The Smoothed Number of Pareto-optimal Solutions in Non-integer Bicriteria Optimization*

Heiko Röglin and Clemens Rösner

Department of Computer Science, University of Bonn, Germany
{roeglin,roesner}@cs.uni-bonn.de

Abstract. Pareto-optimal solutions are one of the most important and well-studied solution concepts in multi-objective optimization. Often the enumeration of all Pareto-optimal solutions is used to filter out unreasonable trade-offs between different criteria. While in practice, often only few Pareto-optimal solutions are observed, for almost every problem with at least two objectives there exist instances with an exponential number of Pareto-optimal solutions. To reconcile theory and practice, the number of Pareto-optimal solutions has been analyzed in the framework of smoothed analysis, and it has been shown that the expected value of this number is polynomially bounded for linear integer optimization problems. In this paper we make the first step towards extending the existing results to non-integer optimization problems. Furthermore, we improve the previously known analysis of the smoothed number of Pareto-optimal solutions in bicriteria integer optimization slightly to match its known lower bound.

1 Introduction

Optimization problems that arise from real-world applications often come with multiple objective functions. Since there is usually no solution that optimizes all objectives simultaneously, trade-offs have to be made. One of the most important solution concept in multi-objective optimization is that of *Pareto-optimal solutions*, where a solution is called Pareto-optimal if there does not exist another solution that is simultaneously better in all objectives. Intuitively Pareto-optimal solutions represent the reasonable trade-offs between the different objectives, and it is a common approach to compute the set of Pareto-optimal solutions to filter out all unreasonable trade-offs.

For many multi-objective optimization problems there exist algorithms that compute the set of Pareto-optimal solutions in polynomial time with respect to the input size and the number of Pareto-optimal solutions. These algorithms are not efficient in the worst case because for almost every problem with two or more objectives there exist instances with an exponential number of Pareto-optimal solutions. Since this does not reflect experimental results, where the number of

* This research was supported by ERC Starting Grant 306465 (BeyondWorstCase).

Pareto-optimal solutions is usually small, there has been a significant interest in probabilistic analyses of multi-objective optimization problems in the last decade.

The analyses in the literature are restricted to *linear integer optimization problems*, in which the solutions can be encoded as integer vectors and there is a constant number of linear objective functions to be optimized. To be more precise, an instance of a linear integer optimization problem is given by a set $\mathcal{S} \subseteq \{-k, -k+1, \ldots, k\}^n$ of feasible solutions for some $k \in \mathbb{N}$ and d linear objective functions $c^1, \ldots, c^d$ for some constant d. The function $c^i : \mathcal{S} \to \mathbb{R}$ is of the form $c^i(x) = c^i_1 x_1 + \ldots + c^i_n x_n$ for $x = (x_1, \ldots, x_n)$. Many well-known optimization problems can be formulated as a linear integer optimization problem. Consider, for example, the bicriteria shortest path problem in which one has to find a path in a graph $G = (V, E)$ from some source node s to some target node t and every edge has a certain length and induces certain costs. Then every s-t-path has a total length and total costs, and ideally one would like to minimize both simultaneously. A given instance of the bicriteria shortest path problem can easily be formulated as an instance of a linear bicriteria integer optimization problem by choosing $\mathcal{S} \subseteq \{0,1\}^{|E|}$ as the set of incidence vectors of s-t-paths. Then the coefficients in the two linear objective functions coincide with the edge lengths and costs.

A particular well-studied case are linear integer optimization problems with two objective functions. In the worst case it is very easy to come up with instances that have an exponential number of Pareto-optimal solutions. On the contrary, it has been proven that the expected number of Pareto-optimal solutions is polynomially bounded if the coefficients of one of the two objective functions are chosen at random, regardless of the choice of $\mathcal{S} \subseteq \{-k, -k+1, \ldots, k\}^n$. This is not only true if the coefficients are chosen uniformly at random but also for more sophisticated probabilistic models like smoothed analysis, in which the coefficients can roughly be determined by an adversary and are only slightly perturbed at random. Furthermore, it suffices if only the coefficients of one of the objective functions are chosen at random; the other objective function can be adversarial and does not even have to be linear. This can be seen as a theoretical explanation for why in experiments usually only few Pareto-optimal solutions exist because already a small amount of random noise in the coefficients suffices to render it very unlikely to encounter an instance with many Pareto-optimal solutions.

The analyses in the literature are restricted to the case that the set of solutions is a subset of a discrete set $\{-k, -k+1, \ldots, k\}^n$. Consider, for example, the binary case $\mathcal{S} \subseteq \{0,1\}^n$, and assume that we allow a little bit more flexibility in choosing the set of feasible solutions as follows: every solution $x \in \mathcal{S} \subseteq \{0,1\}^n$ may be replaced by a solution $\bar{x}$ with $|x_i - \bar{x}_i| \leq \varepsilon$ for every component i for a small ε. This way a new set of feasible solutions $\bar{\mathcal{S}} \subseteq [-\varepsilon, 1+\varepsilon]^n$ is obtained. Intuitively, if ε is very small, then the expected number of Pareto-optimal solutions with respect to $\mathcal{S}$ and with respect to $\bar{\mathcal{S}}$ should be roughly the same. However, this is not covered by the previous analyses and indeed analyzing the expected number of Pareto-optimal solutions with respect to $\bar{\mathcal{S}}$ seems to be a much harder problem.

In this paper we initiate the study of more general sets of feasible solutions. We do not solve the problem in full generality but we will make the first step towards understanding non-discrete sets of feasible solutions. The idea we use to obtain bounds for the expected number of Pareto-optimal solutions for the more general setting allows us also to improve slightly the best known bound for the bicriteria integer case, matching a known lower bound. In the following, we will first give a motivating example and then we will discuss our results and the previous work in more detail.

