The Fortieth AAAI Conference on Artificial Intelligence (AAAI-26)

Center-Outward q-Dominance: A Sample-Computable Proxy for Strong Stochastic Dominance in Stochastic Multi-Objective Optimisation

Robin van der Laag[1], Hao Wang[1], Thomas Bäck[1], Yingjie Fan[1]
[1]LIACS, Leiden University
{r.p.van.der.laag, h.wang, t.h.w.baeck, y.fan}@liacs.leidenuniv.nl

Abstract

Stochastic multi-objective optimization (SMOOP) requires ranking multivariate distributions; yet, most empirical studies perform scalarization, which loses information and is unreliable. Based on the optimal transport theory, we introduce the *center-outward q-dominance relation* and prove it implies strong first-order stochastic dominance (FSD). Also, we develop an empirical test procedure based on q-dominance, and derive an explicit sample size threshold, $n^*(\delta)$, to control the Type I error. We verify the usefulness of our approach in two scenarios: (1) as a ranking method in hyperparameter tuning; (2) as a selection method in multi-objective optimization algorithms. For the former, we analyze the final stochastic Pareto sets of seven multi-objective hyperparameter tuners on the YAHPO-MO benchmark tasks with q-dominance, which allows us to compare these tuners when the expected hypervolume indicator (HVI, the most common performance metric) of the Pareto sets becomes indistinguishable. For the latter, we replace the mean value-based selection in the NSGA-II algorithm with q-dominance, which shows a superior convergence rate on noise-augmented ZDT benchmark problems. These results establish center-outward q-dominance as a principled, tractable foundation for seeking truly stochastically dominant solutions for SMOOPs.

Code — https://github.com/RvdLaag/qDominance
Extended version — https://arxiv.org/abs/2511.12545

1 Introduction

Decision-making often involves balancing conflicting objectives under stochasticity. Hyperparameter optimization (HPO) is a typical example, where each run of an algorithm produces a vector of random outcomes (validation accuracy, training time, memory footprint, etc.) whose joint distribution reflects noise in the data and the learning procedure itself. More generally, each candidate solution of a stochastic multi-objective optimization problem (SMOOP) is associated with a distribution of multi-objective performance. Comparing different solutions, therefore, requires ranking these distributions rather than comparing single deterministic points.

Studies on SMOOPs have typically resorted to one of the following pragmatic strategies:

- Scalarization: turns the problem into a single-objective stochastic optimization problem by applying some predetermined scalarization function to the objectives (Caballero et al. 2004; Abdelaziz 2012; Drugan and Nowé 2013; Trappler, Helbert, and Riche 2025);
- Single-sample iterations: solving the optimization problem by evaluating candidate solutions using a single Monte-Carlo draw per iteration (Jin and Branke 2005);
- Moment surrogates: replacing each random objective by a summary statistic, such as the mean (Fliege and Xu 2011), variance, (C)VaR (Daulton et al. 2022), or its worst-case scenario, e.g., minimax robustness (Ehrgott, Ide, and Schöbel 2014), multi-objective multi-armed bandits (Lu et al. 2019; Xu and Klabjan 2023), thereby converting the problem into a deterministic one;
- Weak stochastic dominance testing: comparing the empirical cumulative distribution functions (CDF) of the candidate solutions (Teich 2001);

While these strategies ease computation, they also discard some parts of the distributional structure and can lead to suboptimal solutions.

A principled alternative to compare distributions is *strong first-order stochastic dominance* (FSD): a distribution A is preferred to B if every non-decreasing utility considers samples from A at least as good as samples from B in expectation. Strong FSD is attractive; it preserves distributional information and imposes no arbitrary scalarization. Section 2 contains the necessary background on (stochastic) multi-objective optimization problems and stochastic dominance.

Testing for strong FSD through samples can be computationally difficult in higher dimensions. Recent work by Rioux et al., 2025 proposes a statistic that assesses multivariate *almost* stochastic dominance by constructing an optimal coupling between two distributions.

In this paper, we take a different route. We introduce *center-outward q-dominance* (formalized in Section 4): utilizing center-outward distributions and quantile functions (Hallin 2017), which we cover in Section 3, we construct maps from each distribution to the same uniform distribution U_d on the unit ball. With this construction, all pairwise comparisons reuse a single common reference frame; we thus only compare samples with identical center-outward rank and sign, ruling out the possibility of incoherent couplings when comparing more than two distributions, such as $X_i \leftrightarrow Y_j$, $X_i \leftrightarrow Z_k$, but not $Y_j \leftrightarrow Z_k$. Consequently, we only need to compute one coupling per distribution, independent of the number of distributions we wish to compare, in contrast to the method in (Rioux et al. 2025), which requires the computation of a coupling for each order pair of distributions. We show that if q-dominance holds for all quantiles, we obtain strong FSD; if it only holds up to some quantile, we get a natural relaxation. Because the center-outward ranks are computed once, the entire set of relaxations is obtained at no extra cost. Rioux et al. can similarly vary their threshold ε_0 for free once an entropic-OT coupling is fixed, but exploring different regularization levels λ or functions h still requires a new coupling to be obtained, whereas our method has no tunable hyperparameters. Furthermore, our finite-sample test admits an explicit minimum sample size threshold $n^*(\delta)$ that guarantees a user-specified Type-I error δ. Lastly, we propose a sorting procedure that enables practical ranking of solutions through pairwise q-dominance and its relaxation.

We verify the usefulness of our approach with two complementary experiments in Section 5. Firstly, we re-analyze the final stochastic Pareto sets of seven multi-objective HPO methods on the YAHPO-MO benchmark tasks (Pfisterer et al. 2022) with q-dominance, as opposed to the expected Hypervolume Indicator (HVI). We find that, with q-dominance, we can still meaningfully compare the different methods even when the expected HVI of the Pareto sets becomes indistinguishable. Secondly, we apply q-dominance to the well-known NSGA-II algorithm (Deb et al. 2002) for multi-objective optimization, where q-dominance is used to compare different solutions under uncertainty of the objectives. Compared to the default mean value-based or single-sample-based selection, q-dominance enables the algorithm to converge significantly faster on noise-augmented ZDT benchmark problems (Zitzler, Deb, and Thiele 2000).

2 Background

In this section, we will provide an introduction to the background of multi-objective optimization and stochastic dominance.

2.1 Multi-Objective Optimization Problems

A multi-objective optimization problem (MOOP) is typically formulated as

$$\begin{aligned} \max_{x} \quad & (f_1(x), \ldots, f_m(x)) \\ \text{subject to} \quad & x \in \mathcal{X}, \end{aligned} \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^d$ is the solution space, and $f : \mathcal{X} \to \mathbb{R}^m$, with $m \geq 2$, the vector-valued objective function.

