems are determined by the score, which is written formally in the following Theorem 4.9.

Theorem 4.9. *Under Assumptions 5–9, as $|\mathcal{P}^c(i)| \to \infty$ for all $i \in \mathcal{P}$, then*

$$\frac{\tilde{\alpha}^*_{j'}}{\tilde{\alpha}^*_j} = \frac{\mathbb{S}_j}{\mathbb{S}_{j'}} = \frac{\min_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i}\left(\inf_{x_j \le g_{i+1},\ y_j \le h_i} I_j(x_j, y_j)\right)}{\min_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i}\left(\inf_{x_{j'} \le g_{i+1},\ y_{j'} \le h_i} I_j(x_j, y_j)\right)} \quad \text{for all } j, j' \in \mathcal{P}^c.$$

Under Assumption 6, finding each $\inf_{x_j \le g_{i+1},\ y_j \le h_i} I_j(x_j, y_j)$ is a quadratic program with box constraints. The following proposition presents its closed form solution. We do not provide a proof of this result since it is a special case of previous results.

Proposition 4.10. *Under Assumption 6, for all non-Pareto systems $j \in \mathcal{P}^c$, the score is calculated as $\mathbb{S}_j = \min_{\ell \in \mathcal{P}^{\text{ph}}} \mathbb{S}_j(\ell)$, where recalling that $\ell = i$,*

$$\mathbb{S}_j(\ell) := \begin{cases} J_j(g_{i+1}) & \text{iff } \ell \ne p,\ \frac{(g_j - g_{i+1})}{\sigma_j} > 0,\ \frac{(h_j - h_i)}{\sigma_{h_j}} \le \rho_j \frac{(g_j - g_{i+1})}{\sigma_{g_j}}; \\ K_j(h_i) & \text{iff } \ell \ne 0,\ \frac{(h_j - h_i)}{\sigma_{h_j}} > 0,\ \frac{(g_j - g_{i+1})}{\sigma_{g_j}} \le \rho_j \frac{(h_j - h_i)}{\sigma_{h_j}}; \\ I_j(g_{i+1}, h_i) & \text{iff } \ell \notin \{0, p\},\ \frac{(g_j - g_{i+1})}{\sigma_{g_j}} > \rho_j \frac{(h_j - h_i)}{\sigma_{h_j}},\ \frac{(h_j - h_i)}{\sigma_{h_j}} > \rho_j \frac{(g_j - g_{i+1})}{\sigma_{g_j}}. \end{cases}$$

Thus in our asymptotic regime, the relative allocations between the non-Pareto systems can be expressed in closed form, where the allocation to a particular non-Pareto system is inversely proportional to its scaled distance from the Pareto frontier in the objective function space. Notice that the value of $\mathbb{S}_j(\ell)$ in Proposition 4.10 is identical to the value of $R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1})$ in Proposition 4.1 when $w_g = 1$ and $w_h = 1$.

4.3 Allocation to Pareto Systems

While we express the relative allocations between the non-Pareto systems in closed form, we also require a sense of how much sample to allocate to the Pareto systems. The following Theorem 4.11 states that as the number of non-Pareto systems tends to infinity, the allocations to the Pareto systems also tend to zero, but at a much slower rate than the allocations to the non-Pareto systems.

Theorem 4.11. *Under Assumptions 5–9, as $|\mathcal{P}^c(i)| \to \infty$ for all $i \in \mathcal{P}$, $\tilde{\alpha}^*_i = \Theta(1/\sqrt{|\mathcal{P}^c|})$ for all Pareto systems $i \in \mathcal{P}$.*

Theorem 4.11 only gives us a sense of the allocation to the Pareto systems in the limit. To solve for a specific allocation to each of the Pareto systems, we require heuristics, discussed in §5.

4.4 Equivalence of Allocations When the Number of Non-Pareto Systems is Large

Recall that all results presented in §4.2 and §4.3 pertain to Problem $\widetilde{Q}$ and not to the original characterization of the optimal allocation as the solution to Problem Q. The following Theorem 4.12 states that as the number of non-Pareto systems grows, the optimal allocation provided by Problem $\widetilde{Q}$ is equal to that provided by Problem Q.

Theorem 4.12. *Under Assumptions 5–9, for large enough $|\mathcal{P}^c|$, $\tilde{\boldsymbol{\alpha}}^* = \boldsymbol{\alpha}^*$.*

Intuitively, Theorem 4.12 holds because in the limiting regime, the Pareto systems receive so many more samples than the non-Pareto systems that MCE events between Pareto systems cannot be the unique minimum in the rate of decay of the $\mathbf{P}\{\text{MC}\}$.

4.5 Numerical Evaluation of the Limiting Regime

We have shown that under some conditions, as the number of non-Pareto systems tends to infinity, the rate of decay of $\mathbf{P}\{\text{MCE}_{\mathcal{P}}\}$ becomes non-binding in Problem Q. Now, we numerically evaluate this effect on a set of randomly-generated test problems.

To create the test problems, first, we place five Pareto systems at equally spaced angles on a circle of radius six. This spacing guarantees the minimum separation between Pareto systems on both objectives is greater than 0.5, so that $|g_i - g_{i'}| > 0.5$, $|h_i - h_{i'}| > 0.5$ for all $i, i' \in \mathcal{P}$. Then we generate non-Pareto systems by one of two methods: uniform or normal. In the uniform method, non-Pareto systems are generated uniformly in a circle centered at (100,100) with radius six. In the normal method, non-Pareto systems are generated according to an independent bivariate normal distribution with both means equal to 100 and both standard deviations equal to three. Thus the majority of systems are within six units of the mean. In both methods, non-Pareto systems less than 0.25 units away from the Pareto frontier are rejected. This condition ensures Assumption 7 is satisfied, and that the rate is large enough for us to obtain the optimal allocation numerically from Problem Q. Figures 3 and 6 show example problem instances in which 445 non-Pareto systems are added according to the uniform and normal methods, respectively. All systems have bivariate normal rate functions under Assumption 6 with independent objectives and unit variance.

As non-Pareto systems are added to fifty problem instances of each type, uniform and normal, we solve Problem Q for the optimal allocation. We then create two types of plots: Figures 4 and 7, which show the percent of problem instances with binding MCE constraints, and Figures 5 and 8, which show box plots of the percent of the dual variable values associated with MCE constraints at optimality in Problem Q. To better understand what we mean by the percent of dual variable values associated with MCE constraints, in Problem Q, let $\lambda_{ii'}^{\mathcal{P}}$ for all $i, i' \in \mathcal{P}$ be the dual

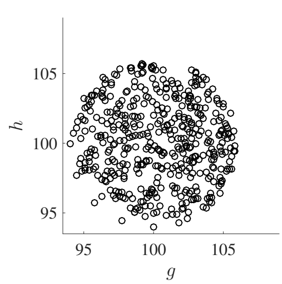

Fig. 3. The figure shows an example uniform problem with 445 non-Paretos generated uniformly in a circle of radius six.

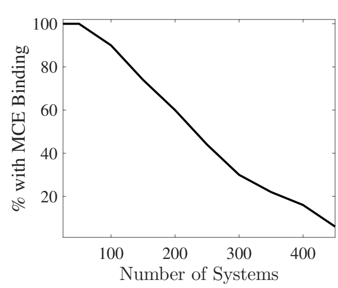

Fig. 4. The percent of 50 uniform problems with MCE constraints binding in Problem Q decreases as systems increase.

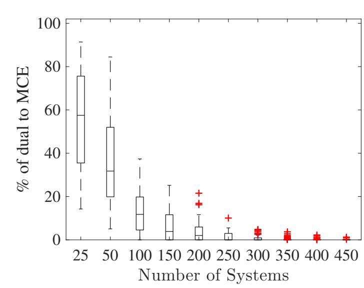

Fig. 5. The percent of dual variable value associated with MCE constraints, across 50 uniform problems, decreases as systems increase.

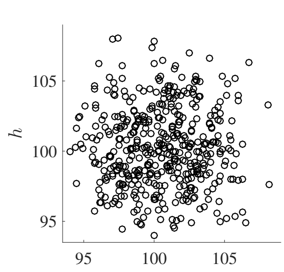

Fig. 6. The figure shows an example normal problem with 445 non-Paretos generated via a bivariate normal.

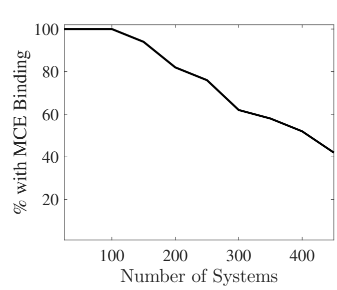

Fig. 7. The percent of 50 normal problems with MCE constraints binding in Problem Q decreases as systems increase.

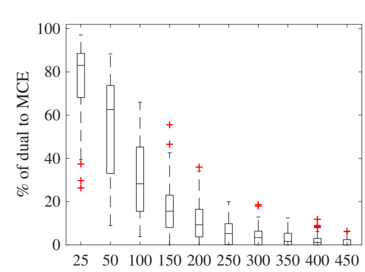

Fig. 8. The percent of dual variable value associated with MCE constraints, across 50 normal problems, decreases as systems increase.

variables corresponding to MCE constraints and $\lambda_{j\ell}^{\mathcal{P}^c}$ for all $j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}$ be the dual variables corresponding to MCI constraints. Then the percent of dual variable values associated with MCE constraints is $\sum_{i\in\mathcal{P}} \sum_{i'\in\mathcal{P},i'\neq i} \lambda_{ik}^{\mathcal{P}} / (\sum_{i\in\mathcal{P}} \sum_{i'\in\mathcal{P},i'\neq i} \lambda_{ii'}^{\mathcal{P}} + \sum_{\ell\in\mathcal{P}^{\text{ph}}} \sum_{j\in\mathcal{P}^c} \lambda_{j\ell}^{\mathcal{P}^c})$ at optimality.

While one could argue that our problems are somewhat artificial, given the nicely spaced Pareto frontier and its buffer away from the non-Pareto systems, we believe there is an important message in Figures 3–8. If the systems can be viewed as coming from a distribution, the distribution will affect the rate at which the limiting regime kicks in. When the systems are generated according to the normal method, they cluster together near the mean at (100,100). Thus many of the non-Pareto systems are some distance away from the Pareto frontier, and the limit kicks in slower. When the systems are generated according to the uniform method, the systems are more dispersed in their "allegiances" to Pareto systems and closer to the Pareto frontier, and the limit kicks in faster.

5 THE SCORE ALLOCATION

In this section, we describe a heuristic allocation for implementation based on the theory in §4 called the SCORE allocation. As in Pasupathy et al. [2015], we use Theorem 4.9 to determine the relative allocations among the non-Pareto systems. We now describe the method by which we determine the remainder of the allocations.

First, notice that if we let the rate of decay of $\mathbf{P}\{\text{MCI}_{\text{ph}}\}$ be determined by the limiting rates in Theorem 4.8, and let the primary phantom Pareto be determined by the scores as $\ell^* = \text{argmin}_{\ell\in\mathcal{P}^{\text{ph}}} \mathbb{S}_j(\ell)$, then there exists a version of Problem $\widetilde{Q}$ in which there is exactly one binding constraint corresponding to MCI for each non-Pareto system $j \in \mathcal{P}^c$. Thus the dual variables $\lambda_{j\ell} = 0$ for all $\ell \in \mathcal{P}^{\text{ph}}$ such that $\ell \neq \ell^*$. Then for all non-Pareto systems $j, j' \in \mathcal{P}^c$, the KKT condition in (4) implies that as $|\mathcal{P}^c(i)| \to \infty$, for all Pareto systems $i \in \mathcal{P}$, $\lambda_{j\ell^*}/\lambda_{j'\ell'^*} = I_{j'}(\mathfrak{z}_{j'}^*(\alpha_{j'}^*, \alpha_{i'^*}^*, \alpha_{i'^*+1}^*))/I_j(\mathfrak{z}_j^*(\alpha_j^*, \alpha_{i^*}^*, \alpha_{i^*+1}^*)) = \mathbb{S}_{j'}/\mathbb{S}_j = \alpha_j^*/\alpha_{j'}^*$, so that $\lambda_j^{\mathbb{S}} := \alpha_j^*/\sum_{j'\in\mathcal{P}^c} \alpha_{j'}^* = \mathbb{S}_j^{-1}/\sum_{j'\in\mathcal{P}^c} \mathbb{S}_{j'}^{-1}$ is the proportion of the "non-Pareto simulation budget" allocated to system j for all non-Pareto systems $j \in \mathcal{P}^c$. Then we write the allocation to a non-Pareto system j as $\alpha_j^* = \lambda_j^{\mathbb{S}}(1 - \sum_{i=1}^{p} \alpha_i^*)$.

It is tempting to create a heuristic allocation by solving a reduced version of Problem $\widetilde{Q}$ that only includes one constraint for each non-Pareto system j and its primary phantom Pareto ℓ^*. However, there are some drawbacks of this approach. First, the assumptions of the limiting score regime may not be satisfied, and some Pareto systems may receive a falsely low allocation by not including constraints corresponding to MCE. Since constraints corresponding to MCE involve only Pareto systems, including these constraints in the allocation heuristic may yield better allocations without adding much computational complexity. Second, while such a version of Problem $\widetilde{Q}$ has reduced complexity, when the number of Pareto systems is large, the number of constraints still grows with the number of non-Pareto systems. We avoid these issues by creating a new reduced version of Problem Q, called Problem $Q_{\mathbb{S}}$, that includes at least one constraint corresponding to MCI for each Pareto system and includes all constraints corresponding to MCE, as follows.

For each phantom Pareto system $\ell \in \mathcal{P}^{\text{ph}}, \ell = i$, find the "closest" non-Pareto systems

$$j_i^*(\ell) := \text{argmin}_{j\in\mathcal{P}^c}\{\mathbb{S}_j(\ell) : \mathbb{S}_j(\ell) \in \{K_j(h_i), I_j(g_{i+1}, h_i)\}\} \text{ if } \ell \neq 0,$$
$$j_{i+1}^*(\ell) := \text{argmin}_{j\in\mathcal{P}^c}\{\mathbb{S}_j(\ell) : \mathbb{S}_j(\ell) \in \{J_j(g_{i+1}), I_j(g_{i+1}, h_i)\}\} \text{ if } \ell \neq p,$$

and let $\mathcal{J}^*(\ell) := \{j_i^*(\ell)\} \cup \{j_{i+1}^*(\ell)\}$, where $\{j_0^*(0)\} := \emptyset$ and $\{j_{p+1}^*(p)\} := \emptyset$. The set $\mathcal{J}^*(\ell)$ contains up to two of the "closest" non-Pareto systems to phantom Pareto system ℓ, as determined by the scores. Because it is possible for a non-Pareto system j to bind with only one of the Pareto systems i or $i + 1$ through the phantom Pareto system ℓ, we ensure that we retain at least one non-Pareto system that binds with each Pareto system i and $i + 1$ in the set $\mathcal{J}^*(\ell)$. Then the SCORE allocation we recommend results from solving the following reduced problem for the Pareto system allocations:

$$\begin{aligned}
\text{Problem } Q_{\mathbb{S}}: \quad & \text{maximize } z \text{ s.t.} \\
& R_i(\alpha_i, \alpha_{i'}) \geq z \text{ for all } i, i' \in \mathcal{P} \text{ such that } i' \neq i, \\
& R_{j^*\ell}(\lambda_{j^*}^{\mathbb{S}}(1 - \textstyle\sum_{i=1}^{p} \alpha_i), \alpha_i, \alpha_{i+1}) \geq z \text{ for all } j^* \in \mathcal{J}^*(\ell), \ell \in \mathcal{P}^{\text{ph}}, \ell = i, \\
& \textstyle\sum_{i=1}^{p} \alpha_i \leq 1, \ \alpha_i \geq 0 \text{ for all } i \in \mathcal{P}.
\end{aligned}$$

Since we have at most two constraints corresponding to MCI for each phantom Pareto system, the complexity of Problem $Q_{\mathbb{S}}$ depends only on the number of Pareto systems.

To speed up the computation in Problem $Q_{\mathbb{S}}$, we use closed form expressions of the rate functions corresponding to MCI, presented in Proposition 4.1. The following Proposition 5.1 provides corresponding closed form expressions for the rate functions corresponding to MCE, written without weights. Notice that the correlations for both systems, ρ_i and $\rho_{i'}$, appear in the rate.

Proposition 5.1. *Under Assumption 6, the rate function corresponding to the* $\text{MCE}_{\mathcal{P}}$ *event for systems* $i, i' \in \mathcal{P}$ *is*

$$R_i(\alpha_i, \alpha_{i'}) = \begin{cases} \frac{(g_{i'}-g_i)^2}{2(\sigma_{g_i}^2/\alpha_i+\sigma_{g_{i'}}^2/\alpha_{i'})} & \textit{iff } g_{i'} > g_i, \ h_{i'} \leq h_i + (g_{i'} - g_i)\left(\frac{\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i+\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'}}{\sigma_{g_i}^2/\alpha_i+\sigma_{g_{i'}}^2/\alpha_{i'}}\right) \\ \frac{(h_{i'}-h_i)^2}{2(\sigma_{h_i}^2/\alpha_i+\sigma_{h_{i'}}^2/\alpha_{i'})} & \textit{iff } h_{i'} > h_i, \ g_{i'} \leq g_i + (h_{i'} - h_i)\left(\frac{\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i+\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'}}{\sigma_{h_i}^2/\alpha_i+\sigma_{h_{i'}}^2/\alpha_{i'}}\right) \\ \frac{\left[\frac{\sigma_{g_i}^2}{\alpha_i}+\frac{\sigma_{g_{i'}}^2}{\alpha_{i'}}\right](h_{i'}-h_i)^2-2\left[\frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}+\frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\right](g_{i'}-g_i)(h_{i'}-h_i)+\left[\frac{\sigma_{h_i}^2}{\alpha_i}+\frac{\sigma_{h_{i'}}^2}{\alpha_{i'}}\right](g_{i'}-g_i)^2}{2\left[(\sigma_{g_i}^2/\alpha_i+\sigma_{g_{i'}}^2/\alpha_{i'})(\sigma_{h_i}^2/\alpha_i+\sigma_{h_{i'}}^2/\alpha_{i'})-(\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i+\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2\right]} \\ \quad \textit{otherwise.} \end{cases}$$

6 TIME TO SOLVE FOR THE SCORE ALLOCATION VERSUS OPTIMALITY GAP

In practice, a decision-maker's choice of simulation budget allocation method is influenced by the amount of time it takes to solve for the allocation, as well as how close that allocation is to the optimal allocation. We now give a sense of how our proposed allocations perform on these metrics as the number of systems increases. (In this section, we assume all rate functions are known.)

For a population of ten problems generated according to the uniform method from §4.5, the following Table 2 reports the average wall-clock time to solve for each allocation, the average rate z achieved by the resulting allocation, and the average optimality gap. We keep the same Pareto systems as in §4.5, but instead of rejecting non-Pareto systems that are less than 0.25 units away, we reject non-Pareto systems that are less than 0.05 units away. Thus the problems are more realistic while keeping the full Problem Q solvable for up to a thousand systems. We also let all systems have bivariate normal rate functions with correlated objectives and unit variances. Within each of the 10 problem instances, all systems share the same correlation between the objectives. The correlations between the objectives for the ten problems, rounded to the second digit, are −0.81, −0.51, −0.36, −0.21, −0.08, 0.23, 0.26, 0.46, 0.55, 0.80. The specified allocation models in Table 2 are BVN True, in which we solve the full Problem Q for the asymptotically optimal allocation $\boldsymbol{\alpha}^*$; BVN Independent, in which we solve the full Problem Q, except we model the correlation between the objectives for all systems as $\rho_k = 0$ for all $k \in \mathcal{S}$; SCORE; the non-sequential MOCBA allocations [Lee et al. 2010, p. 661, Lemmas 4 and 5]; M-MOBA; and equal allocation.

The timings reported in Table 2 approximate how long it takes to perform one sample allocation update in the sequential algorithm (see §7, Algorithm 1, Step 3). Note that M-MOBA is a myopic procedure for which asymptotic allocations are not provided, so we are unable to report its rate. Further, the MOCBA allocations in Lee et al. [2010, p. 661, Lemmas 4 and 5] do not always ensure

Table 2. For ten problems randomly generated via the uniform method of §4.5, the table reports the average wall-clock time to solve for each allocation, as well as the average rate of decay of the probability of misclassification, z, and the average optimality gap, $z(\boldsymbol{\alpha}^*) - z(\boldsymbol{\alpha})$, for $\boldsymbol{\alpha}$ specified by each allocation.

r	Metric	BVN True	BVN Indep.	SCORE	MOCBA[†]	M-MOBA[‡]	Equal
20	Time	0.07 sec	0.06 sec	0.04 sec	0.01 sec	0.06 sec	0 sec
	Rate $z \times 10^4$	54.96	44.63	52.52	0.04		13.43
	Opt. Gap $\times 10^4$	0 [a]	10.33	2.44	54.92		41.53
100	Time	0.53 sec	0.47 sec	0.06 sec	0.02 sec	0.25 sec	0 sec
	Rate $z \times 10^4$	12.97	11.83	11.31	2.47		0.66
	Opt. Gap $\times 10^4$	0	1.14	1.66	10.50		12.31
500	Time	43.16 sec	25.06 sec	0.08 sec	0.42 sec	1.30 sec	0 sec
	Rate $z \times 10^4$	1.65	1.37	1.46	0.13		0.021
	Opt. Gap $\times 10^4$	0	0.28	0.19	1.52		1.63
1,000	Time	18.09 min	12.31 min	0.13 sec	1.67 sec	2.64 sec	0 sec
	Rate $z \times 10^4$	0.95	0.81	0.82	0.02		0.009
	Opt. Gap $\times 10^4$	0	0.14	0.13	0.93		0.94
2,000	Time	> 6 hr	> 6 hr	0.22 sec	6.71 sec	5.46 sec	0 sec
	Rate $z \times 10^4$	— [b]	—	0.45	0.005		0.004
	Opt. Gap $\times 10^4$	0	—	—	—		—
5,000	Time	> 6 hr	> 6 hr	0.48 sec	42.49 sec	14.80 sec	0 sec
	Rate $z \times 10^4$	—	—	0.22	0.0004		0.001
	Opt. Gap $\times 10^4$	0	—	—	—		—
10,000	Time	> 6 hr	> 6 hr	0.92 sec	2.90 min	33.68 sec	0 sec
	Rate $z \times 10^4$	—	—	0.11	0.0002		0.0006
	Opt. Gap $\times 10^4$	0	—	—	—		—

Note: We perform all computing on a 2.5 GHz Intel Core i7 processor with 16GB 1600MHz DDR3 memory. The algorithms for BVN True, BVN Independent, SCORE, and M-MOBA are written in MATLAB and run in MATLAB R2015b. The algorithm that calculates MOCBA is written in C++.

[†]We compare with Lee et al. [2010, p. 661, Lemmas 4 and 5], which may allocate $\alpha_k = 0$ for some $k \in \mathcal{S}$, implying $z = 0$.

[‡]We do not report a rate for M-MOBA, since it is a myopic procedure.

[a]The optimality gap is to the precision of the solver.

[b]The symbol '—' indicates that data is unavailable due to the large computational time.

the allocations to all systems are positive. If there exists a system k such that $\alpha_k = 0$, the rate is $z = 0$. This fact may be responsible for its relatively large theoretical optimality gap.

Interestingly, from Table 2, the BVN Independent allocation is much slower to calculate than SCORE. Further, it often yields an average optimality gap larger than SCORE, which emphasizes the usefulness of incorporating correlation into the allocation model. Since the BVN Independent allocation is not a competitive allocation relative to the others, we do not include the BVN Independent allocation in further numerical experiments. Table 2 also seems to show that SCORE is an extremely competitive allocation scheme whether the number of systems is small, e.g. on the order of 20 systems, or very large, e.g. on the order of 10,000 systems. Further, SCORE is fast — on average, it takes less than a second to solve for the SCORE allocation in a problem with 10,000 systems.

7 A SEQUENTIAL ALGORITHM FOR IMPLEMENTATION

Since the SCORE allocation framework requires knowledge of rate functions that we do not know in advance, we present sequential Algorithm 1 for implementation. The broad idea of Algorithm 1 is:

(a) obtain an initial sample of size $\delta_0 \geq 2$ from each system to estimate the SCORE allocation; (b) use the estimated SCORE allocation as a probability distribution from which to obtain the next $\delta \geq 1$ samples; (c) update the estimated optimal allocation and return to step (b). The minimum-sample proportion $\alpha_\varepsilon > 0$, which should be small relative to $1/r$, ensures that each system is sampled infinitely often as the sequential algorithm progresses. This algorithm proceeds until some total sampling budget specified by the user has been expended. Since implementing such a stopping rule is trivial, we write the sequential algorithm as non-terminating.