1.1 Knapsack Problem with Substitutes

In the knapsack problem, a set of n items with profits $p_1, \ldots, p_n$ and weights $w_1, \ldots, w_n$ is given. The goal is to find a vector $x \in \{0,1\}^n$ such that the total profit $p(x) = p_1x_1 + \ldots + p_nx_n$ is maximized under the constraint that the total weight $w(x) = w_1x_1 + \ldots + w_nx_n$ does not exceed a given capacity B. If one disregards the capacity, one can view the knapsack problem as a bicriteria optimization problem in which one seeks a solution from the set $\mathcal{S} = \{0,1\}^n$ with large profit and small weight. The assumption that profit and weight are linear functions is not always justified. If some items are substitute or complementary goods, then their joint profit can be smaller or larger than the sum of their single profits. Also if the weights represent costs and one gets a volume discount, the weight function is not linear.

In order to take this into account, we consider a more general version of the knapsack problem. We allow an arbitrary weight function $w : \{0,1\}^n \to \mathbb{R}$ that assigns a weight to every subset of items. Furthermore, we assume that some function $\alpha : \{0,1\}^n \to [0,1]^n$ is given and that the profit of a solution $x \in \{0,1\}^n$ is given as $p(x) = \alpha(x)_1 p_1 + \ldots + \alpha(x)_n p_n$. Hence, the function α determines for each item and each solution the fraction of the item's value that it contributes. One could, for example, encode rules like "if the second item is present, the first item counts only half, and if the second and third item are present, then the first item counts only a third".

The question we study in this paper is how many solutions from $\{0,1\}^n$ are Pareto-optimal with respect to the objective functions p and w. Formally, a solution $x \in \{0,1\}^n$ is Pareto-optimal if there does not exist a solution $y \in \{0,1\}^n$ that *dominates* x in the sense that y is at least as good as x in all criteria and strictly better than x in at least one criterion. Observe that we can reformulate the model in the following way so that it fits to our discussion above: We define $\bar{\mathcal{S}} = \{\alpha(x) \mid x \in \{0,1\}^n\} \subseteq [0,1]^n$, $\bar{w} : \bar{\mathcal{S}} \to \mathbb{R}$ by $\bar{w}(x) = w(\alpha^{-1}(x))$ for $x \in \bar{\mathcal{S}}$, assuming that α is injective, and $\bar{p}(x) = p_1x_1 + \ldots + p_nx_n$. Now the goal is to minimize the arbitrary objective function $\bar{w}$ and to maximize the linear objective function $\bar{p}$ over the set $\bar{\mathcal{S}} \subseteq [0,1]^n$.

As even for the simple linear case one can easily find instances in which every solution from $\{0,1\}^n$ is Pareto-optimal, it does not make sense to study this question in a classical worst-case analysis. We will instead assume that the profits $p_1, \dots, p_n$ are chosen at random and we will prove polynomial bounds for this case under the assumption that $\alpha(x)_i = 0$ for $x_i = 0$ and $\alpha(x)_i \geq \delta$ for $x_i = 1$ for some $\delta > 0$ for every $x \in \{0,1\}^n$ and every i (i.e., any item that is not part of a solution does not contribute any of its profit and any item that is part of a solution contributes at least some small fraction δ of its profit). In the literature only the case that α is the identity has been studied.

Since, we are only interested in the expected number of Pareto-optimal solutions, we do not care how the functions w and α are encoded. Our results are true for all functions, regardless of whether or not they can be encoded and evaluated efficiently.

1.2 Smoothed Analysis

Smoothed analysis has been introduced by Spielman and Teng [15] to explain why the simplex algorithm is efficient in practice despite its exponential worst-case behavior. We use the framework of smoothed analysis to study the number of Pareto-optimal solutions, and we will use the following model, which has already been used in the literature for the analysis of multi-objective linear optimization problems.

In our model, we assume that an arbitrary set $\mathcal{S} \subseteq [0,1]^n$ of feasible solutions that satisfies a certain property, which we define below, can be chosen by an adversary. Furthermore, the adversary can also choose an arbitrary objective function $w : \mathcal{S} \to \mathbb{R}$, which is to be minimized. Finally a second linear objective function $p : \mathcal{S} \to \mathbb{R}$ is given, which is to be maximized. This function is of the form $p(x) = p_1x_1 + \ldots + p_nx_n$ and in contrast to a worst-case analysis we do not allow the adversary to choose the coefficients $p_1, \dots, p_n$ exactly but we assume that they are chosen at random. For this, let $\phi \geq 1$ be some parameter and assume that the adversary can choose, for each coefficient p_i, a probability density function $f_i : [0,1] \to [0,\phi]$ according to which p_i is chosen independently of the other profits.

The smoothing parameter ϕ can be seen as a measure specifying how close the analysis is to a worst-case analysis. The larger ϕ, the more concentrated the probability mass can be: the adversary could for example define for each coefficient a uniform distribution on an interval of length $\frac{1}{\phi}$ from which it is chosen uniformly at random. This shows that for $\phi \to \infty$ our analysis approaches a worst-case analysis.

In the following, we will even allow a different parameter ϕ_i for each coefficient c_i, i.e., the density f_i is bounded from above by ϕ_i. Then $\phi = \max_{i \in [n]} \phi_i$, where $[n]$ denotes the set $\{1, \dots, n\}$. We define the *smoothed number of Pareto-optimal solutions* as the largest expected number of Pareto-optimal solutions that the adversary can achieve by choosing the set $\mathcal{S}$, the objective function w, and the densities $f_1, \dots, f_n$.

1.3 Previous Results

Multi-objective optimization is a well studied research area. There exist several algorithms to generate Pareto sets of various optimization problems like, e.g., the (bounded) knapsack problem [8, 12], the bicriteria shortest path problem [5, 14], and the bicriteria network flow problem [6, 11]. The running time of these algorithms depends crucially on the number of Pareto-optimal solutions and, hence, none of them runs in polynomial time in the worst case. In practice, however, generating the Pareto set is tractable in many situations [7, 10].