Definition 1. A feasible solution $x \in \mathcal{X}$ is said to *Pareto dominate* another solution $y \in \mathcal{X}$, if

(i) For all $i \in \{1, \ldots, m\}$, $f_i(x) \geq f_i(y)$, and
(ii) There exists an $i \in \{1, \ldots, m\}$, such that $f_i(x) > f_i(y)$.

A solution $x^* \in \mathcal{X}$ along with its image point $f(x^*)$ is called *Pareto optimal* if there does not exist another solution that dominates it. The set of Pareto optimal points is called the *Pareto front*.

Assume a probability space $(\Omega, \mathcal{A}, P)$. Stochastic multi-objective optimization problems (SMOOP) arise when the objective function is stochastic, i.e., $f : \mathcal{X} \times \Omega \to \mathbb{R}^m$. With $\omega \in \Omega$, an SMOOP can be written as

$$\begin{aligned} \max_{x} \quad & (f_1(x, \omega), \ldots, f_m(x, \omega)) \\ \text{subject to} \quad & x \in \mathcal{X}, \end{aligned} \tag{2}$$

Notably, the Pareto optimal solution to an SMOOP depends on the realization ω. Two typical approaches to interpret the maximization, and thus the comparison, of random vectors are the *multi-objective method* and the *stochastic method* (Abdelaziz 2012).

The *multi-objective method* defines, for each component of the objective function f_i, a vector $(\mathcal{F}_i^{(1)}(f_i(x, \omega)), \ldots, \mathcal{F}_i^{(r_i)}(f_i(x, \omega)))$ of one or more statistical functionals $\mathcal{F}_i^{(j)}$ of the random variable $f_i(x, \omega)$, common functionals are the expectation and the variance. With this, we reformulate the SMOOP (2) as an MOOP with the $r_1 + \cdots + r_m$ functionals as the deterministic objectives.

The *stochastic method* scalarizes the random objectives $f_1(x, \omega), \ldots, f_m(x, \omega)$ using a function $u : \mathbb{R}^m \to \mathbb{R}$, and reformulates SMOOP (2) as a single-objective stochastic optimization problem.

Both of these methods have obvious downsides. The multi-objective method reduces the joint distribution of the objectives to a finite set of summary statistics (the functionals), losing information about the distribution in the process, particularly because the dependency between the random objectives is not taken into account. The stochastic method does not lose this dependency between objectives; however, it does require that the scalarization function is known, and the solution does not generalize to other scalarization functions.

A different approach is to utilize the concept of multivariate stochastic dominance, which allows us to compare random vectors directly without losing any information.

2.2 Stochastic Dominance

Firstly, let us introduce the concepts of first-order stochastic dominance (FSD) in the scalar case.

Definition 2. Let X and Y be real-valued random variables with cumulative distribution functions F_X and F_Y. We say that X first-order stochastically dominates Y, written $X \succeq_1 Y$, if and only if any (and hence all) of the following equivalent conditions hold:

(i) $F_X(z) \leq F_Y(z)$ for all $z \in \mathbb{R}$;
(ii) $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all non-decreasing utility functions $u : \mathbb{R} \to \mathbb{R}$.

A natural first step is to lift Definition 2 to $\mathbb{R}^d$ by replacing the scalar CDF with the joint CDF

$$F_{\mathbf{X}}(\mathbf{z}) = \mathbb{P}(\mathbf{X}_1 \leq \mathbf{z}_1, \ldots, \mathbf{X}_d \leq \mathbf{z}_d)$$

and by letting the test set be all non-decreasing utility functions $u : \mathbb{R}^d \to \mathbb{R}$. Crucially, in dimension $d > 1$, these two characterizations diverge, giving rise to the weak (via CDFs) and strong (via utility functions) notions of multivariate FSD.

Definition 3 (Weak FSD)**.** Let $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^d$ be real-valued random vectors with joint CDFs $F_{\mathbf{X}}$ and $F_{\mathbf{Y}}$. We say that $\mathbf{X}$ weakly stochastically dominates $\mathbf{Y}$, written $\mathbf{X} \succeq_w \mathbf{Y}$, if and only if

$$F_{\mathbf{X}}(\mathbf{z}) \leq F_{\mathbf{Y}}(\mathbf{z}), \qquad \forall \mathbf{z} \in \mathbb{R}^d.$$

Definition 4 (Strong FSD)**.** We say that $\mathbf{X}$ strongly stochastically dominates $\mathbf{Y}$, written $\mathbf{X} \succeq_1 \mathbf{Y}$, if and only if

$$\mathbb{E}[u(\mathbf{X})] \geq \mathbb{E}[u(\mathbf{Y})],$$

for all non-decreasing utility functions $u : \mathbb{R}^d \to \mathbb{R}$.

To relate the two orders, we first recall an equivalent characterization of strong FSD from (Sriboonchita et al. 2009) that replaces utilities by probabilities on so-called upper sets.

Proposition 1. *For random vectors $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^d$ the following are equivalent:*

(i) $\mathbf{X} \succeq_1 \mathbf{Y}$ *(Definition 4).*

(ii) For every upper set $M \subseteq \mathbb{R}^d$,

$$\mathbb{P}(\mathbf{X} \in M) \geq \mathbb{P}(\mathbf{Y} \in M).$$

Here an upper set is any $M \subseteq \mathbb{R}^d$ such that for any $\mathbf{z}, \mathbf{w} \in \mathbb{R}^d$ with $\mathbf{w} \geq \mathbf{z}$ we have that $\mathbf{w} \in M$ whenever $\mathbf{z} \in M$.

Proof. We refer to (Sriboonchita et al. 2009). □

The upper set view makes the hierarchy between the two orders apparent, yielding the following result.

Theorem 1. *For every dimension $d \geq 1$ and every pair of random vectors $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^d$,*

$$\mathbf{X} \succeq_1 \mathbf{Y} \implies \mathbf{X} \succeq_w \mathbf{Y}.$$

Proof. For $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$ we have that $1 - F_{\mathbf{X}}(\mathbf{m}) = \mathbb{P}(\mathbf{X} \geq \mathbf{m}) = \mathbb{P}(\mathbf{X} \in [m_1, \infty) \times \cdots \times [m_d, \infty))$ and similarly for $\mathbf{Y}$. The set $[m_1, \infty) \times \cdots \times [m_d, \infty)$ is an upper set in $\mathbb{R}^d$, thus if $\mathbb{P}(\mathbf{X} \in M) \geq \mathbb{P}(\mathbf{Y} \in M)$ for all upper sets $M \in \mathcal{M}$, then it also holds for the special form above. This gives us that $1 - F(\mathbf{m}) \geq 1 - G(\mathbf{m})$, implying $\mathbf{X} \succeq_w \mathbf{Y}$. □

The converse does not hold for dimensions $d > 1$, as the next counterexample from (Kopa and Petrová 2018) demonstrates.