ALGORITHM 1: A sequential algorithm to sample from systems using the proposed allocations

Require: Initial sample size $\delta_0 \geq 2$; sample size between allocation vector updates $\delta \geq 1$; and a minimum-sample proportion $0 < \alpha_\varepsilon < 1/r$ that is small relative to $1/r$.

1: *Initialize:* collect δ_0 simulation replications from each system $k \in \mathcal{S}$; set $n = r \times \delta_0$, $n_k = \delta_0$ for all $k \in \mathcal{S}$.
2: *Estimate:* Update the parameter estimators $\hat{H}_k, \hat{G}_k, \hat{\sigma}^2_{g_k}, \hat{\sigma}^2_{h_k}$, and $\hat{\rho}_k$ for all $k \in \mathcal{S}$; use these estimators to construct estimated rate functions $\hat{I}_k(x_k, y_k)$ for all $k \in \mathcal{S}$;
3: *Calculate:* Solve an estimated version of Problem Q or $Q_{\mathbb{S}}$ using Step 2 estimators to obtain estimated optimal or SCORE allocations, $\hat{\boldsymbol{\alpha}}^*_n$.
4: **for** $m = 1, 2, \ldots, \delta$ **do**
5:     *Sample:* Select a system K_m from which to obtain the next simulation replication, where each K_m is an i.i.d. random variable with probability mass function $\hat{\boldsymbol{\alpha}}^*_n$ and support $\mathcal{S}$.
6:     *Simulate:* Collect one simulation replication from system K_m and set $n_{K_m} = n_{K_m} + 1$.
7: **end for**
8: Set $n = n + \delta$ and update $\bar{\boldsymbol{\alpha}}_n = \{n_1/n, n_2/n, \ldots, n_r/n\}$. Set $\delta^+ = 0$.
9: **for** $k = 1, 2, \ldots, r$ **do**
10:     If $n_k/n < \alpha_\varepsilon$, collect one simulation replication from system k. Then set $n_k = n_k + 1$ and $\delta^+ = \delta^+ + 1$.
11: **end for**
12: Set $n = n + \delta^+$ and go to Step 2.

8 NUMERICAL PERFORMANCE OF SEQUENTIAL ALLOCATIONS

In this section, we evaluate the performance of sequential versions of the proposed allocations on several test problems.

8.1 Test Problems

We construct six problems to test our algorithm. First, we generate two problem instances, 1 and 2, of true system performances by uniformly generating 100 systems in a circle of radius six, centered at $(100, 100)$. A listing of the (g_k, h_k) values for all $k \in \mathcal{S}$ is provided in Online Appendix N. We create sub-problems A, B, and C by setting the variances to one and the correlations to $\rho_k = -0.8$, $\rho_k = 0$, and $\rho_k = 0.8$ for all $k \in \mathcal{S}$, respectively. The system objective values in the first test problem set correspond to the circle centers in Figures 9–11. This test problem set has a high percent of dual variable values associated with MCE constraints. The second test problem set has a low percent of dual variable values associated with MCE constraints. Due to space constraints, the second test problem set and results appear in Online Appendix O.

Note that in Figures 9–11, the asymptotically optimal allocations are proportional to the size of the circle. While there is no obvious visible difference in the optimal allocations with different correlations, the allocations do differ slightly.

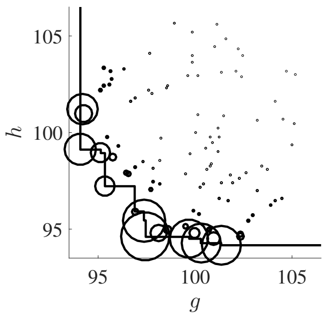

Fig. 9. Test 1A: $r = 100, |\mathcal{P}| = 9$, $\rho_k = -0.8$ for all $k \in \mathcal{S}$, % dual to MCE = 74.5, $z^* = 3.46 \times 10^{-4}$.

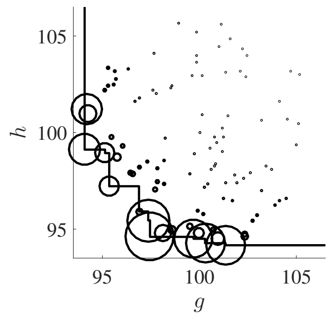

Fig. 10. Test 1B: $r = 100, |\mathcal{P}| = 9$, $\rho_k = 0$ for all $k \in \mathcal{S}$, % dual to MCE = 74.0, $z^* = 3.44 \times 10^{-4}$.

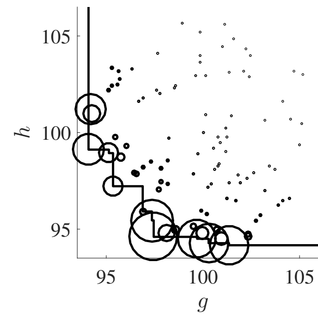

Fig. 11. Test 1C: $r = 100, |\mathcal{P}| = 9$, $\rho_k = 0.8$ for all $k \in \mathcal{S}$, % dual to MCE = 73.4, $z^* = 3.42 \times 10^{-4}$.

8.2 Estimated Expected Number of Misclassifications

For each algorithm BVN True, SCORE, MOCBA, M-MOBA, and equal allocation, we run 10, 000 independent sample paths on each of the test problems 1A, 1B, and 1C. For each algorithm, we calculate the average number of misclassifications, false exclusions, and false inclusions across the sample paths, as a function of sample size. For a particular sample path, the sequence containing the number of misclassifications as a function of the sample size n is autocorrelated.

In all implementations of Algorithm 1, which include all sample paths of the BVN True and SCORE allocations, we use parameter settings $\delta_0 = 5$, $\delta = 20$, and $\alpha_\varepsilon = 10^{-8}$. In our implementation of MOCBA [Lee et al. 2010], we use parameter settings $N_0 = \delta_0 = 5$, $\Delta = \delta = 20$, and $\tau = \Delta/2 = 10$, where τ is the maximum number of samples one system can receive in a given iteration. In our implementation of M-MOBA [Branke and Zhang 2015], we set $n_0 = \delta_0 = 5$ and $\tau = \delta = 20$, where here, τ is the amount of sample given to the alternative with the largest probability of changing the set of Pareto systems. Our stopping rule in M-MOBA is the sampling budget rule. We have chosen $\delta = 20$ as a reasonable sampling update schedule that is computationally feasible for all algorithms. M-MOBA, however, is designed for $\delta = 1$, since all of the samples between updates are allocated to a single system. Ideally we would run all algorithms with $\delta = 1$, unfortunately, doing so would require significant computational resources. The resulting performances are reported in Figures 12–14.

Considering the overall percent of systems misclassified, all algorithms exhibit close performance in Figures 12–14. (They exhibit even closer performance on the second test problem set in Figures 20–22 of Online Appendix O.) We notice that MOCBA seems to perform particularly well at preventing false exclusions of Pareto systems, but performs less well at preventing false inclusions of non-Pareto systems.

Since the optimality guarantees on the BVN True allocation are asymptotic, it is not clear that allocating according to BVN True will perform better than other allocation schemes for finite n. However, BVN True seems to perform about as well as its peers, and the performance of the SCORE allocation tracks the BVN True allocation closely. Importantly, since Test Problems 1A–1C have a high percent of dual constraints to MCE — implying the assumptions required in the limiting SCORE framework may not hold — we do not notice a loss of quality in the SCORE allocation relative to BVN True in Figures 12–14. We remind the reader that these test problems were randomly-generated; performance of the algorithms on other problems may vary.

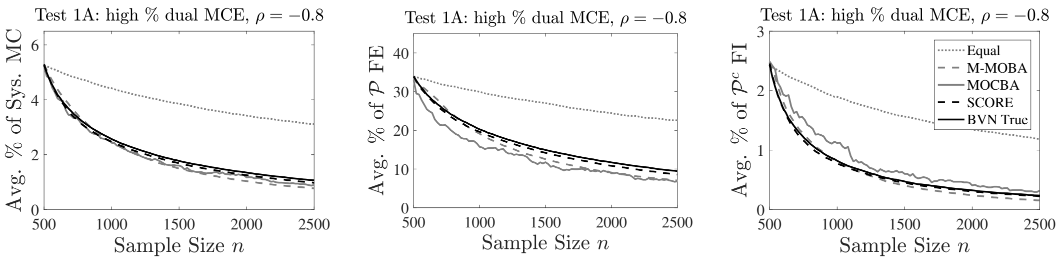

Fig. 12. Test 1A: For 10,000 sample paths per algorithm, the graphs show the average % of systems misclassified (MC), % of Paretos falsely excluded (FE), and % of non-Paretos falsely included (FI), respectively.

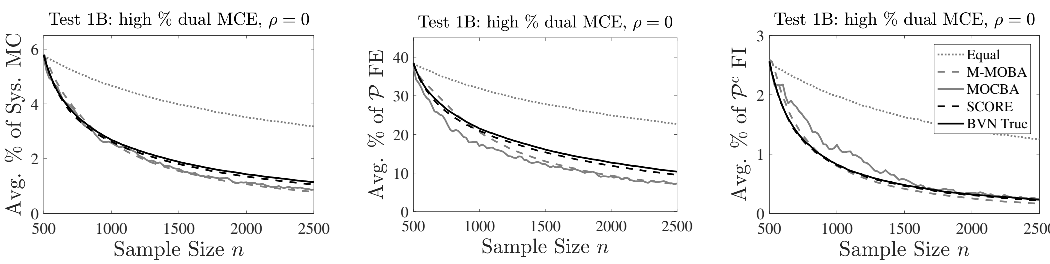

Fig. 13. Test 1B: For 10,000 sample paths per algorithm, the graphs show the average % of systems misclassified (MC), % of Paretos falsely excluded (FE), and % of non-Paretos falsely included (FI), respectively.

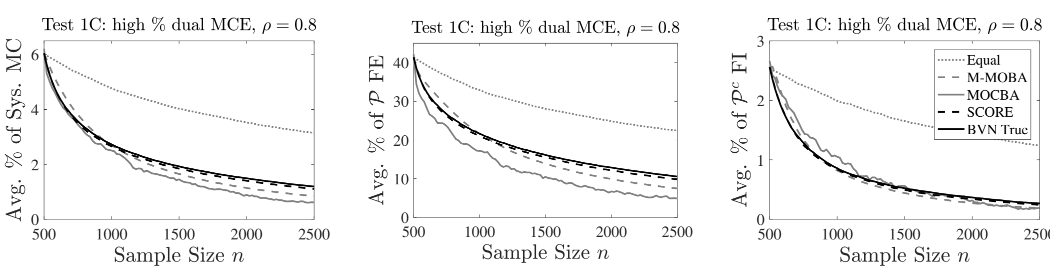

Fig. 14. Test 1C: For 10,000 sample paths per algorithm, the graphs show the average % of systems misclassified (MC), % of Paretos falsely excluded (FE), and % of non-Paretos falsely included (FI), respectively.

9 CONCLUDING REMARKS

SCORE is a fast, approximately optimal allocation for bi-objective R&S that accounts for correlation between the objectives and is derived from an asymptotically optimal allocation framework. We are aware of issues with estimating rate functions in a general context [Glynn and Juneja 2015, 2011]. However, our numerical experience in the case of normal rate functions has been overwhelmingly positive [Hunter and McClosky 2016; Pasupathy et al. 2015]. Finally, it remains to be seen whether our methods for bi-objective R&S extend cleanly to multi-objective R&S. We rely heavily on the phantom Pareto systems, which are easily constructed only in two objectives. Feldman et al. [2015] and Feldman [2017] provide further insights to the multi-objective case.

SUPPLEMENTARY MATERIALS

Our supplementary materials include the following Online Appendices A–O.

A PROOF OF THEOREM 3.1

For use in the proof, we expand the MC and MC_{ph} events. First, expand the MCE event, so that $\text{MCE} = \text{MCE}_{\mathcal{P}} \cup \text{MCE}_{\mathcal{P}^c}$, where

$$\text{MCE}_{\mathcal{P}^c} := \bigcup_{i\in\mathcal{P}} \bigcup_{j\in\mathcal{P}^c} (\hat{G}_j \le \hat{G}_i) \cap (\hat{H}_j \le \hat{H}_i).$$

Then we can write the MC event as

$$\begin{aligned}\text{MC} &= \text{MCE}_{\mathcal{P}} \cup \text{MCE}_{\mathcal{P}^c} \cup \text{MCI}\\ &= \text{MCE}_{\mathcal{P}} \cup [(\text{MCI} \cup \text{MCE}_{\mathcal{P}^c}) \cap \text{MCE}^c_{\mathcal{P}}]\\ &= \text{MCE}_{\mathcal{P}} \cup [(\text{MCI} \cap \text{MCE}^c_{\mathcal{P}}) \cup (\text{MCE}_{\mathcal{P}^c} \cap \text{MCE}^c_{\mathcal{P}})].\end{aligned}$$

Then, notice that the MC_{ph} event can be written as

$$\text{MC}_{\text{ph}} = \text{MCE}_{\mathcal{P}} \cup \text{MCI}_{\text{ph}} = \text{MCE}_{\mathcal{P}} \cup (\text{MCI}_{\text{ph}} \cap \text{MCE}^c_{\mathcal{P}}).$$

(MC implies MC_{ph}). We show MC implies MC_{ph} in two parts. First, we show $\text{MCI} \cap \text{MCE}^c_{\mathcal{P}}$ implies MCI_{ph}, then we show $\text{MCE}_{\mathcal{P}^c} \cap \text{MCE}^c_{\mathcal{P}}$ implies MCI_{ph}.

Suppose $\text{MCI} \cap \text{MCE}^c_{\mathcal{P}}$ occurs. Since MCI occurs, let $j \in \mathcal{P}^c$ and $j \in \hat{\mathcal{P}}$ be a non-Pareto system falsely estimated as Pareto. Then for each $i \in \mathcal{P}$, $\hat{G}_j \le \hat{G}_i$ or $\hat{H}_j \le \hat{H}_i$. Thus $(\hat{G}_j, \hat{H}_j) \in \cap_{i\in\mathcal{P}}\{(\hat{G}, \hat{H}) : (\hat{G} \le \hat{G}_i) \cup (\hat{H} \le \hat{H}_i)\}$. Since $\text{MCE}^c_{\mathcal{P}}$ occurs, no Pareto systems dominate other Pareto systems, and we have $p + 1$ estimated phantom Pareto systems, indexed by $\ell \in \mathcal{P}^{\text{ph}}$, $\ell = i$. Therefore

$$(\hat{G}_j, \hat{H}_j) \in \bigcup_{\ell\in\mathcal{P}^{\text{ph}}, \ell=i} \{(\hat{G}, \hat{H}) : (\hat{G} \le \hat{G}_{[i+1]}) \cap (\hat{H} \le \hat{H}_{[i]})\},$$

that is, j lies in the union of the southwest quadrants defined by origins at the estimated phantom systems. Therefore MCI_{ph} occurs.

Now we show $\text{MCE}_{\mathcal{P}^c} \cap \text{MCE}^c_{\mathcal{P}}$ implies MCI_{ph}. Suppose $\text{MCE}_{\mathcal{P}^c} \cap \text{MCE}^c_{\mathcal{P}}$ occurs. Since $\text{MCE}_{\mathcal{P}^c}$ occurs, there exists a non-Pareto system $j \in \mathcal{P}^c$ such that for some Pareto system $i \in \mathcal{P}$, $(\hat{G}_j \le \hat{G}_i) \cap (\hat{H}_j \le \hat{H}_i)$. Suppose that, among the Pareto systems, $i \in \mathcal{P}$ is estimated as being in the $(i' + 1)$th place on the g objective for some $i' \in \{0, 1, \ldots, p\}$, so that $\hat{G}_j \le \hat{G}_{[i'+1]}$. Since $\text{MCE}^c_{\mathcal{P}}$ occurs, system i is also in the $(i' + 1)$th place on the h objective, thus $\hat{H}_j \le \hat{H}_{[i'+1]} \le \hat{H}_{[i']}$. Therefore MCI_{ph} occurs.

(MC_{ph} implies MC). We show $\text{MCI}_{\text{ph}} \cap \text{MCE}^c_{\mathcal{P}}$ implies MC. Suppose $\text{MCI}_{\text{ph}} \cap \text{MCE}^c_{\mathcal{P}}$ occurs. Then no Pareto systems are estimated as dominating other Pareto systems, and at least one non-Pareto system is estimated as dominating one of the $p + 1$ phantom Pareto systems. From the set of all $j \in \mathcal{P}^c$ dominating some estimated phantom Pareto system, there exists $j^* \in \mathcal{P}^c$ such that $j^* \in \hat{\mathcal{P}}$. (Otherwise, if there exists no such j^*, then each $j \in \mathcal{P}^c$ is dominated by some $i \in \mathcal{P}$, and MCI_{ph} does not occur.) Therefore $\hat{\mathcal{P}} \ne \mathcal{P}$, which implies MC.

B PROOF OF LEMMA 3.3

Let $\mathcal{S}$ be an ordered set of p elements of the form $\{(x_1, y_1), (x_2, y_2), \ldots, (x_p, y_p)\}$ where the x and y coordinates are separately drawn without replacement from the set of Pareto indices $\{1, 2, \ldots, p\}$. The set $\mathcal{S}$ corresponds to a (fixed) realized instance of $\hat{\mathcal{O}}$.

By the law of total probability,

$$\mathbf{P}\{\text{MCI}_{\text{ph}}\} = \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\} + \textstyle\sum_{\text{all } \mathcal{S}\ne\mathcal{O}} \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\}.$$

Consider the first term on the right hand side, $\mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\}$. Since $\hat{\mathcal{O}} = \mathcal{O}$ occurs, we may write MCI_{ph} without order statistics so that

$$\mathbf{P}\{\text{MCI}_{\text{ph}}\} = \mathbf{P}\{\text{MCI}^*_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\} + \textstyle\sum_{\mathcal{S} \neq \mathcal{O}} \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\}.$$

Assuming the limits exist, by the principle of the slowest term [Ganesh et al. 2004, Lemma 2.1],

$$\begin{aligned}
&- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}_{\text{ph}}\} \\
&= - \lim_{n\to\infty} \tfrac{1}{n} \log\big(\mathbf{P}\{\text{MCI}^*_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\} + \textstyle\sum_{\mathcal{S} \neq \mathcal{O}} \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\}\big) \\
&= \min\big(- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}^*_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\}, \min_{\mathcal{S}\neq\mathcal{O}} - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\}\big).
\end{aligned}$$

Then from equation (1),

$$\begin{aligned}
- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MC}\} = \min\big(&- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCE}_{\mathcal{P}}\}, - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}^*_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\}, \\
&\min_{\mathcal{S}\neq\mathcal{O}} - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\}\big). \qquad (14)
\end{aligned}$$

We now show that $\min_{\mathcal{S}\neq\mathcal{O}} - \lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\}$ is never the binding minimum in equation (14). First, it is clear that any value of $\mathcal{S}$ that results in an $\text{MCE}_{\mathcal{P}}$ event will have a rate function that is greater than or equal to the corresponding rate function for $\text{MCE}_{\mathcal{P}}$. Now consider values of $\mathcal{S}$ that do *not* result in $\text{MCE}_{\mathcal{P}}$, such that all Pareto systems are estimated as Pareto, but may be estimated in the wrong *order*, e.g., $\mathcal{S} = \{(2,2),(1,1),(3,3),\ldots,(p,p)\}$, where Pareto systems 1 and 2 have exchanged positions. In this case, it is sufficient to consider only instances of $\mathcal{S}$ that contain pairwise exchanges, and further, it is sufficient to consider instances of $\mathcal{S}$ in which there is exactly one "pair exchange." For any i_1, i_2 indexing the Pareto systems such that $i_1 < i_2$, a pair exchange occurs if $(\hat{G}_{i_2} \leq \hat{G}_{i_1}) \cap (\hat{H}_{i_1} \leq \hat{H}_{i_2})$. Let $\mathcal{S}_{(i_1,i_2)}$ denote an ordering with exactly one pair exchange where i_1 and i_2, $i_1 < i_2$, are the Pareto systems whose places have been exchanged. Thus

$$\begin{aligned}
&\min_{\mathcal{S}\neq\mathcal{O}} - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}\} \\
&\geq \min\big(- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCE}_{\mathcal{P}}\}, - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCI}_{\text{ph}} \cap \hat{\mathcal{O}} = \mathcal{S}_{(i_1,i_2)}\}\big) \\
&\geq \min\big(- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\text{MCE}_{\mathcal{P}}\}, - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\hat{\mathcal{O}} = \mathcal{S}_{(i_1,i_2)}\}\big).
\end{aligned}$$

Now notice that

$$\begin{aligned}
&- \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\hat{\mathcal{O}} = \mathcal{S}_{(i_1,i_2)}\} \geq \min_{i_1,i_2\in\mathcal{P}} - \lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{(\hat{G}_{i_2} \leq \hat{G}_{i_1}) \cap (\hat{H}_{i_1} \leq \hat{H}_{i_2})\} \\
&= \min_{i_1,i_2\in\mathcal{P}} \inf_{x_{i_2}\leq x_{i_1},\ y_{i_1}\leq y_{i_2}} \alpha_{i_1} I_{i_1}(x_{i_1}, y_{i_1}) + \alpha_{i_2} I_{i_2}(x_{i_2}, y_{i_2}). \qquad (15)
\end{aligned}$$

Since the infimum in equation (15) is a strictly convex minimization problem with a nonempty interior, the KKT conditions [Boyd and Vandenberghe 2004] are necessary and sufficient for global optimality. In addition to primary feasibility conditions, for $\lambda_x \geq 0$ and $\lambda_y \geq 0$ we have the complementary slackness conditions $\lambda_x(x^*_{i_2} - x^*_{i_1}) = 0$ and $\lambda_y(y^*_{i_1} - y^*_{i_2}) = 0$, and the stationarity conditions

$$\begin{aligned}
\alpha_{i_1} \frac{\partial I_{i_1}(x^*_{i_1}, y^*_{i_1})}{\partial x_{i_1}} - \lambda_x &= 0, & \alpha_{i_1} \frac{\partial I_{i_1}(x^*_{i_1}, y^*_{i_1})}{\partial y_{i_1}} + \lambda_y &= 0, \\
\alpha_{i_2} \frac{\partial I_{i_2}(x^*_{i_2}, y^*_{i_2})}{\partial x_{i_2}} + \lambda_x &= 0, & \alpha_{i_2} \frac{\partial I_{i_2}(x^*_{i_2}, y^*_{i_2})}{\partial y_{i_2}} - \lambda_y &= 0.
\end{aligned}$$

Since $\lambda_x = \lambda_y = 0$ implies that $(x^*_{i_1}, y^*_{i_1}) = (g_{i_1}, h_{i_1})$ and $(x^*_{i_2}, y^*_{i_2}) = (g_{i_2}, h_{i_2})$, which is infeasible, then it must be the case that $\lambda_x > 0$ or $\lambda_y > 0$. Recall that both i_1 and i_2 are Pareto systems. First, suppose that $\lambda_y > 0$ so that $y^*_{i_1} = y^*_{i_2}$. Then continuing from line (15),

$$(15) = \min_{i_1, i_2 \in \mathcal{P}} \inf_{x_{i_2} \le x_{i_1},\ y_{i_1} = y_{i_2}} \alpha_{i_1} I_{i_1}(x_{i_1}, y_{i_1}) + \alpha_{i_2} I_{i_2}(x_{i_2}, y_{i_2})$$
$$\ge \min_{i_1, i_2 \in \mathcal{P}} \inf_{x_{i_2} \le x_{i_1},\ y_{i_2} \le y_{i_1}} \alpha_{i_1} I_{i_1}(x_{i_1}, y_{i_1}) + \alpha_{i_2} I_{i_2}(x_{i_2}, y_{i_2}) \ge -\lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\},$$

and the result follows. A similar proof holds for the case in which $\lambda_x > 0$.

C PROOF OF THEOREM 3.4

Starting from Lemma 3.3, we now show that

$$-\lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\mathrm{MCI}^*_{\mathrm{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\}$$
$$\ge \min\left(-\lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}, \min_{j \in \mathcal{P}^c} \min_{\ell \in \mathcal{P}^{\mathrm{ph}}, \ell = i} R_j(\alpha_j, \alpha_i, \alpha_{i+1})\right),$$

which, together with Lemma 3.2, implies the result. Recall that, without loss of generality, we reserve the indices $1, \ldots, p$ for the Pareto systems, and the indices $p+1, \ldots, r$ for the non-Pareto systems.