Beier and Vöcking initiated the study of the expected number of Pareto-optimal solutions for binary optimization problems [2]. They consider the model described in Section 1.2 with $\mathcal{S} \subseteq \{0,1\}^n$ and show that the expected number of Pareto-optimal solutions is bounded from above by $O(n^4\phi)$ and from below by $\Omega(n^2)$ even for $\phi = 1$. In [1] Beier, Röglin, and Vöcking analyze the smoothed complexity of bicriteria integer optimization problems. They show that the smoothed number of Pareto-optimal solutions is bounded from above by $O(n^2k^2\log(k)\phi)$ if $\mathcal{S} \subseteq \{0,\dots,k-1\}^n$. This improved the upper bound for the binary case to $O(n^2\phi)$. They also present a lower bound of $\Omega(n^2k^2)$ on the expected number of Pareto-optimal solutions for profits that are chosen uniformly from the interval $[-1,1]$.

Röglin and Teng generalized the binary setting $\mathcal{S} \subseteq \{0,1\}^n$ to an arbitrary constant number $d \geq 1$ of linear objective functions with random coefficients plus one arbitrary objective function [13]. They showed that the smoothed number of Pareto-optimal solutions is in $O((n^2\phi)^{f(d)})$, for a function f that is roughly $f(d) = 2^d d!$. In [9] this bound was significantly improved to $O(n^{2d}\phi^{d(d+1)/2})$ by Moitra and O'Donnel. Brunsch et al. proved in [3] a lower bound of $\Omega(n^{d-1.5}\phi^d)$ for the same setting. Instead of binary optimization problems Brunsch and Röglin analyze the smoothed number of Pareto-optimal solutions for multi-objective integer optimization problems [4]. They consider $\mathcal{S} \subseteq \{0,\dots,k\}^n$ and show that the expected number of Pareto-optimal solutions is in $k^{2(d+1)^2} \cdot O(n^{2d}\phi^{d(d+1)})$.

None of these analyses applies to the case that the set $\mathcal{S}$ of feasible solutions is a non-integral subset of $[0,1]^n$.

1.4 Our Results

We study bicriteria optimization problems in which the set $\mathcal{S}$ of feasible solutions is a finite subset of $[0,1]^n$ and one wants to optimize one arbitrary objective function $w : \mathcal{S} \to \mathbb{R}$ and one linear objective function $p : \mathcal{S} \to \mathbb{R}$. We call w weight and p profit. We do not care about the exact values of w and will therefore assume that w is given as a ranking on $\mathcal{S}$ where solutions with a lower weight have a higher ranking. In order to obtain interesting results about the number of Pareto-optimal solutions, it is necessary to restrict the set $\mathcal{S}$. We define the (k,δ)-*property* as follows.

Definition 1. *For given $k \in \mathbb{N}$ and $\delta \in (0,1]$, a set of solutions $\mathcal{S} \subseteq [0,1]^n$ satisfies the (k,δ)-property if there exist finite sets $K_i \subseteq [0,1]$ with $|K_i| \leq k$ for*

$i \in [n]$, such that for each pair of solutions $s \neq s' \in \mathcal{S}$ either $|\{i \in [n] \mid s_i \in K_i\}| \neq |\{i \in [n] \mid s'_i \in K_i\}|$ or there exist indices $i, j \in [n]$ such that $s_i \in K_i$, $|s_i - s'_i| \geq \delta$, and $s'_j \in K_j$, $|s_j - s'_j| \geq \delta$.

Let $\mathcal{S} \subseteq [0,1]^n$ be an instance of the Knapsack Problem with Substitutes, as described in Section 1.1. Recall that different solutions $s \neq s' \in \mathcal{S}$ differ in the coordinates with a value of 0, i.e., there exists $i \in [n]$ such that either $s_i = 0 \neq s'_i$ or $s_i \neq 0 = s'_i$. Since the value of each coordinate has to be 0 or at least δ we can set $K_i = \{0\}$ for every $i \in [n]$ and see that $\mathcal{S}$ has the $(1, \delta)$-property.

For finite bicriteria integer optimization problems we have $\mathcal{S} \subseteq \{-k, -k+1, \ldots, k-1, k\}^n$ for some $k \in \mathbb{N}$. For such sets the definition of the (k, δ)-property does not apply immediately. Instead we can first shift and then scale $\mathcal{S}$ to obtain $\hat{\mathcal{S}} \subseteq \{0 = \frac{0}{2k}, \frac{1}{2k}, \ldots, \frac{2k}{2k} = 1\}^n \subseteq [0,1]^n$ (First add k to every coordinate of every solution and then divide the result by $2k$). This shifting and scaling does not change the profit order and with the same ranking as before the shift, $\hat{\mathcal{S}}$ and $\mathcal{S}$ have the same number of Pareto-optimal solutions. With $K_i = \{0, \frac{1}{2k}, \ldots, 1\}$ it is easy to see that $\hat{\mathcal{S}}$ has the $(2k+1, \frac{1}{2k})$-property.

In this paper, we present an approach for bounding the smoothed number of Pareto-optimal solutions for bicriteria optimization problems that have a finite set $\mathcal{S} \subseteq [0,1]^n$ of feasible solutions with the (k, δ)-property. The general idea underlying our analysis is similar to the one used by Beier, Röglin, and Vöcking [1] to analyze integer problems. The basic idea is to partition the Pareto-optimal solutions into different classes and to analyze the expected number in each class separately. Roughly the class of a Pareto-optimal solution x is determined by the indices in which it differs from the next Pareto-optimal solution, i.e., the Pareto-optimal solution with smallest weight among all Pareto-optimal solutions with larger profit than x.