Example 1. Consider the two random vectors

$$\mathbf{X} = \begin{cases} (0,1), & \text{w.p. } 1/2, \\ (1,0), & \text{w.p. } 1/2. \end{cases} \quad \mathbf{Y} = \begin{cases} (0,0), & \text{w.p. } 1/2, \\ (1,1), & \text{w.p. } 1/2. \end{cases}$$

Then $F_{\mathbf{X}}(\mathbf{z}) \leq F_{\mathbf{Y}}(\mathbf{z})$ for all $\mathbf{z} \in \mathbb{R}^2$ and thus $\mathbf{X} \succeq_w \mathbf{Y}$. Consider the upper set $M = \{\mathbf{m} \in \mathbb{R}^2 : m_1 + m_2 \geq \frac{3}{2}\}$, then $0 = \mathbb{P}(\mathbf{X} \in M) < \mathbb{P}(\mathbf{Y} \in M) = \frac{1}{2}$, and thus $\mathbf{X} \not\succeq_1 \mathbf{Y}$.

Example 1 makes the dilemma explicit: in dimensions $d > 1$ weak FSD can be too permissive, declaring $\mathbf{X} \succeq_w \mathbf{Y}$ even though strong FSD rejects that claim, i.e. $\mathbf{X} \not\succeq_1 \mathbf{Y}$. Equivalently, there exists a non-decreasing utility u—for instance $u(x_1, x_2) = \exp(x_1 + x_2)$—for which $\mathbb{E}[u(\mathbf{X})] < \mathbb{E}[u(\mathbf{Y})]$.

As noted in the introduction, testing for strong FSD can be computationally difficult without full knowledge of the joint distribution. Rioux et al. use the following theorem from (Shaked and Shanthikumar 2007) for their optimal transport approach.

Theorem 2. *The random vectors $\mathbf{X}, \mathbf{Y}$ satisfy $\mathbf{X} \succeq_1 \mathbf{Y}$, if and only if there exists a coupling $(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})$ of $(\mathbf{X}, \mathbf{Y})$ satisfying $\mathbb{P}(\widehat{\mathbf{X}} \geq \widehat{\mathbf{Y}}) = 1$.*

Rioux et al. then provides the following lemma, casting this theorem into the context of optimal transport.

Lemma 3. *Let $P_{\mathbf{X}}, P_{\mathbf{Y}}$ denote the distributions of the random vectors $\mathbf{X}$ and $\mathbf{Y}$, respectively. Then $\mathbf{X} \succeq_1 \mathbf{Y}$ if*

$$\inf_{\pi \in \Pi(P_{\mathbf{X}}, P_{\mathbf{Y}})} \int c \, \mathrm{d}\pi = 0,$$

where $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ is the cost function $c(x, y) = \mathbf{1}_{\{x \leq y\}}$ and $\Pi(P_{\mathbf{X}}, P_{\mathbf{Y}})$ denotes the set of all couplings of $(P_{\mathbf{X}}, P_{\mathbf{Y}})$.

When more than two distributions $P_1, P_2, \ldots$ must be compared, this method requires solving an optimal transport problem for every ordered pair (P_i, P_j). Because each coupling is optimized in isolation a sample $\mathbf{X} \sim P_1$ that is matched to $\mathbf{Y} \sim P_2$ under $\pi_{1,2}$ and to $\mathbf{Z} \sim P_3$ under $\pi_{1,3}$ need not have $\mathbf{Y}$ matched to $\mathbf{Z}$ under $\pi_{2,3}$. The resulting family of couplings, therefore, lacks a common reference frame and is difficult to interpret jointly.

In the next section, we introduce *center-outward ranks and signs*, providing the missing common reference frame that underpins our definition of q-dominance.

3 Center-Outward Ranks and Signs

In this section, we define the center-outward distribution and quantile function, which are rooted in the main result of (McCann 1995) and further developed in (Hallin 2017) and (Hallin et al. 2021).

Throughout this section, we make use of the following notation. Let μ_d denote the Lebesgue measure over $\mathbb{R}^d$ equipped with its Borel σ-field $\mathcal{B}_d$. Furthermore, denote by $\mathcal{P}_d$ the family of Lebesgue-absolutely continuous distributions over $(\mathbb{R}^d, \mathcal{B}_d)$. We use the notation $T\#P_1 = P_2$ for the distribution P_2 of $T(X)$, where $X \sim P_1$, and say that T is pushing P_1 forward to P_2. Lastly, let $\mathcal{S}_{d-1}$, $\mathbb{S}_d$, and $\overline{\mathbb{S}}_d$ denote the unit sphere, the open unit ball, and the closed unit ball in $\mathbb{R}^d$, respectively.

3.1 Theoretical Definition

We restate the main result of (McCann 1995).

Theorem 4 (McCann 1995)**.** *For two distributions $P_1, P_2 \in \mathcal{P}_d$ the following statements hold:*

(i) *the class of functions*

$$\nabla\Psi_{P_1;P_2} := \{\nabla\psi \mid \psi : \mathbb{R}^d \to \mathbb{R},\ \nabla\psi\#P_1 = P_2\},$$

where ψ is convex and lower semi-continuous, is non-empty;

(ii) *if $\nabla\psi', \nabla\psi'' \in \nabla\Psi_{P_1;P_2}$, they coincide P_1-a.s., that is $P_1(\{\mathbf{x} \mid \nabla\psi'(\mathbf{x}) \neq \nabla\psi''(\mathbf{x})\}) = 0$;*

(iii) *if P_1 and P_2 have finite moments of order two, any element of $\nabla\Psi_{P_1;P_2}$ is an optimal quadratic transport pushing P_1 forward to P_2, i.e. it is a solution to*

$$\inf_{\gamma\in\Pi(P_1,P_2)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \|x-y\|^2 \,\mathrm{d}\gamma(x,y),$$

where $\Pi(P_1,P_2)$ denotes the set of joint distributions with marginals P_1 and P_2.

Denote by U_d the uniform distribution over $\mathbb{S}_d$, which is the product of the uniform over the unit sphere and the uniform over the unit interval. The center-outward distribution and quantile functions are then defined as follows.

