Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

PROPERTIES OF SEVERAL PERFORMANCE INDICATORS FOR GLOBAL MULTI-OBJECTIVE SIMULATION OPTIMIZATION

Susan R. Hunter
Burla E. Ondes

School of Industrial Engineering
Purdue University
315 N Grant Street
West Lafayette, IN 47906, USA

ABSTRACT

We discuss the challenges in constructing and analyzing performance indicators for multi-objective simulation optimization (MOSO), and we examine properties of several performance indicators for assessing algorithms designed to solve MOSO problems to global optimality. Our main contribution lies in the definition and analysis of a modified coverage error; the modification to the coverage error enables us to obtain an upper bound that is the sum of deterministic and stochastic error terms. Then, we analyze each error term separately to obtain an overall upper bound on the modified coverage error that is a function of the dispersion of the visited points in the compact feasible set and the sampling error of the objective function values at the visited points. The upper bound provides a foundation for future mathematical analyses that characterize the rate of decay of the modified coverage error.

1 INTRODUCTION

Consider the context of evaluating one or more algorithms designed to solve the multi-objective simulation optimization (MOSO) problem to global optimality by characterizing the entire efficient and Pareto sets. That is, we wish to solve the MOSO problem

$$\text{minimize} \quad \{f(x) = (f_1(x), \ldots, f_d(x)) \coloneqq (\mathrm{E}[F_1(x,\xi)], \ldots, \mathrm{E}[F_d(x,\xi)])\} \quad \text{s.t.} \quad x \in \mathcal{X}, \tag{1}$$

where $f\colon \mathcal{D} \to \mathbb{R}^d, \mathcal{D} \subseteq \mathbb{R}^q, q \geq 2$ is a vector-valued function composed of $d \geq 2$ unknown, real-valued, continuous and conflicting objective functions $f_k\colon \mathcal{D} \to \mathbb{R}, k = 1, \ldots, d$. Each objective is formulated as the expected value

$$f_k(x) = \mathrm{E}[F_k(x,\xi)] = \int_{\Xi} F_k(x,z) P(dz)$$

where $F_k\colon \mathcal{D} \times \Xi \to \mathbb{R}$, P is the probability measure induced by the random object $\xi\colon \Omega \to \Xi$, and $F\colon \mathcal{D} \times \Xi \to \mathbb{R}^d$ is the vector-valued function composed of $F_1, \ldots, F_d$. We wish to solve (1) by characterizing its global solution, called the efficient set $\mathcal{E}$, and the efficient set's image, the global Pareto set $\mathcal{P}$, where

$$\mathcal{E} \coloneqq \{x^* \in \mathcal{X}\colon \nexists x \in \mathcal{X} \text{ such that } f(x) \leq f(x^*)\}, \qquad \mathcal{P} \coloneqq \{f(x^*)\colon x^* \in \mathcal{E}\}$$

(see §2.1 for notation). We remark here that formulating the objectives using an expected value is extremely general, in the sense that the objectives can be the expected value of some loss function of the performance measure of interest. A famous example is the mean-variance portfolio optimization problem of Markowitz (1952), in which the performance measure of interest is the (random) return of the portfolio over some time horizon. The objectives are the expected return and variance of the return; thus, each objective is formulated

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

as an expected value. MOSO problems of the form in (1) arise in many other real-world application areas; see Hunter et al. (2019), Fu (2015).

Central to the task of evaluating MOSO algorithm performance is the fact that the objective function values of (1) at each feasible point can only be observed with stochastic error, e.g. as the output of a Monte Carlo simulation oracle. As in Ragavan et al. (2022), for each objective $k = 1, \ldots, d$, let ε_j denote the vector-valued error random field

$$\varepsilon_j(\cdot) := F(\cdot, \xi_j) - f(\cdot) \tag{2}$$

for independent and identically distributed (iid) random objects $\xi_j, j = 1, \ldots, n$. The vector-valued objective estimator is

$$\bar{F}(\cdot, n) := f(\cdot) + n^{-1} \textstyle\sum_{j=1}^{n} \varepsilon_j(\cdot),$$

where $\bar{F}(\cdot, n) = (\bar{F}_1(\cdot, n), \ldots, \bar{F}_d(\cdot, n))$. We emphasize that vector-valued objective estimator is observed *point-wise*; that is, we do not observe the entire field $\bar{F}(\cdot, n)$ at once. Rather, at each visited point $x \in \mathcal{X}$, we must expend new effort to obtain n_x observations from the stochastic oracle at x. Letting $\mathcal{X}_m \subset \mathcal{X}$ be a finite set of m feasible points visited by a MOSO algorithm, suppose that after expending a total simulation effort $N = \sum_{x \in \mathcal{X}_m} n_x$, the algorithm returns an estimated efficient set as

$$\hat{\mathcal{E}}_m(N) := \{X^* \in \mathcal{X}_m \colon \nexists X \in \mathcal{X}_m \text{ such that } \bar{F}(X, n_X) \leq \bar{F}(X^*, n_{X^*})\}.$$

The estimated image of $\hat{\mathcal{E}}_m(N)$ is the returned estimated Pareto set

$$\hat{\mathcal{P}}_m(N) := \{\bar{F}(X^*, n_{X^*}) \in \mathbb{R}^d \colon X^* \in \hat{\mathcal{E}}_m(N)\}.$$

Thus, the MOSO algorithm returns two random sets. This fact complicates assessment relative to the single-objective simulation optimization context, in which the algorithm usually returns only a random vector X^* and its estimated image $\bar{F}_1(X^*, n_{X^*})$.

There are at least three contexts in which we might like to assess the performance of one or more MOSO algorithms designed to solve the problem defined above: (a) on a testbed problem for which the true efficient and Pareto sets are *known*, in which case all algorithms can be evaluated separately with respect to the known Pareto set, (b) on a testbed problem for which the true efficient and Pareto sets are *unknown*, in which case we might evaluate algorithms' returned solutions relative to each other, and (c) in a theoretical context, where we assume knowledge of the true efficient and Pareto sets for the purposes of mathematical analysis leading to efficient algorithmic design. In this paper, we consider contexts (a) and (c) exclusively; that is, we assume the true efficient and Pareto sets are known in either a testbed or theoretical context. We remark that context (b), in which we compare algorithms to each other on problems with unknown efficient and Pareto sets, provides important motivation for the further development of confidence regions on efficient and Pareto sets — an area that has seen little work to date; see Hunter and Pasupathy (2022), Vogel (2017). (For recent advances in simulation optimization testbed problems, also see recent work on the SimOpt library by Eckman et al. (2023a), Eckman et al. (2023b).)

1.1 Challenges in Constructing and Analyzing Performance Indicators for MOSO

In deterministic multi-objective optimization, *performance indicators* (Audet et al. 2021) measure the performance of an algorithm by mapping the returned Pareto set approximation to a real number. Different performance indicators capture different desirable characteristics for an algorithm's returned solution, such as convergence (how "close" it is to the true Pareto set), spread or coverage, and uniformity (Li and Yao 2019). Each performance indicator has benefits and drawbacks; therefore, depending on the context, one or more performance indicators may be appropriate. The literature on performance indicators is extensive: Reviews including Li and Yao (2019), Audet et al. (2021), Faulkenberg and Wiecek (2010) collectively identify over 100 different performance indicators for use when solving deterministic multi-objective optimization problems.