First, notice that we may write the rate function for $\mathbf{P}\{\mathrm{MCI}^*_{\mathrm{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\}$ as

$$-\lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\mathrm{MCI}^*_{\mathrm{ph}} \cap \hat{\mathcal{O}} = \mathcal{O}\} = \min_{j \in \mathcal{P}^c} \min_{\ell \in \mathcal{P}^{\mathrm{ph}}, \ell = i} \Big(
-\lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{\underbrace{(\hat{G}_j \le \hat{G}_{i+1}) \cap (\hat{H}_j \le \hat{H}_i)}_{j \in \mathcal{P}^c \text{ dom. a phantom}} \cap \underbrace{[\cap_{i'=1}^{p-1} (\hat{G}_{i'} \le \hat{G}_{i'+1}) \cap (\hat{H}_{i'+1} \le \hat{H}_{i'})]}_{\text{all Paretos estimated ``in order''}}\}\Big). \tag{16}$$

Let $j \in \mathcal{P}^c$, $\ell \in \mathcal{P}^{\mathrm{ph}}$, $\ell = i$, and consider the inner rate function from line (16),

$$-\lim_{n\to\infty} \tfrac{1}{n} \log \mathbf{P}\{(\hat{G}_j \le \hat{G}_{i+1}) \cap (\hat{H}_j \le \hat{H}_i) \cap [\cap_{i'=1}^{p-1} (\hat{G}_{i'} \le \hat{G}_{i'+1}) \cap (\hat{H}_{i'+1} \le \hat{H}_{i'})]\}. \tag{17}$$

Then it can be shown that

$$(17) = \begin{cases} \inf\limits_{\substack{x_j \le x_1 \\ x_1 \le \ldots \le x_p,\ y_p \le \ldots \le y_1}} \alpha_j I_j(x_j, y_j) + \sum_{i'=1}^{p} \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) & \text{if } \ell = 0 \\ \inf\limits_{\substack{x_j \le x_{i+1},\ y_j \le y_i \\ x_1 \le \ldots \le x_p,\ y_p \le \ldots \le y_1}} \alpha_j I_j(x_j, y_j) + \sum_{i'=1}^{p} \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) & \text{if } \ell \in \{1, \ldots, p-1\} \\ \inf\limits_{\substack{y_j \le y_p \\ x_1 \le \ldots \le x_p,\ y_p \le \ldots \le y_1}} \alpha_j I_j(x_j, y_j) + \sum_{i'=1}^{p} \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) & \text{if } \ell = p. \end{cases} \tag{18}$$

Since all problems in (18) are convex minimization problems where Slater's condition holds [Boyd and Vandenberghe 2004], the KKT conditions are necessary and sufficient for global optimality. Then from (18), for $\lambda_{x_j} \ge 0$; $\lambda_{y_j} \ge 0$; $\lambda_{x_{i'}} \ge 0$ for all $i' = 1, \ldots, p-1$; and $\lambda_{y_{i'}} \ge 0$ for all $i' = 2, \ldots, p$, the KKT conditions for complementary slackness are

$$\lambda_{x_j}(x^*_j - x^*_{i+1}) = 0 \text{ if } \ell \ne p; \qquad \lambda_{y_j}(y^*_j - y^*_i) = 0 \text{ if } \ell \ne 0;$$
$$\lambda_{x_{i'}}(x^*_{i'} - x^*_{i'+1}) = 0 \ \ \forall i' = 1, \ldots, p-1; \qquad \lambda_{y_{i'+1}}(y^*_{i'+1} - y^*_{i'}) = 0 \ \ \forall i' = 1, \ldots, p-1.$$

Let $\lambda_{x_0} := 0, \lambda_{x_p} := 0, \lambda_{y_1} := 0, \lambda_{y_{p+1}} := 0$. For all $i' \in \{2, \ldots, p-1\}$, the stationarity conditions are

$$\alpha_j \frac{\partial I_j(x_j^*, y_j^*)}{\partial x_j} + \lambda_{x_j} \mathbb{I}_{[\ell \neq p]} = 0, \qquad \alpha_j \frac{\partial I_j(x_j^*, y_j^*)}{\partial y_j} + \lambda_{y_j} \mathbb{I}_{[\ell \neq 0]} = 0,$$

$$\text{if } \ell \neq p,\ \alpha_{i+1} \frac{\partial I_{i+1}(x_{i+1}^*, y_{i+1}^*)}{\partial x_{i+1}} - \lambda_{x_j} - \lambda_{x_i} + \lambda_{x_{i+1}} = 0, \quad \text{if } \ell \neq 0,\ \alpha_i \frac{\partial I_i(x_i^*, y_i^*)}{\partial y_i} - \lambda_{y_j} - \lambda_{y_{i+1}} + \lambda_{y_i} = 0,$$

$$\text{if } \ell \neq 0,\ \alpha_1 \frac{\partial I_1(x_1^*, y_1^*)}{\partial x_1} + \lambda_{x_1} = 0, \qquad \text{if } \ell \neq 1,\ \alpha_1 \frac{\partial I_1(x_1^*, y_1^*)}{\partial y_1} - \lambda_{y_2} = 0,$$

$$\text{if } i' \neq i+1,\ \alpha_{i'} \frac{\partial I_{i'}(x_{i'}^*, y_{i'}^*)}{\partial x_{i'}} - \lambda_{x_{i'-1}} + \lambda_{x_{i'}} = 0, \qquad \text{if } i' \neq i,\ \alpha_{i'} \frac{\partial I_{i'}(x_{i'}^*, y_{i'}^*)}{\partial y_{i'}} - \lambda_{y_{i'+1}} + \lambda_{y_{i'}} = 0,$$

$$\text{if } \ell \neq p-1,\ \alpha_p \frac{\partial I_p(x_p^*, y_p^*)}{\partial x_p} - \lambda_{x_{p-1}} = 0, \qquad \text{if } \ell \neq p,\ \alpha_p \frac{\partial I_p(x_p^*, y_p^*)}{\partial y_p} + \lambda_{y_p} = 0.$$

Now suppose that at optimality in (18), there exists $i' \in \{1, \ldots, p-1\}$ such that $x_{i'}^* = x_{i'+1}^*$. Then removing constraints in (18),

$$\begin{aligned}(18) &\geq \inf_{x_{i'}^* = x_{i'+1}^*,\ y_{i'+1} \leq y_{i'}} \alpha_{i'} I_{i'}(x_{i'}^*, y_{i'}) + \alpha_{i'+1} I_{i'+1}(x_{i'+1}^*, y_{i'+1}) \\ &\geq \inf_{x_{i'} \geq x_{i'+1},\ y_{i'+1} \leq y_{i'}} \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) + \alpha_{i'+1} I_{i'+1}(x_{i'+1}, y_{i'+1}) \geq -\lim_{n \to \infty} \tfrac{1}{n} \log \mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}.\end{aligned}$$

By a similar argument, it also can be shown that if there exists $i' \in \{2, \ldots, p\}$ such that $y_{i'}^* = y_{i'-1}^*$, then the rate in (18) is bounded below by the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$. Therefore it is sufficient to consider only $\lambda_{x_{i'}} = 0$ for all $i' = 1, \ldots, p-1$ and $\lambda_{y_{i'}} = 0$ for all $i' = 2, \ldots, p$; otherwise, the rate would never be the unique minimum in the overall rate of decay of $\mathbf{P}\{\mathrm{MC}\}$.

Using this information in the KKT conditions, we consider only the case $(x_{i'}^*, y_{i'}^*) = (h_{i'}, g_{i'})$ for all $i' = 1, \ldots, p, i' \neq i, i' \neq i+1$. Then it is sufficient to simplify (18) and consider only

$$\begin{cases} \inf\limits_{\substack{x_j \leq x_1 \\ x_1 < g_2,\ h_2 < y_1}} \alpha_j I_j(x_j, y_j) + \alpha_1 I_1(x_1, y_1) & \text{if } \ell = 0 \\ \inf\limits_{\substack{x_j \leq x_{i+1},\ y_j \leq y_i \\ g_{i-1} < x_i < x_{i+1} < g_{i+2}, \\ h_{i+2} < y_{i+1} < y_i < h_{i-1}}} \alpha_j I_j(x_j, y_j) + \sum_{i'=i}^{i+1} \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) & \text{if } \ell \in \{1, \ldots, p-1\} \\ \inf\limits_{\substack{y_j \leq y_p \\ g_{p-1} < x_p,\ y_p < h_{p-1}}} \alpha_j I_j(x_j, y_j) + \alpha_p I_p(x_p, y_p) & \text{if } \ell = p. \end{cases} \tag{19}$$

As before, if any of the constraints that ensure the Pareto systems are estimated in order are violated, then rate is bounded below by the rate of decay of $\mathbf{P}\{\mathrm{MCE}_{\mathcal{P}}\}$. Then we further simplify the rate in (19) to consider only

$$\begin{cases} \inf\limits_{x_j \leq x_1} & \alpha_j I_j(x_j, y_j) + \alpha_1 I_1(x_1, y_1) & \text{if } \ell = 0 \\ \inf\limits_{x_j \leq x_{i+1},\ y_j \leq y_i} & \alpha_j I_j(x_j, y_j) + \sum_{i'=i}^{i+1} \alpha_{i'} I_{i'}(x_{i'}, y_{i'}) & \text{if } \ell \in \{1, \ldots, p-1\} \\ \inf\limits_{y_j \leq y_p} & \alpha_j I_j(x_j, y_j) + \alpha_p I_p(x_p, y_p) & \text{if } \ell = p \end{cases}$$

$$= \begin{cases} \inf\limits_{x_j \leq x_1} & \alpha_j \inf\limits_{y_j} I_j(x_j, y_j) + \alpha_1 \inf\limits_{y_1} I_1(x_1, y_1) & \text{if } \ell = 0 \\ \inf\limits_{\substack{x_j \leq x_{i+1}, \\ y_j \leq y_i}} & \alpha_j I_j(x_j, y_j) + \alpha_i \inf\limits_{x_i} I_i(x_i, y_i) + \alpha_{i+1} \inf\limits_{y_{i+1}} I_{i+1}(x_{i+1}, y_{i+1}) & \text{if } \ell \in \{1, \ldots, p-1\} \\ \inf\limits_{y_j \leq y_p} & \alpha_j \inf\limits_{x_j} I_j(x_j, y_j) + \alpha_p \inf\limits_{x_p} I_p(x_p, y_p) & \text{if } \ell = p \end{cases} \tag{20}$$

$$= \begin{cases} \inf\limits_{x_j \leq x_1} & \alpha_j J_j(x_j) + \alpha_1 J_1(x_1) & \text{if } \ell = 0 \\ \inf\limits_{x_j \leq x_{i+1},\ y_j \leq y_i} & \alpha_j I_j(x_j, y_j) + \alpha_i K_i(y_i) + \alpha_{i+1} J_{i+1}(x_{i+1}) & \text{if } \ell \in \{1, \ldots, p-1\} \\ \inf\limits_{y_j \leq y_p} & \alpha_j K_j(y_j) + \alpha_p K_p(y_p) & \text{if } \ell = p \end{cases} \tag{21}$$

$$= \begin{cases} \inf_{x} & \alpha_j J_j(x) + \alpha_1 J_1(x) & \text{if } \ell = 0 \\ \inf_{x_j \le x_{i+1},\ y_j \le y_i} & \alpha_j I_j(x_j, y_j) + \alpha_i K_i(y_i) + \alpha_{i+1} J_{i+1}(x_{i+1}) & \text{if } \ell \in \{1, \dots, p-1\} \\ \inf_{y} & \alpha_j K_j(y) + \alpha_p K_p(y) & \text{if } \ell = p \end{cases} \quad (22)$$
$$= R_{j\ell}(\alpha_j, \alpha_i, \alpha_{i+1}),$$

where equality in (20) holds by Boyd and Vandenberghe [2004, p. 133]. Since $I_k(x, y)$ is a good, strictly convex rate function for all systems $k \in \mathcal{S}$, equality in (21) holds by letting $x = f(x, y)$ in the contraction principle [Dembo and Zeitouni 1998, p. 126]. Equality in (22) follows from Glynn and Juneja [2004].

D STATEMENT OF LEMMA D.1 WITH PROOF

Lemma D.1. *Suppose $\alpha_j > 0, \alpha_i > 0, \alpha_{i+1} > 0$. At optimality in Problem $R^{\text{MCI}}_{j\ell}$, if $\ell \in \{1, \dots, p-1\}$, then $g_{i+1} \le x^*_{i+1}$, $h_i \le y^*_i$, and $x^*_j \le g_j$ or $y^*_j \le h_j$. Further, if $\lambda_x = 0$, then $y^*_j \le h_j$, and if $\lambda_y = 0$, then $x^*_j \le g_j$.*

Proof. Since $J_{i+1}(\cdot)$ is convex, $\left(\frac{\partial J_{i+1}(g_{i+1})}{\partial x_{i+1}} - \frac{\partial J_{i+1}(x^*_{i+1})}{\partial x_{i+1}}\right)(g_{i+1} - x^*_{i+1}) \ge 0$, which, together with the KKT conditions, implies $-\frac{\partial J_{i+1}(x^*_{i+1})}{\partial x_{i+1}}(g_{i+1} - x^*_{i+1}) \ge 0$. Thus $g_{i+1} \le x^*_{i+1}$. A similar proof shows that $-\frac{\partial K_i(y^*_i)}{\partial y_i}(h_i - y^*_i) \ge 0$. Thus $h_i \le y^*_i$. Using similar logic, by the convexity of $I_j(\cdot)$, we have $-\frac{\partial I_j(x^*_j, y^*_j)}{\partial x_j}(g_j - x^*_j) - \frac{\partial I_j(x^*_j, y^*_j)}{\partial y_j}(h_j - y^*_j) \ge 0$. Since the KKT conditions imply both partial derivatives are non-positive, then $x^*_j \le g_j$ or $y^*_j \le h_j$, and the implications when $\lambda_x = 0$ or $\lambda_y = 0$ follow. □

E PROOF OF PROPOSITION 4.1

First, recall that

$$\text{Problem } R^{\text{MCI}}_{j\ell}: \quad \text{minimize } \alpha_j I_j(x_j, y_j) + \alpha_i K_i(y_i)\mathbb{I}_{[\ell \ne 0]} + \alpha_{i+1} J_{i+1}(x_{i+1})\mathbb{I}_{[\ell \ne p]}$$
$$\text{s.t. } (x_j - x_{i+1})\mathbb{I}_{[\ell \ne p]} \le 0, \quad (y_j - y_i)\mathbb{I}_{[\ell \ne 0]} \le 0.$$

Under Assumption 6, the stationarity conditions from lines (2)–(3) are

$$\frac{\alpha_j}{(1-\rho_j^2)}\left[\frac{(x^*_j - g_j)}{\sigma^2_{g_j}} - \rho_j \frac{(y^*_j - h_j)}{\sigma_{g_j}\sigma_{h_j}}\right] + \lambda_x \mathbb{I}_{[\ell \ne p]} = 0, \qquad \frac{\alpha_j}{(1-\rho_j^2)}\left[\frac{(y^*_j - h_j)}{\sigma^2_{h_j}} - \rho_j \frac{(x^*_j - g_j)}{\sigma_{g_j}\sigma_{h_j}}\right] + \lambda_y \mathbb{I}_{[\ell \ne 0]} = 0,$$
$$\text{if } \ell \ne p,\ \alpha_{i+1}\left[\frac{(x^*_{i+1} - g_{i+1})}{\sigma^2_{g_{i+1}}}\right] - \lambda_x = 0, \qquad \text{if } \ell \ne 0,\ \alpha_i\left[\frac{(y^*_i - h_i)}{\sigma^2_{h_i}}\right] - \lambda_y = 0,$$

and complementary slackness implies $\lambda_x(x^*_j - x^*_{i+1}) = 0$ if $\ell \ne p$, $\lambda_y(y^*_j - y^*_i) = 0$ if $\ell \ne 0$. Notice that if $\lambda_x \mathbb{I}_{[\ell \ne p]} = \lambda_y \mathbb{I}_{[\ell \ne 0]} = 0$, then we have primal infeasibility.

In the sections that follow, we provide forward and backward proofs by considering each case.

E.1 Proof of Proposition 4.1: Forward

We consider the each possible value of λ_x and λ_y as follows.

E.1.1 Problem $R^{\text{MCI}}_{j\ell}$: $\lambda_x \mathbb{I}_{[\ell \ne p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell \ne 0]} = 0$. Suppose $\lambda_x \mathbb{I}_{[\ell \ne p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell \ne 0]} = 0$.

$$x^*_j = x^*_{i+1} = (1 - w_g)g_j + w_g g_{i+1}, \quad y^*_j = h_j + \rho_j \frac{\sigma_{h_j}}{\sigma_{g_j}}(g_{i+1} - g_j)w_g, \quad y^*_i = h_i,$$

which implies the results in Parts (1) and (2) in this case. Primal feasibility of y^*_j implies $h_j \le h_i + \rho_j \frac{\sigma_{h_j}}{\sigma_{g_j}}(g_j - g_{i+1})w_g$, and $x^*_{i+1} > g_{i+1}$ implies $g_j > g_{i+1}$.

E.1.2 Problem $R^{\text{MCI}}_{j\ell}$: $\lambda_x \mathbb{I}_{[\ell\neq p]} = 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$. Suppose $\lambda_x \mathbb{I}_{[\ell\neq p]} = 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$.

$$x^*_{i+1} = g_{i+1}, \quad x^*_j = g_j + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_i - h_j)w_h, \quad y^*_j = y^*_i = (1-w_h)h_j + w_h h_i,$$

which implies the results in Parts (1) and (2) in this case. Primal feasibility of x^*_j implies $g_j \leq g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$, and $y^*_i > h_i$ implies $h_j > h_i$.

E.1.3 Problem $R^{\text{MCI}}_{j\ell}$: $\lambda_x \mathbb{I}_{[\ell\neq p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$. Suppose $\lambda_x \mathbb{I}_{[\ell\neq p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$.

$$x^*_j = x^*_{i+1} = \tfrac{1}{1-\rho_j^2 w_g w_h}\left[(1-w_g)(g_j + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_i - h_j)w_h) + w_g(1-\rho_j^2 w_h)g_{i+1}\right],$$
$$y^*_j = y^*_i = \tfrac{1}{1-\rho_j^2 w_g w_h}\left[(1-w_h)(h_j + \rho_j \tfrac{\sigma_{h_j}}{\sigma_{g_j}}(g_{i+1} - g_j)w_g) + w_h(1-\rho_j^2 w_g)h_i\right],$$

which implies the results in Parts (1) and (2) in this case. Algebra reveals that $x^*_{i+1} > g_{i+1}$ and $y^*_i > h_i$ imply $g_j > g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$ and $h_j > h_i + \rho_j \tfrac{\sigma_{h_j}}{\sigma_{g_j}}(g_j - g_{i+1})w_g$.

E.2 Proof of Proposition 4.1: Backward

We consider each possible system location, as follows.

E.2.1 Problem $R^{\text{MCI}}_{j\ell}$: $\ell \neq p, g_j > g_{i+1}, h_j \leq h_i + \rho_j \tfrac{\sigma_{h_j}}{\sigma_{g_j}}(g_j - g_{i+1})w_g$. Suppose $\ell \neq p, g_j > g_{i+1}, h_j \leq h_i + \rho_j \tfrac{\sigma_{h_j}}{\sigma_{g_j}}(g_j - g_{i+1})w_g$. First, since $\ell \neq p, g_j > g_{i+1}$, it follows that $x^*_j = g_j$, $x^*_{i+1} = g_{i+1}$ is infeasible. If we nonetheless have $x^*_{i+1} = g_{i+1}$, then $\lambda_x \mathbb{I}_{[\ell\neq p]} = 0$, $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$. However, from §E.1.2, this implies $g_j \leq g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$ and $h_j > h_i$, which provides a contradiction. Therefore it must hold that $x^*_{i+1} > g_{i+1}$, and $\lambda_x \mathbb{I}_{[\ell\neq p]} > 0$. From §E.1.3, if $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$, we also have a contradiction. Therefore $\lambda_x \mathbb{I}_{[\ell\neq p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} = 0$.

E.2.2 Problem $R^{\text{MCI}}_{j\ell}$: $\ell \neq 0, h_j > h_i, g_j \leq g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$. Suppose $\ell \neq 0, h_j > h_i, g_j \leq g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$. First, since $\ell \neq 0, h_j > h_i$, then $y^*_j = h_j, y^*_i = h_i$ is infeasible. If we nonetheless have $y^*_i = h_i$, then $\lambda_x \mathbb{I}_{[\ell\neq p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} = 0$. However, the results of §E.1.1 provide a contradiction. Therefore it must be the case that $y^*_i > h_i$, and hence $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$. Since the results of §E.1.3 also provide a contradiction, we have $\lambda_x \mathbb{I}_{[\ell\neq p]} = 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$.

E.2.3 Problem $R^{\text{MCI}}_{j\ell}$: $\ell \notin \{0,p\}, g_j > g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$, $h_j > h_i + \rho_j \tfrac{\sigma_{h_j}}{\sigma_{g_j}}(g_j - g_{i+1})w_g$. Suppose $\ell \notin \{0,p\}, g_j > g_{i+1} + \rho_j \tfrac{\sigma_{g_j}}{\sigma_{h_j}}(h_j - h_i)w_h$, and $h_j > h_i + \rho_j \tfrac{\sigma_{h_j}}{\sigma_{g_j}}(g_j - g_{i+1})w_g$. After considering the combined results above, the only remaining possibility is that $\lambda_x \mathbb{I}_{[\ell\neq p]} > 0$ and $\lambda_y \mathbb{I}_{[\ell\neq 0]} > 0$.

E.3 Problem $R^{\text{MCI}}_{j\ell}$: Values of the Dual Variables λ_x, λ_y

Finally, the values of the dual variables in Problem $R^{\text{MCI}}_{j\ell}$ are $\lambda_x = -\alpha_j \frac{\partial I_j(x^*_j, y^*_j)}{\partial x_j} = \alpha_{i+1} \frac{\partial J_{i+1}(x^*_{i+1})}{\partial x_{i+1}}$ and $\lambda_y = -\alpha_j \frac{\partial I_j(x^*_j, y^*_j)}{\partial y_j} = \alpha_i \frac{\partial K_i(y^*_i)}{\partial y_i}$, where

$$\tfrac{\epsilon}{\sigma_b^2} w_g \mathbb{I}^g_{j\ell} \leq \frac{-\partial I_j(x^*_j, y^*_j)}{\partial x_j} = \frac{(1/\sigma_{g_j}) w_g \mathbb{I}^g_{j\ell}}{(1-\rho_j^2 w_g w_h \mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell})}\left[\frac{(g_j - g_{i+1})}{\sigma_{g_j}} - \rho_j \frac{(h_j - h_i)}{\sigma_{h_j}} w_h \mathbb{I}^h_{j\ell}\right] \leq \frac{\beta(1+\rho_b)}{\sigma_a^2(1-\rho_b)} w_g \mathbb{I}^g_{j\ell},$$
$$\tfrac{\epsilon}{\sigma_b^2} w_h \mathbb{I}^h_{j\ell} \leq \frac{-\partial I_j(x^*_j, y^*_j)}{\partial y_j} = \frac{(1/\sigma_{h_j}) w_h \mathbb{I}^h_{j\ell}}{(1-\rho_j^2 w_g w_h \mathbb{I}^g_{j\ell}\mathbb{I}^h_{j\ell})}\left[\frac{(h_j - h_i)}{\sigma_{h_j}} - \rho_j \frac{(g_i - g_{i+1})}{\sigma_{g_j}} w_g \mathbb{I}^g_{j\ell}\right] \leq \frac{\beta(1+\rho_b)}{\sigma_a^2(1-\rho_b)} w_h \mathbb{I}^h_{j\ell}.$$

F PROOF OF LEMMA 4.2

We do not provide a proof when $i \in \{1, p\}$. To prove the result when $i \in \{2, \ldots, p-1\}$, for a contradiction, suppose that $J_i(x_i^*(\alpha_j, \alpha_{i-1}, \alpha_i)) = 0$ and $K_i(y_i^*(\alpha_j, \alpha_i, \alpha_{i+1})) = 0$. Let $\lambda_x(\ell-1), \lambda_y(\ell-1)$ be the dual variables for Problem $R_{j\ell-1}^{\text{MCI}}$ and let $\lambda_x(\ell), \lambda_y(\ell)$ be the dual variables for Problem $R_{j\ell}^{\text{MCI}}$; recall that $\ell = i$.

