MULTI-OBJECTIVE PROBABILISTIC PREFERENCE LEARNING WITH SOFT AND HARD BOUNDS

Edward Chen, Sang T. Truong, Natalie Dullerud, Sanmi Koyejo & Carlos Guestrin
Department of Computer Science
Stanford University
{edjchen,sttruong,ndulleru,sanmi,guestrin}@cs.stanford.edu

ABSTRACT

Complex decision-making tasks across healthcare, drug discovery, engineering design, and deep learning frequently involve optimizing multiple competing objectives while navigating computationally expensive evaluations. Decision-makers (DMs) must select Pareto-optimal solutions that align with their implicit preferences, yet the computational burden of evaluating solutions and the cognitive load of validating tradeoffs make exhaustive Pareto frontier exploration infeasible. While DMs often possess domain knowledge that helps constrain the initial search space – for instance, clinicians typically have priors for specific regions of the tradeoff surface to explore – existing methods lack systematic approaches for iteratively refining these regions of interest. Critically, in high-stakes domains like healthcare, DMs must develop confidence that they have not overlooked superior solutions before committing to a final decision. We present Active-MoSH, a novel interactive framework that formalizes this exploration process by integrating soft-hard utility functions with probabilistic preference learning. Our framework maintains distributions over both the DM's latent preference vector and feedback, by way of soft and hard bounds, enabling adaptive refinement of the explored Pareto frontier subset. We develop an active sampling strategy that optimizes the exploration-exploitation tradeoff while minimizing cognitive burden. To address the fundamental need for solution validation, we propose T-MoSH, which leverages local multi-objective sensitivity analysis to systematically build DM trust by quantifying the robustness of solutions and identifying potentially overlooked regions of the Pareto frontier. Through extensive experiments on synthetic benchmarks and real-world applications in engineering design and cervical cancer brachytherapy treatment planning, we demonstrate that our framework efficiently guides DMs toward optimal tradeoff points while providing rigorous validation of solution quality.

1 INTRODUCTION

Countless critical decision-making tasks within the domains of healthcare, drug discovery, engineering design, and deep learning involve determining the optimal tradeoff point among multiple competing, and often computationally expensive, objectives $f_1(x), ..., f_L(x)$ (Fromer & Coley, 2023; Luukkonen et al., 2023; Xie et al., 2021; Yu et al., 2000; Papadimitriou & Yannakakis, 2001). In general, such decision-makers (DMs) typically aim to select a Pareto-optimal point — one that possesses an ideal set of tradeoffs — that matches their hidden, internal set of preferences among the multiple objectives. However, having the DM manually search through the entire set of tradeoff points along the Pareto frontier is infeasible due to the often computationally expensive objectives, lengthy DM validation of each point, and, in general, limited amounts of human attention. As a result, DMs typically navigate the Pareto frontier through an interactive process involving the validation of points along different subsets of the Pareto frontier before selecting the point matching their ideal internal set of objectives Paria et al. (2019); Ozaki et al. (2023); Wilde et al. (2021).

Within the healthcare domain, clinicians must determine patient-specific treatment tradeoffs that maximize effectiveness while minimizing side effects. Brachytherapy, internal radiation therapy for cancer treatment, exemplifies this challenge. Clinicians must devise treatment plans that maximize radiation damage to cancer tumors while minimizing exposure to nearby healthy organs. However, the time-consuming nature of treatment planning and the vast space of radiation dosage tradeoffs make it infeasible for clinicians to navigate and validate each potential plan, reducing the time available for direct patient care.

Oftentimes, however, based on prior clinical experience, the clinician has a general idea of initial radiation dosage tradeoff regions that they would like to explore. This typically is in the form of initial bounds for the different objectives, e.g. need to cover at least 90% of the cancer tumor, but ideally more than 95%, and preferably less than 513 centigrays (cGY) of radiation to the healthy bladder, but no more than 601 cGY (Chen et al., 2024). This initiates an iterative, interactive process where clinicians refine their preferred region of the Pareto frontier based on observed treatment plans until selecting an ideal solution. Building trust in this process is crucial; decision-makers must be confident that they have not overlooked potentially superior Pareto-optimal points. Despite significant advances in multi-objective preference learning, methods that explicitly support such interactive refinement and validation within Pareto frontier subsets remain underexplored.

This paper operationalizes the above intuitions and formalizes the Pareto frontier exploration process through a novel interactive framework that iteratively learns a DM's optimal tradeoff point via systematic feedback on specific regions of interest. We extend the soft-hard utility function framework of Chen et al. (2024), originally developed for static Pareto frontier exploration, to a dynamic preference learning setting. Our proposed method, Active-MoSH, introduces a probabilistic framework that integrates these utility functions with interactive preference learning. The framework maintains and updates probabilistic distributions over both the DM's latent preference vector and their soft-hard bounds through Plackett-Luce and Gaussian likelihood functions (Luce, 1959). This probabilistic approach enables adaptive refinement of the explored Pareto frontier subset, concentrating computational resources on regions that align with the DM's evolving preferences. Building on these updated preference distributions, we develop an active sampling strategy that optimizes the exploration-exploitation tradeoff while minimizing both computational cost and cognitive burden on the DM. To address the critical issue of solution trustworthiness in complex multi-objective optimization, we further propose T-MoSH, which leverages local multi-objective sensitivity analysis to help DMs develop trust in their selected solutions through a rigorous validation process.

Although there exists an extensive line of work on interactive multi-objective decision making frameworks, most works focus on feedback mechanisms that may lack the expressiveness to intuitively explore certain sub-regions of the tradeoff surface (Wilde et al., 2021; Bıyık et al., 2019; Astudillo & Frazier, 2020; Ozaki et al., 2023). Several recent works allows for the DM to impose priors on specific subsets of the Pareto frontier, however, they lack focus on the interactive feedback setting, which we believe poses several additional challenges such as building DM trust in the framework (Paria et al., 2019; Malkomes et al., 2021). In summary, our contributions comprise the following:

1. We propose a novel interactive framework using multi-objective soft and hard bounds, allowing for the DM to intuitively explore and continually refine different regions of the Pareto frontier for exploration.
2. We introduce multi-objective probabilistic learning and active sampling steps for more efficient refinement of the DM's hidden preferences along with T-MoSH, a method for helping build DM trust in such interactive multi-objective decision-making scenarios.
3. We conduct experiments on both synthetic functions and real-world problems in engineering design and healthcare, specifically, cervical cancer treatment planning, demonstrating the usefulness of our framework. Our evaluation setup takes into consideration the complexity of various feedback mechanisms, allowing us to illustrate the feedback efficiency of Active-MoSH, compared to baselines.