Definition 5. The center-outward quantile function $\mathbf{Q}^{\pm}$ of $P \in \mathcal{P}_d$ is the a.e. unique element $\nabla\psi \in \nabla\Psi_{U_d;P}$, such that ψ satisfies

$$\psi(\mathbf{u}) = \infty, \text{ for } \|\mathbf{u}\| > 1$$

and

$$\psi(\mathbf{u}) = \liminf_{\mathbb{S}_d\ni\mathbf{v}\to\mathbf{u}} \psi(\mathbf{v}), \text{ for } \|\mathbf{u}\| = 1.$$

Definition 6. Call $\mathbf{F}^{\pm} := \nabla\phi$ the center-outward distribution function, where ϕ is defined as the Legendre transform

$$\phi(\mathbf{x}) := \psi^*(\mathbf{x}) := \sup_{\mathbf{u}\in\mathbb{S}_d} (\langle\mathbf{u},\mathbf{x}\rangle - \psi(\mathbf{u})), \quad \mathbf{x}\in\mathbb{R}^d.$$

The following propositions from (Hallin et al. 2021) summarize the main properties of these functions.

Proposition 2. *Let $\mathbf{Z} \sim P \in \mathcal{P}_d$ and let $\mathbf{F}^{\pm}$ be the center-outward distribution function of P, then*

(i) *$\mathbf{F}^{\pm}$ takes values in $\overline{\mathbb{S}}_d$ and $\mathbf{F}^{\pm}\#P = U_d$. Thus $\mathbf{F}^{\pm}$ is a probability-integral transformation;*

(ii) *$\|\mathbf{F}^{\pm}(\mathbf{Z})\|$ is uniform over $[0,1]$, $\mathbf{S}(\mathbf{Z}) = \mathbf{F}^{\pm}(\mathbf{Z})/\|\mathbf{F}_{\pm}(\mathbf{Z})\|$ is uniform over $\mathcal{S}_{d-1}$, and they are mutually independent;*

(iii) *$\mathbf{F}^{\pm}$ entirely characterizes P;*

(iv) *for $d = 1$, $\mathbf{F}^{\pm}$ coincides with $2F - 1$, where F is the tradition distribution function.*

Definition 7. For $q \in (0,1)$ we call

$$\mathcal{C}(q) := \mathbf{Q}^{\pm}(q\mathcal{S}_{d-1}) = \{\mathbf{z}\in\mathbb{R}^d \mid \|\mathbf{F}^{\pm}(\mathbf{z})\| = q\}$$

the center-outward quantile counter, and

$$\mathbb{C}(q) := \mathbf{Q}^{\pm}(q\overline{\mathbb{S}}_d) = \{\mathbf{z}\in\mathbb{R}^d \mid \|\mathbf{F}^{\pm}(\mathbf{z})\| \leq q\}$$

the center-outward quantile region of order q. Furthermore,

$$\mathbb{C}(0) := \bigcap_{0<q<1} \mathbb{C}(q) = \partial\psi(\mathbf{0}).$$

Proposition 3. *Let $P \in \mathcal{P}_d$ have center-outward quantile function $\mathbf{Q}_{\pm}$, then*

(i) *$\mathbf{Q}^{\pm}\#U_d = P$, and hence $\mathbf{Q}^{\pm}$ entirely characterizes P;*

(ii) *the center-outward quantile region $\mathbb{C}(q)$, for $0 < q < 1$, has P-probability content q;*

3.2 Empirical Estimation

We now examine how to construct empirical counterparts to $\mathbf{F}^{\pm}$ and $\mathbf{Q}^{\pm}$ using samples.

Denote by $\mathbf{Z}^{(n)} := (\mathbf{Z}_1,\dots,\mathbf{Z}_n)$ an n-tuple of i.i.d. random vectors from distribution $P \in \mathcal{P}$, with density f and center-outward distribution function $\mathbf{F}^{\pm}$. For the empirical center-outward distribution function $\widehat{\mathbf{F}}^{\pm}$ (Hallin et al. 2021) propose the following construction.

Assume that $d \geq 2$, and factorize $n = n_R n_S + n_0$ with $n_R, n_S, n_0 \in \mathbb{N}$ and $0 \leq n_0 < \min\{n_R, n_S\}$, such that $n_R \to \infty$ and $n_S \to \infty$ as $n \to \infty$. Next, we define a sequence of regular grids over the unit ball $\mathbb{S}_d$ as the intersection between

- a regular n_S-tuple $\mathfrak{S}^{(n_s)} := (\mathbf{u}_1,\dots,\mathbf{u}_{n_S})$ of unit vectors, and
- n_R hyperspheres centered at $\mathbf{0}$, with radii $\left\{\frac{1}{n_R+1},\dots,\frac{n_R}{n_R+1}\right\}$,

along with n_0 copies of the origin. The discrete distribution with probability masses $1/n$ at each of the $n_R n_S$ grid points and probability mass n_0/n at the origin converges weakly to the uniform U_d over the ball $\mathbb{S}_d$. We refer to this grid as the *augmented grid*.

Definition 8. The empirical center-outward distribution function is any mapping $\widehat{\mathbf{F}}^{\pm} : (\mathbf{Z}_1,\dots,\mathbf{Z}_n) \mapsto \left(\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_1),\dots,\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_n)\right)$ satisfying

$$\sum_{i=1}^{n} \|\mathbf{Z}_i - \widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i)\|^2 = \min_{\pi} \sum_{i=1}^{n} \|\mathbf{Z}_{\pi(i)} - \widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i)\|^2, \quad (3)$$

where the set $\{\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i) \mid i = 1,\dots,n\}$ consists of the n points of the augmented grid and π ranges over the $n!$ permutations of $\{1,2,\dots,n\}$.

The assignment problem in Equation (3) can be solved by, for example, the Hungarian algorithm in $\mathcal{O}(n^3)$.

Along with the definition of the empirical center-outward distribution function $\widehat{\mathbf{F}}^{\pm}$, we define the following concepts:

- center-outward ranks $\widehat{R}_i := (n_R+1)\|\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i)\|$;
- empirical center-outward quantile contours $\widehat{\mathcal{C}}(\frac{j}{n_R+1}) := \{\mathbf{Z}_i \mid \widehat{R}_i = j\}$ and regions $\widehat{\mathbb{C}}(\frac{j}{n_R+1}) := \{\mathbf{Z}_i \mid \widehat{R}_i \leq j\}$, where $j/(n_R+1)$, $j \in \{0,\dots,n_R\}$, are empirical probability contents, to be interpreted as a quantile order;
- center-outward signs: $\widehat{\mathbf{S}}_i := \mathbf{1}_{\{\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i)\neq\mathbf{0}\}} \frac{\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i)}{\|\widehat{\mathbf{F}}^{\pm}(\mathbf{Z}_i)\|}$, and sign curves $\{\mathbf{Z}_i \mid \widehat{\mathbf{S}}_i = \mathbf{u}\}$, for $\mathbf{u} \in \mathfrak{S}^{(n_S)}$.