In Problem $R_{j\ell-1}^{\text{MCI}}$, $x_i^*(\alpha_j, \alpha_{i-1}, \alpha_i) = g_i$ and $\lambda_x(\ell-1) = 0,\ \lambda_y(\ell-1) > 0$. Thus

$$x_j^*(\alpha_j, \alpha_{i-1}, \alpha_i) \le g_i, \qquad h_i < h_{i-1} \le y_j^*(\alpha_j, \alpha_{i-1}, \alpha_i) = y_{i-1}^*(\alpha_j, \alpha_{i-1}, \alpha_i) \le h_j. \tag{23}$$

In Problem $R_{j\ell}^{\text{MCI}}$, $y_i^*(\alpha_j, \alpha_i, \alpha_{i+1}) = h_i$, and $\lambda_x(\ell) > 0,\ \lambda_y(\ell) = 0$. Thus

$$g_i < g_{i+1} \le x_j^*(\alpha_j, \alpha_i, \alpha_{i+1}) = x_{i+1}^*(\alpha_j, \alpha_i, \alpha_{i+1}) \le g_j, \qquad y_j^*(\alpha_j, \alpha_i, \alpha_{i+1}) \le h_i. \tag{24}$$

Putting (23) and (24) together, along with Assumption 2, we have that $g_{i-1} < g_i < g_{i+1} < g_j$ and $h_{i+1} < h_i < h_{i-1} < h_j$, which implies Pareto systems $i-1$, i, and $i+1$ dominate non-Pareto system j (see Figure 2).

From Proposition 4.1, $\mathbb{I}_{j\ell-1}^{g} = 0$ in Problem $R_{j\ell-1}^{\text{MCI}}$, $\mathbb{I}_{j\ell}^{h} = 0$ in Problem $R_{j\ell}^{\text{MCI}}$, and under Assumption 6, we have

$$\tfrac{(g_j-g_i)}{\sigma_{g_j}} \le \rho_j \tfrac{(h_j-h_{i-1})w_h(\alpha_j,\alpha_{i-1})}{\sigma_{h_j}} \le \rho_j \tfrac{(h_j-h_{i-1})}{\sigma_{h_j}}, \quad \tfrac{(h_j-h_i)}{\sigma_{h_j}} \le \rho_j \tfrac{(g_j-g_{i+1})w_g(\alpha_i,\alpha_{i+1})}{\sigma_{g_j}} \le \rho_j \tfrac{(g_j-g_{i+1})}{\sigma_{g_j}};$$

simplifying these inequalities with the fact that $g_{i-1} < g_i < g_{i+1} < g_j$ and $h_{i+1} < h_i < h_{i-1} < h_j$ yields

$$\tfrac{(g_j-g_{i+1})}{(h_j-h_{i-1})} \tfrac{\sigma_{h_j}}{\sigma_{g_j}} < \tfrac{(g_j-g_i)}{(h_j-h_i)} \tfrac{\sigma_{h_j}}{\sigma_{g_j}} \le \rho_j \quad \text{and} \quad \tfrac{(h_j-h_{i-1})}{(g_j-g_{i+1})} \tfrac{\sigma_{g_j}}{\sigma_{h_j}} < \tfrac{(h_j-h_i)}{(g_j-g_{i+1})} \tfrac{\sigma_{g_j}}{\sigma_{h_j}} \le \rho_j. \tag{25}$$

Combining the two inequalities in (25) implies

$$1 \le \max\left(\tfrac{(g_j-g_{i+1})}{(h_j-h_{i-1})} \tfrac{\sigma_{h_j}}{\sigma_{g_j}}, \tfrac{(h_j-h_{i-1})}{(g_j-g_{i+1})} \tfrac{\sigma_{g_j}}{\sigma_{h_j}} \right) \le \rho_j,$$

which cannot hold. Thus we have a contradiction.

G PROOF OF THEOREM 4.3

We prove the theorem in parts.

G.1 Proof that All Allocations are Strictly Positive at Optimality

Proof that $\tilde{z}^* > 0$ and $\tilde{\alpha}_j^* > 0$ for all $j \in \mathcal{P}^c$. First, notice that $\tilde{\alpha}_k = 1/r$ for all $k \in \mathcal{S}$ is a feasible solution for Problem $\widetilde{Q}$ that results in

$$r\tilde{z} = \min_{j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i} \Big(\inf_{\substack{(x_j - x_{i+1})\mathbb{I}_{[\ell \ne p]} \le 0 \\ (y_j - y_i)\mathbb{I}_{[\ell \ne 0]} \le 0}} I_j(x_j, y_j) + K_i(y_i)\mathbb{I}_{[\ell \ne 0]} + J_{i+1}(x_{i+1})\mathbb{I}_{[\ell \ne p]} \Big) > 0$$

under Lemma 3.5; therefore $\tilde{z}^* > 0$. Further, notice that if $\tilde{\alpha}_j^* = 0$ for some $j \in \mathcal{P}^c$ in Problem $\widetilde{Q}$, then $\tilde{z} = 0$. Thus it must be the case that $\tilde{\alpha}_j^* > 0$ for all $j \in \mathcal{P}^c$.

Proof that $\tilde{\alpha}_1^* > 0$, $\tilde{\alpha}_p^* > 0$, and $\max\{\tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*\} > 0$ for all $i \in \{1, \ldots, p-1\}$. Now let $j \in \mathcal{P}^c$, and consider $R_j(\tilde{\alpha}_j^*, \tilde{\alpha}_i, \tilde{\alpha}_{i+1})$. If $\ell = 0$ and $\tilde{\alpha}_1 = 0$, then $\tilde{z} = 0$. If $\ell = p$ and $\tilde{\alpha}_p = 0$, then $\tilde{z} = 0$. Further, if there exists $i \in \{1, \ldots, p-1\}$ such that $\max\{\tilde{\alpha}_i, \tilde{\alpha}_{i+1}\} = 0$, then $\tilde{z} = 0$. Since $\tilde{z}^* > 0$, the result holds.

Proof that for all $i \in \{2,\ldots,p-1\}$, if there exists $j \in \mathcal{P}^c$ such that $h_j \le h_{i-1}$ or $g_j \le g_{i+1}$, then $\tilde{\alpha}_i^* > 0$. Suppose $p \ge 3$ and there exists $i \in \{2,\ldots,p-1\}$ such that $\tilde{\alpha}_i^* = 0$. Then $\tilde{\alpha}_{i-1}^* > 0$, $\tilde{\alpha}_{i+1}^* > 0$, and $\tilde{\alpha}_j^* > 0$ for all $j \in \mathcal{P}^c$. Thus

$$\begin{aligned}
\tilde{z}^* &\le \min_{j\in\mathcal{P}^c} \min(R_j(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*), R_j(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))\\
&= \min_{j\in\mathcal{P}^c} \min\Big(\inf_{\substack{x_j\le x_i\\ y_j\le y_{i-1}}} \tilde{\alpha}_j^* I_j(x_j,y_j) + \tilde{\alpha}_{i-1}^* K_{i-1}(y_{i-1}), \inf_{\substack{x_j\le x_{i+1}\\ y_j\le y_i}} \tilde{\alpha}_j^* I_j(x_j,y_j) + \tilde{\alpha}_{i+1}^* J_{i+1}(x_{i+1})\Big)\\
&= \min_{j\in\mathcal{P}^c} \min\Big(\inf_{y_j\le y_{i-1}} \tilde{\alpha}_j^* K_j(y_j) + \tilde{\alpha}_{i-1}^* K_{i-1}(y_{i-1}), \inf_{x_j\le x_{i+1}} \tilde{\alpha}_j^* J_j(x_j) + \tilde{\alpha}_{i+1}^* J_{i+1}(x_{i+1})\Big).\\
&\le \min_{j\in\mathcal{P}^c} \min\Big(\inf_{y_j\le h_{i-1}} \tilde{\alpha}_j^* K_j(y_j), \inf_{x_j\le g_{i+1}} \tilde{\alpha}_j^* J_j(x_j)\Big). \qquad (26)
\end{aligned}$$

The expression in line (26) equals zero if there exists $j \in \mathcal{P}^c$ such that $h_j \le h_{i-1}$ or $g_j \le g_{i+1}$. Since $\tilde{z}^* > 0$, we have a contradiction, and the result follows.

Proof that $\tilde{\alpha}_i^* > 0$ for all $i \in \{2,\ldots,p-1\}$. This result holds under Assumption 8.

G.2 KKT Conditions for Problem $\widetilde{Q}$

Let $\nu, \lambda_{j\ell} \ge 0$ for all $j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}$ be dual variables. Then the complementary slackness conditions are $\lambda_{j\ell}(R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*) - \tilde{z}^*) = 0$ for all $j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i$; and the stationarity conditions are

$$\textstyle\sum_{j\in\mathcal{P}^c} \left(\lambda_{j\ell-1} \frac{\partial R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)}{\partial \tilde{\alpha}_i} + \lambda_{j\ell} \frac{\partial R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)}{\partial \tilde{\alpha}_i} \right) = \nu \quad \forall i \in \mathcal{P}; \qquad (27)$$

$$\textstyle\sum_{\ell\in\mathcal{P}^{\text{ph}}, \ell=i} \lambda_{j\ell} \frac{\partial R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)}{\partial \tilde{\alpha}_j} = \nu \quad \forall j \in \mathcal{P}^c; \qquad (28)$$

$$\textstyle\sum_{j\in\mathcal{P}^c} \sum_{\ell\in\mathcal{P}^{\text{ph}}} \lambda_{j\ell} = 1. \qquad (29)$$

G.3 Proof that the dual variable $\nu > 0$

Proof that $\frac{\partial R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)}{\partial \tilde{\alpha}_j} > 0$ for all $j \in \mathcal{P}^c$ and all $\ell \in \mathcal{P}^{\text{ph}}, \ell = i$ in (28). First, let $j \in \mathcal{P}^c$ and $\ell \in \mathcal{P}^{\text{ph}}, \ell = i$, and recall $\tilde{\alpha}_j^* > 0, \tilde{\alpha}_i^* > 0$, and $\tilde{\alpha}_{i+1}^* > 0$. Then using the KKT conditions for Problem $R_{j\ell}^{\text{MCI}}$,

$$\begin{aligned}
\frac{\partial R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)}{\partial \tilde{\alpha}_j} &= I_j(x_j^*, y_j^*) + \tilde{\alpha}_j^* \left(\frac{\partial I_j(x_j^*, y_j^*)}{\partial x_j^*} \frac{\partial x_j^*}{\partial \tilde{\alpha}_j} + \frac{\partial I_j(x_j^*, y_j^*)}{\partial y_j^*} \frac{\partial y_j^*}{\partial \tilde{\alpha}_j} \right)\\
&\quad + \tilde{\alpha}_i^* \left(\frac{\partial K_i(y_i^*)}{\partial y_i^*} \frac{\partial y_i^*}{\partial \tilde{\alpha}_j} \right) \mathbb{I}_{[\ell\ne 0]} + \tilde{\alpha}_{i+1}^* \left(\frac{\partial J_{i+1}(x_{i+1}^*)}{\partial x_{i+1}^*} \frac{\partial x_{i+1}^*}{\partial \tilde{\alpha}_j} \right) \mathbb{I}_{[\ell\ne p]}\\
&= I_j(x_j^*, y_j^*) + \lambda_y \mathbb{I}_{[\ell\ne 0]} \left(\frac{\partial y_i^*}{\partial \tilde{\alpha}_j} - \frac{\partial y_j^*}{\partial \tilde{\alpha}_j} \right) + \lambda_x \mathbb{I}_{[\ell\ne p]} \left(\frac{\partial x_{i+1}^*}{\partial \tilde{\alpha}_j} - \frac{\partial x_j^*}{\partial \tilde{\alpha}_j} \right) \qquad (30)\\
&= I_j(x_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*), y_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) = I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)), \qquad (31)
\end{aligned}$$

where the equality from (30) to (31) holds because if $\lambda_x \mathbb{I}_{[\ell\ne p]} > 0$, then $x_{i+1}^* = x_j^*$, implying $\frac{\partial x_{i+1}^*}{\partial \tilde{\alpha}_j} = \frac{\partial x_j^*}{\partial \tilde{\alpha}_j}$. Likewise, if $\lambda_y \mathbb{I}_{[\ell\ne 0]} > 0$, then $y_i^* = y_j^*$, implying $\frac{\partial y_i^*}{\partial \tilde{\alpha}_j} = \frac{\partial y_j^*}{\partial \tilde{\alpha}_j}$. By Lemma 3.5, $\forall j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i$, we have $I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0$, implying the result.

Proof that $\nu > 0$. From (29), at least one of the dual variables $\lambda_{j\ell}$ in the sum must be strictly positive. Combining this fact and the previous result, we have $\nu > 0$.

G.4 Proof of Theorem 4.3 Parts (1)–(5)

Note that previously, we showed $\frac{\partial R_{j\ell}(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)}{\partial \tilde{\alpha}_j} = I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0$ for all $j \in \mathcal{P}^c, \ell \in \mathcal{P}^{\text{ph}}, \ell = i$, which we use in the sequel.

Proof of Part (1). Since $\nu > 0$, in (28) we have $\sum_{\ell \in \mathcal{P}^{\text{ph}}, \ell = i} \lambda_{j\ell} I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0$ for all $j \in \mathcal{P}^c$. Thus for every $j \in \mathcal{P}^c$, there exists a phantom Pareto system $\ell^* \in \mathcal{P}^{\text{ph}}$, ℓ^* dependent on j, such that $\lambda_{j\ell^*} > 0$, and Part (1) of the Theorem holds.

Proof of Part (2). In line (27), using the KKT conditions for Problem $R_{j\ell-1}^{\text{MCI}}$, we have

$$\begin{aligned}\frac{\partial R_{j\ell}(\tilde{\alpha}_j, \tilde{\alpha}_{i-1}, \tilde{\alpha}_i)}{\partial \tilde{\alpha}_i} &= J_i(x_i^*) + \tilde{\alpha}_i \left(\frac{\partial J_i(x_i^*)}{\partial x_i^*} \frac{\partial x_i^*}{\partial \tilde{\alpha}_i} \right) + \tilde{\alpha}_{i-1} \left(\frac{\partial K_{i-1}(y_{i-1}^*)}{\partial y_{i-1}^*} \frac{\partial y_{i-1}^*}{\partial \tilde{\alpha}_i} \right) \mathbb{I}_{[\ell \neq 1]} \\ &\quad + \tilde{\alpha}_j \left(\frac{\partial I_j(x_j^*, y_j^*)}{\partial x_j^*} \frac{\partial x_j^*}{\partial \tilde{\alpha}_i} + \frac{\partial I_j(x_j^*, y_j^*)}{\partial y_j^*} \frac{\partial y_j^*}{\partial \tilde{\alpha}_i} \right) \\ &= J_i(x_i^*) + \lambda_x \left(\frac{\partial x_i^*}{\partial \tilde{\alpha}_i} - \frac{\partial x_j^*}{\partial \tilde{\alpha}_i} \right) + \lambda_y \mathbb{I}_{[\ell \neq 1]} \left(\frac{\partial y_{i-1}^*}{\partial \tilde{\alpha}_i} - \frac{\partial y_j^*}{\partial \tilde{\alpha}_i} \right) = J_i(x_i^*(\tilde{\alpha}_j, \tilde{\alpha}_{i-1}, \tilde{\alpha}_i)),\end{aligned}$$

where the penultimate equality holds by noticing that if $\lambda_x > 0$, then $x_j^* = x_i^*$, implying $\frac{\partial x_i^*}{\partial \tilde{\alpha}_i} = \frac{\partial x_j^*}{\partial \tilde{\alpha}_i}$. Likewise, if $\lambda_y > 0$, then $y_j^* = y_{i-1}^*$, implying $\frac{\partial y_{i-1}^*}{\partial \tilde{\alpha}_i} = \frac{\partial y_j^*}{\partial \tilde{\alpha}_i}$.

Using similar logic, from the KKT conditions for Problem $R_{j\ell}^{\text{MCI}}$, we have

$$\begin{aligned}\frac{\partial R_{j\ell}(\tilde{\alpha}_j, \tilde{\alpha}_i, \tilde{\alpha}_{i+1})}{\partial \tilde{\alpha}_i} &= K_i(y_i^*) + \tilde{\alpha}_i \left(\frac{\partial K_i(y_i^*)}{\partial y_i^*} \frac{\partial y_i^*}{\partial \tilde{\alpha}_i} \right) + \tilde{\alpha}_{i+1} \left(\frac{\partial J_{i+1}(x_{i+1}^*)}{\partial x_{i+1}^*} \frac{\partial x_{i+1}^*}{\partial \tilde{\alpha}_i} \right) \mathbb{I}_{[\ell \neq p]} \\ &\quad + \tilde{\alpha}_j \left(\frac{\partial I_j(x_j^*, y_j^*)}{\partial x_j^*} \frac{\partial x_j^*}{\partial \tilde{\alpha}_i} + \frac{\partial I_j(x_j^*, y_j^*)}{\partial y_j^*} \frac{\partial y_j^*}{\partial \tilde{\alpha}_i} \right) \\ &= K_i(y_i^*) + \lambda_y \left(\frac{\partial y_i^*}{\partial \tilde{\alpha}_i} - \frac{\partial y_j^*}{\partial \tilde{\alpha}_i} \right) + \lambda_x \mathbb{I}_{[\ell \neq p]} \left(\frac{\partial x_{i+1}^*}{\partial \tilde{\alpha}_i} - \frac{\partial x_j^*}{\partial \tilde{\alpha}_i} \right) = K_i(y_i^*(\tilde{\alpha}_j, \tilde{\alpha}_i, \tilde{\alpha}_{i+1})).\end{aligned}$$

Then since $\nu > 0$, from (27), for each Pareto system $i \in \mathcal{P}$,

$$\textstyle\sum_{j \in \mathcal{P}^c} [\lambda_{j\ell-1} J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell} K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))] > 0.$$

Thus at least one constraint is binding for each Pareto system at optimality, and the result in Part (2) follows.

Proof of Parts (3)–(5). Updating the stationarity conditions in (27)–(29) yields the KKT conditions in equations (7)–(9). Substituting equation (8) into equation (7) yields equation (6); dividing yields equations (4) and (5).

H PROOF OF LEMMA 4.4

We prove the lemma in parts, using the following notation. Under Assumptions 7–8,

$$\text{if } \ell \neq p,\ \tilde{w}_g^{\text{L}} := \frac{\sigma_a^2}{\sigma_b^2} \frac{1}{1 + \tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*} \leq w_g(\tilde{\alpha}_j^*, \tilde{\alpha}_{i+1}^*) = \frac{\sigma_{g_j}^2/\tilde{\alpha}_j^*}{\sigma_{g_j}^2/\tilde{\alpha}_j^* + \sigma_{g_{i+1}}^2/\tilde{\alpha}_{i+1}^*} \leq \tilde{w}_g^{\text{U}} := \frac{\sigma_b^2}{\sigma_a^2} \frac{1}{1 + \tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*},$$

$$\text{if } \ell \neq 0,\ \tilde{w}_h^{\text{L}} := \frac{\sigma_a^2}{\sigma_b^2} \frac{1}{1 + \tilde{\alpha}_j^*/\tilde{\alpha}_i^*} \leq w_h(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*) = \frac{\sigma_{h_j}^2/\tilde{\alpha}_j^*}{\sigma_{h_j}^2/\tilde{\alpha}_j^* + \sigma_{h_i}^2/\tilde{\alpha}_i^*} \leq \tilde{w}_h^{\text{U}} := \frac{\sigma_b^2}{\sigma_a^2} \frac{1}{1 + \tilde{\alpha}_j^*/\tilde{\alpha}_i^*}.$$

For compactness, we denote $w_g(\tilde{\alpha}_j^*, \tilde{\alpha}_{i+1}^*)$ and $w_h(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*)$ as $\tilde{w}_g^*$ and $\tilde{w}_h^*$, respectively. The values of κ_R^{L} and κ_R^{U} are

$$\kappa_R^{\text{L}} = \min\{\kappa_{R1}^{\text{L}}, \kappa_{R2}^{\text{L}}, \kappa_{R3}^{\text{L}}, \kappa_{R4}^{\text{L}}\} \quad \text{and} \quad \kappa_R^{\text{U}} = \max\{\kappa_{R1}^{\text{U}}, \kappa_{R2}^{\text{U}}, \kappa_{R3}^{\text{U}}, \kappa_{R4}^{\text{U}}\},$$

where the values of $\kappa_{R1}^{\text{L}}, \kappa_{R2}^{\text{L}}, \kappa_{R3}^{\text{L}}, \kappa_{R4}^{\text{L}}$ and $\kappa_{R1}^{\text{U}}, \kappa_{R2}^{\text{U}}, \kappa_{R3}^{\text{U}}, \kappa_{R4}^{\text{U}}$ are defined in each subsection.

H.1 Proof of the Bounds in Equation (10)

We prove the result in cases, where $\kappa_{R1}^{\text{L}} = \frac{\epsilon^2 \sigma_a^2}{2 c_b \sigma_b^4}$ and $\kappa_{R1}^{\text{U}} = \frac{\beta^2 \sigma_b^2}{c_a \sigma_a^4}$.

H.1.1 $\mathbb{I}^g_{j\ell} > 0, \mathbb{I}^h_{j\ell} = 0$ *in* $R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})$. In this case,

$$\text{lower: } R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) = \tfrac{\tilde{\alpha}^*_j}{2}\big(\tfrac{(g_j-g_{i+1})^2}{\sigma^2_{g_j}}\tilde{w}^*_g\big) \geq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\epsilon^2}{\sigma^2_b}\tilde{w}^{\mathrm{L}}_g = \tilde{\alpha}^*_j\tfrac{\epsilon^2\sigma^2_a}{2\sigma^4_b}\tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_{i+1}};$$

$$\text{upper: } R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) = \tfrac{\tilde{\alpha}^*_j}{2}\big(\tfrac{(g_j-g_{i+1})^2}{\sigma^2_{g_j}}\tilde{w}^*_g\big) \leq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\beta^2}{\sigma^2_a}\tilde{w}^{\mathrm{U}}_g = \tilde{\alpha}^*_j\tfrac{\beta^2\sigma^2_b}{2\sigma^4_a}\tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_{i+1}}.$$

H.1.2 $\mathbb{I}^g_{j\ell} = 0, \mathbb{I}^h_{j\ell} > 0$ *in* $R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})$. In this case,

$$\text{lower: } R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) = \tfrac{\tilde{\alpha}^*_j}{2}\big(\tfrac{(h_j-h_i)^2}{\sigma^2_{h_j}}\tilde{w}^*_h\big) \geq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\epsilon^2}{\sigma^2_b}\tilde{w}^{\mathrm{L}}_h = \tilde{\alpha}^*_j\tfrac{\epsilon^2\sigma^2_a}{2\sigma^4_b}\tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_i};$$

$$\text{upper: } R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) = \tfrac{\tilde{\alpha}^*_j}{2}\big(\tfrac{(h_j-h_i)^2}{\sigma^2_{h_j}}\tilde{w}^*_h\big) \leq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\beta^2}{\sigma^2_a}\tilde{w}^{\mathrm{U}}_h = \tilde{\alpha}^*_j\tfrac{\beta^2\sigma^2_b}{2\sigma^4_a}\tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_i}.$$

H.1.3 $\mathbb{I}^g_{j\ell} > 0, \mathbb{I}^h_{j\ell} > 0$ *in* $R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})$. In this case, recall that

$$R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) = \frac{\tilde{\alpha}^*_j}{2}\begin{bmatrix} g_j - g_{i+1} \\ h_j - h_i \end{bmatrix}^{\intercal} \begin{bmatrix} \sigma^2_{g_j}/\tilde{w}^*_g & \rho_j\sigma_{g_j}\sigma_{h_j} \\ \rho_j\sigma_{g_j}\sigma_{h_j} & \sigma^2_{h_j}/\tilde{w}^*_h \end{bmatrix}^{-1} \begin{bmatrix} g_j - g_{i+1} \\ h_j - h_i \end{bmatrix}.$$

We bound this rate using standard bounds for quadratics. First, define the matrices

$$\Sigma_{j\ell} := \begin{bmatrix} \sigma^2_{g_j}/\tilde{w}^*_g & \rho_j\sigma_{g_j}\sigma_{h_j} \\ \rho_j\sigma_{g_j}\sigma_{h_j} & \sigma^2_{h_j}/\tilde{w}^*_h \end{bmatrix} \quad \text{and} \quad W_{j\ell} := \begin{bmatrix} 1/\tilde{w}^*_g & 1 \\ 1 & 1/\tilde{w}^*_h \end{bmatrix},$$

so that $\Sigma_{j\ell}$ is the element-wise (Hadamard) product of the matrices Σ_j and $W_{j\ell}$, that is, $\Sigma_{j\ell} = W_{j\ell} \circ \Sigma_j$. Since $\sigma^2_{g_j}/\tilde{w}^*_g > 0$ and $\sigma^2_{g_j}\sigma^2_{h_j}[(1/\tilde{w}^*_g)(1/\tilde{w}^*_h) - \rho^2_j] > 0$, then $\Sigma_{j\ell}$ is positive definite [Strang 1988, p. 330]; likewise, $W_{j\ell}$ is also positive definite. By Assumption 6, Σ_j is positive definite.