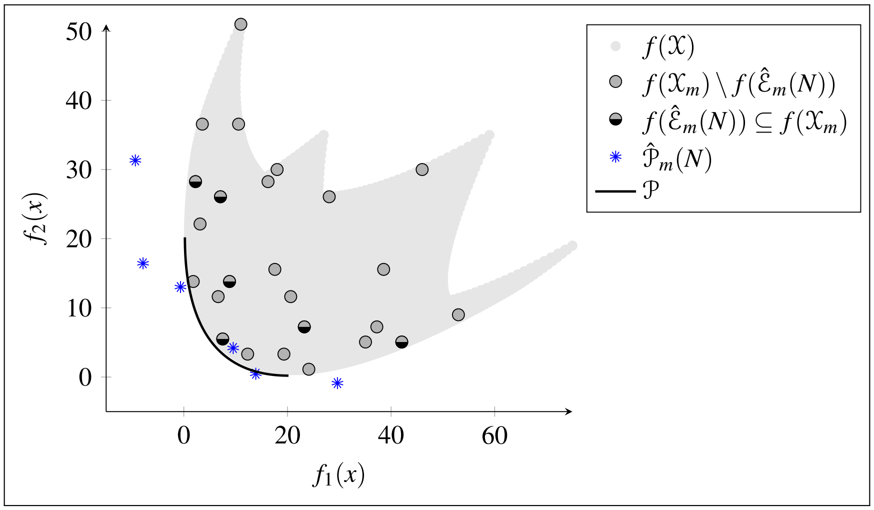

Figure 1: The figure shows the image space $f(\mathcal{X})$ of a bi-objective convex quadratic problem (light gray region) with true Pareto set $\mathcal{P}$ (black curve) and rectangular feasible set $\mathcal{X} \subset \mathbb{R}^2$. The discretization $\mathcal{X}_m$ consists of $m = 25$ points selected on a grid, and $f(\mathcal{X}_m)$ is represented by circular points (dark gray and half gray, half black points). For one instance of additive, mean-zero normally distributed noise, the blue stars indicate the estimated Pareto set $\hat{\mathcal{P}}_m(N)$ and the half gray, half black points represent $f(\hat{\mathcal{E}}_m(N))$. In this case, $f(\hat{\mathcal{E}}_m(N))$ contains points that are dominated by other points in $f(\hat{\mathcal{E}}_m(N))$.

Far less work has been done on creating and determining the properties of performance indicators specifically for MOSO (Hunter et al. 2019; Branke 2023), which is a newer field with unique challenges. To explore these challenges in detail, recall that we are assessing MOSO algorithms in a context where we can evaluate — either in truth or in theory — the objective functions both with and without error, as in Figure 1. Consistent with the philosophies regarding assessment for simulation optimization algorithms described in Pasupathy and Henderson (2006), we use the random set

$$f(\hat{\mathcal{E}}_m(N)) := \{f(X^*) \in \mathbb{R}^d : X^* \in \hat{\mathcal{E}}_m(N)\}$$

as input to the performance indicator that measures how well the algorithm approximates the true Pareto set, $\mathcal{P}$. The set $f(\hat{\mathcal{E}}_m(N))$ contains the true objective function values of the estimated efficient set and is usually unknown, except in this testbed problem or theoretical context. Specifically, $f(\hat{\mathcal{E}}_m(N))$ is always a random set, in the sense that if we ran the MOSO algorithm again with a different starting seed $\omega \in \Omega$ for the simulation oracle (Law 2015), we would obtain a different value for $f(\hat{\mathcal{E}}_m(N))$ due to the random object ξ, regardless of whether the algorithm is a *randomized algorithm* (Karp 1991). (For simplicity, we assume the estimator $\hat{\mathcal{E}}_m(N)$ is constructed as a finite set, without interpolation. We consider assessment of interpolated or continuous estimators as future work.) With this example in mind, we consider two specific challenges in constructing and determining the properties of performance indicators for a MOSO context:

C.1 A MOSO algorithm may return two points in $\hat{\mathcal{E}}_m(N)$ whose estimated objective function values belong to the estimated Pareto set $\hat{\mathcal{P}}_m(n)$, but where one point truly dominates the other; that is, $f(\hat{\mathcal{E}}_m(N))$ may contain points that are dominated by other points in $f(\hat{\mathcal{E}}_m(N))$.

C.2 The set $f(\hat{\mathcal{E}}_m(N))$ is a random set, which implies that the performance indicator operating on this set is also a random object.

As stated in C.1 and shown by the example in Figure 1, a MOSO algorithm can return two points estimated as Pareto, but in which one point dominates the other. Since most deterministic multi-objective optimization algorithms do not return dominated points (and if they do, the dominated points are usually straightforward to identify), performance indicators for the deterministic multi-objective optimization setting typically do not account for this possibility. For example, the *coverage error* under the L_p-norm (Sayin 2000) for $f(\hat{\mathcal{E}}_m(N))$ is the random variable

$$\text{dist}_p(\mathcal{P}, f(\hat{\mathcal{E}}_m(N))) = \sup_{y^* \in \mathcal{P}} \inf_{f(X^*) \in f(\hat{\mathcal{E}}_m(N))} \|y^* - f(X^*)\|_p, \tag{3}$$

where $p = 1, 2, \ldots, \infty$. The value of the coverage error is the same for an algorithm that returns only points that do not dominate each other and an algorithm that returns the same set plus an arbitrary number of "extra" dominated points. While coverage error is certainly informative in a MOSO context, it may need to be considered alongside other performance indicators that account for the possibility of returning dominated points. However, few such performance indicators exist.

Challenge C.2 states that $f(\hat{\mathcal{E}}_m(N))$ is a random set and the performance indicator is a random object. Thus, regardless of which performance indicator we choose, we need to consider properties of its distribution. When conducting numerical experiments on testbed problems, we should report several quantiles of the random performance indicator across many independent algorithm runs. When performing mathematical analysis on a performance indicator that assesses convergence properties, we may wish to analyze a quantity such as the probability that performance indicator value exceeds a certain threshold. For example, given a value $a \in \mathbb{R}$, we may wish to determine the probability that the coverage error exceeds a, $\mathbb{P}\{\text{dist}_p(\mathcal{P}, f(\hat{\mathcal{E}}_m(N))) > a\}$ as a function of m and N. The difficulty of such analyses depends on the properties of the chosen performance indicator. To the best of our knowledge, relevant properties of performance indicators for MOSO remain mostly unexplored.

1.2 Summary of Contributions

We ultimately consider the properties of four performance indicators formulated for a MOSO context: the Hausdorff distance (§3), the coverage error and modified coverage error (§4), and the hypervolume difference (§5). Each performance indicator has benefits and drawbacks which we briefly discuss here, along with our contributions.