In Figure 1, samples from a bi-variate distribution along with selected quantile contours and signs in the sample space $\mathbb{R}^2$ (left) and in the unit ball $\mathbb{S}_2$ (right) are shown, along with the mappings $\mathbf{F}_{\pm}$ and $\mathbf{Q}_{\pm}$.

4 Center-Outward Dominance Relation

We now introduce our main contribution: *center-outward q-dominance*, a dominance relation between two distributions

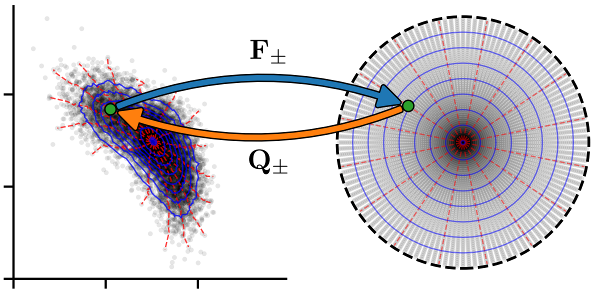


Figure 1: Samples of a bi-variate distribution (left) and the points on the augmented grid (right). Selected center-outward quantile contours are shown in blue and signs in red.

using their center-outward distribution and quantile functions.

Let $P_1, P_2 \in \mathcal{P}_d$ be two probability distributions with center-outward distribution and quantile functions $\mathbf{F}_1^\pm, \mathbf{Q}_1^\pm$ and $\mathbf{F}_2^\pm, \mathbf{Q}_2^\pm$.

Definition 9 (q-dominance)**.** For $q \in [0, 1)$ we say that P_1 dominates P_2 at quantile q, writing $P_1 \succeq_q P_2$, if for every $\mathbf{y} \in \mathbb{C}_{P_2}(q)$ we have that

$$\mathbf{x} := \mathbf{Q}_1^\pm \left(\mathbf{F}_2^\pm(\mathbf{y})\right) \geq \mathbf{y}. \tag{4}$$

When $\mathbf{F}_2^\pm(\mathbf{y}) = 0$ (meaning $\mathbf{y} \in \mathbb{C}_{P_2}(0)$), we interpret $\mathbf{Q}_1^\pm(0) \geq \mathbf{y}$ as "every $\mathbf{x} \in \mathbb{C}_{P_1}(0)$ satisfies $\mathbf{x} \geq \mathbf{y}$."

4.1 Theoretical Properties

Definition 9 yields three immediate consequences, which we collect below as Corollary 1, and is connected to strong FSD, as we will show in Theorem 5.

Corollary 1. *If P_1 dominates P_2 for a quantile $q \in [0, 1)$; $P_1 \succeq_q P_2$, then:*

(i) For any $q' \leq q$ we also have that $P_1 \succeq_{q'} P_2$, since $\mathbb{C}_{P_i}(q') \subseteq \mathbb{C}_{P_i}(q)$.

(ii) For any non-decreasing utility $u : \mathbb{R}^d \to \mathbb{R}$,

$$\mathbb{E}[u(\mathbf{X}) \mid \mathbf{X} \in \mathbb{C}_{P_1}(q)] \geq \mathbb{E}[u(\mathbf{Y}) \mid \mathbf{Y} \in \mathbb{C}_{P_2}(q)].$$

If u is bounded from above and below by constants M and m respectively, then

$$\mathbb{E}[u(\mathbf{X})] - \mathbb{E}[u(\mathbf{Y})] \geq \Delta_q - (1-q)(M-m),$$

where $\Delta_q := \int_{\|v\| \leq q} u(\mathbf{Q}_1^\pm(v)) - u(\mathbf{Q}_2^\pm(v))\, \mathrm{d}v \geq 0$.

(iii) There exists a coupling $(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})$ with marginals P_1, P_2 such that

$$\mathbb{P}(\widehat{\mathbf{X}} \geq \widehat{\mathbf{Y}}) \geq q.$$

The bound is attained by the monotone coupling $\widehat{\mathbf{X}} = \mathbf{Q}_1^\pm(\mathbf{Z})$ and $\widehat{\mathbf{Y}} = \mathbf{Q}_2^\pm(\mathbf{Z})$, with $\mathbf{Z} \sim U_d$.

Theorem 5. *If $P_1 \succeq_q P_2$ for all $q \in [0, 1)$ then P_1 strongly (and thus also weakly) stochastically dominates P_2, i.e. $P_1 \succeq P_2$.*

Proof. By Corollary 1(iii), the map $\mathbf{Z} \mapsto \left(\mathbf{Q}_1^\pm(\mathbf{Z}), \mathbf{Q}_2^\pm(\mathbf{Z})\right)$ provides a coupling with $\mathbb{P}(\mathbf{X} \geq \mathbf{Y}) = 1$, since $P_1 \succeq_q P_2$ for all $q \in [0, 1)$, and Theorem 2 then yields $P_1 \succeq_1 P_2$. $\square$

Thus, q-dominance can be seen as a natural relaxation of strong FSD, where we obtain strong FSD when q-dominance holds for all quantiles.

4.2 Empirical Test

To transform the theory into a finite-sample decision rule, we discretize (4) on the augmented grid defined in Section 3.2. Let $\mathbf{X}^{(n)} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ and $\mathbf{Y}^{(n)} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ be n i.i.d. samples from P_1 and P_2 respectively, with empirical center-outward distribution and quantile functions $\widehat{\mathbf{F}}_1^\pm, \widehat{\mathbf{Q}}_1^\pm$ and $\widehat{\mathbf{F}}_2^\pm, \widehat{\mathbf{Q}}_2^\pm$.

Definition 10. For $q \in \left\{0, \frac{1}{n_R+1}, \ldots, \frac{n_R}{n_R+1}\right\}$, we have that $\mathbf{X}^{(n)} \succeq_q \mathbf{Y}^{(n)}$ if

$$\widehat{\mathbf{Q}}_1^\pm(q'\mathbf{u}) \geq \widehat{\mathbf{Q}}_2^\pm(q'\mathbf{u}),$$

for all $q' \in \left\{0, \frac{1}{n_R+1}, \ldots, \frac{\lfloor q(n_R+1) \rfloor}{n_R+1}\right\}$ and all $\mathbf{u} \in \mathfrak{S}^{(n_S)}$. When $n_0 \geq 2$ the terms $\widehat{\mathbf{Q}}_1^\pm(0)$ and $\widehat{\mathbf{Q}}_2^\pm(0)$ are multi-valued, in this case we interpret the defining inequality as "for every $\mathbf{x} \in \widehat{\mathbf{Q}}_1^\pm(0)$ and $\mathbf{y} \in \widehat{\mathbf{Q}}_2^\pm(0)$, we have that $\mathbf{x} \geq \mathbf{y}$."