To analyze the expected number of Pareto-optimal solutions in one class, we first partition the interval $[0, n]$ of possible profits of solutions from $\mathcal{S}$ uniformly into small subintervals. Then, by linearity of expectation, it suffices to bound for each subinterval I the expected number of Pareto-optimal solutions with a profit in I. Let $I = [t-\varepsilon, t)$ for some t and $\varepsilon > 0$ be such a subinterval. If ε is very small, then with high probability I contains either none or exactly one Pareto-optimal solution. Hence, the expected number of Pareto-optimal solutions in I equals almost exactly the probability that there exists a Pareto-optimal solution whose profit lies in I. In order to bound this probability, we use the principle of deferred decisions as follows. First we uncover all profits except for the profit p_i for one of the positions i in which x differs from its next Pareto-optimal solution. This information suffices to identify a set of candidates for a Pareto-optimal solution in I. That is, if there exists a Pareto-optimal solution in I, then it must come from this set of candidate solutions. Beier, Röglin and Vöcking [1] treated each of these candidates separately and used linearity of expectation to bound the probability that any of them becomes a Pareto-optimal solution with profit in I. This is not possible anymore in our more general setting because there could be an exponential number of candidates. We instead use a new method, in which we exploit dependencies between the different candidates. This allows us to treat the set of candidates as a whole and to obtain a better bound on the probability that one of them becomes a Pareto-optimal solution with profit in I.

Theorem 2. *Let $k \in \mathbb{N}$, $\delta \in (0,1]$, and let $\mathcal{S} \subseteq [0,1]^n$ be a set of feasible solutions with the (k,δ)-property and some arbitrary ranking w. Assume that each profit p_i is a random variable with density function $f_i : [0,1] \to [0,\phi_i]$ and let $\phi = \max_{i\in[n]} \phi_i$. Let q denote the number of Pareto-optimal solutions in $\mathcal{S}$. Then*

$$\boldsymbol{E}[q] = O\left(\frac{n^2}{\delta}\sum_{i=1}^{n} k_i\phi_i\right) = O\left(\frac{n^3 k\phi}{\delta}\right).$$

We will show that every set of solutions $\mathcal{S} \subseteq \{0,\dots,k-1\}^n$ can be scaled into a set of solutions $\mathcal{S}' \in [0,1]^n$ with the $(k,\frac{1}{k})$-property. For bicriteria integer optimization problems we then further improve our analysis to improve the best previous result [1] and match the known lower bound $\Omega(n^2k^2)$ [1] for constant ϕ.

Theorem 3. *Let $\mathcal{S} \subseteq \mathcal{D}^n$ be a set feasible solutions with a finite domain $\mathcal{D} = \{0,\dots,k-1\} \subseteq \mathbb{Z}$ and an arbitrary ranking w. Assume that each profit p_i is a random variable with density function $f_i : [0,1] \to [0,\phi_i]$ and let $\phi = \max_{i\in[n]} \phi_i$. Let q denote the number of Pareto-optimal solutions in $\mathcal{S}$. Then*

$$\boldsymbol{E}[q] = O\left(nk^2\sum_{i=1}^{n} \phi_i\right) = O\left(n^2k^2\phi\right).$$

We also show a lower bound of $\Omega(\min\{(\frac{1}{\delta})^{\log_3(2)}, 2^n\})$ for the expected number of Pareto-optimal solutions in an instance with the $(1,\delta)$-property, where all profits are drawn according to a uniform distribution on the interval $[\frac{1}{2},1]$. This shows that the dependence on δ in Theorem 2 is necessary.

2 Upper Bound on the Expected Number of Pareto-optimal Solutions

As discussed above, the methods and ideas we use in this chapter are inspired by the analysis of Beier, Röglin and Vöcking [1]. We adapt their approach to the non-integer setting and also improve their analysis of the integer case.

Lemma 4. *Let $k \in \mathbb{N}$, $\ell \in [n]$, $\delta \in (0,1]$, and let $\mathcal{S} \subseteq [0,1]^n$ be a set of solutions with the (k,δ)-property. Assume $K_i \subseteq [0,1]$ with $|K_i| = k_i \le k$ for $i \in [n]$ to be corresponding sets for the (k,δ)-property of $\mathcal{S}$. Also let $|\{i \in [n] \mid s_i \in K_i\}| = \ell$ for every solution $s \in \mathcal{S}$. Assume that each profit p_i is a random variable with density function $f_i : [0,1] \to [0,\phi_i]$ and let $\phi = \max_{i\in[n]} \phi_i$. Let q denote the number of Pareto-optimal solutions in $\mathcal{S}$. Then*

$$\boldsymbol{E}[q] \le \left(\sum_{i=1}^{n} \frac{4nk_i\phi_i}{\delta}\right) + 1 \le \frac{4n^2k\phi}{\delta} + 1.$$

Proof. The idea of the proof is to partition the set of Pareto-optimal solutions into different classes and to compute the expected number of Pareto-optimal solutions in each of these classes separately. Let $\mathcal{P} \subseteq \mathcal{S}$ denote the set of all Pareto-optimal solutions. For each Pareto-optimal solution $s \in \mathcal{P}$, except for the one with largest profit, let $\text{next}(s) := \text{argmin}\{p(s') \mid s' \in \mathcal{P} \text{ and } p(s') > p(s)\}$ denote the Pareto-optimal solution with the next larger profit than s. Now let $s \in \mathcal{P}$ be an arbitrary Pareto-optimal solution that is not the one with the largest profit. By definition of the set $\mathcal{S}$, there has to be an $i \in [n]$ such that $\text{next}(s)_i = v$ for some $v \in K_i$ and $|s_i - \text{next}(s)_i| \geq \delta$. We then say that s belongs to the class (i, v). With the Pareto-optimal solution with the largest profit being a separate class by itself, every Pareto-optimal solutions is part of at least one of the classes. Note that a Pareto-optimal solution can belong to several different classes.