Of the three performance indicators, only the Hausdorff distance addresses challenge C.1 by accounting for the fact that $f(\hat{\mathcal{E}}_m(N))$ may contain points that dominate other points in $f(\hat{\mathcal{E}}_m(N))$. The Hausdorff distance is also a metric, which enables the use of the triangle inequality to upper bound the Hausdorff distance between $\mathcal{P}$ and $f(\hat{\mathcal{E}}_m(N))$ by a deterministic error term plus a stochastic error term, each of which can be analyzed separately. Such an upper bound is a common tool for error analysis in stochastic optimization algorithms (see, e.g., Yakowitz et al. (2000), Shapiro et al. (2009)). Here, the deterministic error results from only considering a finite set of m points, while the stochastic error results from the fact that the objective function values can only be estimated. However as discussed in §3, under the Hausdorff distance, each of these error terms is itself difficult to analyze.

The coverage error, which is one of the two one-way distances whose maximum equals the Hausdorff distance, is simpler to analyze. However, unfortunately, it is not a metric. Therefore, we cannot naïvely invoke the triangle inequality to obtain an upper bound as the sum of the two types of error. Our main contribution regards a modification to the coverage error that enables us to obtain such an upper bound. Then, we analyze each error term separately to obtain an overall upper bound on the *modified coverage error* that is the sum of an upper bound on the deterministic error term plus an upper bound on the stochastic error term. The upper bound on the modified coverage error is stated in Theorem 1 and constitutes our main contribution; it enables future mathematical analyses that address C.2, such as determining an upper bound on the probability the modified coverage error exceeds some value a. Loosely speaking, the upper bound is a function of the *dispersion* of the discretized feasible set under the L_r-norm in the decision space,

$$t := \text{dist}_r(\mathcal{X}, \mathcal{X}_m) = \sup_{x \in \mathcal{X}} \inf_{\tilde{x} \in \mathcal{X}_m} \|x - \tilde{x}\|_r = \sup_{x \in \mathcal{X}} \min_{1 \le i \le m} \|x - \tilde{x}_i\|_r, \tag{4}$$

and the extent to which the true objective function values are well-approximated by the estimated objective function values at each sampled point. (We use the L_r-norm in the decision space and the L_p-norm in the objective space.) The dispersion is the smallest radius a such that for $\mathcal{X}_m = \{\tilde{x}_1, \ldots, \tilde{x}_m\}$, the union of the q-balls, $\cup_{i=1}^m B_r(\tilde{x}_i, a)$, covers the compact feasible set $\mathcal{X}$. For more details on dispersion and related results, see Yakowitz et al. (2000), Niederreiter (1992).

Finally, we briefly consider the hypervolume difference in a MOSO context. While it does not address C.1, this performance metric has many desirable qualities for deterministic multi-objective optimization and has been the subject of considerable analysis in that context. For MOSO, we show that it is straightforward to split this performance indicator into the sum of a deterministic term and a stochastic term; we leave further analyses of each term to future work.

The remainder of the paper is organized as follows: §2 contains preliminaries including notation, terminology, and assumptions. We discuss the three performance indicators in §3–§5, with the main theorem appearing in §4. Concluding remarks appear in §6.

2 PRELIMINARIES

We discuss notation and terminology, including dominance concepts, in §2.1. Then, in §2.2, we discuss the required assumptions that may be invoked in the remainder of the paper.

2.1 Notation and Terminology

The q-by-q identity matrix is I_q; the origin is $0_q \in \mathbb{R}^q$ and $1_q \in \mathbb{R}^q$ is a q-dimensional vector of ones. For $x = (x_1, \ldots, x_q) \in \mathbb{R}^q$ and $1 \leq p < \infty$, the L_p-norm is $\|x\|_p := (|x_1|^p + |x_2|^p + \ldots + |x_d|^p)^{1/p}$, the L_2-norm is the Euclidean norm, and the L_∞-norm is $\|x\|_\infty := \max\{|x_1|, |x_2|, \ldots, |x_d|\}$. Note that for any $r > p \geq 1$,

$$\|x\|_r \leq \|x\|_p \leq q^{(1/p-1/r)}\|x\|_r. \tag{5}$$

For $x \in \mathbb{R}^q$ and $1 \leq p \leq \infty$, let $B_p(x,t) := \{x' \in \mathbb{R}^q \colon \|x' - x\|_p \leq t\}$. Henceforth, we drop the $p = 2$ subscript on all instances of the Euclidean norm.

Let $\mathcal{A} \subseteq \mathbb{R}^q$ be a set. Its t-expansion is $B_p(\mathcal{A},t) := \cup_{x \in \mathcal{A}} B_p(x,t)$. Its complement is $\mathcal{A}^c$, convex hull is conv($\mathcal{A}$), boundary is bd($\mathcal{A}$), interior is int($\mathcal{A}$), and diameter is $\text{diam}(\mathcal{A}) = \sup_{x,y \in \mathcal{A}} \|x - y\|$. $\mathcal{A}$ is bounded if $\exists c < \infty$ such that $\forall x, y \in \mathcal{A}$, $\|x - y\| \leq c$. The sum of two sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^q$ is the Minkowski sum, $\mathcal{A} + \mathcal{B} := \{a + b \colon a \in \mathcal{A}, b \in \mathcal{B}\}$. Let $\mathcal{A} \subset \mathbb{R}^q, \mathcal{B} \subset \mathbb{R}^q$ be nonempty sets. For $x \in \mathbb{R}^q$, $\text{dist}_p(x, \mathcal{B}) := \inf_{y \in \mathcal{B}} \|x - y\|_p$; $\text{dist}_p(\mathcal{A}, \mathcal{B}) := \sup_{x \in \mathcal{A}} \text{dist}_p(x, \mathcal{B})$. The Hausdorff distance under the L_p-norm is $\mathbb{H}_p(\mathcal{A}, \mathcal{B}) := \max\{\text{dist}_p(\mathcal{A}, \mathcal{B}), \text{dist}_p(\mathcal{B}, \mathcal{A})\}$.

For sequences of reals $\{a_n\}$ and $\{b_n\}$, $a_n = O(1)$ if there exists $c \in (0, \infty)$ with $|a_n| < c$ for large enough n. For positive-valued sequences of reals $\{a_n\}$ and $\{b_n\}$, we say $a_n = O(b_n)$ if $a_n/b_n = O(1)$ as $n \to \infty$.

Let $f \colon \mathcal{D} \to \mathbb{R}^d, \mathcal{D} \subseteq \mathbb{R}^q$ be a vector-valued function. For a set $\mathcal{S} \subseteq \mathcal{D}$, $f(\mathcal{S}) := \{f(x) \colon x \in \mathcal{S}\}$. For $y, \tilde{y} \in \mathbb{R}^d$, $y \leqq \tilde{y}$ (y weakly dominates $\tilde{y}$) if $y_k \leq \tilde{y}_k$ for all $k = 1, 2$; $y \leq \tilde{y}$ (y dominates $\tilde{y}$) if $y \leqq \tilde{y}$ and $y \neq \tilde{y}$; $y < \tilde{y}$ (y strictly dominates $\tilde{y}$) if $y_k < \tilde{y}_k$ for all $k = 1, 2$. The set $\mathbb{R}^d_{\geqq}$ denotes $\{y \in \mathbb{R}^d \colon 0 \leqq y\}$. See Hunter et al. (2019), p. 7 for commentary on our conventions regarding the terms *efficient* and *Pareto*. We use *Pareto set* instead of *Pareto front* for mathematical clarity.