Then using Schur's theorems [Horn 1990, p. 95], it follows that

$$\min\{\tfrac{1}{\tilde{w}^*_g}, \tfrac{1}{\tilde{w}^*_h}\}\lambda_{\min}(\Sigma_j) \leq \lambda_{\min}(\Sigma_{j\ell}) \leq \lambda_{\max}(\Sigma_{j\ell}) \leq \max\{\tfrac{1}{\tilde{w}^*_g}, \tfrac{1}{\tilde{w}^*_h}\}\lambda_{\max}(\Sigma_j). \tag{32}$$

We find the eigenvalues of $\Sigma^{-1}_{j\ell}$ as $\lambda_{\max}(\Sigma^{-1}_{j\ell}) = 1/\lambda_{\min}(\Sigma_{j\ell})$ and $\lambda_{\min}(\Sigma^{-1}_{j\ell}) = 1/\lambda_{\max}(\Sigma_{j\ell})$ [Strang 1988, p. 258]. From line (32) and using Assumption 6, we have

$$\tfrac{\min\{\tilde{w}^{\mathrm{L}}_g, \tilde{w}^{\mathrm{L}}_h\}}{c_b} \leq \tfrac{\min\{\tilde{w}^*_g, \tilde{w}^*_h\}}{\lambda_{\max}(\Sigma_j)} \leq \lambda_{\min}(\Sigma^{-1}_{j\ell}) \leq \lambda_{\max}(\Sigma^{-1}_{j\ell}) \leq \tfrac{\max\{\tilde{w}^*_g, \tilde{w}^*_h\}}{\lambda_{\min}(\Sigma_j)} \leq \tfrac{\max\{\tilde{w}^{\mathrm{U}}_g, \tilde{w}^{\mathrm{U}}_h\}}{c_a}.$$

Then using bounds on quadratic forms [Boyd and Vandenberghe 2004, p. 647], we have

$$\tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\min\{\tilde{w}^{\mathrm{L}}_g, \tilde{w}^{\mathrm{L}}_h\}}{c_b}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix}^{\intercal}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix} \leq R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) \leq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\max\{\tilde{w}^{\mathrm{U}}_g, \tilde{w}^{\mathrm{U}}_h\}}{c_a}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix}^{\intercal}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix},$$

which implies

$$\text{lower: } R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) \geq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\min\{\tilde{w}^{\mathrm{L}}_g, \tilde{w}^{\mathrm{L}}_h\}}{c_b}\epsilon^2 \geq \tilde{\alpha}^*_j\tfrac{\epsilon^2\sigma^2_a}{2c_b\sigma^2_b}\min\{\tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_{i+1}}, \tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_i}\},$$

$$\text{upper: } R_{j\ell}(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}) \leq \tfrac{\tilde{\alpha}^*_j}{2}\tfrac{\max\{\tilde{w}^{\mathrm{U}}_g, \tilde{w}^{\mathrm{U}}_h\}}{c_a}2\beta^2 \leq \tilde{\alpha}^*_j\tfrac{\beta^2\sigma^2_b}{c_a\sigma^2_a}\max\{\tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_{i+1}}, \tfrac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_i}\}.$$

H.2 Proof of the Bounds in Equation (11)

We prove the result in cases, where $\kappa^{\mathrm{L}}_{R2} = \frac{\epsilon^2(1-\rho^2_b)\sigma^4_a}{2c_b\sigma^6_b}$ and $\kappa^{\mathrm{U}}_{R2} = \frac{\beta^2\sigma^4_b}{c_a\sigma^6_a}$.

H.2.1 $\mathbb{I}^g_{j\ell} > 0, \mathbb{I}^h_{j\ell} = 0$ *in* $I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}))$. In this case,

$$\text{lower: } I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})) = \frac{(g_j-g_{i+1})^2}{2\sigma^2_{g_j}}(\tilde{w}^*_g)^2 \geq \frac{\epsilon^2}{2\sigma^2_b}(\tilde{w}^{\mathrm{L}}_g)^2 = \frac{\epsilon^2\sigma^4_a}{2\sigma^6_b}\left(\frac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_{i+1}}\right)^2;$$

$$\text{upper: } I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})) = \frac{(g_j-g_{i+1})^2}{2\sigma^2_{g_j}}(\tilde{w}^*_g)^2 \leq \frac{\beta^2}{2\sigma^2_a}(\tilde{w}^{\mathrm{U}}_g)^2 = \frac{\beta^2\sigma^4_b}{2\sigma^6_a}\left(\frac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_{i+1}}\right)^2.$$

H.2.2 $\mathbb{I}^g_{j\ell} = 0, \mathbb{I}^h_{j\ell} > 0$ *in* $I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}))$. In this case,

$$\text{lower: } I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})) = \frac{(h_j-h_i)^2}{2\sigma^2_{h_j}}(\tilde{w}^*_h)^2 \geq \frac{\epsilon^2}{2\sigma^2_b}(\tilde{w}^{\mathrm{L}}_h)^2 = \frac{\epsilon^2\sigma^4_a}{2\sigma^6_b}\left(\frac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_i}\right)^2;$$

$$\text{upper: } I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})) = \frac{(h_j-h_i)^2}{2\sigma^2_{h_j}}(\tilde{w}^*_h)^2 \leq \frac{\beta^2}{2\sigma^2_a}(\tilde{w}^{\mathrm{U}}_h)^2 = \frac{\beta^2\sigma^4_b}{2\sigma^6_a}\left(\frac{1}{1+\tilde{\alpha}^*_j/\tilde{\alpha}^*_i}\right)^2.$$

H.2.3 $\mathbb{I}^g_{j\ell} > 0, \mathbb{I}^h_{j\ell} > 0$ *in* $I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}))$. In this case, we write $I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}))$ as a quadratic form using matrix notation,

$$I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1})) =$$

$$\frac{1}{2}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix}^\intercal \begin{bmatrix} \sigma^2_{g_j}\left[\frac{1}{(\tilde{w}^*_g)^2} + \left(\frac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\frac{1}{\tilde{w}^*_g}\right)^2\right] & \rho_j\sigma_{g_j}\sigma_{h_j}\left[1 - \frac{(1-\frac{1}{\tilde{w}^*_g})(1-\frac{1}{\tilde{w}^*_h})}{1-\rho^2_j}\right] \\ \rho_j\sigma_{g_j}\sigma_{h_j}\left[1 - \frac{(1-\frac{1}{\tilde{w}^*_g})(1-\frac{1}{\tilde{w}^*_h})}{1-\rho^2_j}\right] & \sigma^2_{h_j}\left[\frac{1}{(\tilde{w}^*_h)^2} + \left(\frac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\frac{1}{\tilde{w}^*_h}\right)^2\right] \end{bmatrix}^{-1} \begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix}.$$

Define the matrix

$$M_{j\ell} := \begin{bmatrix} \frac{1}{(\tilde{w}^*_g)^2} + \left(\frac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\frac{1}{\tilde{w}^*_g}\right)^2 & 1 - \frac{(1-\frac{1}{\tilde{w}^*_g})(1-\frac{1}{\tilde{w}^*_h})}{1-\rho^2_j} \\ 1 - \frac{(1-\frac{1}{\tilde{w}^*_g})(1-\frac{1}{\tilde{w}^*_h})}{1-\rho^2_j} & \frac{1}{(\tilde{w}^*_h)^2} + \left(\frac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\frac{1}{\tilde{w}^*_h}\right)^2 \end{bmatrix},$$

which is positive definite. Then define the element-wise (Hadamard) product of the matrices $M_{j\ell}$ and Σ_j as $\tilde{\Sigma}_{j\ell}$, that is, $\tilde{\Sigma}_{j\ell} := M_{j\ell} \circ \Sigma_j$. By Schur's theorems [Horn 1990, p. 95], we have

$$\begin{aligned} &\min\{\tfrac{1}{(\tilde{w}^*_g)^2}, \tfrac{1}{(\tilde{w}^*_h)^2}\}\lambda_{\min}(\Sigma_j) \\ &\leq \min\{\tfrac{1}{(\tilde{w}^*_g)^2} + \left(\tfrac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\tfrac{1}{\tilde{w}^*_g}\right)^2, \tfrac{1}{(\tilde{w}^*_h)^2} + \left(\tfrac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\tfrac{1}{\tilde{w}^*_h}\right)^2\}\lambda_{\min}(\Sigma_j) \leq \lambda_{\min}(\tilde{\Sigma}_{j\ell}) \\ &\leq \lambda_{\max}(\tilde{\Sigma}_{j\ell}) \leq \max\{\tfrac{1}{(\tilde{w}^*_g)^2} + \left(\tfrac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\tfrac{1}{\tilde{w}^*_g}\right)^2, \tfrac{1}{(\tilde{w}^*_h)^2} + \left(\tfrac{\rho^2_j}{1-\rho^2_j}\right)\left(1-\tfrac{1}{\tilde{w}^*_h}\right)^2\}\lambda_{\max}(\Sigma_j) \\ &\leq \max\{\tfrac{1}{(\tilde{w}^*_g)^2}\left(\tfrac{1}{1-\rho^2_j}\right), \tfrac{1}{(\tilde{w}^*_h)^2}\left(\tfrac{1}{1-\rho^2_j}\right)\}\lambda_{\max}(\Sigma_j) \\ &\leq \left(\tfrac{1}{1-\rho^2_b}\right)\max\{\tfrac{1}{(\tilde{w}^*_g)^2}, \tfrac{1}{(\tilde{w}^*_h)^2}\}\lambda_{\max}(\Sigma_j). \end{aligned} \tag{33}$$

We find the eigenvalues of the matrix $\tilde{\Sigma}^{-1}_{j\ell}$ as $\lambda_{\max}(\tilde{\Sigma}^{-1}_{j\ell}) = 1/\lambda_{\min}(\tilde{\Sigma}_{j\ell})$ and $\lambda_{\min}(\tilde{\Sigma}^{-1}_{j\ell}) = 1/\lambda_{\max}(\tilde{\Sigma}_{j\ell})$ [Strang 1988, p. 258]. From the inequalities that end in line (33) and using Assumption 6, we have

$$\begin{aligned} (1-\rho^2_b)\frac{\min\{(\tilde{w}^{\mathrm{L}}_g)^2, (\tilde{w}^{\mathrm{L}}_h)^2\}}{c_b} &\leq (1-\rho^2_b)\frac{\min\{(\tilde{w}^*_g)^2, (\tilde{w}^*_h)^2\}}{\lambda_{\max}(\Sigma_j)} \leq \lambda_{\min}(\tilde{\Sigma}^{-1}_{j\ell}) \\ &\leq \lambda_{\max}(\tilde{\Sigma}^{-1}_{j\ell}) \leq \frac{\max\{(\tilde{w}^*_g)^2, (\tilde{w}^*_h)^2\}}{\lambda_{\min}(\Sigma_j)} \leq \frac{\max\{(\tilde{w}^{\mathrm{U}}_g)^2, (\tilde{w}^{\mathrm{U}}_h)^2\}}{c_a}. \end{aligned}$$

Then using bounds on quadratic forms [Boyd and Vandenberghe 2004, p. 647], we have

$$\frac{1}{2}(1-\rho^2_b)\frac{\min\{(\tilde{w}^{\mathrm{L}}_g)^2, (\tilde{w}^{\mathrm{L}}_h)^2\}}{c_b}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix}^\intercal \begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i \end{bmatrix} \leq I_j(\mathfrak{z}^*_j(\tilde{\alpha}^*_j, \tilde{\alpha}^*_i, \tilde{\alpha}^*_{i+1}))$$

$$\leq \tfrac{1}{2}\frac{\max\{(\tilde{w}_g^{\mathrm{U}})^2,(\tilde{w}_h^{\mathrm{U}})^2\}}{c_a}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i\end{bmatrix}^{\intercal}\begin{bmatrix} g_j-g_{i+1} \\ h_j-h_i\end{bmatrix},$$

which implies

$$\begin{aligned}\text{lower: } I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*)) &\geq \tfrac{(1-\rho_b^2)}{2}\frac{\min\{(\tilde{w}_g^{\mathrm{L}})^2,(\tilde{w}_h^{\mathrm{L}})^2\}}{c_b}\epsilon^2 \\ &\geq \frac{\epsilon^2(1-\rho_b^2)\sigma_a^4}{2c_b\sigma_b^4}\min\{\tfrac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*)^2},\tfrac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*)^2}\},\\ \text{upper: } I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*)) &\leq \tfrac{1}{2}\frac{\max\{(\tilde{w}_g^{\mathrm{U}})^2,(\tilde{w}_h^{\mathrm{U}})^2\}}{c_a}2\beta^2 \leq \frac{\beta^2\sigma_b^4}{c_a\sigma_a^4}\max\{\tfrac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*)^2},\tfrac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*)^2}\}.\end{aligned}$$

H.3 Proof of the Bounds in Equation (12)

We prove the result in cases, where $\kappa_{R3}^{\mathrm{L}} = \frac{\epsilon^2\sigma_a^6}{2\sigma_b^6}$ and $\kappa_{R3}^{\mathrm{U}} = \frac{\beta^2\sigma_b^6(1+\rho_b)^2}{2\sigma_a^8(1-\rho_b^2)^2}$.

H.3.1 $\mathbb{I}_{j\ell}^{g} = 0, \mathbb{I}_{j\ell}^{h} > 0$ *in* $K_i(y_i^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*))$. In this case,

$$\begin{aligned}\text{lower: } K_i(y_i^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*)) &= \frac{(h_j-h_i)^2}{2\sigma_{h_i}^2}(1-\tilde{w}_h^*)^2 \geq \frac{\epsilon^2}{2\sigma_b^2}(1-\tilde{w}_h^*)^2 \geq \frac{\epsilon^2\sigma_a^4}{2\sigma_b^6}\frac{1}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2};\\ \text{upper: } K_i(y_i^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*)) &= \frac{(h_j-h_i)^2}{2\sigma_{h_i}^2}(1-\tilde{w}_h^*)^2 \leq \frac{\beta^2}{2\sigma_a^2}(1-\tilde{w}_h^*)^2 \leq \frac{\beta^2\sigma_b^4}{2\sigma_a^6}\frac{1}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2}.\end{aligned}$$

H.3.2 $\mathbb{I}_{j\ell}^{g} > 0, \mathbb{I}_{j\ell}^{h} > 0$ *in* $K_i(y_i^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*))$. In this case, under Assumption 7(b), the systems cannot systematically approach the boundary of the region in which $\mathbb{I}_{j\ell}^{g} > 0$, $\mathbb{I}_{j\ell}^{h} > 0$ from the interior. Thus for all $i, i+1 \in \mathcal{P}$, the term $\big[\frac{(h_j-h_i)}{\sigma_{h_j}} - \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}\tilde{w}_g^*\big]^2$ is bounded away from zero, uniformly in j. Then we have the bounds

$$\begin{aligned}\text{lower: } K_i(y_i^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*)) &= \frac{(1-\tilde{w}_h^*)^2}{2[1-\rho_j^2\tilde{w}_g^*\tilde{w}_h^*]^2}\frac{\sigma_{h_j}^2}{\sigma_{h_i}^2}\big[\frac{(h_j-h_i)}{\sigma_{h_j}} - \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}\tilde{w}_g^*\big]^2\\ &\geq \frac{\sigma_a^2}{2\sigma_b^2}\big[\frac{1}{[1-\rho_j^2\tilde{w}_g^*\tilde{w}_h^*]}\big(\frac{(h_j-h_i)}{\sigma_{h_j}} - \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}\tilde{w}_g^*\big)\big]^2(1-\tilde{w}_h^*)^2\\ &\geq \frac{\sigma_a^2}{2\sigma_b^2}\big[\frac{(h_j-h_i)}{\sigma_{h_j}} - \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}\tilde{w}_g^*\big]^2(1-\tilde{w}_h^*)^2\\ &\geq \frac{\epsilon^2\sigma_a^2}{2\sigma_b^2}(1-\tilde{w}_h^*)^2 \geq \frac{\epsilon^2\sigma_a^6}{2\sigma_b^6}\frac{1}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2};\\ \text{upper: } K_i(y_i^*(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*)) &= \frac{(1-\tilde{w}_h^*)^2}{2[1-\rho_j^2\tilde{w}_g^*\tilde{w}_h^*]^2}\frac{\sigma_{h_j}^2}{\sigma_{h_i}^2}\big[\frac{(h_j-h_i)}{\sigma_{h_j}} - \rho_j\frac{(g_j-g_{i+1})}{\sigma_{g_j}}\tilde{w}_g^*\big]^2\\ &\leq \frac{\sigma_b^2}{2\sigma_a^2}\frac{1}{[1-\rho_b^2]^2}\big[\frac{(h_j-h_i)}{\sigma_{h_j}} + \rho_b\frac{(g_j-g_{i+1})}{\sigma_{g_j}}\tilde{w}_g^*\big]^2(1-\tilde{w}_h^*)^2\\ &\leq \frac{\sigma_b^2}{2\sigma_a^2}\frac{1}{[1-\rho_b^2]^2}\big[\frac{\beta}{\sigma_a}(1+\rho_b)\big]^2(1-\tilde{w}_h^*)^2 \leq \frac{\beta^2\sigma_b^6(1+\rho_b)^2}{2\sigma_a^8(1-\rho_b^2)^2}\frac{1}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2}.\end{aligned}$$

H.4 Proof of the Bounds in Equation (13)

We omit the proof for the bounds on $J_{\ell+1}(x_{\ell+1}^*(\tilde{\alpha}_j^*,\tilde{\alpha}_\ell^*,\tilde{\alpha}_{\ell+1}^*))$, since it is similar to the proof for $K_\ell(y_\ell^*(\tilde{\alpha}_j^*,\tilde{\alpha}_\ell^*,\tilde{\alpha}_{\ell+1}^*))$. The constants in this case are $\kappa_{R4}^{\mathrm{L}} = \kappa_{R3}^{\mathrm{L}}$ and $\kappa_{R4}^{\mathrm{U}} = \kappa_{R3}^{\mathrm{U}}$.

I PROOF OF PROPOSITION 4.5

From the upper bound in (10), at optimality in Problem $\widetilde{Q}$, it also holds that

$$\tilde{z}^* \leq R_{j\ell}(\tilde{\alpha}_j^*,\tilde{\alpha}_i^*,\tilde{\alpha}_{i+1}^*) \leq \tilde{\alpha}_j^*\kappa_R^{\mathrm{U}}\big[\tfrac{1}{1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i+1}^*} + \tfrac{1}{1+\tilde{\alpha}_j^*/\tilde{\alpha}_i^*}\big] \leq \tilde{\alpha}_j^* 2\kappa_R^{\mathrm{U}}$$

for all $j \in \mathcal{P}^c$ and all $\ell \in \mathcal{P}^{\text{ph}}, \ell = i$. Then by Theorem 4.3 Part (1), for each non-Pareto system $j \in \mathcal{P}^c$, there exists a phantom Pareto system $\ell^*(j) \in \mathcal{P}^{\text{ph}}$, $\ell(j)^* = i^*(j)$ such that $\tilde{z}^* = R_{j\ell^*(j)}(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*)$. Then it follows that

$$|\mathcal{P}^c|\tilde{z}^* = \textstyle\sum_{j\in\mathcal{P}^c} \tilde{z}^* = \sum_{j\in\mathcal{P}^c} R_{j\ell^*(j)}(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*) \leq 2\kappa_R^{\text{U}} \sum_{j\in\mathcal{P}^c} \tilde{\alpha}_j^* \leq 2\kappa_R^{\text{U}},$$

which implies $\tilde{z}^* \leq 2\kappa_R^{\text{U}}/|\mathcal{P}^c|$. Thus $\tilde{z}^* = O(1/|\mathcal{P}^c|)$.

J PROOF OF THEOREM 4.7

We prove the theorem in parts. In what follows, notice that under Assumption 9, there exists $\kappa_0 \in (0, \infty)$ such that $\lambda_{j\ell}(r) \leq \kappa_0 \lambda_{j\ell^*}(r)$ for all $\ell \in \mathcal{P}^{\text{ph}}$, all $r \geq r_{j0}$.

J.1 Proof of Theorem 4.7 Part (1)

First, recall that for all $i \in \mathcal{P}, j \in \mathcal{P}^c$, we have $I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) > 0$ for all $r \geq r_{j0}$.

Now for a contradiction, suppose there exists $j^* \in \mathcal{P}^c$ such that $\ell^* \in \mathcal{P}^{\text{ph}}, \ell^* = i^*$ is its primary phantom Pareto system, and $\liminf I_{j^*}(\mathfrak{z}_{j^*}^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) = 0$ as $|\mathcal{P}^c(i)| \to \infty$ for all $i \in \mathcal{P}$. Then there exists a subsequence $\{r_k : k = 1, 2, \ldots; r_1 \geq r_{j^*0}\}$ such that

$$\lim_{k\to\infty} I_{j^*}(\mathfrak{z}_{j^*}^*(\tilde{\alpha}_{j^*}^*(r_k), \tilde{\alpha}_{i^*}^*(r_k), \tilde{\alpha}_{i^*+1}^*(r_k))) = 0.$$

Henceforth, we consider only this subsequence.

From the bounds on the derivatives in §E.3, it must be the case that $\tilde{w}_g^* \to 0$ if $\lim_{k\to\infty} \mathbb{I}_{j^*\ell^*}^g(r_k) > 0$, and $\tilde{w}_h^* \to 0$ if $\lim_{k\to\infty} \mathbb{I}_{j^*\ell^*}^h(r_k) > 0$. (See §H for a definition of the $\tilde{w}_g^*, \tilde{w}_h^*$ notation.) Since the variances are uniformly bounded, $\tilde{\alpha}_{j^*}^*/\tilde{\alpha}_{i^*+1}^* \to \infty$ if $\lim_{k\to\infty} \mathbb{I}_{j^*\ell^*}^g(r_k) > 0$ and $\tilde{\alpha}_{j^*}^*/\tilde{\alpha}_{i^*}^* \to \infty$ if $\lim_{k\to\infty} \mathbb{I}_{j^*\ell^*}^h(r_k) > 0$.

From the bounds in equations (12) and (13), there exist constants $\tau_K > 0, \tau_J > 0$, not dependent on r, such that $K_{i^*}(y_{i^*}^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) \geq \mathbb{I}_{j^*\ell^*}^h \tau_K$ and $J_{i^*+1}(x_{i^*+1}^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) \geq \mathbb{I}_{j^*\ell^*}^g \tau_J$ for all $k = 1, 2, \ldots$ indexing the subsequence. From the KKT conditions in equation (7), for all r,

$$2\nu \geq \textstyle\sum_{j\in\mathcal{P}^c}[\lambda_{j\ell^*} K_{i^*}(y_{i^*}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*))] + \sum_{j\in\mathcal{P}^c}[\lambda_{j\ell^*} J_{i^*+1}(x_{i^*+1}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*))].$$

Then we have $2\nu \geq \lambda_{j^*\ell^*}(\mathbb{I}_{j^*\ell^*}^h \tau_K + \mathbb{I}_{j^*\ell^*}^g \tau_J)$ for all $k = 1, 2, \ldots$. Letting $\tau_1 := (1/2)(\mathbb{I}_{j^*\ell^*}^h \tau_K + \mathbb{I}_{j^*\ell^*}^g \tau_J) > 0$, we have the lower bound $\nu \geq \tau_1 \lambda_{j^*\ell^*}$ for all $k = 1, 2, \ldots$.