Let $i \in [n]$ and $v \in K_i$. We will now analyze the expected number of Pareto-optimal solutions in class (i, v). For this we consider the set

$$\mathcal{S}_{i,v} := \{s' \in [-1,1]^n \mid \exists s \in \mathcal{S} \text{ such that } s'_i = s_i - v \text{ and } \forall j \in [n] \setminus \{i\} : s_j = s'_j\}.$$

For each solution $s \in \mathcal{S}$ the set $\mathcal{S}_{i,v}$ contains a corresponding solution s', which is identical to s except for the i-th coordinate, where it is smaller by v. This does not change the profit order of the solutions because the profit of each solution in $\mathcal{S}_{i,v}$ is smaller by exactly $v \cdot p_i$ than the profit of its corresponding solution in $\mathcal{S}$. Given the same ranking (i.e., the same weight function) on $\mathcal{S}_{i,v}$ as on $\mathcal{S}$, there is a one-to-one correspondence between the Pareto-optimal solutions in $\mathcal{S}$ and $\mathcal{S}_{i,v}$. Hence, instead of analyzing the number of class (i, v) Pareto-optimal solution in $\mathcal{S}$, we can also analyze the number of class $(i, 0)$ Pareto-optimal solutions in $\mathcal{S}_{i,v}$. Instead of class $(i, 0)$ Pareto-optimal solutions, we will use the term *class i Pareto-optimal solutions* in the following. The following lemma concludes the proof by summing over the $\sum_{i=1}^{n} k_i$ different choices for the pair (i, v). Note that the term $+1$ in the lemma accounts for the Pareto-optimal solution with the largest profit. □

Lemma 5. *Consider the setting described in Lemma 4 and let $i \in [n]$ and $v \in K_i$ be arbitrary. Let q' denote the number of class i Pareto-optimal solutions in $\mathcal{S}_{i,v}$. Then*

$$\boldsymbol{E}[q'] \leq \frac{4n\phi_i}{\delta}.$$

Proof. The key idea is to prove an upper bound on the probability that there exists a class i Pareto-optimal solution in $\mathcal{S}_{i,v}$, whose profit falls into a small interval $[t - \varepsilon, t)$, for arbitrary t and $\varepsilon > 0$. We use the principle of deferred decisions and will assume in the following that all profits p_j for $j \neq i$ are already fixed arbitrarily. We will only exploit the randomness of p_i.

We want to bound the probability that there exists a class i Pareto optimal solution, whose profit lies in the interval $[t-\varepsilon, t)$. Let $\mathcal{S}_{x_i=0} := \{s \in \mathcal{S}_{i,v} \mid s_i = 0\}$. Define x_1^* to be the highest ranked solution $x \in \mathcal{S}_{x_i=0}$ satisfying $p(x) \geq t$ and define $X^* := \{x_1^*, x_2^*, \ldots, x_{m_{t,\varepsilon}}^*\}$ to be the set containing x_1^* and all solutions

$x \in \mathcal{S}_{x_i=0}$ that are Pareto-optimal with respect to the set $\mathcal{S}_{x_i=0}$ and that satisfy $t-\varepsilon \leq p(x) < t$. We assume X^* to be ordered such that for all $x_j^* \in X^*$ we have $p(x_j^*) < p(x_{j-1}^*)$ (see Figure 1).

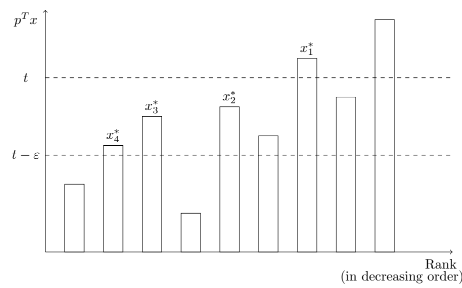

Fig. 1: Example of $\mathcal{S}_{x_i=0}$, with solutions in $X^* = \{x_1^*, x_2^*, x_3^*, x_4^*\}$ marked as such.

Note that the solutions in X^* do not have to be Pareto-optimal in $\mathcal{S}_{i,v}$ (they could be dominated by solutions outside of $\mathcal{S}_{x_i=0}$) and that X^* does not have to contain any solutions. If $\mathcal{S}_{i,v}$ contains a class i Pareto-optimal solution x, whose profit falls into the interval $[t-\varepsilon, t)$, we have $\text{next}(x) \in \mathcal{S}_{x_i=0}$. Since $\text{next}(x)$ is Pareto-optimal in $\mathcal{S}_{i,v}$, it has to be Pareto-optimal in $\mathcal{S}_{x_i=0}$ as well. We claim $\text{next}(x) \in X^*$. Assume $\text{next}(x) \notin X^*$, then $p(\text{next}(x))$ must be at least t and therefore $\text{next}(x)$ must have a lower rank than x_1^*. Since $\text{next}(x)$ is Pareto-optimal we must have $p(\text{next}(x)) > p(x_1^*)$. By the definition of $\text{next}(x)$ there can be no Pareto-optimal solution x' in $\mathcal{S}_{i,v}$ with $p(x) < p(x') < p(\text{next}(x))$, which means that there can be no Pareto-optimal solution in $\mathcal{S}_{i,v}$ that dominates x_1^* but not $\text{next}(x)$, which is a contradiction. Analogously it follows that $\text{next}(x)$ is the solution with the highest rank among all $x_j^* \in X^*$ with $p(x_j^*) > p(x)$. If $\text{next}(x) = x_j^*$ for some $j < m_{t,\varepsilon}$, then $p(x) \in [p(x_{j+1}^*), p(x_j^*))$, and if $\text{next}(x) = x_{m_{t,\varepsilon}}^*$, then $p(x) \in [t-\varepsilon, p(x_{m,\varepsilon}^*))$.

In order to analyze the probability that there exists a class i Pareto-optimal solution, whose profit lies in the interval $[t-\varepsilon, t)$, we look at each $x_j^* \in X^*$. Let $r_1 = t$, $r_{m_{t,\varepsilon}+1} = t-\varepsilon$ and $r_j = p(x_j^*)$ for $j \in \{2, \ldots, m_{t,\varepsilon}\}$. We will, for each $j \in [m_{t,\varepsilon}]$, bound the probability that $\mathcal{S}_{i,v}$ contains a class i Pareto-optimal solution, whose profit lies in the interval $[r_{j+1}, r_j)$.