2.2 Assumptions

In the results that follow, we invoke one or more of the assumptions presented in this section. Given that our goal is to solve the MOSO problem to global optimality, the following Assumption 1 ensures that relevant quantities, such as the feasible set, are properly bounded.

Assumption 1 The feasible set $\mathcal{X} \subseteq \mathcal{D} \subseteq \mathbb{R}^q$, $q \geq 2$ is compact, convex, and $\mathcal{E} \subset \text{int}(\mathcal{X})$.

On this bounded set, we further require Lipschitz continuity of the vector-valued random function $F(\cdot,\xi)$ for a.e. $\xi \in \Xi$. For generality, we allow different norms in the decision space and the objective space.

Assumption 2 The vector-valued function $F(\cdot,\xi) = (F_1(\cdot,\xi),\ldots,F_d(\cdot,\xi))$ is $L(\xi)$-Lipschitz continuous for a.e. $\xi \in \Xi$; that is, given $p,r \in \{1,\ldots,\infty\}$, for all $x_1,x_2 \in \mathcal{D}$ and a.e. $\xi \in \Xi$,

$$\|F(x_1,\xi) - F(x_2,\xi)\|_p \le L(\xi)\|x_1 - x_2\|_r$$

where $0 < \mathrm{E}[L(\xi)^2] < \infty$ and $\ell := \mathrm{E}[L(\xi)] > 0$.

Assumption 2 implies that f and $\bar{F}(\cdot,n)$ are also Lipschitz continuous, as stated in the following Lemma 1.

Lemma 1 Under Assumption 2, letting $\bar{L} := n^{-1}\sum_{i=1}^{n} L(\xi_i)$,

1. f is ℓ-Lipschitz continuous; that is, for all $x_1,x_2 \in \mathcal{D}$, $\|f(x_1) - f(x_2)\|_p \le \ell\|x_1 - x_2\|_r$, and
2. $\bar{F}(\cdot,n)$ is $\bar{L}$-Lipschitz continuous a.s.; that is, for all $x_1,x_2 \in \mathcal{D}$, $\|\bar{F}(x_1,n) - \bar{F}(x_2,n)\|_p \le \bar{L}\|x_1 - x_2\|_r$.

Proof. Hunter and Pasupathy (2022) contains a similar result and proof, which we extend here in a straightforward fashion. Let $x_1,x_2 \in \mathcal{D}$. For Part 1, use Jensen's inequality to see that

$$\|f(x_1) - f(x_2)\|_p = \|\mathrm{E}[F(x_1,\xi) - F(x_2,\xi)]\|_p \le \mathrm{E}[\|F(x_1,\xi) - F(x_2,\xi)\|_p] \le \mathrm{E}[L(\xi)]\|x_1 - x_2\|_r.$$

For Part 2, we have

$$\|\bar{F}(x_1,n) - \bar{F}(x_2,n)\|_p \le n^{-1}\textstyle\sum_{i=1}^{n}\|F(x_1,\xi_i) - F(x_2,\xi_i)\|_p \le (n^{-1}\sum_{i=1}^{n} L(\xi_i))\|x_1 - x_2\|_r. \qquad \square$$

In addition to Assumption 2, for certain results, we also require the following Assumption 3. Assumption 3 implies that as the objective function values between two points become close on one objective, they must become close on all objectives, regardless of whether the relevant decision variable values are close in the decision space.

Assumption 3 There exists a positive constant $c \in (0,\infty)$ such that for all $\delta > 0$ and all $x_1,x_2 \in \mathcal{X}$, if $|f_k(x_1) - f_k(x_2)| < \delta$ for at least one objective $k \in \{1,\ldots,d\}$, then $\|f(x_1) - f(x_2)\|_p < c\delta$.

Assumption 3 prevents "flat spots" in the objective space. More specifically, in some of the results that follow, we consider the problem in (1) solved only on the discretized feasible set $\mathcal{X}_m$. This assumption prevents the existence of weakly Pareto points in the discretized problem, regardless of the choice of $\mathcal{X}_m$. Weakly Pareto points cannot be classified as Pareto or non-Pareto with a finite number of simulation replications n because they have objective function values equal to those of Pareto points on at least one objective. Thus, in MOSO, we have two choices for controlling the error with regards to weakly Pareto points: assume there are no weakly Pareto points, or define the error such that we do not penalize an algorithm for incorrectly classifying weakly Pareto points (also see the discussions in Hunter et al. (2019), Applegate et al. (2020), Cooper et al. (2020)). Here, as in Applegate et al. (2020), we take the approach of assuming the weakly Pareto points do not exist.

3 HAUSDORFF DISTANCE

We begin our discussion with the Hausdorff distance, which addresses the challenge in C.1 by penalizing algorithms that return extra dominated points. In our context, the relevant Hausdorff distance is the random variable

$$\mathbb{H}_p(\mathcal{P}, f(\hat{\mathcal{E}}_m(N))) = \max\{\mathrm{dist}_p(\mathcal{P}, f(\hat{\mathcal{E}}_m(N))), \mathrm{dist}_p(f(\hat{\mathcal{E}}_m(N)),\mathcal{P})\}. \quad (6)$$

It can be considered a worst-case performance indicator, in the sense that the distance returned is the maximum of the one-way distances between the Pareto set and the points in $f(\hat{\mathcal{E}}_m(N))$. The Hausdorff distance is a natural choice for analyzing the convergence and convergence rates of MOSO algorithms since (a) it equals zero if and only if $\mathcal{P}$ and $f(\hat{\mathcal{E}}_m(N))$ have the same closure, (b) it does not require a reference point or any parameters, and (c) it is a metric. The Hausdorff distance in (6) is useful for numerical analysis; for example, Cooper et al. (2020) use sample quantiles of the Hausdorff distance to compare MOSO algorithms throughout their numerical section.

For mathematical analysis, the fact that the Hausdorff distance is a metric implies that the triangle inequality can be used to obtain an upper bound as the sum of a deterministic error term and a stochastic error term. Henceforth, let

$$\mathcal{E}_m := \{\tilde{x}^* \in \mathcal{X}_m \colon \nexists \tilde{x} \in \mathcal{X}_m \text{ such that } f(\tilde{x}) \leq f(\tilde{x}^*)\}, \qquad \mathcal{P}_m := \{f(\tilde{x}^*)\colon x^* \in \mathcal{E}_m\}$$

denote the respective global efficient and Pareto sets that result from solving (1) on only the discretized feasible set $\mathcal{X}_m$. Then a.s.,

$$\begin{aligned} \mathbb{H}_p(\mathcal{P}, f(\hat{\mathcal{E}}_m(N))) &\leq \mathbb{H}_p(\mathcal{P},\mathcal{P}_m) + \mathbb{H}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))) \qquad (7) \\ &= \underbrace{\max\big\{\mathrm{dist}_p(\mathcal{P},\mathcal{P}_m), \mathrm{dist}_p(\mathcal{P}_m,\mathcal{P})\big\}}_{\text{deterministic error}} + \underbrace{\max\big\{\overbrace{\mathrm{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N)))}^{>0 \Rightarrow \text{ false exclusion}}, \overbrace{\mathrm{dist}_p(f(\hat{\mathcal{E}}_m(N)),\mathcal{P}_m)}^{>0 \Rightarrow \text{ false inclusion}}\big\}}_{\text{stochastic error}}, \end{aligned}$$

where the terminology of false exclusion and false inclusion comes from the multi-objective ranking and selection literature (Hunter and McClosky 2016; Hunter et al. 2019). However, (7) currently has limited usefulness, since it is rather difficult to analyze. To the best of our knowledge, there is no known upper bound on the right-hand side of (7).