Now under Assumptions 5 and 6, there exists a constant $b_1 < \infty$ such that $I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) \leq b_1$ for all $r \geq r_{j0}$. From the KKT conditions in equation (8), for all r, we have

$$\nu = \sum_{\ell\in\mathcal{P}^{\text{ph}}, \ell=i} \lambda_{j^*\ell} I_{j^*}(\mathfrak{z}_{j^*}^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) \leq \lambda_{j^*\ell^*} I_{j^*}(\mathfrak{z}_{j^*}^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) + \sum_{\ell\in\mathcal{P}^{\text{ph}}, \ell\neq\ell^*} \lambda_{j^*\ell} b_1$$

Combining the upper and lower bounds on ν results in

$$0 < \tau_1 \leq \frac{\nu}{\lambda_{j^*\ell^*}} \leq I_{j^*}(\mathfrak{z}_{j^*}^*(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) + \sum_{\ell\in\mathcal{P}^{\text{ph}}, \ell\neq\ell^*} \frac{\lambda_{j^*\ell}}{\lambda_{j^*\ell^*}} b_1$$

for all $k = 1, 2, \ldots$, which provides a contradiction since the limit of the right-hand side is zero on the subsequence under Assumption 9.

J.2 Proof of Theorem 4.7 Part (2)

Proposition 4.5 provides the lower bound $\tilde{\alpha}_j^* \geq \tilde{z}^*/(2\kappa_R^{\text{U}})$. For the upper bound, by Theorem 4.7 Part (1), for all $j \in \mathcal{P}^c$ with primary phantom Pareto system $\ell^*(j) \in \mathcal{P}^{\text{ph}}, \ell^*(j) = i^*(j)$, we have $\tilde{z}^* \geq \tilde{\alpha}_j^* I_j(\mathfrak{z}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*)) \geq \tilde{\alpha}_j^* \kappa_1$ for all $r \geq r_{j0}$. Since there exists a primary phantom Pareto system for every $j \in \mathcal{P}^c$, the result holds for all $j \in \mathcal{P}^c$ and all $r \geq r_{j0}$.

J.3 Proof of Theorem 4.7 Part (3)

Let $j \in \mathcal{P}^c$ have primary phantom Pareto $\ell^* \in \mathcal{P}^{\text{ph}}$, $\ell^* = i^*$. Under Assumptions 5 and 6, recall that there exists a constant $b_1 < \infty$ such that $I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*}^*, \tilde{\alpha}_{i^*+1}^*)) \leq b_1$ for all $r \geq r_{j0}$. From equation (8) and using Theorem 4.7 Part (1) and Assumption 9, we have

$$\lambda_{j\ell^*}\kappa_1 \leq v = \textstyle\sum_{\ell \in \mathcal{P}^{\text{ph}}, \ell=i} \lambda_{j\ell} I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) \leq (p+1)\kappa_0 \lambda_{j\ell^*} b_1. \tag{34}$$

Flipping these inequalities around results in $v[(p+1)\kappa_0 b_1]^{-1} \leq \lambda_{j\ell^*} \leq v\kappa_1^{-1}$, which implies $\lambda_{j\ell^*} = \Theta(v)$ for all $j \in \mathcal{P}^c$.

Summing over $j \in \mathcal{P}^c$ in the right-hand side of (34) and applying equation (9) implies $|\mathcal{P}^c|v \leq (p+1)\kappa_0 b_1 \sum_{j \in \mathcal{P}^c} \lambda_{j\ell^*} \leq (p+1)\kappa_0 b_1$, and hence $v = O(1/|\mathcal{P}^c|)$. For the lower bound, from equation (8), using Theorem 4.7 Part (1) and applying equation (9),

$$|\mathcal{P}^c|v = \textstyle\sum_{j \in \mathcal{P}^c} \sum_{\ell \in \mathcal{P}^{\text{ph}}, \ell=i} \lambda_{j\ell} I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)) \geq \kappa_1 \sum_{j \in \mathcal{P}^c} \sum_{\ell \in \mathcal{P}^{\text{ph}}} \lambda_{j\ell} = \kappa_1.$$

Combining this lower bound with the upper bound implies $v = \Theta(1/|\mathcal{P}^c|)$.

J.4 Proof of Theorem 4.7 Part (4)

Recall that $\tilde{\alpha}_i^*(r)/\tilde{\alpha}_j^*(r) > 0$ for all $i \in \mathcal{P}, j \in \mathcal{P}^c$, and all $r \geq r_{j0}$. For a contradiction, let $i \in \mathcal{P}$ be a Pareto system, and suppose there exists a non-Pareto system $j^* \in \mathcal{P}^c(i)$ such that $\liminf_{r\to\infty} \tilde{\alpha}_i^*(r)/\tilde{\alpha}_{j^*}^*(r) = 0$. Then there exists a subsequence $\{r_k : k = 1, 2, \ldots; r_1 \geq r_{j^*0}\}$ such that $\lim_{k\to\infty} \tilde{\alpha}_i^*(r_k)/\tilde{\alpha}_{j^*}^*(r_k) = 0$; henceforth, we consider only this subsequence.

J.4.1 Pareto System $i \in \{1, p\}$. If $i \in \{1, p\}$ and $j^* \in \mathcal{P}^c(i)$, then for all $r \geq r_{j0}$,

$$
\begin{aligned}
\tilde{z}^*/\tilde{\alpha}_{j^*}^* &= \min\big(R_{j^*\ell-1}(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)/\tilde{\alpha}_{j^*}^*, R_{j^*\ell}(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)/\tilde{\alpha}_{j^*}^*\big) \\
&= \begin{cases} \min\big(\inf_x J_{j^*}(x) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} J_i(x), \inf_{\substack{x_{j^*} \leq x_{i+1}, \\ y_{j^*} \leq y_i}} I_{j^*}(x_{j^*}, y_{j^*}) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} K_i(y_i) + \frac{\tilde{\alpha}_{i+1}^*}{\tilde{\alpha}_{j^*}^*} J_{i+1}(x_{i+1})\big) & \text{if } i = 1 \\ \min\big(\inf_{\substack{x_{j^*} \leq x_i, \\ y_{j^*} \leq y_{i-1}}} I_{j^*}(x_{j^*}, y_{j^*}) + \frac{\tilde{\alpha}_{i-1}^*}{\tilde{\alpha}_{j^*}^*} K_{i-1}(y_{i-1}) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} J_i(x_i), \inf_y K_{j^*}(y) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} K_i(y)\big) & \text{if } i = p \end{cases} \\
&\leq [\inf_x J_{j^*}(x) + (\tilde{\alpha}_i^*/\tilde{\alpha}_{j^*}^*) J_i(x)]\mathbb{I}_{[i=1]} + [\inf_y K_{j^*}(y) + (\tilde{\alpha}_i^*/\tilde{\alpha}_{j^*}^*) K_i(y)]\mathbb{I}_{[i=p]}.
\end{aligned}
$$

Since $\lim_{k\to\infty} \tilde{\alpha}_i^*(r_k)/\tilde{\alpha}_{j^*}^*(r_k) = 0$, the right-hand side of the inequality above goes to zero on the subsequence, which contradicts Theorem 4.7 Part (2).

J.4.2 Pareto System $i \in \{2, \ldots, p-1\}$ and $g_{j^*} < g_{i+1}$ or $h_{j^*} < h_{i-1}$. Now suppose that $j^* \in \mathcal{P}^c(i)$ for $i \in \{2, \ldots, p-1\}$, and $g_{j^*} < g_{i+1}$ or $h_{j^*} < h_{i-1}$. Under Assumptions 5 and 6, there exists a constant $b_2 < \infty$ such that $\max_{i \in \mathcal{P}, j \in \mathcal{P}^c}(J_i(x_i(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)), K_i(y_i(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))) \leq b_2$ for all r. Then for all $r \geq r_{j^*0}$,

$$
\begin{aligned}
\tilde{z}^*/\tilde{\alpha}_{j^*}^* &= \min\big(R_{j^*\ell-1}(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)/\tilde{\alpha}_{j^*}^*, R_{j^*\ell}(\tilde{\alpha}_{j^*}^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*)/\tilde{\alpha}_{j^*}^*\big) \\
&= \min\Big(\inf_{x_{j^*} \leq x_i,\ y_{j^*} \leq y_{i-1}} I_{j^*}(x_{j^*}, y_{j^*}) + \frac{\tilde{\alpha}_{i-1}^*}{\tilde{\alpha}_{j^*}^*} K_{i-1}(y_{i-1}) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} J_i(x_i), \\
&\qquad \inf_{x_{j^*} \leq x_{i+1},\ y_{j^*} \leq y_i} I_{j^*}(x_{j^*}, y_{j^*}) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} K_i(y_i) + \frac{\tilde{\alpha}_{i+1}^*}{\tilde{\alpha}_{j^*}^*} J_{i+1}(x_{i+1})\Big) \\
&\leq \min\Big(\inf_{x_{j^*} \leq x_i,\ y_{j^*} \leq y_{i-1}} I_{j^*}(x_{j^*}, y_{j^*}) + \frac{\tilde{\alpha}_{i-1}^*}{\tilde{\alpha}_{j^*}^*} K_{i-1}(y_{i-1}) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} b_2, \\
&\qquad \inf_{x_{j^*} \leq x_{i+1},\ y_{j^*} \leq y_i} I_{j^*}(x_{j^*}, y_{j^*}) + \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} b_2 + \frac{\tilde{\alpha}_{i+1}^*}{\tilde{\alpha}_{j^*}^*} J_{i+1}(x_{i+1})\Big). \\
&= \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} b_2 + \min\Big(\inf_{y_{j^*} \leq y_{i-1}} K_{j^*}(y_{j^*}) + \frac{\tilde{\alpha}_{i-1}^*}{\tilde{\alpha}_{j^*}^*} K_{i-1}(y_{i-1}), \inf_{x_{j^*} \leq x_{i+1}} J_{j^*}(x_{j^*}) + \frac{\tilde{\alpha}_{i+1}^*}{\tilde{\alpha}_{j^*}^*} J_{i+1}(x_{i+1})\Big)
\end{aligned}
$$

$$= \frac{\tilde{\alpha}_i^*}{\tilde{\alpha}_{j^*}^*} b_2 + \min\big([\inf_y K_{j^*}(y) + \tfrac{\tilde{\alpha}_{i-1}^*}{\tilde{\alpha}_{j^*}^*} K_{i-1}(y)]\mathbb{I}_{[h_{j^*} > h_{i-1}]}, [\inf_x J_{j^*}(x) + \tfrac{\tilde{\alpha}_{i+1}^*}{\tilde{\alpha}_{j^*}^*} J_{i+1}(x)]\mathbb{I}_{[g_{j^*} > g_{i+1}]}\big).$$

Then the limit on the right-hand side of the above inequality is zero on the subsequence, which contradicts Theorem 4.7 Part (2). (Notice that the indicator functions in the last line of the inequality above are not functions of r; that is, they evaluate to zero or one for all $r \geq r_{j0}$. Thus we need not consider $\tilde{\alpha}_{i-1}^*/\tilde{\alpha}_{j^*}^* \to \infty$ or $\tilde{\alpha}_{i+1}^*/\tilde{\alpha}_{j^*}^* \to \infty$ when at least one indicator function evaluates to zero.)

J.5 Proof of Theorem 4.7 Part (5)

We prove the upper bound first, then the lower bound.

J.5.1 Theorem 4.7 Part (5): Upper Bound. From the proof of Proposition 4.5, we have

$$1/\tilde{\alpha}_j^* \leq 2\kappa_R^{\mathrm{U}}/\tilde{z}^* \quad \text{for all } j \in \mathcal{P}^c, r \geq r_{j0}. \tag{35}$$

Now from Theorem 4.7 Part (2), there exists $\tau_2^{\mathrm{L}} \in (0, \infty)$ such that $\tilde{\alpha}_j^* \geq \tau_2^{\mathrm{L}} \tilde{z}^*$ for all $j \in \mathcal{P}^c$ and all $r \geq r_{j0}$. From Theorem 4.7 Part (4), if $j^* \in \mathcal{P}^c(i)$ is such that $g_{j^*} < g_{i+1}$ or $h_{j^*} < h_{i-1}$, then $\tilde{\alpha}_i^* \geq \kappa_4 \tilde{\alpha}_{j^*}^* \geq \kappa_4 \tau_2^{\mathrm{L}} \tilde{z}^*$. Since there exists such a $j^* \in \mathcal{P}^c(i)$ for all $i \in \mathcal{P}$ and all r under Assumption 9, it follows that

$$1/\tilde{\alpha}_i^* \leq (\kappa_4 \tau_2^{\mathrm{L}})^{-1}/\tilde{z}^* \quad \text{for all } i \in \mathcal{P} \text{ and all } r. \tag{36}$$

Let $\kappa_2 := 2\kappa_R^{\mathrm{U}} + (\kappa_4 \tau_2^{\mathrm{L}})^{-1}$. Taking equations (35) and (36) together, we have

$$1/\tilde{\alpha}_j^* + 1/\tilde{\alpha}_i^* \leq \kappa_2/\tilde{z}^* \quad \text{for all } i \in \mathcal{P}, j \in \mathcal{P}^c, r \geq r_{j0}. \tag{37}$$

J.5.2 Theorem 4.7 Part (5): Lower Bound. From Theorem 4.7 Part (2), $1/\tilde{\alpha}_j^* \geq \kappa_1/\tilde{z}^*$ for all $j \in \mathcal{P}^c, r \geq r_{j0}$. Combining this fact with (37), $\kappa_1/\tilde{z}^* \leq 1/\tilde{\alpha}_j^* \leq 1/\tilde{\alpha}_j^* + 1/\tilde{\alpha}_i^* \leq \kappa_2/\tilde{z}^*$.

J.6 Proof of Theorem 4.7 Part (6)

From the KKT conditions in equation (5) and using the bounds in Lemma 4.4, for all $i, i' \in \mathcal{P}$,

$$1 = \frac{\sum_{j \in \mathcal{P}^c} \lambda_{j\ell-1}(r) J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell}(r) K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))}{\sum_{j \in \mathcal{P}^c} \lambda_{j\ell'-1}(r) J_{i'}(x_{i'}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'-1}^*, \tilde{\alpha}_{i'}^*)) + \lambda_{j\ell'}(r) K_{i'}(y_{i'}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'}^*, \tilde{\alpha}_{i'+1}^*))}$$

$$\leq \frac{\sum_{j \in \mathcal{P}^c} \kappa_R^{\mathrm{U}} \Big[\frac{\lambda_{j\ell-1}(r)\mathbb{I}_{j\ell-1}^g(r) + \lambda_{j\ell}(r)\mathbb{I}_{j\ell}^h(r)}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2}\Big]}{\sum_{j \in \mathcal{P}^c} \kappa_R^{\mathrm{L}} \Big[\frac{\lambda_{j\ell'-1}(r)\mathbb{I}_{j\ell'-1}^g(r) + \lambda_{j\ell'}(r)\mathbb{I}_{j\ell'}^h(r)}{(1+\tilde{\alpha}_{i'}^*/\tilde{\alpha}_j^*)^2}\Big]} = \frac{\kappa_R^{\mathrm{U}}}{\kappa_R^{\mathrm{L}}} \frac{\tilde{\alpha}_{i'}^{*2}}{\tilde{\alpha}_i^{*2}} \frac{\sum_{j \in \mathcal{P}^c} \Big[\frac{\lambda_{j\ell-1}(r)\mathbb{I}_{j\ell-1}^g(r) + \lambda_{j\ell}(r)\mathbb{I}_{j\ell}^h(r)}{(1/\tilde{\alpha}_i^* + 1/\tilde{\alpha}_j^*)^2}\Big]}{\sum_{j \in \mathcal{P}^c} \Big[\frac{\lambda_{j\ell'-1}(r)\mathbb{I}_{j\ell'-1}^g(r) + \lambda_{j\ell'}(r)\mathbb{I}_{j\ell'}^h(r)}{(1/\tilde{\alpha}_{i'}^* + 1/\tilde{\alpha}_j^*)^2}\Big]}. \tag{38}$$

From Theorem 4.7 Part (5), it follows that

$$\kappa_1^2/\tilde{z}^{*2} \leq (1/\tilde{\alpha}_j^* + 1/\tilde{\alpha}_i^*)^2 \leq \kappa_2^2/\tilde{z}^{*2} \quad \text{for all } i \in \mathcal{P}, j \in \mathcal{P}^c, r \geq r_{j0}. \tag{39}$$

Continuing from (38) and using line (39) and equation (9), respectively, we have

$$\frac{\tilde{\alpha}_i^{*2}}{\tilde{\alpha}_{i'}^{*2}} \leq \frac{\kappa_R^{\mathrm{U}}}{\kappa_R^{\mathrm{L}}} \frac{\sum_{j \in \mathcal{P}^c} \Big[\frac{\lambda_{j\ell-1}(r)\mathbb{I}_{j\ell-1}^g(r) + \lambda_{j\ell}(r)\mathbb{I}_{j\ell}^h(r)}{\kappa_1^2/\tilde{z}^{*2}}\Big]}{\sum_{j \in \mathcal{P}^c} \Big[\frac{\lambda_{j\ell'-1}(r)\mathbb{I}_{j\ell'-1}^g(r) + \lambda_{j\ell'}(r)\mathbb{I}_{j\ell'}^h(r)}{\kappa_2^2/\tilde{z}^{*2}}\Big]} = \frac{\kappa_R^{\mathrm{U}}\kappa_2^2}{\kappa_R^{\mathrm{L}}\kappa_1^2} \frac{\Big[\sum_{j \in \mathcal{P}^c} \lambda_{j\ell-1}(r)\mathbb{I}_{j\ell-1}^g(r)\Big] + \Big[\sum_{j \in \mathcal{P}^c} \lambda_{j\ell}(r)\mathbb{I}_{j\ell}^h(r)\Big]}{\sum_{j \in \mathcal{P}^c} \Big[\lambda_{j\ell'-1}(r)\mathbb{I}_{j\ell'-1}^g(r) + \lambda_{j\ell'}(r)\mathbb{I}_{j\ell'}^h(r)\Big]}$$

$$\leq \frac{\kappa_R^{\mathrm{U}}\kappa_2^2}{\kappa_R^{\mathrm{L}}\kappa_1^2} \frac{2}{\sum_{j \in \mathcal{P}^c(i')} \Big[\lambda_{j\ell'-1}(r)\mathbb{I}_{j\ell'-1}^g(r)\mathbb{I}_{[\ell^*(j)=\ell'-1]} + \lambda_{j\ell'}(r)\mathbb{I}_{j\ell'}^h(r)\mathbb{I}_{[\ell^*(j)=\ell']}\Big]}. \tag{40}$$

From Theorem 4.7 Part (3), for all $i \in \mathcal{P}$, there exists $\tau_3^{\mathrm{L}} \in (0, \infty)$ such that

$$\textstyle\sum_{j \in \mathcal{P}^c(i)} \lambda_{j\ell-1}(r)\mathbb{I}_{j\ell-1}^g(r)\mathbb{I}_{[\ell^*(j)=\ell-1]} + \lambda_{j\ell}(r)\mathbb{I}_{j\ell}^h(r)\mathbb{I}_{[\ell^*(j)=\ell]}$$

$$\geq \textstyle\sum_{j\in\mathcal{P}^c(i)} \tau_3^{\rm L} \frac{1}{|\mathcal{P}^c|} = \tau_3^{\rm L} \frac{|\mathcal{P}^c(i)|}{|\mathcal{P}^c|} \geq \tau_3^{\rm L}\kappa, \tag{41}$$

where the last step follows by Assumption 9. Combining this bound with (40), we have

$$\frac{\tilde{\alpha}_i^{*2}}{\tilde{\alpha}_{i'}^{*2}} \leq \frac{2\kappa_R^{\rm U}\kappa_2^2}{\kappa_R^{\rm L}\kappa_1^2\tau_3^{\rm L}\kappa},$$

and the result follows.

J.7 Proof of Theorem 4.7 Part (7)

First we prove the lower bound, then we prove the upper bound. In what follows, notice that from Theorem 4.7 Part (2), there exists $\tau_2 \in (0,\infty)$ such that $\tilde{\alpha}_j^*/\tilde{\alpha}_{j'}^* < \tau_2$ for all $j, j' \in \mathcal{P}^c$ and all $r \geq \max\{r_{j0}, r_{j'0}\}$.

J.7.1 Proof of Theorem 4.7 Part (7): Lower Bound. Let $i \in \mathcal{P}, j' \in \mathcal{P}^c$. From the KKT conditions in equation (6), the bounds in Lemma 4.4, the inequality in (39), and Assumption 9, respectively,

$$\begin{aligned}
1 &= \sum_{j\in\mathcal{P}^c} \frac{\lambda_{j\ell-1}J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell}K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))}{\sum_{\ell'\in\mathcal{P}^{\rm ph},\ell'=i'} \lambda_{j\ell'}I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'}^*, \tilde{\alpha}_{i'+1}^*))} \\
&\leq \sum_{j\in\mathcal{P}^c} \frac{\kappa_R^{\rm U}\left[\frac{\lambda_{j\ell-1}\mathbb{I}_{j\ell-1}^g + \lambda_{j\ell}\mathbb{I}_{j\ell}^h}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2}\right]}{\sum_{\ell'\in\mathcal{P}^{\rm ph},\ell'=i'} \lambda_{j\ell'}\kappa_R^{\rm L}\min\left[\frac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i'}^*)^2}, \frac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i'+1}^*)^2}\right]} \\
&= \frac{\kappa_R^{\rm U}}{\kappa_R^{\rm L}}\frac{1}{\tilde{\alpha}_i^{*2}}\sum_{j\in\mathcal{P}^c}\tilde{\alpha}_j^{*2} \frac{\left[\frac{\lambda_{j\ell-1}\mathbb{I}_{j\ell-1}^g + \lambda_{j\ell}\mathbb{I}_{j\ell}^h}{(1/\tilde{\alpha}_i^*+1/\tilde{\alpha}_j^*)^2}\right]}{\sum_{\ell'\in\mathcal{P}^{\rm ph},\ell'=i'} \lambda_{j\ell'}\min\left[\frac{1}{(1/\tilde{\alpha}_j^*+1/\tilde{\alpha}_{i'}^*)^2}, \frac{1}{(1/\tilde{\alpha}_j^*+1/\tilde{\alpha}_{i'+1}^*)^2}\right]} \\
&\leq \frac{\kappa_R^{\rm U}\kappa_2^2}{\kappa_R^{\rm L}\kappa_1^2}\frac{1}{\tilde{\alpha}_i^{*2}}\sum_{j\in\mathcal{P}^c}\tilde{\alpha}_j^{*2}\frac{\left[\lambda_{j\ell-1}\mathbb{I}_{j\ell-1}^g + \lambda_{j\ell}\mathbb{I}_{j\ell}^h\right]}{\sum_{\ell'\in\mathcal{P}^{\rm ph}}\lambda_{j\ell'}} \\
&\leq \frac{\kappa_R^{\rm U}\kappa_2^2}{\kappa_R^{\rm L}\kappa_1^2}\frac{1}{\tilde{\alpha}_i^{*2}}\sum_{j\in\mathcal{P}^c}\tilde{\alpha}_j^{*2}\frac{\left[2\kappa_0\lambda_{j\ell^*}\right]}{\lambda_{j\ell^*}} = \frac{\kappa_R^{\rm U}\kappa_2^2 2\kappa_0}{\kappa_R^{\rm L}\kappa_1^2}\frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}}\sum_{j\in\mathcal{P}^c}\frac{\tilde{\alpha}_j^{*2}}{\tilde{\alpha}_{j'}^{*2}} \leq \frac{\kappa_R^{\rm U}\kappa_2^2 2\kappa_0}{\kappa_R^{\rm L}\kappa_1^2}\frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}}|\mathcal{P}^c|\tau_2^2.
\end{aligned}$$

Then letting $\tau_7^{\rm L} := \frac{\kappa_R^{\rm L}\kappa_1^2}{2\kappa_R^{\rm U}\kappa_0\kappa_2^2\tau_2^2}$, it follows that

$$\frac{\tau_7^{\rm L}}{|\mathcal{P}^c|} \leq \frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} \quad \text{for all } i\in\mathcal{P}, j'\in\mathcal{P}^c, r\geq r_{j'0}. \tag{42}$$