In the following theorem, we prove that, once the sample size is large enough, theoretical q-dominance of the distributions $P_1 \succeq_q P_2$ carries over to the empirical maps with high probability.

Theorem 6. *Let distributions P_1 and P_2 have center-outward quantile functions $\mathbf{Q}_1^\pm, \mathbf{Q}_2^\pm$ that are bi-Lipschitz continuous with constants L_1 and L_2. Choose $n_R = n^\theta$ and $n_S = n^{1-\theta}$, with $\theta \in \left(\frac{1}{2d}, \frac{d+1}{2d}\right)$ for $d \leq 4$ and $\theta \in \left(\frac{d-2}{d^2}, \frac{2d-3}{d^2}\right)$ for $d \geq 5$. If we assume that $P_1 \succeq_q P_2$ for every $q \in [0, 1)$, then for every confidence level $0 < \delta < 1$ there exists an explicit threshold $n^*(\delta)$ such that for all $n \geq n^*(\delta)$*

$$\mathbf{X}^{(n)} \succeq_q \mathbf{Y}^{(n)}, \quad \textit{for every } q \in \left\{\frac{j}{n_R+1}\right\}_{j=0}^{n_R},$$

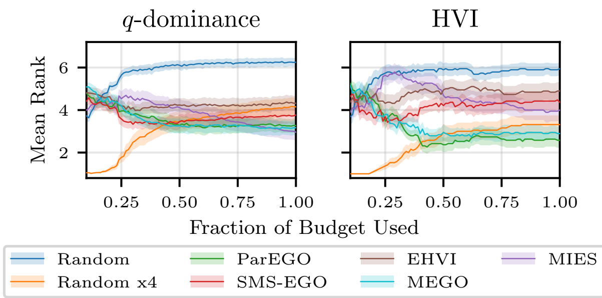
with probability at least $1 - \delta$.

A complete proof, including an explicit formula for $n^*(\delta)$, is provided in Appendix A.

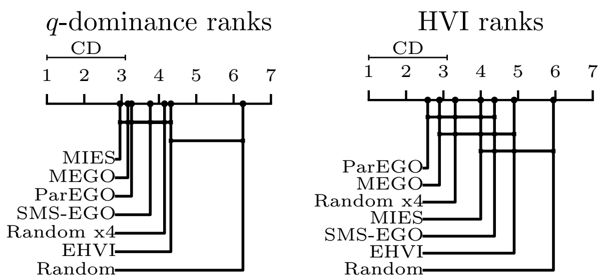
Lastly, we provide a straightforward procedure, outlined in Appendix B (Algorithm 1), that ranks a collection of empirical distributions by iteratively applying center-outward q-dominance on a shared augmented grid. We first build non-dominated fronts for successively smaller q values, and then, within each front, order the remaining distributions by a measure of how close they are to being dominated.

5 Experiments

To demonstrate the usefulness of q-dominance, we conduct two numerical studies: (1) to help improve the robust ranking of optimizers in multi-objective hyperparameter optimization tasks; (2) to assist the selection in multi-objective evolutionary algorithms (MOEAs) under noise.



(a) Mean-rank trajectories (±1 s.d.).



(b) Critical difference diagrams for mean ranks after 100% of budget used.

Figure 2: Results of YAHPO-MO benchmarks based on q-dominance (left) and HVI (right).

5.1 Multi-Objective Hyperparameter Optimization (HPO) Rankings

Pfisterer et al. compare seven optimizers: Random Search, Random Search 4x (Random Search with quadrupled budget), ParEGO (Knowles 2006), SMS-EGO (Ponweiser et al. 2008), EHVI (Emmerich, Giannakoglou, and Naujoks 2006), MEGO (Jeong and Obayashi 2005), and MIES (Li et al. 2013) on 25 different multi-objective hyperparameter optimization problem instances with 2 to 4 objectives. They compare these optimizers by computing the normalized Hypervolume Indicator (HVI) of their found Pareto fronts after specific fractions of their budget have been used. For our approach, we sample $k \leq 5$ points uniformly at random from the same Pareto fronts at each replication, resulting in $30k$ samples from the optimizer's underlying stochastic Pareto set for each problem instance and at each considered fractional budget. With these samples, we then rank the optimizers using our q-dominance relation, by computing their empirical center-outward quantile maps and sorting them using Algorithm 1. We repeat this procedure 100 times to minimize the influence of the random sampling. More complete experimental details can be found in Appendix C.1.

In Figure 2a, we show the mean rank, based on q-dominance (left) and HVI (right), of the different HPO methods as a function of the fraction of budget used. Figure 2b shows the critical difference plot, with $\alpha = 0.05$, after 100% of the budget has been used.

The first difference we note is the spread of the mean

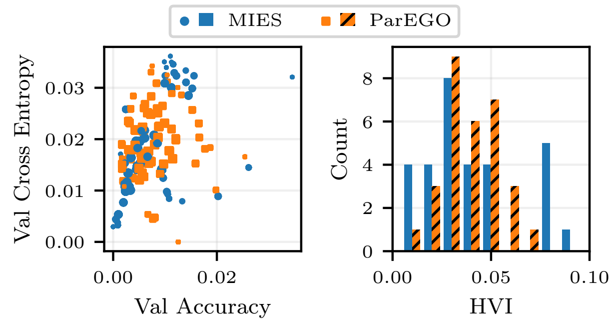


Figure 3: Left: pooled samples from the final stochastic Pareto sets for MIES and ParEGO on *lcbench 167152* (marker size $\propto$ quantile^{-1}). Right: histogram of the HVI values for those same Pareto sets.

ranks between the optimizers after around 50% of the budget has been used. In Figure 2a (left), we see that the mean ranks, based on q-dominance, of all the methods aside from Random Search are bunched quite close together, and in Figure 2b (left), we see that none of these methods significantly outperform the others. Whilst for the mean ranks based on HVI, we have a much larger spread on the right side of Figure 2a and in Figure 2b.

Furthermore, in Figure 2b we see that, based on the q-dominance ranking, only EHVI does not significantly improve on Random Search after the entire budget has been used. Contrastingly, based on the HVI ranking, in addition to EHVI, SMS-EGO and MIES also do not significantly improve on Random Search——a notable difference, specifically for MIES, which on average performs the best according to q-dominance.