Even for the deterministic error term $\mathbb{H}_p(\mathcal{P},\mathcal{P}_m)$, we do not know of an upper bound except in the case of a deterministic bi-objective convex quadratic program under the Euclidean norm. In this case, working mostly in the decision space, Ondes and Hunter (2023) show that $\mathbb{H}(\mathcal{E},\mathcal{E}_m) = O(\sqrt{t})$ and, subsequently, that $\mathbb{H}(\mathcal{P},\mathcal{P}_m) = O(\sqrt{t})$. When the maximum condition number across both objectives is 1 so that the level sets are hyper-spheres, they also show that $\mathbb{H}(\mathcal{E},\mathcal{E}_m) \leq \sqrt{t\theta + t^2}$ is a least upper bound, where $\theta := \mathrm{diam}(\mathcal{E})$ is the length of the efficient set. However, whether the corresponding $O(\sqrt{t})$ result in the objective space is tight remains an open question.

4 COVERAGE ERROR AND MODIFIED COVERAGE ERROR

Given the difficulty in analyzing the Hausdorff distance, we next consider the coverage error, which is one of the two distances whose maximum equals the Hausdorff distance in (6). Recall from (3) that the coverage error is the random variable

$$\mathrm{dist}_p(\mathcal{P}, f(\hat{\mathcal{E}}_m(N))) = \sup_{y^* \in \mathcal{P}} \inf_{f(X^*) \in f(\hat{\mathcal{E}}_m(N))} \|y^* - f(X^*)\|_p.$$

The coverage error measures the maximum distance from a point in $\mathcal{P}$ to its nearest representative in $f(\hat{\mathcal{E}}_m(N))$. Notice that the coverage error would equal zero if $\mathcal{P}$ were a subset of the closure of $f(\hat{\mathcal{E}}_m(N))$. Further, the one-way set distance used in the coverage error is not a metric, and in general, the triangle inequality does not hold for one-way set distances.

Nevertheless, the coverage error is easier to handle than the Hausdorff distance, and we can modify it to obtain a triangle-like inequality due to the special properties of the sets under consideration. Letting $\mathbb{R}^d_{\geqq} := \{y \in \mathbb{R}^d \colon 0 \leqq y\}$ be the nonnegative orthant and the sum of two sets be the Minkowski sum, define

$$\mathcal{P}^+ := \mathcal{P} + \mathbb{R}^d_{\geqq}, \qquad \mathcal{P}^+_m := \mathcal{P}_m + \mathbb{R}^d_{\geqq}, \qquad f(\hat{\mathcal{E}}_m(N))^+ := f(\hat{\mathcal{E}}_m(N)) + \mathbb{R}^d_{\geqq}, \qquad f(\mathcal{X}_m)^+ := f(\mathcal{X}_m) + \mathbb{R}^d_{\geqq}.$$

These sets contain all points in $\mathcal{P}, \mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))$, and $f(\mathcal{X}_m)$ respectively, along with all points in $\mathbb{R}^d$ that they dominate. Now define the *modified coverage error* as the random variable

$$\text{dist}_p(\mathcal{P}^+, f(\hat{\mathcal{E}}_m(N))^+) = \sup_{y \in \mathcal{P}^+} \inf_{Y \in f(\hat{\mathcal{E}}_m(N))^+} \|y - Y\|_p. \tag{8}$$

We obtain the desired upper bound for the modified coverage error in the following proposition.

Proposition 1 The following inequality holds almost surely:

$$\text{dist}_p(\mathcal{P}^+, f(\hat{\mathcal{E}}_m(N))^+) \leq \underbrace{\text{dist}_p(\mathcal{P}^+, \mathcal{P}_m^+)}_{\text{deterministic term}} + \underbrace{\text{dist}_p(\mathcal{P}_m^+, f(\hat{\mathcal{E}}_m(N))^+)}_{\text{stochastic term}}. \tag{9}$$

Proof. The special structure of the sets $\mathcal{P}, \mathcal{P}_m$, and $f(\hat{\mathcal{E}}_m(N))$ implies $f(\hat{\mathcal{E}}_m(N))^+ \subseteq \mathcal{P}_m^+ \subseteq \mathcal{P}^+$. Therefore a.s., we have $\text{dist}_p(f(\hat{\mathcal{E}}_m(N))^+, \mathcal{P}^+) = 0$, $\text{dist}_p(f(\hat{\mathcal{E}}_m(N))^+, \mathcal{P}_m^+) = 0$, and $\text{dist}_p(\mathcal{P}_m^+, \mathcal{P}^+) = 0$. Then the relevant one-way distances equal the Hausdorff distances such that a.s.,

$$\begin{aligned}\text{dist}_p(\mathcal{P}^+, f(\hat{\mathcal{E}}_m(N))^+) = \mathbb{H}_p(\mathcal{P}^+, f(\hat{\mathcal{E}}_m(N))^+) &\leq \mathbb{H}_p(\mathcal{P}^+, \mathcal{P}_m^+) + \mathbb{H}_p(\mathcal{P}_m^+, f(\hat{\mathcal{E}}_m(N))^+) \\ &= \text{dist}_p(\mathcal{P}^+, \mathcal{P}_m^+) + \text{dist}_p(\mathcal{P}_m^+, f(\hat{\mathcal{E}}_m(N))^+). \qquad \square\end{aligned}$$

Next, we analyze each term in (9) separately, to determine how these terms change as a function of the chosen discretization $\mathcal{X}_m$ and the amount of sampling N.

4.1 An Upper Bound on the Deterministic Term

In the following Lemma 2, we provide an upper bound on the deterministic error term in (9) that ultimately yields $\text{dist}_p(\mathcal{P}^+, \mathcal{P}_m^+) = O(t)$, where t is the dispersion of the discretized feasible set $\mathcal{X}_m$ from equation (4).