2 Multi-Objective Preference Learning with Soft-Hard Bounds

2.1 Background

Multi-objective optimization (MOO) involves the optimization of multiple objective functions, and can be formulated as the joint maximization of L objective functions over some input space $X \subset \mathbb{R}^d$,

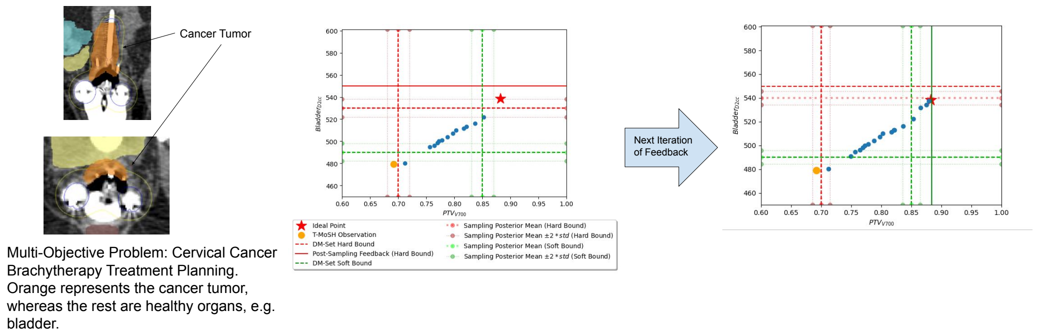

Figure 1: Example iteration of feedback with our proposed interactive framework, Active-MoSH. We evaluated Active-MoSH on a real clinical case for cervical cancer brachytherapy treatment planning, where the objectives are to balance between the radiation dosage levels to the cancer tumor and to the nearby healthy organs – bladder, rectum, and bowel. For the plots, we only showed two of the four dimensions in the multi-criterion objective for this task. After observing the sampled points in the first iteration, the decision-maker relaxes the hard bound for the Bladder$_{D2cc}$ dimension, which leads the decision-maker to being closer to their ideal treatment plan in the next iteration.

$$\max_{x \in X}(f_1(x), ..., f_L(x))$$

in which each f_ℓ, $\ell \in [L]$, defines a noisy and expensive-to-query black-box function $f_\ell : X \rightarrow Y_\ell \subset \mathbb{R}$. As there does not typically exist a feasible solution that marginally optimizes each objective function simultaneously, we focus on points on the Pareto frontier. MOO often uses scalarization to combine the L objectives into a single scalar value, enabling the use of standard optimization techniques. These scalarization functions, such as linear combinations of objectives, are of the form $s_{\boldsymbol{\lambda}}(y) : \mathbb{R}^L \rightarrow \mathbb{R}$ and are parameterized by $\boldsymbol{\lambda}$ from some set Λ in L-dimensional space Roijers et al. (2013). The parameters $\boldsymbol{\lambda} \in \Lambda$ define the relative preferences among the objectives. Optimizing this scalarized objective, for a given set of $\boldsymbol{\lambda} \in \Lambda$, results in a solution along the Pareto frontier.

The DM behind countless practical applications of MOO often has priors on the ideal regions of the multi-dimensional objective space, in the form of soft and hard bounds Chen et al. (2024). We follow the notation described in Chen et al. (2024), where $f_\ell(x) \geq \alpha_{\ell,S}$ describes the ideal region for objective ℓ according to its respective soft bound, $\alpha_{\ell,S}$. Given the competing nature of MOO problems, $\alpha_{\ell,S}$ may not be reachable, which leads the DMs to also have a hard bound $\alpha_{\ell,H}$ for which $f_\ell(x)$ should not drop below. To operationalize such priors, Chen et al. (2024) proposes *soft-hard* utility functions, or SHFs, which encode the above constraints by mapping each f_ℓ to a bounded utility space. We use $u_{\boldsymbol{\alpha}}(x)$ to denote the SHF utility function according to a set of soft and hard bounds $\{\alpha_{1,S},\alpha_{1,H},...,\alpha_{L,S},\alpha_{L,H}\}$. Additional details are in the Appendix.

Given a selected class of scalarization functions $s_{\boldsymbol{\lambda}}$ parameterized by $\boldsymbol{\lambda} \in \Lambda$ and an SHF utility function as defined above, we obtain points $y \in Y$ on the SHF-defined Pareto frontier by solving

$$\max_{x \in X} s_{\boldsymbol{\lambda}}(u_{\boldsymbol{\alpha}}(x)), \tag{1}$$

where $u_{\boldsymbol{\alpha}} := [u_{\alpha_1}, ..., u_{\alpha_L}]$, u_{α_ℓ} subsumes objective function f_ℓ and corresponds to the SHF parameterized by $\alpha_{\ell,S}$ and $\alpha_{\ell,H}$, and we assume that $s_{\boldsymbol{\lambda}}(u_{\boldsymbol{\alpha}}(x))$ is monotonically increasing in all coordinates $u_{\boldsymbol{\alpha}}(x)$. We propose to extend the conceptual framework of SHFs into an interactive preference learning setting and formally describe the setting next.

2.2 Problem Definition

We now introduce the notation we use in this paper and formulate the problem.

Problem Setting As DM evaluation of each point $\boldsymbol{y} \in Y$ is expensive, it is infeasible to observe the full Y on the entire Pareto frontier. We therefore assume the DM has an ideal hidden set of soft and hard bounds which best characterizes the set of tradeoffs most appropriate for their decision-making task. We denote this ideal hidden set as $\{\alpha^*_{1,H}, \alpha^*_{1,S}, ..., \alpha^*_{L,H}, \alpha^*_{L,S}\}$, which we abbreviate to $\{\boldsymbol{\alpha}^*_S, \boldsymbol{\alpha}^*_H\}$. We then assume $\exists\, x^* \in X$ such that $\boldsymbol{y}^* = f(x^*)$, where $\boldsymbol{y}^*$ is the DM's ideal hidden point on the Pareto frontier defined by the SHF, from $\{\boldsymbol{\alpha}^*_S, \boldsymbol{\alpha}^*_H\}$. Accordingly, $\boldsymbol{y}^*$ is a point which trades off among the L objectives according to the DM's set of hidden preferences, $\boldsymbol{\lambda}^*$. Formally,

$$x^* = \arg\max_{x \in X} s_{\boldsymbol{\lambda}^*}(u_{\boldsymbol{\alpha}^*}(x)) \text{ and } \forall \boldsymbol{y} \in Y, \boldsymbol{y}^* \not\succcurlyeq \boldsymbol{y}$$

where $y = f(x)$. *Note: In the naive case, if $\boldsymbol{\alpha}^*_S = \boldsymbol{\alpha}^*_H = 0$, the SHF-defined Pareto frontier is just the original, overall Pareto frontier. In this case, discovering the optimal tradeoff point y^*, which trades off among the L objectives according to λ^*, corresponds to more traditional multi-objective preference learning settings Roijers et al. (2013); Astudillo & Frazier (2020)*

Feedback Mechanism We assume that $\boldsymbol{\lambda}$ and $\{\alpha_{1,S}, \alpha_{1,H}, ..., \alpha_{L,S}, \alpha_{L,H}\}$, which we abbreviate to $\{\boldsymbol{\alpha}_S, \boldsymbol{\alpha}_H\}$, are random variables. We aim to learn $\boldsymbol{\lambda}^*$ and $\{\boldsymbol{\alpha}_S, \boldsymbol{\alpha}_H\}$ through M iterations of interactions from the DM by way of feedback in the form of soft and hard bounds, which we denote with $\alpha_{\ell,S,m}$, $\alpha_{\ell,H,m}$, respectively, for objective ℓ and iteration m Chen et al. (2024). As exact numerical feedback is often imprecise, we express DM uncertainty over their provided soft and hard bounds by imposing the following probabilistic notion: $\boldsymbol{\alpha}_{S,m} \sim N(\mu_{S,m}, \sigma^2_{S,m})$, $\boldsymbol{\alpha}_{H,m} \sim N(\mu_{H,m}, \sigma^2_{H,m})$ to denote the set of soft and hard bounds at iteration m, respectively, where the index for the objective is removed for clarity.