J.7.2 Proof of Theorem 4.7 Part (7): Upper Bound. Let $i \in \mathcal{P}, j' \in \mathcal{P}^c$. By Theorem 4.7 Part (3), there exists $\tau_3^{\rm U} \in (0,\infty)$ such that $\lambda_{j\ell^*(j)} \leq \tau_3^{\rm U}/|\mathcal{P}^c|$ for all $j \in \mathcal{P}^c, r \geq r_{j0}$. Then from the KKT conditions in equation (6), the bounds in Lemma 4.4, the inequality in (39), Assumption 9, and line (41), respectively,

$$\begin{aligned}
1 &= \sum_{j\in\mathcal{P}^c} \frac{\lambda_{j\ell-1}J_i(x_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i-1}^*, \tilde{\alpha}_i^*)) + \lambda_{j\ell}K_i(y_i^*(\tilde{\alpha}_j^*, \tilde{\alpha}_i^*, \tilde{\alpha}_{i+1}^*))}{\sum_{\ell'\in\mathcal{P}^{\rm ph},\ell'=i'} \lambda_{j\ell'}I_j(\mathfrak{z}_j^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i'}^*, \tilde{\alpha}_{i'+1}^*))} \\
&\geq \sum_{j\in\mathcal{P}^c} \frac{\kappa_R^{\rm L}\left[\frac{\lambda_{j\ell-1}\mathbb{I}_{j\ell-1}^g + \lambda_{j\ell}\mathbb{I}_{j\ell}^h}{(1+\tilde{\alpha}_i^*/\tilde{\alpha}_j^*)^2}\right]}{\sum_{\ell'\in\mathcal{P}^{\rm ph},\ell'=i'} \lambda_{j\ell'}\kappa_R^{\rm U}\left[\frac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i'}^*)^2} + \frac{1}{(1+\tilde{\alpha}_j^*/\tilde{\alpha}_{i'+1}^*)^2}\right]}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\kappa_R^{\mathrm{L}}}{\kappa_R^{\mathrm{U}}} \frac{1}{\tilde{\alpha}_i^{*2}} \sum_{j \in \mathcal{P}^c} \tilde{\alpha}_j^{*2} \frac{\left[\frac{\lambda_{j\ell-1}\mathbb{I}^g_{j\ell-1} + \lambda_{j\ell}\mathbb{I}^h_{j\ell}}{(1/\tilde{\alpha}_i^* + 1/\tilde{\alpha}_j^*)^2}\right]}{\sum_{\ell' \in \mathcal{P}^{\mathrm{ph}}, \ell' = i'} \lambda_{j\ell'}\left[\frac{1}{(1/\tilde{\alpha}_j^* + 1/\tilde{\alpha}_{i'}^*)^2} + \frac{1}{(1/\tilde{\alpha}_j^* + 1/\tilde{\alpha}_{i'+1}^*)^2}\right]} \\
&\geq \frac{\kappa_R^{\mathrm{L}}}{\kappa_R^{\mathrm{U}}} \frac{\kappa_1^2}{2\kappa_2^2} \frac{1}{\tilde{\alpha}_i^{*2}} \sum_{j \in \mathcal{P}^c} \tilde{\alpha}_j^{*2} \frac{[\lambda_{j\ell-1}\mathbb{I}^g_{j\ell-1} + \lambda_{j\ell}\mathbb{I}^h_{j\ell}]}{\sum_{\ell' \in \mathcal{P}^{\mathrm{ph}}, \ell' = i'} \lambda_{j\ell'}} \geq \frac{\kappa_R^{\mathrm{L}}}{\kappa_R^{\mathrm{U}}} \frac{\kappa_1^2}{2\kappa_2^2} \frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} \sum_{j \in \mathcal{P}^c(i)} \frac{\tilde{\alpha}_j^{*2}}{\tilde{\alpha}_{j'}^{*2}} \frac{[\lambda_{j\ell-1}\mathbb{I}^g_{j\ell-1} + \lambda_{j\ell}\mathbb{I}^h_{j\ell}]}{(p+1)\kappa_0 \lambda_{j\ell^*(j)}} \\
&\geq \frac{\kappa_R^{\mathrm{L}}}{\kappa_R^{\mathrm{U}}} \frac{\kappa_1^2}{2\kappa_2^2} \frac{\tau_2^{-2}}{(p+1)\kappa_0} \frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} \sum_{j \in \mathcal{P}^c(i)} \frac{\lambda_{j\ell-1}\mathbb{I}^g_{j\ell-1}\mathbb{I}_{[\ell^*(j) = \ell-1]} + \lambda_{j\ell}\mathbb{I}^h_{j\ell}\mathbb{I}_{[\ell^*(j) = \ell]}}{\lambda_{j\ell^*(j)}} \\
&\geq \frac{\kappa_R^{\mathrm{L}}}{\kappa_R^{\mathrm{U}}} \frac{\kappa_1^2}{2\kappa_2^2} \frac{\tau_2^{-2}}{(p+1)\kappa_0} \frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} \sum_{j \in \mathcal{P}^c(i)} \frac{\tau_3^{\mathrm{L}}/|\mathcal{P}^c|}{\tau_3^{\mathrm{U}}/|\mathcal{P}^c|} = \frac{\kappa_R^{\mathrm{L}}\kappa_1^2\tau_3^{\mathrm{L}}}{2(p+1)\kappa_R^{\mathrm{U}}\kappa_0\kappa_2^2\tau_2^2\tau_3^{\mathrm{U}}} \frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} |\mathcal{P}^c(i)| \\
&\geq \frac{\kappa_R^{\mathrm{L}}\kappa_1^2\tau_3^{\mathrm{L}}\kappa}{2(p+1)\kappa_R^{\mathrm{U}}\kappa_0\kappa_2^2\tau_2^2\tau_3^{\mathrm{U}}} \frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} |\mathcal{P}^c|.
\end{aligned}
$$

Then letting $\tau_7^{\mathrm{U}} := \frac{2(p+1)\kappa_R^{\mathrm{U}}\kappa_0\kappa_2^2\tau_2^2\tau_3^{\mathrm{U}}}{\kappa_R^{\mathrm{L}}\kappa_1^2\tau_3^{\mathrm{L}}\kappa}$, it follows that

$$
\frac{\tilde{\alpha}_{j'}^{*2}}{\tilde{\alpha}_i^{*2}} \leq \frac{\tau_7^{\mathrm{U}}}{|\mathcal{P}^c|} \quad \text{for all } i \in \mathcal{P}, j' \in \mathcal{P}^c, r \geq r_{j'0}. \tag{43}
$$

J.8 Proof of Theorem 4.7 Part (8)

Proposition 4.5 provides the upper bound. For the lower bound, for all $j \in \mathcal{P}^c$ with primary phantom Pareto system $\ell^*(j) \in \mathcal{P}^{\mathrm{ph}}, \ell^*(j) = i^*(j)$, we have

$$
\tilde{z}^* \geq \tilde{\alpha}_j^* I_{j\ell^*(j)}(\mathfrak{z}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*)) \geq \tilde{\alpha}_j^* \kappa_1.
$$

Summing over $j \in \mathcal{P}^c$ yields

$$
\begin{aligned}
|\mathcal{P}^c|\tilde{z}^* &= \textstyle\sum_{j \in \mathcal{P}^c} R_{j\ell^*(j)}(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*) \geq \sum_{j \in \mathcal{P}^c} \tilde{\alpha}_j^* I_{j\ell^*(j)}(\mathfrak{z}^*(\tilde{\alpha}_j^*, \tilde{\alpha}_{i^*(j)}^*, \tilde{\alpha}_{i^*(j)+1}^*)) \\
&\geq \kappa_1 \textstyle\sum_{j \in \mathcal{P}^c} \tilde{\alpha}_j^* = \kappa_1(1 - \sum_{i \in \mathcal{P}} \tilde{\alpha}_i^*).
\end{aligned} \tag{44}
$$

From Proposition 4.5 and Theorem 4.7 Parts (2) and (7), it must be the case that $\tilde{\alpha}_i^* \to 0$ for all $i \in \mathcal{P}$. Thus $(1 - \sum_{i \in \mathcal{P}} \tilde{\alpha}_i^*)$ is a constant, and we have the result for $\tilde{z}^*$. The result for $\tilde{\alpha}_j^*$ follows from the result for $\tilde{z}^*$ and Theorem 4.7 Part (2).

J.9 Proof of Theorem 4.7 Part (9)

This result follows from Theorem 4.7 Part (7) and the bounds in Lemma 4.4.

J.10 Proof of Theorem 4.7 Part (10)

This result follows by noticing that Theorem 4.7 Part (7) implies $w_g \to 0$ and $w_h \to 0$ in the proof of Proposition 4.1.

K PROOF OF THEOREM 4.11

From (42) and (43), for all $i \in \mathcal{P}, j \in \mathcal{P}^c, r \geq r_{j0}$, $\tilde{\alpha}_j^{*2}/\tau_7^{\mathrm{U}} \leq \tilde{\alpha}_i^{*2}/|\mathcal{P}^c| \leq \tilde{\alpha}_j^{*2}/\tau_7^{\mathrm{L}}$. Summing across $j \in \mathcal{P}^c$ yields $\sqrt{(1/\tau_7^{\mathrm{U}})\sum_{j \in \mathcal{P}^c} \tilde{\alpha}_j^{*2}} \leq \tilde{\alpha}_i^* \leq \sqrt{(1/\tau_7^{\mathrm{L}})\sum_{j \in \mathcal{P}^c} \tilde{\alpha}_j^{*2}}$. Applying that $\tilde{\alpha}_j^* = \Theta(1/|\mathcal{P}^c|)$ from Theorem 4.7 Part (8) yields the result.

L PROOF OF THEOREM 4.12

The only difference between Problems Q and $\widetilde{Q}$ is that Problem Q includes constraints corresponding to the rate functions for MCE events between all pairs of Pareto systems. The assertion will hold if we show that for large enough $|\mathcal{P}^c|$, $R_i(\tilde{\alpha}_i^*, \tilde{\alpha}_{i'}^*) > \tilde{z}^*$ for all $i, i' \in \mathcal{P}$, $i' \neq i$.

Recall that for all $i, i' \in \mathcal{P}$, $i' \neq i$, we have

$$R_i(\alpha_i, \alpha_{i'}) = \inf_{x_{i'} \leq x_i,\ y_{i'} \leq y_i} \alpha_i I_i(x_i, y_i) + \alpha_{i'} I_{i'}(x_{i'}, y_{i'}).$$

Under our assumptions, notice that there exists a constant $a > 0$ such that $\inf_{x_{i'} \leq x_i,\ y_{i'} \leq y_i} I_i(x_i, y_i) + I_{i'}(x_{i'}, y_{i'}) > a$ for all $i, i' \in \mathcal{P}, i' \neq i$. From Theorem 4.7 Part (6), $\tilde{\alpha}_i^*/\tilde{\alpha}_{i'}^* \leq \kappa_6$ for all $i, i' \in \mathcal{P}, i' \neq i$. Supposing without loss of generality that $\kappa_6 > 1$,

$$\begin{aligned} R_i(\tilde{\alpha}_i^*, \tilde{\alpha}_{i'}^*)/\tilde{\alpha}_i^* &= \inf_{x_{i'} \leq x_i,\ y_k i' \leq y_i} I_i(x_i, y_i) + (\tilde{\alpha}_{i'}^*/\tilde{\alpha}_i^*) I_{i'}(x_{i'}, y_{i'}) \\ &\geq \inf_{x_{i'} \leq x_i,\ y_{i'} \leq y_i} I_i(x_i, y_i) + \kappa_6^{-1} I_k(x_{i'}, y_{i'}) \geq \kappa_6^{-1} a > 0. \end{aligned}$$

Thus letting $\eta := \kappa_6^{-1} a > 0$, $R_i(\tilde{\alpha}_i^*, \tilde{\alpha}_{i'}^*)/\tilde{z}^* \geq \eta(\tilde{\alpha}_i^*/\tilde{z}^*)$. From Theorem 4.7 Part (8), we have $\tilde{z}^* = \Theta(1/|\mathcal{P}^c|)$, and from Theorem 4.11, we have $\tilde{\alpha}_i^* = \Theta(1/\sqrt{|\mathcal{P}^c|})$. Then it follows that $\tilde{\alpha}_i^*/\tilde{z}^* \to \infty$. Thus for large enough $|\mathcal{P}^c|$, the Pareto systems receive a large enough proportion of the allocation that the constraints corresponding to MCE in Problem Q are not binding, and the result holds.

M PROOF OF PROPOSITION 5.1

Under Assumption 6, Problem $R_{ii'}^{\text{MCE}}$ is a quadratic program with linear constraints. The KKT conditions are necessary and sufficient for global optimality. Let $\lambda_x^{\mathcal{P}} \geq 0$ and $\lambda_y^{\mathcal{P}} \geq 0$ be dual variables. In addition to primal feasibility, we have the complementary slackness conditions $\lambda_x^{\mathcal{P}}(x_{i'} - x_i) = 0$ and $\lambda_y^{\mathcal{P}}(y_{i'} - y_i) = 0$, and the stationarity conditions

$$\alpha_i \frac{\partial I_i(x_i, y_i)}{\partial x_i} - \lambda_x^{\mathcal{P}} = 0 \qquad \alpha_i \frac{\partial I_i(x_i, y_i)}{\partial y_i} - \lambda_y^{\mathcal{P}} = 0$$

$$\alpha_{i'} \frac{\partial I_{i'}(x_{i'}, y_{i'})}{\partial x_{i'}} + \lambda_x^{\mathcal{P}} = 0 \qquad \alpha_{i'} \frac{\partial I_{i'}(x_{i'}, y_{i'})}{\partial y_{i'}} + \lambda_y^{\mathcal{P}} = 0,$$

which simplify to

$$\frac{\alpha_i}{(1-\rho_i^2)}\left(\frac{x_i - g_i}{\sigma_{g_i}^2} - \frac{\rho_i(y_i - h_i)}{\sigma_{g_i}\sigma_{h_i}}\right) = \lambda_x^{\mathcal{P}} \qquad \frac{\alpha_i}{(1-\rho_i^2)}\left(\frac{y_i - h_i}{\sigma_{h_i}^2} - \frac{\rho_i(x_i - g_i)}{\sigma_{g_i}\sigma_{h_i}}\right) = \lambda_y^{\mathcal{P}}$$

$$\frac{\alpha_{i'}}{(1-\rho_{i'}^2)}\left(\frac{x_{i'} - g_{i'}}{\sigma_{g_{i'}}^2} - \frac{\rho_{i'}(y_{i'} - h_{i'})}{\sigma_{g_{i'}}\sigma_{h_{i'}}}\right) = -\lambda_x^{\mathcal{P}} \qquad \frac{\alpha_{i'}}{(1-\rho_{i'}^2)}\left(\frac{y_{i'} - h_{i'}}{\sigma_{h_{i'}}^2} - \frac{\rho_{i'}(x_{i'} - g_{i'})}{\sigma_{g_{i'}}\sigma_{h_{i'}}}\right) = -\lambda_y^{\mathcal{P}}.$$

Since $\lambda_x^{\mathcal{P}} \geq 0$ and $\lambda_x^{\mathcal{P}} \geq 0$, then

$$x_i \geq g_i + \rho_i \frac{\sigma_{g_i}}{\sigma_{h_i}}(y_i - h_i) \qquad y_i \geq h_i + \rho_i \frac{\sigma_{h_i}}{\sigma_{g_i}}(x_i - g_i)$$

$$x_{i'} \leq g_{i'} + \rho_{i'} \frac{\sigma_{g_{i'}}}{\sigma_{h_{i'}}}(y_{i'} - h_{i'}) \qquad y_{i'} \leq h_{i'} + \rho_{i'} \frac{\sigma_{h_{i'}}}{\sigma_{g_{i'}}}(x_{i'} - g_{i'}).$$

Since $i, k \in \mathcal{P}$, we cannot have $\lambda_x^{\mathcal{P}} = 0$ and $\lambda_x^{\mathcal{P}} = 0$. We consider three cases, as follows.

Case 1: $\lambda_x^{\mathcal{P}} > 0$ and $\lambda_y^{\mathcal{P}} = 0$. Then $x_i^* = x_{i'}^*$; solving, we find

$$x_i^* = x_{i'}^* = \frac{(\alpha_i/\sigma_{g_i}^2)g_i + (\alpha_{i'}/\sigma_{g_{i'}}^2)g_{i'}}{\alpha_i/\sigma_{g_i}^2 + \alpha_{i'}/\sigma_{g_{i'}}^2} \qquad y_i^* = h_i + \rho_i \frac{\sigma_{h_i}}{\sigma_{g_i}}\left(\frac{(\alpha_{i'}/\sigma_{g_{i'}}^2)(g_{i'} - g_i)}{\alpha_i/\sigma_{g_i}^2 + \alpha_{i'}/\sigma_{g_{i'}}^2}\right)$$

$$y_{i'}^* = h_{i'} - \rho_{i'} \frac{\sigma_{h_{i'}}}{\sigma_{g_{i'}}}\left(\frac{(\alpha_i/\sigma_{g_i}^2)(g_{i'} - g_i)}{\alpha_i/\sigma_{g_i}^2 + \alpha_{i'}/\sigma_{g_{i'}}^2}\right),$$

which implies $g_i \leq g_{i'}$. Primal feasibility further implies

$$h_i \geq h_{i'} - (g_{i'} - g_i)\left(\frac{\rho_{i'}(\sigma_{h_{i'}}/\sigma_{g_{i'}})(\alpha_i/\sigma^2_{g_i}) + \rho_i(\sigma_{h_i}/\sigma_{g_i})(\alpha_{i'}/\sigma^2_{g_{i'}})}{\alpha_i/\sigma^2_{g_i} + \alpha_{i'}/\sigma^2_{g_{i'}}}\right).$$

Substituting into the objective function yields the result in this case.

Case 2: $\lambda^{\mathcal{P}}_x = 0$ *and* $\lambda^{\mathcal{P}}_y > 0$. Then $y^*_i = y^*_{i'}$; solving, we find

$$x^*_i = g_i + \rho_i \frac{\sigma_{g_i}}{\sigma_{h_i}}\left(\frac{(\alpha_{i'}/\sigma^2_{h_{i'}})(h_{i'} - h_i)}{\alpha_i/\sigma^2_{h_i} + \alpha_{i'}/\sigma^2_{h_{i'}}}\right) \qquad y^*_i = y^*_{i'} = \frac{(\alpha_i/\sigma^2_{h_i})h_i + (\alpha_{i'}/\sigma^2_{h_{i'}})h_{i'}}{\alpha_i/\sigma^2_{h_i} + \alpha_{i'}/\sigma^2_{h_{i'}}}$$

$$x^*_{i'} = g_{i'} - \rho_{i'} \frac{\sigma_{g_{i'}}}{\sigma_{h_{i'}}}\left(\frac{(\alpha_i/\sigma^2_{h_i})(h_{i'} - h_i)}{\alpha_i/\sigma^2_{h_i} + \alpha_{i'}/\sigma^2_{h_{i'}}}\right),$$

which implies $h_i \leq h_{i'}$. Primal feasibility further implies

$$g_i \geq g_{i'} - (h_{i'} - h_i)\left(\frac{\rho_{i'}(\sigma_{g_{i'}}/\sigma_{h_{i'}})(\alpha_i/\sigma^2_{h_i}) + \rho_i(\sigma_{g_i}/\sigma_{h_i})(\alpha_{i'}/\sigma^2_{h_{i'}})}{\alpha_i/\sigma^2_{h_i} + \alpha_{i'}/\sigma^2_{h_{i'}}}\right).$$

Substituting into the objective function yields the result in this case.

Case 3: $\lambda^{\mathcal{P}}_x > 0$ *and* $\lambda^{\mathcal{P}}_y > 0$. Then $x^*_i = x^*_{i'}$ and $y^*_i = y^*_{i'}$; solving, we find

$$\begin{aligned}
x^*_i = x^*_{i'} = &\frac{\left[\frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\frac{\sigma^2_{g_{i'}}}{\alpha_{i'}} - \frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}\frac{\sigma^2_{g_i}}{\alpha_i}\right](h_{i'} - h_i)}{(\sigma^2_{g_i}/\alpha_i + \sigma^2_{g_{i'}}/\alpha_{i'})(\sigma^2_{h_i}/\alpha_i + \sigma^2_{h_{i'}}/\alpha_{i'}) - (\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i + \rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2} \\
&+ \frac{\left[\frac{\sigma^2_{g_i}\sigma^2_{h_i}(1-\rho^2_i)}{\alpha^2_i} + \frac{\sigma^2_{g_i}}{\alpha_i}\frac{\sigma^2_{h_{i'}}}{\alpha_{i'}} - \frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}\right]g_{i'}}{(\sigma^2_{g_i}/\alpha_i + \sigma^2_{g_{i'}}/\alpha_{i'})(\sigma^2_{h_i}/\alpha_i + \sigma^2_{h_{i'}}/\alpha_{i'}) - (\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i + \rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2} \\
&+ \frac{\left[\frac{\sigma^2_{g_{i'}}\sigma^2_{h_{i'}}(1-\rho^2_{i'})}{\alpha^2_{i'}} + \frac{\sigma^2_{g_{i'}}}{\alpha_{i'}}\frac{\sigma^2_{h_i}}{\alpha_i} - \frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}\right]g_i}{(\sigma^2_{g_i}/\alpha_i + \sigma^2_{g_{i'}}/\alpha_{i'})(\sigma^2_{h_i}/\alpha_i + \sigma^2_{h_{i'}}/\alpha_{i'}) - (\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i + \rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2} \\
y^*_i = y^*_{i'} = &\frac{\left[\frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\frac{\sigma^2_{h_{i'}}}{\alpha_{i'}} - \frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}\frac{\sigma^2_{h_i}}{\alpha_i}\right](g_{i'} - g_i)}{(\sigma^2_{g_i}/\alpha_i + \sigma^2_{g_{i'}}/\alpha_{i'})(\sigma^2_{h_i}/\alpha_i + \sigma^2_{h_{i'}}/\alpha_{i'}) - (\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i + \rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2} \\
&+ \frac{\left[\frac{\sigma^2_{g_i}\sigma^2_{h_i}(1-\rho^2_i)}{\alpha^2_i} + \frac{\sigma^2_{g_{i'}}}{\alpha_{i'}}\frac{\sigma^2_{h_i}}{\alpha_i} - \frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}\right]h_{i'}}{(\sigma^2_{g_i}/\alpha_i + \sigma^2_{g_{i'}}/\alpha_{i'})(\sigma^2_{h_i}/\alpha_i + \sigma^2_{h_{i'}}/\alpha_{i'}) - (\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i + \rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2} \\
&+ \frac{+\left[\frac{\sigma^2_{g_{i'}}\sigma^2_{h_{i'}}(1-\rho^2_{i'})}{\alpha^2_{i'}} + \frac{\sigma^2_{g_i}}{\alpha_i}\frac{\sigma^2_{h_{i'}}}{\alpha_{i'}} - \frac{\rho_i\sigma_{g_i}\sigma_{h_i}}{\alpha_i}\frac{\rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}}{\alpha_{i'}}\right]h_i}{(\sigma^2_{g_i}/\alpha_i + \sigma^2_{g_{i'}}/\alpha_{i'})(\sigma^2_{h_i}/\alpha_i + \sigma^2_{h_{i'}}/\alpha_{i'}) - (\rho_i\sigma_{g_i}\sigma_{h_i}/\alpha_i + \rho_{i'}\sigma_{g_{i'}}\sigma_{h_{i'}}/\alpha_{i'})^2}.
\end{aligned}$$

It can also be shown that the rate functions depend on the locations of the systems, as in the proof of Proposition 4.1. We do not the provide the details of this proof.

N TEST PROBLEMS

We consider test problems having the following objective function values (g_k, h_k) for all systems $k \leq r$, where the values are truncated to the fourth decimal place.