Lemma 2 Under Assumptions 1 and 2, if the dispersion of the sampled points $\text{dist}_r(\mathcal{X}, \mathcal{X}_m)$ equals $t > 0$, then

$$\text{dist}_p(\mathcal{P}^+, \mathcal{P}_m^+) = \sup_{y \in \mathcal{P}^+} \inf_{\tilde{y} \in \mathcal{P}_m^+} \|y - \tilde{y}\|_p \leq \ell t d^{1/p} = O(t).$$

Proof. Since $\mathcal{P}_m^+ = f(\mathcal{X}_m)^+$, it is sufficient to show the upper bound for $\text{dist}_p(\mathcal{P}^+, f(\mathcal{X}_m)^+)$. First, notice that under Assumptions 1, 2 and by Lemma 1, for each $y^* \in \mathcal{P}$, there exists a point $\tilde{y}(y^*) \in f(\mathcal{X}_m)$ such that $\|y^* - \tilde{y}(y^*)\|_p \leq \ell t$ (see also Niederreiter 1992, Ch. 6). Then it follows that

$$\text{dist}_p(\mathcal{P}, f(\mathcal{X}_m)^+) \leq \text{dist}_p(\mathcal{P}, f(\mathcal{X}_m)) = \sup_{y^* \in \mathcal{P}} \inf_{\tilde{y} \in f(\mathcal{X}_m)} \|y^* - \tilde{y}\|_p \leq \sup_{y^* \in \mathcal{P}} \|y^* - \tilde{y}(y^*)\|_p \leq \ell t. \tag{10}$$

Now, it only remains to examine the relationship between $\text{dist}_p(\mathcal{P}, f(\mathcal{X}_m)^+)$ and

$$\text{dist}_p(\mathcal{P}^+, f(\mathcal{X}_m)^+) = \sup_{y \in \mathcal{P}^+} \inf_{\tilde{y} \in f(\mathcal{X}_m)^+} \|y - \tilde{y}\|_p \tag{11}$$

where it is straightforward to see that $\mathcal{P} \subset \mathcal{P}^+$ implies

$$\text{dist}_p(\mathcal{P}, f(\mathcal{X}_m)^+) \leq \text{dist}_p(\mathcal{P}^+, f(\mathcal{X}_m)^+). \tag{12}$$

Let $y_{\sup} \in \mathcal{P}^+ \cap \mathbb{R}^d$ be a (possibly non-unique) finite-valued location of the supremum in the right side of (11). If $y_{\sup} \in \mathcal{P}$, then $\text{dist}_p(\mathcal{P}, f(\mathcal{X}_m)^+) = \text{dist}_p(\mathcal{P}^+, f(\mathcal{X}_m)^+)$. Thus, henceforth, suppose $y_{\sup} \in \mathcal{P}^+ \setminus \mathcal{P}$. Then there exists $y^*_{\sup} \in \mathcal{P}$ such that $y^*_{\sup} \leq y_{\sup}$. Further, there exists a point in $f(\mathcal{X}_m)^+$ nearest to $y^*_{\sup}$ which we denote as

$$\tilde{y}_{\inf}(y^*_{\sup}) = \inf_{\tilde{y} \in f(\mathcal{X}_m)^+} \|y^*_{\sup} - \tilde{y}\|_p,$$

where $(\{\tilde{y}_{\inf}(y^*_{\sup})\}+\mathbb{R}^d_{\geqq}) \subset f(\mathcal{X}_m)^+$. Then it follows that under the $p=\infty$ norm,

$$\begin{aligned}\mathrm{dist}_\infty(\mathcal{P}^+, f(\mathcal{X}_m)^+) = \mathrm{dist}_\infty(\{y_{\sup}\}, f(\mathcal{X}_m)^+) &\le \mathrm{dist}_\infty\left(\{y^*_{\sup}\}+\mathbb{R}^d_{\geqq},\, \{\tilde{y}_{\inf}(y^*_{\sup})\}+\mathbb{R}^d_{\geqq}\right)\\ &= \sup_{z\in(\{y^*_{\sup}\}+\mathbb{R}^d_{\geqq})}\ \inf_{\tilde{z}\in(\{\tilde{y}_{\inf}(y^*_{\sup})\}+\mathbb{R}^d_{\geqq})} \max\{|z_1-\tilde{z}_1|,\ldots,|z_d-\tilde{z}_d|\}\\ &\le \|y^*_{\sup}-\tilde{y}_{\inf}(y^*_{\sup})\|_\infty \le \mathrm{dist}_\infty(\mathcal{P}, f(\mathcal{X}_m)^+).\end{aligned}$$

Therefore, together with (12), concude $\mathrm{dist}_\infty(\mathcal{P}^+, f(\mathcal{X}_m)^+) = \mathrm{dist}_\infty(\mathcal{P}, f(\mathcal{X}_m)^+)$. Now, to obtain a result under the p-norm, use (5) and (10) to obtain

$$\begin{aligned}\mathrm{dist}_p(\mathcal{P}^+, f(\mathcal{X}_m)^+) = \sup_{y\in\mathcal{P}^+}\inf_{\tilde{y}\in f(\mathcal{X}_m)^+}\|y-\tilde{y}\|_p &\le \sup_{y\in\mathcal{P}^+}\inf_{\tilde{y}\in f(\mathcal{X}_m)^+} d^{1/p}\|y-\tilde{y}\|_\infty\\ &= d^{1/p}\,\mathrm{dist}_\infty(\mathcal{P}^+, f(\mathcal{X}_m)^+) = d^{1/p}\,\mathrm{dist}_\infty(\mathcal{P}, f(\mathcal{X}_m)^+)\\ &\le d^{1/p}\,\mathrm{dist}_p(\mathcal{P}, f(\mathcal{X}_m)^+) \le d^{1/p}\ell t. \qquad \square\end{aligned}$$

4.2 An Upper Bound on the Stochastic Term

Next, we turn our attention to determining an upper bound on the stochastic error term from (9). If this term is nonzero, then a point that is truly Pareto under the discretization is falsely excluded from the estimated Pareto set.

Lemma 3 Let $\delta>0$. Under Assumptions 2 and 3, if $|\bar{F}_k(\tilde{x}_i, n_{\tilde{x}_i}) - f_k(\tilde{x}_i)| < \delta$ for all $k=1,\ldots,d$ and all $\tilde{x}_i \in \mathcal{X}_m$, then

$$\mathrm{dist}_p(\mathcal{P}^+_m, f(\hat{\mathcal{E}}_m(N))^+) \le 2c\delta d^{1/p} \quad a.s.$$

Proof. First, we demonstrate that $\mathrm{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))) \le 2c\delta$. Then, we use techniques similar to those used in the proof of Lemma 2 to achieve the final result.