In iteration $m = 0$, the DM inputs their prior distribution for $\{\boldsymbol{\alpha}_S, \boldsymbol{\alpha}_H\}$, p($\boldsymbol{\alpha}_S$, $\boldsymbol{\alpha}_H$), by providing an initial set of soft and hard bounds, $\{\boldsymbol{\alpha}_{S,0}, \boldsymbol{\alpha}_{H,0}\}$. In iteration m, the DM is presented with a set of K points, $Y_m = \{y_1, ..., y_K\}$, where $y_k = f(x_k)$ for $x_k \in X_m$, which we define as the preference query. We let $\mathcal{D}_m$ = $\{(\boldsymbol{\alpha}_{S,0}, \boldsymbol{\alpha}_{H,0}, Y_0, ..., \boldsymbol{\alpha}_{S,m}, \boldsymbol{\alpha}_{H,m}, Y_m, \boldsymbol{\alpha}_{S,m+1}, \boldsymbol{\alpha}_{H,m+1})\}$ be the obtained set of user feedback data at the end of iteration m, where $\boldsymbol{\alpha}_{S,m+1}$ and $\boldsymbol{\alpha}_{H,m+1}$ is the DM feedback on Y_m. The DM returns feedback in the form of soft and hard bounds. We assume that at each iteration m, the DM can modify or keep the soft and hard bounds from the previous iteration, updating the posterior distributions p($\boldsymbol{\alpha}_S$, $\boldsymbol{\alpha}_H \mid \mathcal{D}_m$) and p($\boldsymbol{\lambda} | \mathcal{D}_m$). Due to potentially complex relationships across the multiple objectives overwhelming the DM, we assume a single modification of any of the soft and hard bounds for all objectives to encompass the full iteration. Since we want to eventually select a Pareto-optimal point that optimally trades off among the L objectives according to $\boldsymbol{\lambda}^*$, at the end of m iterations of feedback, our objective is to maximize the SHF utility ratio (Chen et al., 2024):

$$\frac{\max_{x \in X_m} s_{\boldsymbol{\lambda}^*}(u_{\hat{\boldsymbol{\alpha}}}(x))}{\max_{x \in X} s_{\boldsymbol{\lambda}^*}(u_{\boldsymbol{\alpha}^*}(x))} \tag{2}$$

where $\hat{\boldsymbol{\alpha}}$ is sampled from the posterior distribution p($\boldsymbol{\alpha}_S$, $\boldsymbol{\alpha}_H \mid \boldsymbol{\alpha}_{S,m}$, $\boldsymbol{\alpha}_{H,m}$). Intuitively, the SHF utility ratio is maximized when the points in $f(X_m)$ are Pareto optimal and span the high utility regions of the PF, as defined by the SHFs $u_{\boldsymbol{\alpha}^*}$. We use Equation (2) to evaluate our results in Section 4.

3 Active-MoSH: Probabilistic Modeling and Sampling

3.1 Probabilistic Modeling For Preferences and Soft-Hard Bounds

We aim to learn $\boldsymbol{\lambda}^*$ by way of providing the DM with the preference query and obtaining feedback in the form of the soft-hard bounds. To inform our preference query sampling process, we maintain and update a posterior distribution over the preference vectors $\boldsymbol{\lambda}$ that captures our uncertainty about the DM's true preferences. Our goal is to efficiently update this distribution through interactive feedback while accounting for the human's ability to provide reliable responses. At iteration m, given a preference query Y_m and DM feedback, we update our belief over $\boldsymbol{\lambda}$ using Bayes' rule:

$$p(\boldsymbol{\lambda}|\mathcal{D}_m) \propto p(\boldsymbol{\lambda}) \prod_{i=0}^{m} p(Y_i, \boldsymbol{\alpha}_{S,i}, \boldsymbol{\alpha}_{H,i}, \boldsymbol{\alpha}_{S,i+1}, \boldsymbol{\alpha}_{H,i+1}|\boldsymbol{\lambda}) \tag{3}$$

where $p(Y_i, \boldsymbol{\alpha}_{S,i}, \boldsymbol{\alpha}_{H,i}, \boldsymbol{\alpha}_{S,i+1}, \boldsymbol{\alpha}_{H,i+1}|\boldsymbol{\lambda})$ is the likelihood of the DM providing feedback in the form of $\boldsymbol{\alpha}_{S,i+1}, \boldsymbol{\alpha}_{H,i+1}$ given preference query Y_m, current soft and hard bounds $\boldsymbol{\alpha}_{S,i}, \boldsymbol{\alpha}_{H,i}$, and preference vectors $\boldsymbol{\lambda}$, $p(\boldsymbol{\lambda})$ represents our prior beliefs about the preference vectors, and the last equality assumes conditional independence. This formulation allows us to maintain uncertainty over $\boldsymbol{\lambda}$ while incorporating both soft and hard bound feedback through the likelihood function.

As described in Section 2.2, we assume uncertainty in the DM feedback in the form of imprecise modification of the soft and hard bounds. As a result, we maintain and update a posterior distribution over $\boldsymbol{\alpha}_S$ and $\boldsymbol{\alpha}_H$ which captures that uncertainty. Similarly, we update our belief over $\boldsymbol{\alpha}_S, \boldsymbol{\alpha}_H$ using Bayes' rule:

$$p(\boldsymbol{\alpha}_H|\mathcal{D}_m) \propto p(\boldsymbol{\alpha}_{H,0}) \prod_{i=0}^{m} p(\boldsymbol{\alpha}_{H,i+1}|\boldsymbol{\alpha}_{H,i}) \tag{4}$$

We only represent the belief for $\boldsymbol{\alpha}_H$ for clarity. We assume a Gaussian likelihood model for the bound being modified at each iteration. We discuss our formulation for the preference likelihood model next.

3.2 Modeling Soft and Hard Bounds Feedback

Feedback Interpretation We introduce an inductive bias into the sampling process to guide exploration effectively by weighting the posterior distribution of $\boldsymbol{\lambda}$ towards regions near the modified soft and hard bounds. This is motivated by the assumption that, in our interactive setup, the DM is most interested in evaluating tradeoffs in the objective space near the regions that have been explicitly selected – hence also being more actionable. As a result, exploration becomes more informed and focused compared to what occurs in settings that adopt a uniform distribution. For clarity, we highlight the possible scenarios for a single objective dimension for iteration m below (we remove the objective index for clarity):

- $\boldsymbol{\alpha}_{H,m} > \boldsymbol{\alpha}_{H,m-1}$. This scenario describes the DM providing feedback by sliding the hard bound higher, depicted in Figure 2 (middle).
- $\boldsymbol{\alpha}_{H,m} < \boldsymbol{\alpha}_{H,m-1}$. The DM moves the hard bound lower (see Figure 2 (left)), indicating that the previous constraint was too strict and permitting exploration of solutions with lower values.
- $\boldsymbol{\alpha}_{S,m} \neq \boldsymbol{\alpha}_{S,m-1}$. The DM adjusts the soft bound (see Figure 2 (right)), signaling a refined target and refocusing exploration on solutions closer to the updated preferred region.
- $\boldsymbol{\alpha}_{S,m} = \boldsymbol{\alpha}_{S,m-1}, \boldsymbol{\alpha}_{H,m} = \boldsymbol{\alpha}_{H,m-1}$. Higher certainty is placed on the unmodified bounds.