Let $j \in [m_{t,\varepsilon}]$ and let $\hat{x}_j$ denote the solution that has the largest profit among all solutions x with $|x_i| \geq \delta$ that are higher ranked than x_j^*. Assume that there exists a class i solution x with profit in the interval $[r_{j+1}, r_j)$. Then x_j^* has to be a Pareto-optimal solution and x has to be higher ranked than x_j^*, because otherwise x_j^* would dominate x. Let y denote the solution, among all solutions that are higher ranked than x_j^*, that has the largest profit. Since x_j^* is Pareto-optimal, y is Pareto-optimal as well and has less profit than x_j^*. Since we assume x to be a class i solution with profit in the interval $[r_{j+1}, r_j)$ we know $\text{next}(x) = x_j^*$ and therefore $x = y$ and $x = \hat{x}_j$.

We now aim to bound the probability that $\hat{x}_j$ is a class i Pareto-optimal solution and falls into the interval $[r_{j+1}, r_j)$. Define

$$\Lambda(t,j) = \begin{cases} r_j - p(\hat{x}_j) & \text{if } \hat{x}_j \text{ exists} \\ \perp & \text{otherwise.} \end{cases}$$

Let $\mathcal{P}$ denote the set of class i Pareto-optimal solutions and $\varepsilon_j = r_j - r_{j+1}$ for all $j \in [m_{t,\varepsilon}]$. Whenever there exists a class i solution $x \in \mathcal{P}$ with $p(x) \in [r_{j+1}, r_j)$, the choice of $\hat{x}_j$ implies that $x = \hat{x}_j$ and hence $\Lambda(t,j) \in (0, \varepsilon_j]$.

Then

$$\mathbf{Pr}[\exists x \in \mathcal{P} : p(x) \in [r_{j+1}, r_j)] \le \mathbf{Pr}[\Lambda(t,j) \in (0, \varepsilon_j]].$$

Since the expected number of class i Pareto-optimal solutions can be written as

$$\begin{aligned} &\int_{-\infty}^{\infty} \lim_{\varepsilon \to 0} \frac{\mathbf{Pr}[\exists x \in \mathcal{P} : p(x) \in [t-\varepsilon, t)]}{\varepsilon} dt \\ &\le \int_{-\infty}^{\infty} \lim_{\varepsilon \to 0} \frac{\sum_{j=1}^{m_{t,\varepsilon}} \mathbf{Pr}[\exists x \in \mathcal{P} : p(x) \in [r_{j+1}, r_j)]}{\varepsilon} dt \\ &\le \int_{-\infty}^{\infty} \lim_{\varepsilon \to 0} \frac{\sum_{j=1}^{m_{t,\varepsilon}} \mathbf{Pr}[\Lambda(t,j) \in (0, \varepsilon_j]]}{\varepsilon} dt \\ &= \int_{-n}^{n} \lim_{\varepsilon \to 0} \frac{\sum_{j=1}^{m_{t,\varepsilon}} \mathbf{Pr}[\Lambda(t,j) \in (0, \varepsilon_j]]}{\varepsilon} dt, \end{aligned}$$

where the last equality comes from the fact that all solutions have a profit in $[-n, n]$, it remains to analyze the terms $\mathbf{Pr}[\Lambda(t,j) \in (0, \varepsilon_j]]$.

Let $\mathcal{S}^{|x_i| \ge \delta} = \{x \in \mathcal{S}_{i,v} \mid |x_i| \ge \delta\}$ and $\mathcal{S}^{x_i = u} = \{x \in \mathcal{S}_{i,v} \mid x_i = u\}$. For all $j \in [m_{t,\varepsilon}]$ let $\mathcal{L}_j$ consist of all solutions from $\mathcal{S}^{|x_i| \ge \delta}$ that have a higher rank than x_j^* and let $\mathcal{L}_j^u$ consist of all solutions from $\mathcal{S}^{x_i = u}$ that have a higher rank than x_j^*. Let $\hat{x}_j^u$ denote the lowest ranked Pareto-optimal solution from the set $\mathcal{L}_j^u$, i.e., $\hat{x}_j^u$ has the largest profit among all solutions in $\mathcal{L}_j^u$.

The identity of x_j^* is completely determined by the profits p_ℓ, $\ell \ne i$. For all $u \in [-1, 1]$ the set $\mathcal{L}_j^u$ and therefore the existence and identity of $\hat{x}_j^u$ are completely determined by those profits as well. Hence, if we fix all profits except for p_i, then $\hat{x}_j^u$ is fixed and its profit is $c_u + up_i$ for some constant c_u that depends only on the profits already fixed. The identity of $\hat{x}_j$ still depends on the exact value of p_i, but independent of p_i it has to be equal to $\hat{x}_j^u$ for some $u \in [-1, 1]$ with $|u| \ge \delta$. More specifically we have $\hat{x}_j = \operatorname{argmax}\{p(\hat{x}_j^u) \mid \hat{x}_j^u \text{ exists and } |u| \ge \delta\}$, which depends on the exact value of p_i. We can view $\{\hat{x}_j^u \mid \hat{x}_j^u \text{ exists and } |u| \ge \delta\}$ as the set of candidates, which could be a a class i Pareto-optimal solution with profit in the interval $[r_{j+1}, r_j)$.