To further investigate this difference in rankings, we examine the final Pareto sets of MIES and ParEGO, the optimizers that, on average, perform best based on q-dominance and HVI, respectively, on the problem instance lcbench 167152 (See Table 1 in Appendix C.1). On this problem, MIES is preferred over ParEGO based on q-dominance, with a rank of 1.97 ± 0.26 for MIES versus 3.11 ± 0.60 for ParEGO, yet ParEGO is preferred over MIES based on HVI, where ParEGO has an HVI of 0.0399 ± 0.0025 and MIES has an HVI of 0.0400 ± 0.0045 across the 30 replications. In Figure 3, we show samples from the final Pareto sets of MIES (blue circles) and ParEGO (orange squares) in the left plot, and a histogram of the HVI values of both optimizers over the 30 replications in the right plot. From these figures, we can infer that MIES has a high probability of achieving better results than ParEGO, shown by the blue, circular points of MIES in the bottom left of the left plot in Figure 3 and the left-most bars in the right plot. However, this comes at the cost of slightly higher risk, which we can see from the points of MIES in the upper middle part of the left plot in Figure 3 and the right-most bars in the histogram on the right. By taking the expected value of the HVI, we lose this information, resulting in the HVI ranking favoring ParEGO over MIES, as ParEGO's mean HVI value is slightly lower.

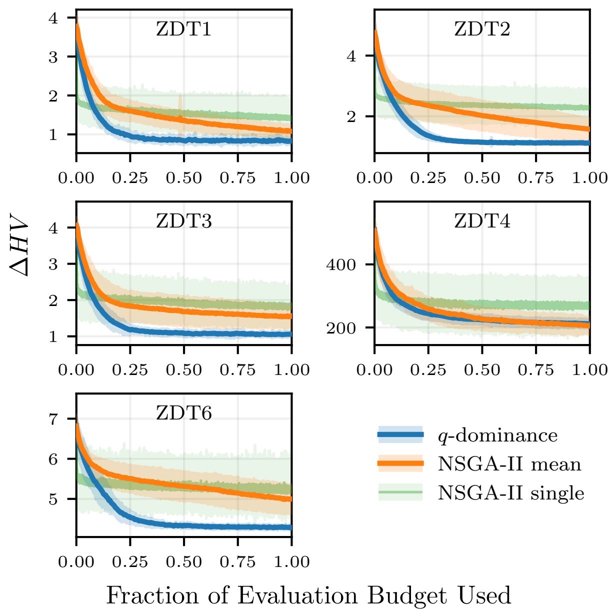


Figure 4: The difference in HV (ΔHV) between the deterministic Pareto front and the expected HV of the solutions at a specific budget used. ZDT5 is excluded as it is a Boolean optimization problem.

5.2 Noise Augmented ZDT

To assess the value of q-dominance when directly applied to SMOOPs, we augment the commonly used ZDT benchmark problems (Zitzler, Deb, and Thiele 2000) with noise in the inputs. Specifically, if f is the objective function, then we consider the optimization problem: $\min_{\mathbf{x}} f(z_1, \ldots, z_k)$, where $z_i \sim \mathcal{N}_{[a_i,b_i]}(x_i, \sigma^2)$. Here a_i, b_i indicate the lower- and upper-bound of the decision variable x_i, such that $x_i \in [a_i, b_i]$, and $\mathcal{N}_{[a_i,b_i]}(x_i, \sigma^2)$ denotes the truncated normal distribution, with mean x_i, variance σ^2, and truncated to lie within the interval $[a_i, b_i]$. With this construction, our objective functions become stochastic, and the underlying deterministic functions are never evaluated outside of their domains.

We consider two simple baselines: NSGA-II with mean-value based selection (denoted by `NSGA-II mean`), i.e., for each candidate solution we evaluate the objective function n times and compute the mean, and NSGA-II with single sample based selection (`NSGA-II single`), where we only evaluate the objective function once (Deb et al. 2002). For our method, we replace the selection with the q-dominance sorting procedure (Algorithm 1), computed on n objective function evaluations per candidate solution. All three methods have identical evaluation budgets and population sizes.

We perform 20 independent runs. Each method has a population size of 20 and proceeds for a total of 200 generations, or $200n$ for `NSGA-II single`. Our q-dominance method and the NSGA-II-mean method use $n = 64$ samples per candidate, where we let $n_R = n_S = 8$ for the center-outward empirical quantile maps. Lastly, for the noise, we take $\sigma = 0.1$. This choice was based on a visual inspection of the stochastic objective functions, with varied σ, evaluated at the deterministic optima. Further details, along with the plots used to determine σ, can be found in Appendix C.2.

In Figure 4 we show the difference in HV between the deterministic Pareto front and the expected HV, including the 95% confidence interval, of the approximated fronts obtained by the different methods. Plots showing the $q \approx 0.55$ center-outward quantile regions of the final Pareto sets are shown in Appendix C.2.

We observe that for all problems, except ZDT4, the q-dominance method converges significantly faster than the two baselines. For ZDT4, we note that none of the solutions are close to the deterministic front in terms of the second objective, f_2, as can be seen in Figure 2 in Appendix C.2. Based on the initial inspection of the stochastic objective functions, we suspect that this is due to the sensitivity of ZDT4 to noise.

6 Conclusion

In this paper, we presented *center-outward q-dominance*, a sample-computable proxy for strong FSD, applicable to stochastic multi-objective optimization problems. Our main theoretical result establishes that q-dominance over the full range of quantiles is *sufficient* for strong FSD:

$$P_1 \succeq_q P_2,\ \forall q \in [0, 1) \implies P_1 \succeq_1 P_2 \text{ (strong FSD).}$$

Because each distribution is mapped once to the common uniform reference on the unit ball, the method provides globally coherent comparisons, scales to many distributions with only one optimal transport map needing to be computed per distribution, and is free of tunable hyperparameters. The same computation yields the entire hierarchy of relaxations, for different quantiles q, at no extra cost, enabling a fast sorting procedure for stochastic multi-objective optimization, and our finite-sample analysis supplies an explicit sample size threshold $n^*(\delta)$ for any desired Type I error.

Our two empirical studies illustrated these properties by demonstrating the potential to enable stable rankings when traditional methods fail, and by providing faster convergence when directly applied to optimization problems. We emphasize that these results are intended to illustrate and validate our theoretical findings, rather than to establish new state-of-the-art performance. Accordingly, we compare only with a small set of widely used stochastic baselines. A thorough benchmarking comparison of q-dominance against a broader range of modern techniques is beyond the scope of this paper and is left for future work.