Likelihood Model Given the DM feedback interpretations listed above, we assume they each induce a complete ranking of the points in the preference query Y_m, where the ranking is determined by proximity to the soft or hard bound being modified – hence being more actionable. As a result, we leverage the Plackett-Luce likelihood (Luce, 1959) to learn the preference vector $\boldsymbol{\lambda}$ at iteration m as:

$$p(\pi \mid \boldsymbol{\lambda}) = \prod_{i=1}^{K-1} Z_j^{-1} \exp\Big(s_{\boldsymbol{\lambda}}(u_{\boldsymbol{\alpha}_{*,m+1}}(x_{\pi(i)}))\Big), Z_j = \sum_{j=i}^{K} \exp\Big(s_{\boldsymbol{\lambda}}(u_{\boldsymbol{\alpha}_{*,m+1}}(x_{\pi(j)}))\Big) \tag{5}$$

where $\pi = \big[\pi(1), \pi(2), \ldots, \pi(K)\big]$ denotes the ranking order of the K points in the current preference query Y_m as provided by the DM, $f(x)_{\pi(i)}$ is the point in Y_m ranked at position i, $u_{\boldsymbol{\alpha}_{*,m+1}}$ is the SHF utility function associated with $\boldsymbol{\alpha}_{S,m+1}$ and $\boldsymbol{\alpha}_{H,m+1}$, and $\boldsymbol{\lambda}$ is the latent preference vector. As a result of our interpretation, we use Equation 5 as the likelihood model represented in Equation 3. We provide additional details in Appendix A.2.1.

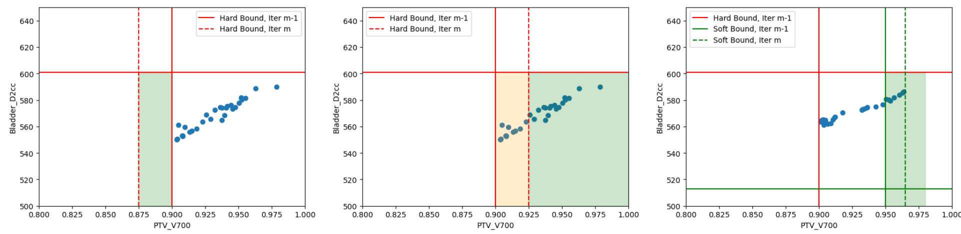

Figure 2: Interpretations of soft and hard bound feedback using brachytherapy as an example. **Left**: Hard bound for PTV$_{V700}$ being moved lower. The green region indicates the preferred region that the DM would like to inspect more closely. **Middle**: Hard bound for PTV$_{V700}$ being moved higher. The green region indicates the preferred region the DM would like to inspect more closely. By placing the updated hard bound at 0.925, the DM is indicating that they desire to observe more samples with PTV$_{V700} > 0.925$, but preferably with a lower tradeoff in the other dimension, Bladder$_{D2cc}$. As a result, ideally additional points closer to the updated hard bound can be obtained. The orange region indicates the region of points that are rejected or dominated with respect to the DM's internal utility function. **Right**: Soft bound for PTV$_{V700}$ being moved higher. The green region indicates the preferred region that the DM would like to inspect more closely. Since the soft bound for PTV$_{V700}$ is the one being shifted forward, the DM would like to prioritize that objective moving forward.

3.3 Active Online Sampling For Preference Queries

This section describes how we obtain the preference queries at each iteration m. Ideally, given that the DM provided feedback in terms of $\boldsymbol{\alpha}_{S,m}$ and $\boldsymbol{\alpha}_{H,m}$, we would like to query the DM with a set of Pareto-optimal points which is easily navigable, and, thus, reducing the cognitive load, and driven by their preferences. As is typical in various science and engineering applications, we assume access to some noisy and expensive black-box function – often modeled with a Gaussian process (GP) (Williams & Rasmussen, 1995) as the surrogate function. As there is uncertainty in the DM's preferences λ, we wish to sample a set of query points that ideally contain the unknown λ^*, given the current set of soft and hard bounds. Since λ^* is unknown to us, we want a preference query set Y_m which is robust to any potential λ^*, weighted according to $p(\boldsymbol{\lambda})$, while also constraining the size of Y_m for the purpose of reducing cognitive load on the DM. As a result, we leverage the formulation for MoSH from (Chen et al., 2024) and actively sample the set Y_m using two steps: (1) obtain a dense set of Pareto-optimal points, Y_{m_D}, weighted according to the current $p(\boldsymbol{\lambda})$, and (2) sparsify Y_{m_D} to ensure it is more easily navigable to the DM, and return as Y_m. We formulate the first step as:

$$\max_{X_{m_D} \subseteq X, |X_{m_D}| \leq k_{\mathcal{D}}} \mathbb{E}_{\boldsymbol{\lambda} \sim p(\boldsymbol{\lambda}|D_m)} \left[\frac{\max_{x \in X_{m_D}} s_{\boldsymbol{\lambda}}(u_{\hat{\boldsymbol{\alpha}}}(x))}{\max_{x \in X} s_{\boldsymbol{\lambda}}(u_{\hat{\boldsymbol{\alpha}}}(x))} \right] \tag{6}$$

where $\hat{\boldsymbol{\alpha}}$ is sampled from the posterior distribution p($\boldsymbol{\alpha}_S$, $\boldsymbol{\alpha}_H \mid \boldsymbol{\alpha}_{S,m}$, $\boldsymbol{\alpha}_{H,m}$) and the term on the right is the SHF utility ratio as described in Section 2.2. As DM validation of each point is often time-consuming, we obtain a sparse set of Y_{m_D} by applying MoSH-Sparse from Chen et al. (2024), and return the result as the preference query for iteration m. Additional details in Appendix A.2.2.

3.4 T-MoSH: Enhancing Decision-Maker Trust With Sensitivity Analysis

We propose T-MoSH, a systematic approach to build DM confidence in multi-objective optimization solutions through sensitivity analysis. A fundamental challenge in preference-based optimization is that DMs often express bounds with inherent uncertainty, leading to questions about solution optimality. Specifically, when DMs identify a seemingly optimal solution, they may lack confidence that small adjustments to their specified bounds wouldn't reveal superior alternatives. To address this challenge, we develop a framework grounded in sensitivity analysis theory that provides qualitative validation of solution quality (Rappaport, 1967).

Overall, we want the DM to be aware of the change of each objective l with respect to some perturbation in another objective l' (i.e. the derivative), to get a sense of how active objective l is with respect to objective l' and whether modifying it will result in a potentially much better point in objective l. We define an objective as *active* if perturbing it results in significant changes in another objective. We still maintain the assumption that evaluating functions f is expensive, so we want to be more conservative and not re-run all of the active sampling step of Active-MoSH, especially since we really only care about *the samples which produce an improvement, and how much of an improvement that is, in the objective f_l*. Let $f_l(x)$ denote the objective function in dimension l, and let x^* be the input sample that obtains the current best solution (in terms of f_l). We define the improvement function in dimension l as

$$I_l(x) \;=\; \max\Big\{\; f_l(x) \;-\; f_l\big(x^*\big),\; 0\Big\}. \tag{7}$$

Intuitively, $I_l(x)$ measures by how much the value of $f_l(x)$ exceeds our current best $f_l(x^*)$, or zero if it does not exceed it at all. We then consider the *expected improvement* $\mathrm{EI}_l(x) = \mathbb{E}\big[I_l(x)\big]$, where the expectation is taken over our posterior model (e.g. a Gaussian process) for $f_l(x)$ (Jones et al., 1998). In practice, $EI_l(x)$ can be computed in closed form. If $EI_l(x) \;>\; 0$, then we consider x to be a worthwhile candidate to evaluate and display $f(x)$ (or to highlight it to the decision-maker). Otherwise, if it is not likely to improve f_l, we do not consider it further. In this case, objective l with respect to objective l', the dimension being perturbed, is not active. We formulate the above problem as:

$$\max_{x\in\mathcal{X}} \mathrm{EI}(x) \quad \text{s.t.} \quad u_{\alpha_\ell}(x) \geq 0 \quad \forall \ell \in [L]/\ell', u_{\alpha_{\ell',H}-\epsilon}(x) \geq 0 \tag{8}$$

This formulation ensures that only samples with positive expected improvement and that satisfy the soft and hard bounds are considered. Consequently, computational resources are focused on promising candidates, effectively guiding exploration in alignment with the decision-maker's priorities. Additional details are provided in Appendix A.2.3.