This means that $\hat{x}_j^u$ takes a profit in the interval $[r_{j+1}, r_j)$ if and only if p_i lies in the interval $[b_u, b_u + \frac{\varepsilon_j}{u}) := [\frac{r_{j+1} - c_u}{u}, \frac{r_j - c_u}{u})$ in case $u > 0$ and $(b_u, b_u + \frac{\varepsilon_j}{-u}] := (\frac{r_j - c_u}{u}, \frac{r_{j+1} - c_u}{u}]$ in case $u < 0$. Let $b = \min\{b_u \mid u \in [\delta, 1] \text{ and } \hat{x}_j^u \text{ exists}\}$ and let $u' = \operatorname{argmin}\{b_u \mid u \in [\delta, 1] \text{ and } \hat{x}_j^u \text{ exists}\}$. Then for $p_i < b$ and all $u \in [\delta, 1]$ we have $p(\hat{x}_j^u) < r_{j+1}$ and for $p_i \geq b + \frac{\varepsilon_j}{u'}$ we have $p(\hat{x}_j^{u'}) \geq r_j$. Let $b' = \max\{b_u \mid u \in [-1, -\delta] \text{ and } \hat{x}_j^u \text{ exists}\}$ and let $u'' = \text{argmax}\{b_u \mid u \in [-1, -\delta] \text{ and } \hat{x}_j^u \text{ exists}\}$. Then for $p_i' \leq b'$ we have $p(\hat{x}_j^{u''}) \geq r_j$ and for $p_i > b' + \frac{\varepsilon_j}{-u''}$ and all $u \in [-1, -\delta]$ we have $p(\hat{x}_j^u) < r_{j+1}$. This implies that for all $p_i \notin [b, b + \frac{\varepsilon_j}{u'}) \cup (b', b' + \frac{\varepsilon_j}{-u''}]$ we have $p(\hat{x}_j) \notin [r_{j+1}, r_j)$. Hence we obtain $\mathbf{Pr}[\Lambda(t,j) \in (0, \varepsilon_j]] \leq \varepsilon_j \left(\frac{\phi_i}{u'} + \frac{\phi_i}{-u''}\right) \leq \varepsilon_j \frac{2\phi_i}{\delta}$.

Now we can bound the expected number of Pareto-optimal solutions:

$$\mathbf{E}[q'] \leq \int_{-n}^{n} \lim_{\varepsilon \to 0} \frac{\sum_{j=1}^{m_{t,\varepsilon}} \varepsilon_j \frac{2\phi_i}{\delta}}{\varepsilon} dt = \int_{-n}^{n} \lim_{\varepsilon \to 0} \frac{\varepsilon \frac{2\phi_i}{\delta}}{\varepsilon} dt = \int_{-n}^{n} \lim_{\varepsilon \to 0} \frac{2\phi_i}{\delta} dt = \frac{4n\phi_i}{\delta}.$$

□

We now show Theorem 2.

Proof (Proof (Theorem 2)). For $\ell \in [n]$ let $\mathcal{S}_\ell = \{s \in \mathcal{S} \mid |\{i \in [n] \mid s_i \in K_i\}| = \ell\}$, where the $K_i \subseteq [0,1]$ with $|K_i| = k_i \leq k$ for $i \in [n]$ denote the corresponding sets for the (k,δ)-property of $\mathcal{S}$. Let $\mathcal{P}_\ell$ denote the set of Pareto-optimal solutions in $\mathcal{S}_\ell$ and let $\mathcal{P}$ be the set of Pareto-optimal solutions in $\mathcal{S}$. Let $s \in \mathcal{P}$ be a Pareto-optimal solution in $\mathcal{S}$. Then there exists no solution $s' \in \mathcal{S}$ that dominates s. For some $\ell \in [n]$ we have $s \in \mathcal{S}_\ell$. $\mathcal{S}_\ell \subseteq \mathcal{S}$ implies that no solution $s' \in \mathcal{S}_\ell$ dominates s. Therefore s is also Pareto-optimal in $\mathcal{S}_\ell$, i.e., $s \in \mathcal{P}_\ell$. This implies $\mathcal{P} \subseteq \bigcup_{\ell \in [n]} \mathcal{P}_\ell$. Let q_ℓ denote the number of Pareto-optimal solutions in $\mathcal{S}_\ell$, i.e., $q_\ell = |\mathcal{P}_\ell|$. Lemma 4 and linearity of expectation yield

$$\begin{aligned}\mathbf{E}[q] &\leq \sum_{\ell \in [n]} \mathbf{E}[q_\ell] + 1 \leq \sum_{\ell \in [n]} \left(\frac{4n}{\delta} \sum_{i=1}^{n} k_i \phi_i + 1\right) + 1 \\ &= \frac{4n^2}{\delta} \sum_{i=1}^{n} k_i \phi_i + n + 1 \leq \frac{4n^3 k \phi}{\delta} + n + 1.\end{aligned}$$

Here the additional 1 comes from a possible solution $s \in \mathcal{S}$ with $|\{i \in [n] \mid s_i \in K_i\}| = 0$. The (k,δ)-property ensures that there can exist at most one such solution. This concludes the proof. □

We now prove Theorem 3.

Proof (Proof (Theorem 3)). We take a look at the scaled version $\mathcal{S}' = \{\frac{1}{k}s \mid s \in \mathcal{S}\}$, where $\frac{1}{k}s$ denotes the solution s' with $s_i' = \frac{1}{k}s_i$ for all $i \in [n]$. Since this scaling operation changes the profit of every solution by a factor of $\frac{1}{k}$, the two sets $\mathcal{S}'$ and $\mathcal{S}$ have the same number of Pareto-optimal solutions. Setting $K = \{\frac{i}{k} \mid i \in \{0, \ldots, k-1\}\}$ and $\delta = \frac{1}{k}$ we can apply Lemma 4 to obtain

$$\mathbf{E}[q] \leq \sum_{i=1}^{n} \frac{4nk_i\phi_i}{1/k} + 1 \leq 4nk^2 \sum_{i=1}^{n} \phi_i + 1 = O\left(nk^2 \sum_{i=1}^{n} \phi_i\right)$$

□

3 A Lower Bound

In this section we will show a simple lower bound for the expected number of Pareto-optimal solutions. For a given $\delta \in (0,1]$ we will show how to find a set of solutions $\mathcal{S} \subseteq (\{0\} \cup [\delta,1])^n$ with the $(1,\delta)$-property such that the number of Pareto-optimal solutions in $\mathcal{S}$ is $\Omega(\min\{2^n, (\frac{1}{\delta})^{\log_3(2)}\})$, assuming that all profits are drawn uniformly at random from the interval $[\frac{1}{2},1]$. Furthermore, the coordinates of the solutions in $\mathcal{S}$ will take at most $n+1$ different values, showing that a bound on the number of different values alone is not sufficient to obtain a polynomial bound on the number of Pareto-optimal solutions.