Furthermore, we plan to compare our construction theoretically to a recent work (Rioux et al. 2025), which is based on the (non-quadratic) optimal transport (OT) between two multivariate distributions, while our method uses the quadratic OT from each distribution to a common reference on the unit ball. In our method, the composition of the OTs $\mathbf{Q}_1^{\pm} \circ \mathbf{F}_2^{\pm}$ is unnecessarily an OT between two distributions. Although both works imply the first-order stochastic dominance, it is unclear whether our q-dominance implies their construction, vice versa, or if they are incompatible.

Acknowledgements

This research is supported by the HyTROS project, funded out of the Dutch Growth Fund Program GroenvermogenNL (Green capacity for the Dutch economy and society) via the NWO call "NGF: Transport en opslag van waterstof Groenvermogen NL - Werkpakket 2". This funding is further complemented with in-kind and cash funding from several of the HyTROS participants.

References

Abdelaziz, F. 2012. Solution approaches for the multiobjective stochastic programming. *European Journal of Operational Research*, 216: 1–16.

Caballero, R.; Cerdá, E.; del Mar Muñoz, M.; and Rey, L. 2004. Stochastic approach versus multiobjective approach for obtaining efficient solutions in stochastic multiobjective programming problems. *European Journal of Operational Research*, 158(3): 633–648.

Daulton, S.; Cakmak, S.; Balandat, M.; Osborne, M. A.; Zhou, E.; and Bakshy, E. 2022. Robust Multi-Objective Bayesian Optimization Under Input Noise. In Chaudhuri, K.; Jegelka, S.; Song, L.; Szepesvári, C.; Niu, G.; and Sabato, S., eds., *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, 4831–4866. PMLR.

Deb, K.; Pratap, A.; Agarwal, S.; and Meyarivan, T. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2): 182–197.

Drugan, M. M.; and Nowé, A. 2013. Designing multi-objective multi-armed bandits algorithms: A study. In *The 2013 International Joint Conference on Neural Networks, IJCNN 2013, Dallas, TX, USA, August 4-9, 2013*, 1–8. IEEE.

Ehrgott, M.; Ide, J.; and Schöbel, A. 2014. Minmax robustness for multi-objective optimization problems. *European Journal of Operational Research*, 239(1): 17–31.

Emmerich, M.; Giannakoglou, K.; and Naujoks, B. 2006. Single- and multiobjective evolutionary optimization assisted by Gaussian random field metamodels. *IEEE Transactions on Evolutionary Computation*, 10(4): 421–439.

Fliege, J.; and Xu, H. 2011. Stochastic Multiobjective Optimization: Sample Average Approximation and Applications. *Journal of Optimization Theory and Applications*, 151(1): 135–162.

Hallin, M. 2017. On Distribution and Quantile Functions, Ranks and Signs in R_d. Working Papers ECARES ECARES 2017-34, ULB – Universite Libre de Bruxelles.

Hallin, M.; del Barrio, E.; Cuesta-Albertos, J.; and Matrán, C. 2021. Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach. *The Annals of Statistics*, 49(2): 1139 – 1165.

Jeong, S.; and Obayashi, S. 2005. Efficient global optimization (EGO) for multi-objective problem and data mining. In *2005 IEEE Congress on Evolutionary Computation*, volume 3, 2138–2145 Vol. 3.

Jin, Y.; and Branke, J. 2005. Evolutionary optimization in uncertain environments-a survey. *IEEE Transactions on Evolutionary Computation*, 9(3): 303–317.

Knowles, J. 2006. ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1): 50–66.

Kopa, M.; and Petrová, B. 2018. Strong and Weak Multivariate First-Order Stochastic Dominance. *Mathematics eJournal*.

Li, R.; Emmerich, M. T.; Eggermont, J.; Bäck, T.; Schütz, M.; Dijkstra, J.; and Reiber, J. 2013. Mixed Integer Evolution Strategies for Parameter Optimization. *Evolutionary Computation*, 21(1): 29–64.

Lu, S.; Wang, G.; Hu, Y.; and Zhang, L. 2019. Multi-Objective Generalized Linear Bandits. In Kraus, S., ed., *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, 3080–3086. ijcai.org.

McCann, R. J. 1995. Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal*, 80(2): 309 – 323.

Pfisterer, F.; Schneider, L.; Moosbauer, J.; Binder, M.; and Bischl, B. 2022. YAHPO Gym - An Efficient Multi-Objective Multi-Fidelity Benchmark for Hyperparameter Optimization. In Guyon, I.; Lindauer, M.; van der Schaar, M.; Hutter, F.; and Garnett, R., eds., *Proceedings of the First International Conference on Automated Machine Learning*, volume 188 of *Proceedings of Machine Learning Research*, 3/1–39. PMLR.

Ponweiser, W.; Wagner, T.; Biermann, D.; and Vincze, M. 2008. Multiobjective Optimization on a Limited Budget of Evaluations Using Model-Assisted $\mathcal{S}$-Metric Selection. In Rudolph, G.; Jansen, T.; Beume, N.; Lucas, S.; and Poloni, C., eds., *Parallel Problem Solving from Nature – PPSN X*, 784–794. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-540-87700-4.

Rioux, G.; Nitsure, A.; Rigotti, M.; Greenewald, K.; and Mroueh, Y. 2025. Multivariate stochastic dominance via optimal transport and applications to models benchmarking. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24. Red Hook, NY, USA: Curran Associates Inc. ISBN 9798331314385.

Shaked, M.; and Shanthikumar, J. G., eds. 2007. *Multivariate Stochastic Orders*, 265–322. New York, NY: Springer New York. ISBN 978-0-387-34675-5.

Sriboonchita, S.; Wong, W.-K.; Dhompongsa, S.; and Nguyen, H. T. 2009. *Stochastic Dominance and Applications to Finance, Risk and Economics*. Chapman and Hall/CRC. ISBN 9781420082678.

Teich, J. 2001. Pareto-Front Exploration with Uncertain Objectives. In Zitzler, E.; Thiele, L.; Deb, K.; Coello Coello, C. A.; and Corne, D., eds., *Evolutionary Multi-Criterion Optimization*, 314–328. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-540-44719-1.

Trappler, V.; Helbert, C.; and Riche, R. L. 2025. Multiobjective Optimization under Uncertainties using Conditional Pareto Fronts. *arXiv preprint arXiv:2504.04944*.

Xu, M.; and Klabjan, D. 2023. Pareto Regret Analyses in Multi-objective Multi-armed Bandit. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, 38499–38517. PMLR.

Zitzler, E.; Deb, K.; and Thiele, L. 2000. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evolutionary Computation*, 8(2): 173–195.