4 Experimental Results

We hypothesize that our proposed interactive framework for soft and hard bounds improves the efficiency of decision-making tasks with multiple objectives. In this section, we comprehensively evaluate several other common feedback mechanisms in a simulated setup across both synthetic and real-world applications in brachytherapy treatment planning and engineering design. We experiment with both synthetic problems and real-world applications and compare our proposed soft and hard bounds interactive framework to other feedback mechanisms: pairwise feedback, complete ranking (Plackett, 1975; Luce, 1959), partial k-ranking (Guiver & Snelson, 2009), and random (Bıyık et al., 2019). The preference queries for each of the baselines, besides random, were selected based on an information gain-based objective (Bıyık et al., 2019).

4.1 Simulation Setup

We performed experiments by simulating DM interactions using a variety of feedback mechanisms. To ensure rigorous evaluation, we implemented several carefully controlled simulation scenarios detailed below.

Ground-truth Values For simulating the ground-truth preferences and objective value, we randomly sampled $\boldsymbol{\lambda}^* \in \Delta(L)$, where $\Delta(L)$ is the probability simplex in L dimensions, and obtain $\boldsymbol{y}^* = \max_{x\in X} s_{\boldsymbol{\lambda}^*}(f(x))$.

Baseline Feedback Simulations To simulate pairwise, ranking, and partial ranking feedback, we utilized $\boldsymbol{\lambda}^*$ with added Gaussian noise $\epsilon \sim N(0, \sigma^2)$ when determining preferences between alternatives. Specifically, for pairwise comparisons, the hidden noisy utilities were computed as $\upsilon_i = s_{\boldsymbol{\lambda}^*}(f(x_i)) + \epsilon_i$, where $f(x_i)$ represents the objective values for alternative i. The alternative with the higher noisy utility value was selected as preferred. Ranking and partial ranking-based feedback follow similarly.

Active-MoSH Feedback Simulations For Active-MoSH feedback, we assumed the DM interaction to be in response to the set of points displayed at iteration m, Y_m. Specifically, we assume the DM to respond to Y_m based on some *reference point* within Y_m. Intuitively, for instance, a DM would observe the values of the points in Y_m at the extreme ends and determine, based on those values, whether some bound should be modified. Details of the Active-MoSH feedback simulations are described in Appendix A.3.1.

4.2 Performance Evaluation

Our primary objective is to assess the interaction efficiency of our proposed method compared to alternative feedback mechanisms. The key metric is the number of DM interactions required to identify the DM's ideal point, $\boldsymbol{y}^*$. To quantify performance at each iteration, we employ Equation 2 with the set of points X_m.

Interaction Units To ensure fair comparison across different feedback mechanisms, we take into consideration the inherent cognitive complexity of providing each of them. To do so, we propose the notion of *interaction units*. As a base case, we assign the pairwise feedback mechanism to be 1 interaction unit. As Active-MoSH requires for the DM to (1) determine the dimension and bound to modify and (2) specify a magnitude and direction of the modification, we assign it to be 2 interaction units. For the ranking-based feedback mechanisms, we assign them to be k interaction units, where k is the number of points displayed at iteration m [1]. This accounting method provides a standardized basis for comparing the efficiency of different preference elicitation approaches, as illustrated in Figure 3.

4.3 Branin-Currin: Synthetic Two-Objective Function

We leverage the Branin-Currin synthetic two-objective optimization problem provided in the BoTorch framework (Balandat et al., 2020), which has a mapping of $[0, 1]^2 \to \mathbb{R}^2$. The results are shown in Figure 3. As shown, Active-MoSH performs similarly to the other baseline feedback mechanisms in this simplified setting.

4.4 Real Clinical Case: Cervical Cancer Brachytherapy Treatment Planning

We validate our methodology using a real clinical case study of cervical cancer brachytherapy treatment planning, which presents a four-objective optimization problem with three continuous decision variables. The objectives comprise maximizing therapeutic radiation dose to the (1) target tumor volume while simultaneously minimizing radiation exposure to three critical organs at risk (OARs): (2) bladder, (3) rectum, and (4) bowel. For computational consistency, we reformulate the minimization objectives for the OARs as maximization problems through appropriate transformations. The decision variables serve as inputs to an epsilon-constraint optimization program (Deufel et al., 2020), which enables systematic exploration of the treatment planning trade-off space. The results are shown in Figure 3.

4.5 Four Bar Truss: Engineering Design

We further evaluate our approach on a multi-objective optimization problem in structural engineering: the four-bar truss design problem from REPROBLEM (Tanabe & Ishibuchi, 2020). This system presents a bi-objective optimization task with four continuous design variables and exhibits a convex Pareto frontier (CHENG & LI, 1999). The objectives involve the simultaneous minimization of structural volume and joint displacement, while the decision variables parameterize the lengths of the individual truss members. The results are shown in Figure 3. As shown, Active-MoSH performs similarly to the other feedback mechanisms. We leave for future work evaluating on human subjects for more nuanced comparisons.

[1] $k \log k$ is also a valid interpretation for the number of interaction units for the ranking-based feedback mechanisms, however use k for several reasons: (1) the cognitive load of ranking points often is not uniform across points, (2) for simplicity.

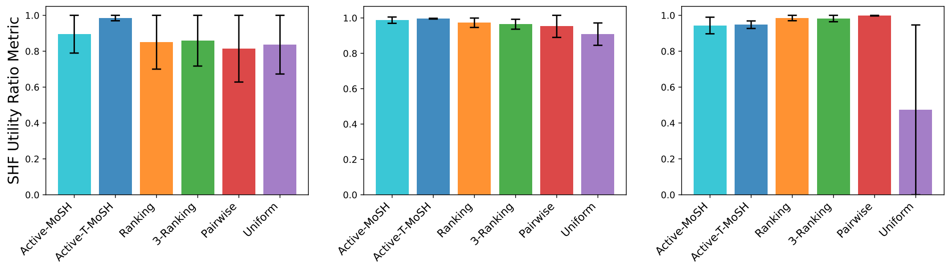

Figure 3: Evaluation results with 10 interaction units for the Branin-Currin synthetic function, Four Bar Truss engineering design, and Brachytherapy Treatment Planning applications (from left to right). The results were all computed over 5 independent trials.

5 Related Works

Interactive Feedback Mechanisms Many feedback mechanisms for interactive settings have been proposed, most of which use pairwise comparisons Chu & Ghahramani (2005); Ozaki et al. (2023); Giovanelli et al. (2024). Wilde et al. (2021) proposed learning reward functions with scale feedback, allowing for slightly more feedback signal. Benavoli et al. (2023) developed a framework for choice functions, which allows the DM to select multiple objects from a given set. When interacting with multiple humans at once, Myers et al. (2021) proposed to learn multimodal rewards. Within MOO, Ozaki et al. (2023) introduces the notion of improvement requests for certain objectives. In contrast, we propose an interactive framework that directly builds soft and hard bound feedback into a unified probabilistic model, allowing for more nuanced, multi-dimensional preferences.