To show that $\mathrm{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))) \le 2c\delta$, let the postulates hold. Then for all $\tilde{x}_i$, the estimated objective vector value $\bar{F}(\tilde{x}_i, n_{\tilde{x}_i})$ is inside a d-dimensional hypercube centered at $f(\tilde{x}_i)$ having edge length 2δ. For $x\in\mathcal{X}$, denote this d-dimensional hypercube as

$$\Delta_d(x,\delta) := \{y\in\mathbb{R}^d\colon |y-f_k(x)|<\delta \text{ for all } k=1,\ldots,d\}.$$

Now, notice that if $\mathrm{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N)))>0$, then $\mathcal{P}_m \setminus f(\hat{\mathcal{E}}_m(N)) \ne \emptyset$. That is, a truly Pareto point (under the discretization) must have been falsely excluded from the estimated Pareto set by some other point. Let $f(\widetilde{X}^*_i) \in \mathcal{P}_m \setminus f(\hat{\mathcal{E}}_m(N))$ be a truly Pareto point under the discretization whose objective function value was excluded. Then there must exist another point $f(\widetilde{X}^*_j) \in f(\hat{\mathcal{E}}_m(N))$ such that $\widetilde{X}^*_j \in \hat{\mathcal{E}}_m(N)$ and $\bar{F}(\widetilde{X}^*_j, n_{\widetilde{X}^*_j}) \le \bar{F}(\widetilde{X}^*_i, n_{\widetilde{X}^*_i})$. Since $\bar{F}(\widetilde{X}^*_j, n_{\widetilde{X}^*_j}) \le \bar{F}(\widetilde{X}^*_i, n_{\widetilde{X}^*_i})$ and $\bar{F}(\widetilde{X}^*_i, n_{\widetilde{X}^*_i}) \in \Delta_d(\widetilde{X}^*_i, \delta)$, we have

$$\bar{F}(\widetilde{X}^*_j, n_{\widetilde{X}^*_j}) \le f(\widetilde{X}^*_i) + \delta 1_d, \tag{13}$$

where 1_d denotes a d-dimensional vector of ones. Since $f(\widetilde{X}^*_i)\in\mathcal{P}_m$ implies $f(\widetilde{X}^*_j) \not\leq f(\widetilde{X}^*_i)$, there exists at least one objective k' such that

$$f_{k'}(\widetilde{X}^*_i) < f_{k'}(\widetilde{X}^*_j). \tag{14}$$

Noting that $\bar{F}(\widetilde{X}^*_j, n_{\widetilde{X}^*_j}) \in \Delta_d(\widetilde{X}^*_j, \delta)$ and combining (13) with (14), it follows that

$$f_{k'}(\widetilde{X}^*_i) - \delta < f_{k'}(\widetilde{X}^*_j) - \delta \le \bar{F}_{k'}(\widetilde{X}^*_j, n_{\widetilde{X}^*_j}) \le f_{k'}(\widetilde{X}^*_i) + \delta, \tag{15}$$

which implies $|f_{k'}(\widetilde{X}_i^*) - f_{k'}(\widetilde{X}_j^*)| \leq 2\delta$. Under Assumption 3, we have $\|f(\widetilde{X}_i^*) - f(\widetilde{X}_j^*)\|_p < 2c\delta$. Thus, $\text{dist}_p(f(\widetilde{X}_i^*), f(\hat{\mathcal{E}}_m(N))) \leq 2c\delta$. Since $f(\widetilde{X}_i^*)$ is an arbitrary point in $\mathcal{P}_m$, then $\text{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))) \leq 2c\delta$.

Now, since $f(\hat{\mathcal{E}}_m(N)) \subset f(\hat{\mathcal{E}}_m(N))^+$, we have $\text{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))^+) \leq \text{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N)))$. Follow a process similar to that in the proof of Lemma 2 to see that

$$\text{dist}_p(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))^+) \leq \text{dist}_p(\mathcal{P}_m^+, f(\hat{\mathcal{E}}_m(N))^+)$$

and $\text{dist}_\infty(\mathcal{P}_m, f(\hat{\mathcal{E}}_m(N))^+) = \text{dist}_\infty(\mathcal{P}_m^+, f(\hat{\mathcal{E}}_m(N))^+)$. Then, the result follows by applying (5). □

4.3 An Upper Bound on the Modified Coverage Error

Combining the inequality in (9) with the upper bounds from Lemmas 2 and 3, we have the following Theorem 1.

Theorem 1 Let $\delta > 0$ and Assumptions 1, 2, and 3 hold. If the dispersion of the sampled points $\text{dist}_r(\mathcal{X}, \mathcal{X}_m)$ equals $t > 0$ and $|\bar{F}_k(\tilde{x}_i, n) - f_k(\tilde{x}_i)| < \delta$ for all $k = 1, \ldots, d$ and all $\tilde{x}_i \in \mathcal{X}_m$, then an upper bound on the modified coverage error is

$$\text{dist}_p(\mathcal{P}^+, f(\hat{\mathcal{E}}_m(N))^+) \leq (\ell t + 2c\delta)d^{1/p} \quad a.s.$$

Further analyses regarding the rate of decay of the modified coverage error can now be conducted using properties of the chosen set $\mathcal{X}_m$ and properties of the underlying probability space; see, e.g., Vershynin (2018). We leave these analyses to future work.

5 HYPERVOLUME DIFFERENCE

Due to their desirable mathematical properties, the hypervolume-related performance indicators are arguably amongst the most popular and well-studied in the deterministic multiobjective optimization community (Audet et al. 2021). Inspired by Branke et al. (2016) and as discussed in Hunter et al. (2019), we consider the hypervolume difference, defined as follows.

First, let the hypervolume of a set $\mathcal{A} \subset \mathbb{R}^d$ with respect to a reference point z be the Lebesgue measure of the set of points that $\mathcal{A}$ dominates and that dominate z; that is,

$$\mathbb{HV}(\mathcal{A}, z) := \lambda\left(\cup_{a \in \mathcal{A}} \{y \in \mathbb{R}^d : a \leq y \leq z\}\right),$$

where λ denotes the Lebesgue measure. The hypervolume difference between the Pareto set and the true objective values of the estimated efficient set equals the random variable

$$\mathbb{HVD}(\mathcal{P}, f(\hat{\mathcal{E}}_m(N)), z) := \mathbb{HV}(\mathcal{P}, z) - \mathbb{HV}(f(\hat{\mathcal{E}}_m(N)), z). \tag{16}$$

This result follows from the more general definition of the hypervolume difference between two sets (see Branke et al. 2016; Hunter et al. 2019) because for every point in $f(\hat{\mathcal{E}}_m(N))$, there exists a point in $\mathcal{P}$ that weakly dominates it. For test problems in which the Pareto set is known and the true objective function values can be evaluated exactly, computing the hypervolume difference in (16) can be accomplished by one of the many algorithms detailed in Guerreiro et al. (2021). Note that $\mathbb{HVD}$ is sensitive to the choice of reference point. The nadir or its approximation is a natural choice (Ehrgott and Tenfelde-Podehl 2003; Kirlik and Sayin 2013; Audet et al. 2021), but there is no consensus on the best choice (Li and Yao 2019).

For mathematical analysis, it is straightforward to split the quantity in (16) into the sum of deterministic and stochastic error terms,

$$\mathbb{HVD}(\mathcal{P}, f(\hat{\mathcal{E}}_m(N)), z) = \underbrace{(\mathbb{HV}(\mathcal{P}, z) - \mathbb{HV}(\mathcal{P}_m, z))}_{\text{deterministic term}} + \underbrace{\left(\mathbb{HV}(\mathcal{P}_m, z) - \mathbb{HV}(f(\hat{\mathcal{E}}_m(N)), z)\right)}_{\text{stochastic term}}.$$

Splitting the analyses in this fashion may enable the use of hypervolume indicator results from the deterministic multi-objective optimization literature; see, for example, Marescaux and Hansen (2021) and the discussion in Shang et al. (2021). We leave specific bounds on the hypervolume difference to future work.