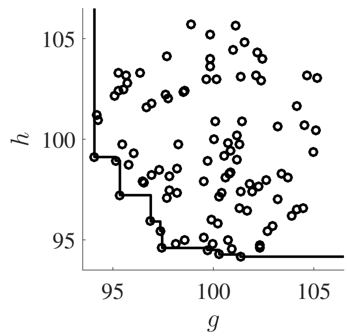

Fig. 15. Test Problems 1A, 1B, 1C: $r = 100, |\mathcal{P}| = 9$

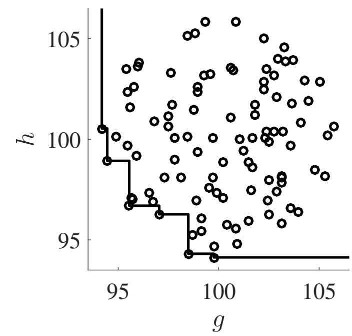

Fig. 16. Test Probelms 2A, 2B, 2C: $r = 100, |\mathcal{P}| = 6$

N.1 Objective Function Values for Test Problems 1A, 1B, and 1C

(100.7668, 98.2846)	(101.5803, 104.8465)	(104.1497, 96.4957)	(98.2325, 99.7302)
(101.6930, 96.4598)	(102.4406, 103.9730)	(99.7297, 98.8975)	(98.5848, 94.9992)
(104.4776, 96.5696)	(100.2660, 97.1450)	(102.3980, 102.8872)	(100.2067, 95.8049)
(98.1913, 98.5706)	(95.3507, 97.2208)	(97.7111, 97.0571)	(97.8318, 97.4372)
(96.3599, 103.2968)	(95.1358, 98.9354)	(105.1037, 100.4531)	(101.3503, 94.1658)
(97.6691, 104.1171)	(100.2951, 94.2726)	(96.0538, 99.3129)	(97.3253, 98.4759)
(100.8665, 98.3810)	(101.3779, 103.1052)	(96.5450, 97.8568)	(104.5639, 100.6942)
(103.1941, 96.9960)	(100.9392, 94.5154)	(96.6498, 101.6032)	(98.5302, 102.3237)
(99.5010, 95.1233)	(97.8254, 98.1723)	(100.7589, 99.7948)	(95.6396, 103.1989)
(104.9940, 99.3581)	(94.0823, 99.1163)	(101.1593, 98.9937)	(97.3611, 95.4519)
(102.1225, 103.2007)	(97.5961, 102.2120)	(102.3316, 94.6044)	(102.1988, 104.3315)
(102.6375, 97.9407)	(100.5365, 99.1881)	(104.6868, 103.2014)	(99.6457, 102.9520)
(95.7051, 102.7686)	(100.0106, 99.9755)	(96.4683, 97.9059)	(99.5199, 97.8854)
(102.2577, 97.6354)	(101.4506, 100.8885)	(95.1119, 102.1749)	(100.4225, 97.3680)
(103.2209, 100.6225)	(95.2964, 102.4307)	(100.1523, 102.9490)	(104.1576, 101.6352)
(101.0120, 104.4635)	(103.4158, 98.2910)	(105.1581, 103.0230)	(100.7040, 94.9867)
(101.9143, 97.3885)	(101.8231, 97.7814)	(102.7163, 95.4334)	(96.9135, 101.7750)
(95.5534, 102.4620)	(95.4512, 99.7769)	(95.7699, 98.7454)	(97.7667, 102.0209)
(99.8642, 103.6219)	(98.1493, 94.7886)	(101.2693, 96.5736)	(98.5932, 102.4351)
(99.8691, 105.1958)	(94.2787, 100.9571)	(100.6072, 98.0980)	(102.9696, 95.7064)
(95.2816, 103.3236)	(98.9207, 105.6910)	(98.1615, 97.3504)	(101.3419, 97.4189)
(96.9166, 98.1973)	(103.9091, 96.2181)	(99.9506, 94.7994)	(103.7129, 98.1209)
(101.1100, 105.6254)	(99.9034, 95.9735)	(99.8464, 103.9968)	(101.2705, 99.7475)
(99.6992, 94.5015)	(94.2172, 101.2350)	(102.3153, 94.7393)	(96.8860, 95.9145)
(97.4538, 94.5998)	(100.1180, 100.9099)	(101.1873, 100.2082)	(100.8755, 99.4588)

N.2 Objective Function Values for Test Problems 2A, 2B, and 2C

(103.6848, 103.9609)	(95.6040, 101.5735)	(97.7876, 100.0430)	(105.3983, 100.1675)
(96.5578, 97.3298)	(100.2110, 98.8719)	(98.8857, 96.9268)	(102.7555, 100.3548)
(102.2526, 102.4639)	(98.4904, 94.3098)	(95.7547, 102.5842)	(103.1232, 97.8702)
(99.5723, 103.2071)	(103.1300, 98.1427)	(100.7443, 103.4310)	(98.9429, 102.6029)

(102.3963, 100.4059)	(100.8767, 95.5375)	(100.8257, 105.8330)	(97.5041, 100.6324)
(99.5536, 97.5641)	(102.4810, 99.8521)	(102.2372, 105.0177)	(101.5048, 95.9317)
(101.7970, 101.6996)	(99.3234, 105.8587)	(103.3371, 103.8918)	(96.0391, 103.7757)
(101.0486, 100.0085)	(98.7336, 101.4354)	(98.4455, 105.1548)	(98.4389, 100.0983)
(103.2828, 104.5445)	(102.5242, 98.0001)	(102.3256, 100.0627)	(105.2779, 98.1816)
(95.6618, 97.0988)	(103.2267, 100.3741)	(100.5905, 103.5784)	(98.7128, 95.2642)
(98.9666, 99.3876)	(104.7891, 98.5051)	(103.1183, 95.8092)	(100.9300, 94.8031)
(104.0700, 100.8312)	(101.4978, 98.8410)	(95.9098, 99.1952)	(94.1848, 100.5359)
(101.6212, 97.4931)	(105.7609, 100.6060)	(95.9841, 103.6136)	(97.5040, 101.1247)
(104.4545, 101.9316)	(99.2888, 103.1690)	(102.8691, 102.1152)	(95.4441, 99.7011)
(98.1023, 98.0927)	(97.7997, 98.9963)	(97.0409, 96.2717)	(99.7799, 94.6877)
(99.7569, 94.1299)	(102.3656, 103.4900)	(96.7093, 96.8682)	(100.4201, 95.7602)
(101.8289, 101.2260)	(95.4270, 103.5057)	(98.9799, 102.3676)	(101.2380, 99.4185)
(98.8547, 105.2610)	(99.7376, 100.0911)	(102.2202, 102.8368)	(95.5491, 96.6968)
(100.1998, 97.0971)	(97.6938, 101.7028)	(99.8989, 97.3402)	(95.7147, 97.0357)
(99.1475, 95.4149)	(103.1432, 98.0707)	(101.6622, 98.5825)	(97.3134, 98.1090)
(101.7079, 100.0766)	(103.5927, 96.5823)	(102.8712, 97.3695)	(97.6245, 103.2716)
(100.6163, 101.0873)	(103.9547, 96.3935)	(99.6997, 98.1131)	(105.0540, 102.8406)
(102.7289, 103.1864)	(95.4602, 102.3155)	(102.4739, 96.2580)	(96.7749, 100.8922)
(102.9739, 103.9954)	(104.2743, 102.9276)	(94.4548, 98.9261)	(103.6188, 101.8041)
(102.4299, 96.9297)	(94.8932, 100.1113)	(99.1705, 96.0906)	(103.5947, 99.6592)

O THE SECOND TEST PROBLEM SET WITH RESULTS

The second set of test problems is shown in Figures 17–19. This set of test problems has a low percent of dual variable values associated with MCE constraints. Note that in Figures 17–19, the asymptotically optimal allocations are proportional to the size of the circle. While there is no obvious visible difference in the optimal allocations with different correlations, the allocations do differ slightly.

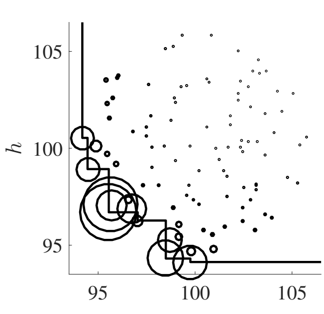

Fig. 17. Test 2A: $r = 100$, $|\mathcal{P}| = 6$, $\rho_k = -0.8$ for all $k \leq r$, % dual to MCE = 26.1, $z^* = 7.71 \times 10^{-4}$.

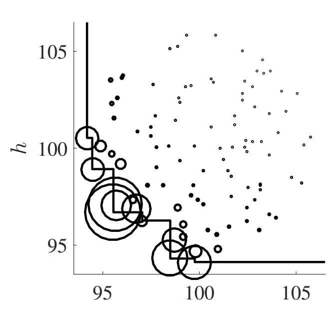

Fig. 18. Test 2B: $r = 100$, $|\mathcal{P}| = 6$, $\rho_k = 0$ for all $k \leq r$, % dual to MCE = 25.4, $z^* = 7.55 \times 10^{-4}$.

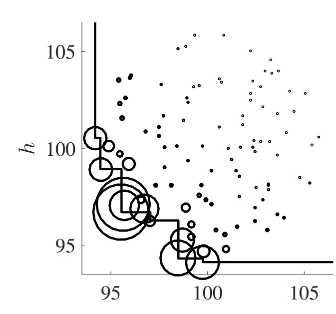

Fig. 19. Test 2C: $r = 100$, $|\mathcal{P}| = 6$, $\rho_k = 0.8$ for all $k \leq r$, % dual to MCE = 25.0, $z^* = 7.47 \times 10^{-4}$.

For each algorithm BVN True, SCORE, MOCBA, M-MOBA, and equal allocation, we run 10,000 independent sample paths on each of the test problems 2A, 2B, and 2C. For each algorithm, we calculate the average number of misclassifications, false exclusions, and false inclusions across the sample paths, as a function of sample size. Note that for a particular sample path, the sequence containing the number of misclassifications as a function of the sample size n is autocorrelated. All algorithm parameters used in the second test problem set are identical to those reported in the main body of the paper. The resulting performance of the algorithms is reported in Figures 20–22.

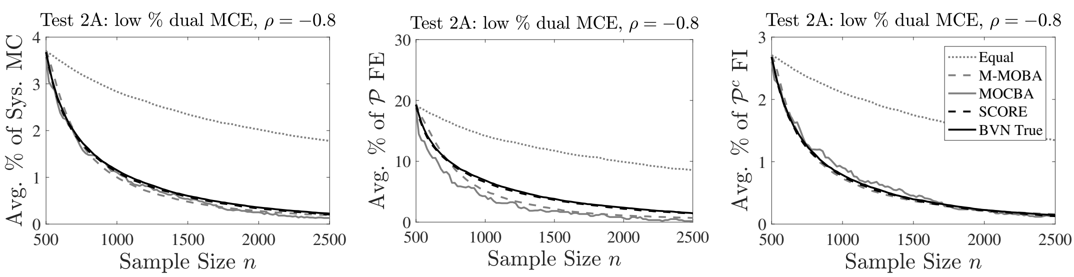

Fig. 20. Test 2A: For 10,000 sample paths per algorithm, the graphs show the average % of systems misclassified (MC), % of Paretos falsely excluded (FE), and % of non-Paretos falsely included (FI), respectively.

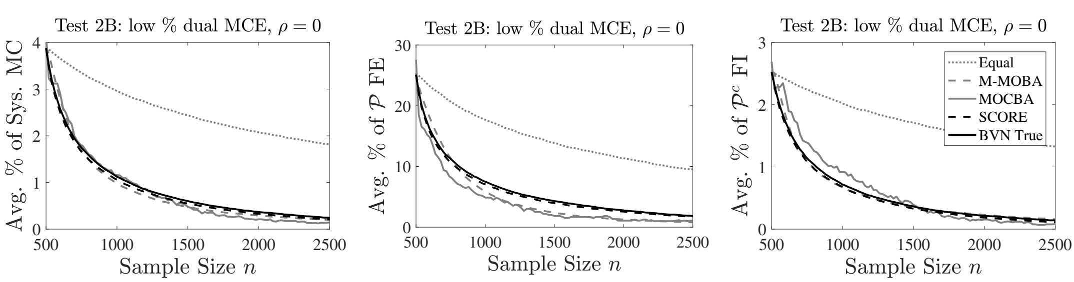

Fig. 21. Test 2B: For 10,000 sample paths per algorithm, the graphs show the average % of systems misclassified (MC), % of Paretos falsely excluded (FE), and % of non-Paretos falsely included (FI), respectively.

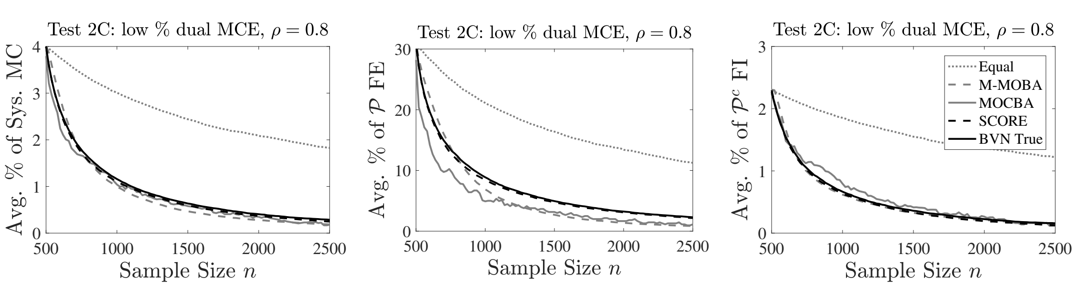

Fig. 22. Test 2C: For 10,000 sample paths per algorithm, the graphs show the average % of systems misclassified (MC), % of Paretos falsely excluded (FE), and % of non-Paretos falsely included (FI), respectively.

ACKNOWLEDGMENTS

The authors would like to thank Juergen Branke and Wen Zhang for providing M-MOBA code, Loo Hay Lee for providing MOCBA code, and Raghu Pasupathy for many interesting discussions relevant to this work. The second author was supported by the National Science Foundation under grant CMMI-1554144.

REFERENCES

S. Andradóttir and S.-H. Kim. 2010. Fully sequential procedures for comparing constrained systems via simulation. *Naval Research Logistics* 57, 5 (2010), 403–421. https://doi.org/10.1002/nav.20408

H. Bonnel and C. Collonge. 2014. Stochastic optimization over a Pareto set associated with a stochastic multi-objective optimization problem. *Journal of Optimization Theory and Applications* 162 (2014), 405–427. https://doi.org/10.1007/s10957-013-0367-8

S. Boyd and L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, New York.

J. Branke and W. Zhang. 2015. A new myopic sequential sampling algorithm for multi-objective problems. In *Proceedings of the 2015 Winter Simulation Conference*, L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 3589–3598.

J. Branke, W. Zhang, and Y. Tao. 2016. Multiobjective ranking and selection based on hypervolume. In *Proceedings of the 2016 Winter Simulation Conference*, T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 859–870.

J. C. Butler, D. J. Morrice, and P. Mullarkey. 2001. A multiple attribute utility theory approach to ranking and selection. *Management Science* 47, 6 (2001), 800–816.

C.-H. Chen, J. Lin, E. Yücesan, and S. E. Chick. 2000. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems* 10, 3 (2000), 251–270. https://doi.org/10.1023/A:1008349927281

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

A. Dembo and O. Zeitouni. 1998. *Large Deviations Techniques and Applications* (2nd ed.). Springer, New York.

M. Ehrgott. 2012. Vilfredo Pareto and multi-objective optimization. In *Optimization stories: 21st International Symposium on Mathematical Programming, Berlin, August 19 – 24, 2012*, M. Grötschel (Ed.). Documenta mathematica, Bielefeld, Germany, 447–453.

G. Feldman. 2017. *Sampling laws for multi-objective simulation optimization on finite sets*. Ph.D. Dissertation. Purdue University.

G. Feldman, S. R. Hunter, and R. Pasupathy. 2015. Multi-objective simulation optimization on finite sets: optimal allocation via scalarization. In *Proceedings of the 2015 Winter Simulation Conference*, L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 3610–3621. https://doi.org/10.1109/WSC.2015.7408520

J. Fliege and H. Xu. 2011. Stochastic multiobjective optimization: sample average approximation and applications. *Journal of Optimization Theory and Applications* 151 (2011), 135–162. https://doi.org/10.1007/s10957-011-9859-6

M. Fu (Ed.). 2015. *Handbook of Simulation Optimization*. International Series in Operations Research & Management Science, Vol. 216. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4939-1384-8

A. Ganesh, N. O'Connell, and D. Wischik. 2004. *Big Queues*. Springer, New York.

P. Glynn and S. Juneja. 2015. Ordinal optimization – empirical large deviations rate estimators, and stochastic multi-armed bandits. (2015). http://arxiv.org/abs/1507.04564

P. W. Glynn and S. Juneja. 2004. A large deviations perspective on ordinal optimization. In *Proceedings of the 2004 Winter Simulation Conference*, R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 577–585. https://doi.org/10.1109/WSC.2004.1371364

P. W. Glynn and S. Juneja. 2011. Ordinal optimization: a nonparametric framework. In *Proceedings of the 2011 Winter Simulation Conference*, S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 4057 – 4064. https://doi.org/10.1109/WSC.2011.6148095

S. G. Henderson and R. Pasupathy. 2017. Simulation Optimization Library. (2017). http://www.simopt.org

T. Homem-de-Mello and G. Bayraksan. 2015. Stochastic constraints and variance reduction techniques. In *Handbook of Simulation Optimization*, M. Fu (Ed.). International Series in Operations Research & Management Science, Vol. 216. Springer-Verlag New York, New York, 245–276. https://doi.org/10.1007/978-1-4939-1384-8_9

R. A. Horn. 1990. The Hadamard product. In *Matrix Theory and Applications*, C. R. Johnson (Ed.). Proceedings of Symposia in Applied Mathematics, Vol. 40. American Mathematical Society, Providence, Rhode Island. https://doi.org/10.1090/psapm/040

H. Huang and Z. B. Zabinsky. 2014. Multiple objective probabilistic branch and bound for Pareto optimal approximation. In *Proceedings of the 2014 Winter Simulation Conference*, A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 3916–3927. http://informs-sim.org/wsc14papers/includes/files/348.pdf

S. R. Hunter. 2011. *Sampling laws for stochastically constrained simulation optimization on finite sets*. Ph.D. Dissertation. Virginia Polytechnic Institute and State University.

S. R. Hunter, E. A. Applegate, V. Arora, B. Chong, K. Cooper, O. Rincón-Guevara, and C. Vivas-Valencia. 2017. An introduction to multi-objective simulation optimization. *Optimization Online* (2017), 33. http://www.optimization-online.org/DB_HTML/2017/03/5903.html

S. R. Hunter and G. Feldman. 2015. Optimal sampling laws for bi-objective simulation optimization on finite sets. In *Proceedings of the 2015 Winter Simulation Conference*, L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 3749–3757. https://doi.org/10.1109/WSC.2015.7408532

S. R. Hunter and B. McClosky. 2016. Maximizing quantitative traits in the mating design problem via simulation-based Pareto estimation. *IIE Transactions* 48, 6 (2016), 565–578. https://doi.org/10.1080/0740817X.2015.1096430

S. R. Hunter and B. L. Nelson. 2017. Parallel ranking and selection. In *Advances in Modeling and Simulation: Seminal Research from 50 Years of Winter Simulation Conferences*, A. Tolk, J. Fowler, G. Shao, and E. Yücesan (Eds.). Springer International, Switzerland, Chapter 12, 249–275. https://doi.org/10.1007/978-3-319-64182-9

S. R. Hunter and R. Pasupathy. 2013. Optimal sampling laws for stochastically constrained simulation optimization on finite sets. *INFORMS Journal on Computing* 25, 3 (2013), 527–542. https://doi.org/10.1287/ijoc.1120.0519

S. Kim and J. Ryu. 2011. The sample average approximation method for multi-objective stochastic optimization. In *Proceedings of the 2011 Winter Simulation Conference*, S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 4026–4037. https://doi.org/10.1109/WSC.2011.6148092

S.-H. Kim and B. L. Nelson. 2006. Selecting the best system. In *Simulation*, S. G. Henderson and B. L. Nelson (Eds.). Elsevier, Amsterdam, The Netherlands, 501–534.

L. H. Lee, E. P. Chew, S. Teng, and D. Goldsman. 2010. Finding the non-dominated Pareto set for multi-objective simulation models. *IIE Transactions* 42 (2010), 656–674. https://doi.org/10.1080/07408171003705367

L. H. Lee, N. A. Pujowidianto, L.-W. Li, C.-H. Chen, and C. M. Yap. 2012. Approximate simulation budget allocation for selecting the best design in the presence of stochastic constraints. *IEEE Trans. Automat. Control* 57, 11 (2012), 2940–2945. https://doi.org/10.1109/TAC.2012.2195931

H. Li, L. H. Lee, E. P. Chew, and P. Lendermann. 2015. MO-COMPASS: A fast convergent search algorithm for multi-objective discrete optimization via simulation. *IIE Transactions* 47, 11 (2015), 1153–1169. https://doi.org/10.1080/0740817X.2015.1005778

J. Li. 2012. *Optimal computing budget allocation for multi-objective simulation optimization*. Ph.D. Dissertation. National University of Singapore.

Y. Luo and E. Lim. 2013. Simulation-based optimization over discrete sets with noisy constraints. *IIE Transactions* 45, 7 (2013), 699–715. https://doi.org/10.1080/0740817X.2012.733580

G. Marmidis, S. Lazarou, and E. Pyrgioti. 2008. Optimal placement of wind turbines in a wind park using Monte Carlo simulation. *Renewable Energy* 33 (2008), 1455–1460.

K. Miettinen. 1999. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston.

K. Nagaraj and R. Pasupathy. 2016. Stochastically constrained simulation optimization on integer-ordered spaces: The cgR-SPLINE algorithm. (2016). http://www.optimization-online.org/DB_HTML/2015/10/5139.html.

C. Osorio and M. Bierlaire. 2013. A simulation-based optimization framework for urban transportation problems. *Operations Research* 61, 6 (2013), 1333–1345. https://doi.org/10.1287/opre.2013.1226

C. Park and S.-H. Kim. 2015. Penalty function with memory for discrete optimization via simulation with stochastic constraints. *Operations Research* 63 (2015), 1195–1212. Issue 5. https://doi.org/10.1287/opre.2015.1417

R. Pasupathy and S. Ghosh. 2013. Simulation optimization: a concise overview and implementation guide. In *TutORials in Operations Research*, H. Topaloglu (Ed.). INFORMS, Catonsville, MD, Chapter 7, 122–150. https://doi.org/10.1287/educ.2013.0118

R. Pasupathy and S. G. Henderson. 2006. A testbed of simulation-optimization problems. In *Proceedings of the 2006 Winter Simulation Conference*, L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 255–263. https://doi.org/10.1109/WSC.2006.323081

R. Pasupathy, S. R. Hunter, N. A. Pujowidianto, L. H. Lee, and C.-H. Chen. 2015. Stochastically constrained ranking and selection via SCORE. *ACM Transactions on Modeling and Computer Simulation* 25, 1, Article 1 (2015), 26 pages. https://doi.org/10.1145/2630066

M. J. Powers, S. M. Sanchez, and T. W. Lucas. 2012. The exponential expansion of simulation in research. In *Proceedings of the 2012 Winter Simulation Conference*, C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher (Eds.).

IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 1552–1563. https://doi.org/10.1109/WSC.2012.6465125

A. Ruszczyński and A. Shapiro (Eds.). 2003. *Stochastic Programming*. Elsevier, New York.

G. Strang. 1988. *Linear Algebra and its Applications* (3 ed.). Harcourt Brace Jovanovich, New York.

K. Subramanyan, U. Diwekar, and S. E. Zitney. 2011. Stochastic modeling and multi-objective optimization for the APECS system. *Computers and Chemical Engineering* 35 (2011), 2667–2679.

R. Szechtman and E. Yücesan. 2008. A new perspective on feasibility determination. In *Proceedings of the 2008 Winter Simulation Conference*, S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler (Eds.). IEEE, Institute of Electrical and Electronics Engineers, Inc., Piscataway, NJ, 273–280. https://doi.org/10.1109/WSC.2008.4736078

J. Xu, B. L. Nelson, and L. J. Hong. 2010. Industrial Strength COMPASS: A comprehensive algorithm and software for optimization via simulation. *ACM Transactions on Modeling and Computer Simulation* 20 (2010), 1–29.

S. Q. Yahyaa, M. M. Drugan, and B. Manderick. 2014a. Knowledge gradient for multi-objective multi-armed bandit algorithms. In *2014 International Conference on Agents and Artificial Intelligence (ICAART)*. SCITEPRESS, Setubal, Portugal, 74–83.

S. Q. Yahyaa, M. M. Drugan, and B. Manderick. 2014b. Multivariate normal distribution based multi-armed bandit pareto algorithm. In *The 7th European Conference on Machine Learning and Principles and Practice of Knowledge and Discovery in Databases (EMCL/PKDD), PhD Track*. Springer, New York, 139–148. https://ai.vub.ac.be/sites/default/files/EpsilonPareto_Normal_ECML_2014.pdf

S. Q. Yahyaa, M. M. Drugan, and B. Manderick. 2014c. The scalarized multi-objective multi-armed bandit problem: An empirical study of its exploration vs. exploitation tradeoff. In *2014 International Joint Conference on Neural Networks (IJCNN)*. IEEE, IEEE Press, Piscataway, NJ, 2290–2297. https://doi.org/10.1109/IJCNN.2014.6889390

Received May 2016; revised December 2016, September 2017; accepted October 2017