Active Sampling and Learning Methods Several active query sampling methods have been proposed (Ozaki et al., 2023). Other works have also explored different characteristics, such as generating easy-to-answer queries (Bıyık et al., 2019). Additionally, approaches that leverage uncertainty-based or diversity-driven strategies to guide query selection have been introduced (Hernández-Lobato et al., 2015). In contrast, our method actively samples an actionable subset of the Pareto frontier while taking into consideration the high cost of DM validation for each additional point.

Pareto Frontier Population Mechanisms Many existing works in MOO focus on populating the entire Pareto frontier (Campigotto et al., 2014; Ponweiser et al., 2008; Emmerich, 2008; Picheny, 2015; Hernández-Lobato et al., 2015; Zhang et al., 2009). Several works use random scalarizations to attempt to recover the entire Pareto frontier (Knowles, 2006; Paria et al., 2019). Other works have also focused on obtaining sparse coverage of the Pareto frontier, perhaps as determined by level sets (Zuluaga et al., 2016; Malkomes et al., 2021). Rather than populating the entire frontier, we use soft and hard bounds to actively direct exploration toward high-utility areas aligned with user preferences, thereby reducing computational cost and cognitive load.

6 Conclusion

This paper introduces a novel interactive framework for multi-objective preference learning that integrates soft-hard bound feedback with probabilistic inference to efficiently identify optimal tradeoffs on the Pareto frontier. The framework's active sampling strategy, guided by updated belief distributions over both preference vectors and bounds, focuses computational resources on promising regions of the objective space while managing cognitive load. Additionally, our proposed T-MoSH extension employs local multi-objective sensitivity analysis to systematically build decision-maker trust by quantifying solution sensitivity and identifying potentially overlooked improvements. Finally, we evaluate across synthetic benchmarks and real-world applications, such as brachytherapy treatment planning, and demonstrate the efficiency of our interactive framework. Future work includes conducting a study with human subjects to more comprehensively evaluate Active-MoSH.

REFERENCES

Raul Astudillo and Peter Frazier. Multi-attribute Bayesian optimization with interactive preference learning. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, pp. 4496–4507. PMLR, June 2020. URL https://proceedings.mlr.press/v108/astudillo20a.html. ISSN: 2640-3498.

Maximilian Balandat, Brian Karrer, Daniel R. Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Advances in Neural Information Processing Systems 33*, 2020. URL http://arxiv.org/abs/1910.06403.

Alessio Benavoli, Dario Azzimonti, and Dario Piga. Learning Choice Functions with Gaussian Processes. In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, pp. 141–151. PMLR, July 2023. URL https://proceedings.mlr.press/v216/benavoli23a.html. ISSN: 2640-3498.

Erdem Bıyık, Malayandi Palan, Nicholas C. Landolfi, Dylan P. Losey, and Dorsa Sadigh. Asking Easy Questions: A User-Friendly Approach to Active Reward Learning, October 2019. URL http://arxiv.org/abs/1910.04365. arXiv:1910.04365 [cs].

Paolo Campigotto, Andrea Passerini, and Roberto Battiti. Active Learning of Pareto Fronts. *IEEE Transactions on Neural Networks and Learning Systems*, 25(3):506–519, March 2014. ISSN 2162-2388. doi: 10.1109/TNNLS.2013.2275918. URL https://ieeexplore.ieee.org/document/6606803. Conference Name: IEEE Transactions on Neural Networks and Learning Systems.

Edward Chen, Natalie Dullerud, Thomas Niedermayr, Elizabeth Kidd, Ransalu Senanayake, Pang Wei Koh, Sanmi Koyejo, and Carlos Guestrin. MoSH: Modeling Multi-Objective Tradeoffs with Soft and Hard Bounds, December 2024. URL http://arxiv.org/abs/2412.06154. arXiv:2412.06154 [cs].

F. Y. CHENG and X. S. LI. Generalized Center Method for Multiobjective Engineering Optimization. *Engineering Optimization*, 31(5):641–661, May 1999. ISSN 0305-215X. doi: 10.1080/03052159908941390. URL https://doi.org/10.1080/03052159908941390. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03052159908941390.

Wei Chu and Zoubin Ghahramani. Preference learning with Gaussian processes. In *Proceedings of the 22nd international conference on Machine learning - ICML '05*, pp. 137–144, Bonn, Germany, 2005. ACM Press. ISBN 978-1-59593-180-1. doi: 10.1145/1102351.1102369. URL http://portal.acm.org/citation.cfm?doid=1102351.1102369.

Christopher L. Deufel, Marina A. Epelman, Kalyan S. Pasupathy, Mustafa Y. Sir, Victor W. Wu, and Michael G. Herman. PNaV: A tool for generating a high-dose-rate brachytherapy treatment plan by navigating the Pareto surface guided by the visualization of multidimensional trade-offs. *Brachytherapy*, 19(4):518–531, July 2020. ISSN 15384721. doi: 10.1016/j.brachy.2020.02.013. URL https://linkinghub.elsevier.com/retrieve/pii/S1538472120300325.

Michael Emmerich. The computation of the expected improvement in dominated hypervolume of Pareto front approximations. January 2008.

Jenna C. Fromer and Connor W. Coley. Computer-aided multi-objective optimization in small molecule discovery. *Patterns*, 4(2):100678, February 2023. ISSN 26663899. doi: 10.1016/j.patter.2023.100678. URL https://linkinghub.elsevier.com/retrieve/pii/S2666389923000016.

Joseph Giovanelli, Alexander Tornede, Tanja Tornede, and Marius Lindauer. Interactive Hyperparameter Optimization in Multi-Objective Problems via Preference Learning, January 2024. URL http://arxiv.org/abs/2309.03581. arXiv:2309.03581 [cs].

John Guiver and Edward Snelson. Bayesian inference for Plackett-Luce ranking models. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pp. 377–384, New York, NY, USA, June 2009. Association for Computing Machinery. ISBN 978-1-60558-516-1. doi: 10.1145/1553374.1553423. URL https://dl.acm.org/doi/10.1145/1553374.1553423.

Daniel Hernández-Lobato, José Miguel Hernández-Lobato, Amar Shah, and Ryan P. Adams. Predictive Entropy Search for Multi-objective Bayesian Optimization, November 2015. URL https://arxiv.org/abs/1511.05467v3.

Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4):455–492, December 1998. ISSN 1573-2916. doi: 10.1023/A:1008306431147. URL https://doi.org/10.1023/A:1008306431147.

J. Knowles. ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1): 50–66, February 2006. ISSN 1941-0026. doi: 10.1109/TEVC.2005.851274. URL https://ieeexplore.ieee.org/document/1583627. Conference Name: IEEE Transactions on Evolutionary Computation.

Andreas Krause, H. Brendan McMahan, Carlos Guestrin, and Anupam Gupta. Robust Submodular Observation Selection. *Journal of Machine Learning Research*, 9(93):2761–2801, 2008. ISSN 1533-7928. URL http://jmlr.org/papers/v9/krause08b.html.

R. Duncan Luce. *Individual choice behavior*. Individual choice behavior. John Wiley, Oxford, England, 1959. Pages: xii, 153.