6 CONCLUDING REMARKS

We discuss several challenges facing the evaluation and assessment of algorithms designed to solve the MOSO problem to global optimality, and we discuss the properties of several performance measures relative to these challenges. We provide a modification to the coverage error that allows us to derive an upper bound enabling future mathematical analysis for this performance indicator. To the best of our knowledge, this result constitutes the first time such analytical upper bounds on a MOSO performance indicator have appeared in the literature.

ACKNOWLEDGMENTS

The authors thank the National Science Foundation for support under grant CMMI-1554144.

REFERENCES

Applegate, E. A., G. Feldman, S. R. Hunter, and R. Pasupathy. 2020. "Multi-Objective Ranking and Selection: Optimal Sampling Laws and Tractable Approximations via SCORE". *Journal of Simulation* 14(1):21–40.

Audet, C., J. Bigeon, D. Cartier, S. Le Digabel, and L. Salomon. 2021. "Performance Indicators in Multiobjective Optimization". *European Journal of Operational Research* 292:397–422.

Branke, J. 2023. "Performance Metrics for Multi-Objective Optimisation Algorithms under Noise". *arXiv:2302.14179*.

Branke, J., W. Zhang, and Y. Tao. 2016. "Multiobjective Ranking and Selection Based on Hypervolume". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 859–870. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Cooper, K., S. R. Hunter, and K. Nagaraj. 2020. "Bi-objective Simulation Optimization on Integer Lattices Using the Epsilon-Constraint Method in a Retrospective Approximation Framework". *INFORMS Journal on Computing* 32(4):1080–1100.

Eckman, D. J., S. G. Henderson, and S. Shashaani. 2023a. "Diagnostic Tools for Evaluating and Comparing Simulation-Optimization Algorithms". *INFORMS Journal on Computing* 35(2):350–367.

Eckman, D. J., S. G. Henderson, and S. Shashaani. 2023b. "SimOpt: A Testbed for Simulation-Optimization Experiments". *INFORMS Journal on Computing* 35(2):495–508.

Ehrgott, M., and D. Tenfelde-Podehl. 2003. "Computation of Ideal and Nadir Values and Implications for Their Use in MCDM Methods". *European Journal of Operational Research* 151:119–139.

Faulkenberg, S. L., and M. M. Wiecek. 2010. "On the Quality of Discrete Representations in Multiple Objective Programming". *Optimization and Engineering* 11:423–440.

Fu, M. C. (Ed.) 2015. *Handbook of Simulation Optimization*, Volume 216 of *International Series in Operations Research & Management Science*. New York: Springer.

Guerreiro, A. P., C. M. Fonseca, and L. Paquete. 2021. "The Hypervolume Indicator: Computational Problems and Algorithms". *ACM Computing Surveys* 54(6):1–42.

Hunter, S. R., E. A. Applegate, V. Arora, B. Chong, K. Cooper, O. Rincón-Guevara, and C. Vivas-Valencia. 2019. "An Introduction to Multi-Objective Simulation Optimization". *ACM Transactions on Modeling and Computer Simulation* 29(1):7:1–7:36.

Hunter, S. R., and B. McClosky. 2016. "Maximizing Quantitative Traits in the Mating Design Problem via Simulation-Based Pareto Estimation". *IIE Transactions* 48(6):565–578.

Hunter, S. R., and R. Pasupathy. 2022. "Central Limit Theorems For Constructing Confidence Regions in Strictly Convex Multi-Objective Simulation Optimization". In *Proceedings of the 2022 Winter Simulation Conference*, edited by B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C. G. Corlu, L. H. Lee, E. P. Chew, T. Roeder, and P. Lendermann, 3015–3026. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Karp, R. M. 1991. "An Introduction to Randomized Algorithms". *Discrete Applied Mathematics* 34(1–3):165–201.

Kirlik, G., and S. Sayin. 2013. "Computing the Nadir Point for Multiobjective Discrete Optimization Problems". *Journal of Global Optimization* 62:79–99.

Law, A. M. 2015. *Simulation Modeling and Analysis*. 5 ed. New York: McGraw Hill Education.

Li, M., and X. Yao. 2019. "Quality Evaluation of Solution Sets in Multiobjective Optimisation: A Survey". *ACM Computing Surveys* 52(2):1–38.

Marescaux, E., and N. Hansen. 2021. "Hypervolume in Biobjective Optimization Cannot Converge Faster Than $\Omega(1/p)$". In *2021 Genetic and Evolutionary Computation Conference*, GECCO '21, 430 – 438. New York, NY, USA: Association for Computing Machinery.

Markowitz, H. 1952. "Portfolio Selection". *The Journal of Finance* 7(1):77–91.

Niederreiter, H. 1992. *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Ondes, B. E., and S. R. Hunter. 2023. "An Upper Bound on the Hausdorff Distance between a Pareto Set and Its Discretization in Bi-objective Convex Quadratic Optimization". *Optimization Letters* 17:45–74.

Pasupathy, R., and S. G. Henderson. 2006. "A Testbed of Simulation-Optimization Problems". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 255–263. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ragavan, P. K., S. R. Hunter, R. Pasupathy, and M. R. Taaffe. 2022. "Adaptive Sampling Line Search for Local Stochastic Optimization with Integer Variables". *Mathematical Programming* 196:775–804.

Sayin, S. 2000. "Measuring the Quality of Discrete Representations of Efficient Sets in Multiple Objective Mathematical Programming". *Mathematical Programming* 87:543–560.

Shang, K., H. Ishibuchi, L. He, and L. M. Pang. 2021. "A Survey on the Hypervolume Indicator in Evolutionary Multiobjective Optimization". *IEEE Transactions on Evolutionary Computation* 25(1):1–20.

Shapiro, A., D. Dentcheva, and A. Ruszczyński. 2009. *Lectures on Stochastic Programming: Modeling and Theory*. MPS-SIAM Series on Optimization. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Vershynin, R. 2018. *High-Dimensional Probability: An Introduction with Applications in Data Science*, Volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge, UK: Cambridge University Press.

Vogel, S. 2017. "Random Approximations in Multiobjective Optimization". *Mathematical Programming* 164:29–53.

Yakowitz, S., P. L'Ecuyer, and F. Vázquez-Abad. 2000. "Global Stochastic Optimization with Low-Dispersion Point Sets". *Operations Research* 48(6):939–950.

AUTHOR BIOGRAPHIES

SUSAN R. HUNTER is an Associate Professor in the School of Industrial Engineering at Purdue University. Her research interests include Monte Carlo methods and stochastic optimization with a specific focus on multi-objective simulation optimization. She currently serves as Co-Proceedings Editor for the 2023 Winter Simulation Conference, Program Chair for the 2024 Winter Simulation Conference, and as an associate editor for Operations Research. Her email address is susanhunter@purdue.edu and her website is https://web.ics.purdue.edu/~hunter63/.

BURLA E. ONDES is a Ph.D. student in the School of Industrial Engineering at Purdue University. She received a B.S. in Industrial and Systems Engineering from the University of Florida in 2019 and a M.S. in Industrial Engineering from Purdue University in 2022. Her primary research interest is multi-objective simulation optimization. Her email address is bondes@purdue.edu and her website is https://burlaondes.github.io/.