Sohvi Luukkonen, Helle W. van den Maagdenberg, Michael T. M. Emmerich, and Gerard J. P. van Westen. Artificial intelligence in multi-objective drug design. *Current Opinion in Structural Biology*, 79:102537, April 2023. ISSN 0959-440X. doi: 10.1016/j.sbi.2023.102537. URL https://www.sciencedirect.com/science/article/pii/S0959440X23000118.

Gustavo Malkomes, Bolong Cheng, Eric H. Lee, and Mike Mccourt. Beyond the Pareto Efficient Frontier: Constraint Active Search for Multiobjective Experimental Design. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 7423–7434. PMLR, July 2021. URL https://proceedings.mlr.press/v139/malkomes21a.html. ISSN: 2640-3498.

Vivek Myers, Erdem Bıyık, Nima Anari, and Dorsa Sadigh. Learning Multimodal Rewards from Rankings, October 2021. URL http://arxiv.org/abs/2109.12750. arXiv:2109.12750 [cs].

G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions—I. *Mathematical Programming*, 14(1):265–294, December 1978. ISSN 0025-5610, 1436-4646. doi: 10.1007/BF01588971. URL http://link.springer.com/10.1007/BF01588971.

Ryota Ozaki, Kazuki Ishikawa, Youhei Kanzaki, Shinya Suzuki, Shion Takeno, Ichiro Takeuchi, and Masayuki Karasuyama. Multi-Objective Bayesian Optimization with Active Preference Learning, November 2023. URL http://arxiv.org/abs/2311.13460. arXiv:2311.13460 [cs].

Christos H. Papadimitriou and Mihalis Yannakakis. Multiobjective query optimization. In *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, PODS '01, pp. 52–59, New York, NY, USA, May 2001. Association for Computing Machinery. ISBN 978-1-58113-361-5. doi: 10.1145/375551.375560. URL https://dl.acm.org/doi/10.1145/375551.375560.

Biswajit Paria, Kirthevasan Kandasamy, and Barnabás Póczos. A Flexible Framework for Multi-Objective Bayesian Optimization using Random Scalarizations, June 2019. URL http://arxiv.org/abs/1805.12168. arXiv:1805.12168 [cs, stat].

Victor Picheny. Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction. *Statistics and Computing*, 25(6):1265–1280, November 2015. ISSN 1573-1375. doi: 10.1007/s11222-014-9477-x. URL https://doi.org/10.1007/s11222-014-9477-x.

R. L. Plackett. The Analysis of Permutations. *Journal of the Royal Statistical Society Series C*, 24(2):193–202, 1975. URL https://ideas.repec.org//a/bla/jorssc/v24y1975i2p193-202.html. Publisher: Royal Statistical Society.

Wolfgang Ponweiser, Tobias Wagner, Dirk Biermann, and Markus Vincze. Multiobjective Optimization on a Limited Budget of Evaluations Using Model-Assisted $\mathcal{S}$-Metric Selection. In Günter Rudolph, Thomas Jansen, Nicola Beume, Simon Lucas, and Carlo Poloni (eds.), *Parallel Problem Solving from Nature – PPSN X*, pp. 784–794, Berlin, Heidelberg, 2008. Springer. ISBN 978-3-540-87700-4. doi: 10.1007/978-3-540-87700-4_78.

Alfred Rappaport. Sensitivity Analysis in Decision Making. *The Accounting Review*, 42(3):441–456, 1967. ISSN 0001-4826. URL https://www.jstor.org/stable/243710. Publisher: American Accounting Association.

D. M. Roijers, P. Vamplew, S. Whiteson, and R. Dazeley. A Survey of Multi-Objective Sequential Decision-Making. *Journal of Artificial Intelligence Research*, 48:67–113, October 2013. ISSN 1076-9757. doi: 10.1613/jair.3987. URL https://www.jair.org/index.php/jair/article/view/10836.

Ryoji Tanabe and Hisao Ishibuchi. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing*, 89:106078, April 2020. ISSN 1568-4946. doi: 10.1016/j.asoc.2020.106078. URL https://www.sciencedirect.com/science/article/pii/S1568494620300181.

Nils Wilde, Erdem Bıyık, Dorsa Sadigh, and Stephen L. Smith. Learning Reward Functions from Scale Feedback, October 2021. URL http://arxiv.org/abs/2110.00284. arXiv:2110.00284 [cs].

Christopher Williams and Carl Rasmussen. Gaussian Processes for Regression. In *Advances in Neural Information Processing Systems*, volume 8. MIT Press, 1995. URL https://proceedings.neurips.cc/paper_files/paper/1995/hash/7cce53cf90577442771720a370c3c723-Abstract.html.

Yutong Xie, Chence Shi, Hao Zhou, Yuwei Yang, Weinan Zhang, Yong Yu, and Lei Li. MARS: Markov Molecular Sampling for Multi-objective Drug Discovery, March 2021. URL http://arxiv.org/abs/2103.10432. arXiv:2103.10432 [cs, q-bio].

Yan Yu, J. B Zhang, Gang Cheng, M. C Schell, and Paul Okunieff. Multi-objective optimization in radiotherapy: applications to stereotactic radiosurgery and prostate brachytherapy. *Artificial Intelligence in Medicine*, 19(1):39–51, May 2000. ISSN 0933-3657. doi: 10.1016/S0933-3657(99)00049-4. URL https://www.sciencedirect.com/science/article/pii/S0933365799000494.

Qingfu Zhang, Aimin Zhou, Shizheng Zhao, Ponnuthurai Nagaratnam Suganthan, Wudong Liu, and Santosh Tiwari. Multiobjective optimization Test Instances for the CEC 2009 Special Session and Competition. 2009.

Marcela Zuluaga, Andreas Krause, and Markus P{ü}schel. e-PAL: An Active Learning Approach to the Multi-Objective Optimization Problem. *Journal of Machine Learning Research*, 17(104):1–32, 2016. ISSN 1533-7928. URL http://jmlr.org/papers/v17/15-047.html.

A Appendix

A.1 Multi-Objective Preference Learning with Soft-Hard Bounds

A.1.1 Soft-Hard Utility Functions

We follow Chen et al. (2024) and use a similar form for the soft-hard utility functions, as follows:

$$u_\alpha(x) = \begin{cases} 1 + 2\beta \times (\tilde{\alpha}_\tau - \tilde{\alpha}_S) & f(x) \geq \alpha_\tau \\ 1 + 2\beta \times (\tilde{f}(x) - \tilde{\alpha}_S) & \alpha_S < f(x) < \alpha_\tau \\ 1 & f(x) = \alpha_S \\ 2 \ \times \tilde{f}(x) & \alpha_H < f(x) < \alpha_S \\ 0 & f(x) = \alpha_H \\ -\infty & f(x) < \alpha_H \end{cases} \tag{9}$$

where $\tilde{f}(x)$ and $\tilde{\alpha}$ are the soft-hard bound normalized values, α_τ, the saturation point, determines where the utility values begin to saturate, and $\beta \in [0, 1]$ determines the fraction of the original rate of utility, in $[\alpha_H, \alpha_S]$, obtained within $[\alpha_S, \alpha_\tau]$. Normalization, for value z, is performed according to the soft and hard bounds, α_S and α_H, respectively, using: $\tilde{z} = ((z - \alpha_H)/(\alpha_S - \alpha_H)) * 0.5$. In practice, we follow Chen et al. (2024) and determine α_τ to be α_H+$\zeta(\alpha_S - \alpha_H)$, for $\zeta = 2.0$. Additional details may be found in (Chen et al., 2024).

A.2 Active-MoSH: Probabilistic Modeling and Sampling

A.2.1 Modeling Soft and Hard Bounds Feedback

In practice, the ranking of points in Y_m is determined by their Euclidean distance to the bound being modified at the end of iteration m, following the intuition described in Section 3.2.

A.2.2 Active Online Sampling for Preference Queries

We solve the first step of MoSH, Equation (6), using the MoSH-Dense algorithm from Chen et al. (2024). In short, MoSH-Dense uses the notion of random scalarizations to sample a $\boldsymbol{\lambda}$ from its posterior distribution at each iteration, which is then used to compute a multi-objective acquisition function (Paria et al., 2019). The maximizer of the acquisition function is then chosen as the next sample input to be evaluated with the expensive black-box function, resulting in a single Pareto-optimal point. For our experiments, we follow Chen et al. (2024) and use the Upper Confidence Bound (UCB) heuristic. We define acq$(u, \lambda_t, x) = s_{\lambda_t}(u_{\varphi(x)})$ where $\varphi(x) = \mu_t(x) + \sqrt{\beta_t}\sigma_t(x)$ and $\beta_t = \sqrt{0.125 \times \log(2 \times t + 1)}$. For β_t, we followed the optimal suggestion in Paria et al. (2019).

This is continued for a total of T iterations, resulting in a dense set of T Pareto-optimal points. The full algorithm is shown below, and more details may be found in Chen et al. (2024).

Algorithm 1 MoSH-Dense: Dense Pareto Frontier Sampling

```
procedure MoSH-Dense
    Init soft and hard bounds {α_{ℓ,H}, α_{ℓ,S}} ∀ ℓ ∈ [L]
    Initialize D^(0) = ∅
    Initialize GP_ℓ^(0) = GP(0, κ) ∀ ℓ ∈ [L]
    for t = 1 → T do
        Obtain λ_t ∼ p(λ|D_m)
        x_t = arg max_{x∈X} acq(u, λ_t, x)
        Obtain y = f(x_t)
        D^(t) = D^(t−1) ∪ {(x_t, y)}
        GP_ℓ^(t) = post(GP_ℓ^{t−1}|(x_t, y)) ∀ ℓ ∈ [L]
    end for
    Return D^(T)
end procedure
```

To reduce the cognitive load on the DM, we use MoSH-Sparse from Chen et al. (2024) on Y_{m_D} from MoSH-Dense. MoSH-Sparse leverages the notion of submodularity, which encapsulates the concept of diminishing returns in utility for each additional point the DM validates (Chen et al., 2024; Nemhauser et al., 1978; Krause et al., 2008). In doing so, MoSH-Sparse theoretically guarantees for our active sampling of preference queries step to obtain a set of points Y_m from Y_{m_D} which achieves the optimal coverage of $p(\boldsymbol{\lambda}|\mathcal{D}_m)$, albeit with a slightly greater number of points than K. Additional details of MoSH-Sparse may be found in Chen et al. (2024).

A.2.3 T-MoSH: Enhancing Decision-Maker Trust with Sensitivity Analysis

In practice, we found it difficult to display proper objective points to highlight, i.e. within the SHF $\boldsymbol{\alpha}$ and exceeding the current best $f_l(x^*)$ for objective l – especially at earlier iterations. As a result, we solve for Equation 8 multiple times (10) and filter out the obtained objective points which violate the SHF or do not exceed the current best point.

A.3 Experimental Results

A.3.1 Simulation Setup

Throughout our experiments, for the scalarization function we used the augmented Chebyshev scalarization function $s_{\boldsymbol{\lambda}}(y) = -\max_{\ell\in[L]}\{\boldsymbol{\lambda}_\ell|y_\ell - z^*_\ell|\} - \gamma\sum_{\ell=1}^{L}|y_\ell - z^*_\ell|$ where z^*_ℓ is the ideal point for objective ℓ.

For Active-MoSH, we implemented the following simulated behavior:

- When $\boldsymbol{y}^*$ falls outside some of the current bounds $\boldsymbol{\alpha}_{H,m}$, the simulated DM selects the bound associated with the objective ℓ that violates the value of $\boldsymbol{y}^*$ the most and adjusts those bounds to bring $\boldsymbol{y}^*$ closer to being within bounds. The adjustment magnitude is determined by the reference point, which we assume to be the point in Y_m closest to $\boldsymbol{y}^*$.
- When $\boldsymbol{y}^*$ is within all bounds $\boldsymbol{\alpha}_{H,m}$, the DM selects the bound furthest from $\boldsymbol{y}^*$ and, we assume, attempts to narrow the feasible space around $\boldsymbol{y}^*$. For this case, we assume the reference point to be the point in Y_m furthest from $\boldsymbol{y}^*$.
- We assume the DM initially leverages the hard bounds, to ensure that $\boldsymbol{y}^*$ is within the desired region. For some objective ℓ that is selected, if the soft and hard bounds are too close in proximity, i.e. $\alpha_{\ell,S,m} - \alpha_{\ell,H,m} < \beta$, we assume the DM then leverages the soft bounds to *fine tune* the desired region points.

Gaussian noise is added to the observations of points in Y_m at each iteration m. The magnitude of bound adjustments is sampled from a Gaussian distribution centered at the deviation between y^* and the reference objective point, with added noise. When y^* is outside the bounds and our proposed method T-MoSH promotes an improved point, we assume the DM has enhanced confidence and increases the magnitude with which they modify the soft or hard bound for that iteration.

Each simulator maintains consistent noise parameters and sampling procedures to ensure fair comparisons across feedback types. This simulation framework allows us to systematically evaluate how different types of preference feedback and varying levels of DM noise affect the convergence and effectiveness of our proposed approach. The setup enables controlled experiments while capturing realistic aspects of human decision-making behavior such as inconsistency and imprecision in feedback. As there is some degree of subjectivity to the simulation setup, we plan to more rigorously evaluate the efficiency of Active-MoSH through a comprehensive study with human subjects in the future.

A.3.2 Performance Evaluation

For mechanisms requiring multiple interaction units to complete a single feedback instance (e.g., ranking), the utility ratio remains constant across the constituent interaction units until the full feedback is processed. This accounting method provides a standardized basis for comparing the efficiency of different preference elicitation approaches, as illustrated in Figure 3. This overall evaluation framework enables systematic assessment of both the quality of the learned preferences and the human effort required to achieve those preferences across different feedback paradigms.

Lastly, we would like to note that there is some degree of subjectivity in the number of interaction units assigned for each of the feedback mechanisms. To help combat against that, we also plan to provide results illustrating the evaluation metrics if each feedback mechanism had an interaction count of 1. As future work, we plan to more rigorously evaluate the efficiency of Active-MoSH, against the baselines, through a study with human subjects.

A.3.3 Real Clinical Case: Cervical Cancer Brachytherapy Treatment Planning

For the cervical cancer brachytherapy treatment planning application, we leveraged a real clinical case which had been performed in the clinic previously. As a result, we had access to the patient CT data (in the form of DICOM files) and the final treatment plan which had been administered to the patient. For the experiments, we obtained a single treatment plan via a linear program formulated as an epsilon-constraint method Deufel et al. (2020). The three parameters to that linear program, which implicitly control the weights of the different objectives (radiation dosage to the bladder, rectum, bowel, and cancer tumor), were employed as the decision variables.