Theorem 6. *Suppose profits are drawn according to a uniform distribution from the interval $[\frac{1}{2},1]$. Then for every $\delta \in (0,1]$ there exists a set $\mathcal{S} \subseteq (\{0\} \cup [\delta,1])^n$ with the $(1,\delta)$-property and a ranking on $\mathcal{S}$ such that the number of Pareto-optimal solutions in $\mathcal{S}$ is $\Omega(\min\{(\frac{1}{\delta})^{\log_3(2)}, 2^n\})$.*

Proof. For $i \in [n]$, let $x_i = \frac{1}{3^{i-1}}$ and let $\mathcal{S}' = \{0,x_1\} \times \{0,x_2\} \times \ldots \times \{0,x_n\}$. The choice of x_i guarantees that for all $i \in [n]$ we get $\frac{x_i}{2} \geq \sum_{j=i+1}^{n} x_j$. This implies that regardless of how the values of the profits p_i for $i \in [n]$ are chosen, the lexicographical order of the solutions is equal to their profit order. When we use the lexicographical order as our ranking as well, this implies that all solutions are Pareto-optimal. We will define $\mathcal{S} = \{s \in \mathcal{S}' \mid \forall i : s_i \in \{0\} \cup [\delta,1]\}$ to be the subset of $\mathcal{S}'$ that contains only the solutions, whose coordinates have values of 0 or at least δ. We get $\mathcal{S} = \{0,x_1\} \times \{0,x_2\} \times \ldots \times \{0,x_{\lfloor \log_3 \frac{1}{\delta} \rfloor + 1}\} \times \{0\} \times \ldots \times \{0\}$ for the case $\lfloor \log_3 \frac{1}{\delta} \rfloor + 1 < n$ and $\mathcal{S} = \mathcal{S}'$ otherwise. With $K_i = \{0\}$ for all $i \in [n]$ we can see that $\mathcal{S}$ has the $(1,\delta)$-property. The set $\mathcal{S}$ contains $\min\{2^n, 2^{\lfloor \log_3 \frac{1}{\delta} \rfloor + 1}\}$ different solutions, and as we have seen, all solutions are Pareto-optimal. The observation that $(\frac{1}{\delta})^{\log_3(2)} = 2^{\log_3(\frac{1}{\delta})}$ concludes the proof. □

4 Conclusion and Open Problems

We defined for bicriteria optimization problems with a finite set of solutions $\mathcal{S} \subseteq [0,1]^n$ the (k,δ)-property and showed how to obtain an upper bound for the smoothed number of Pareto-optimal solutions in instances with the (k,δ)-property. It is easy to see that the (k,δ)-property can be applied to any finite set of solutions. However, in general δ can be arbitrarily small.

Lemma 7. *Let $\mathcal{S} \subseteq [0,1]^n$ be a finite set of solutions. There exist $k \in \mathbb{N}$ and $\delta \in (0,1]$, such that $\mathcal{S}$ has the (k,δ)-property.*

Proof. Let $E_{\mathcal{S}} = \{x \in [0,1] \mid \exists s \in \mathcal{S}, i \in [n] : s_i = x\}$ denote the set of values that the coordinates of the solutions in $\mathcal{S}$ take. Let $k = |E_{\mathcal{S}}|$ and $\delta = \min_{x \neq y \in E_{\mathcal{S}}} |x-y|$. Since $\mathcal{S}$ is a finite set, this is well defined and we get $k \in \mathbb{N}$ and $\delta \in (0,1]$. We choose $K_i = E_{\mathcal{S}}$ for all $i \in [n]$. Let $s \neq s' \in \mathcal{S}$ be two different solutions then there must exist $i \in [n]$ such that $s_i \neq s'_i$. By definition of K_i and δ we have $s_i \in K_i$, $s'_i \in K_i$ and $|s_i - s'_i| \geq \delta$, which yields that $\mathcal{S}$ has the (k,δ)-property. □

As our upper bound on the expected number of Pareto-optimal solutions depends on both k and δ, one can ask if there exists an upper bound that depends only on k or only on δ. Theorem 6 shows there can be no polynomial upper bound only in n, ϕ, and k. On the other hand, we conjecture that there exists an upper bound for the smoothed number of Pareto-optimal solutions that depends polynomially on n, ϕ and the inverse of the minimum distance $\min_{s \neq s' \in \mathcal{S}} \|s - s'\|$ between solutions.

Bibliography

[1] R. Beier, H. Röglin, and B. Vöcking. The smoothed number of pareto optimal solutions in bicriteria integer optimization. *Integer Programming and Combinatorial Optimization (IPCO)*, 12:53–67, 2007.

[2] R. Beier and B. Vöcking. Random knapsack in expected polynomial time. *Journal of Computer and System Sciences*, 69(3):30632, 2004.

[3] T. Brunsch, N. Goyal, L. Rademacher, and H. Röglin. Lower bounds for the average and smoothed number of pareto-optima. *Theory Comput.*, 10(10):237256, 2014.

[4] T. Brunsch and H. Röglin. Improved smoothed analysis of multiobjective optimization. *Journal of the ACM*, 62(1), 2015.

[5] H. Corley and I. Moon. Shortest paths in networks with vector weights. *Journal of Optimization Theory and Application*, 46(1):7986, 1985.

[6] M. Ehrgott. Integer solutions of multicriteria network flow problems. *Investigacao Operacional*, 19:6173, 1999.

[7] M. Ehrgott and X. Gandibleux. Multiple criteria optimization. In *Lecture Notes in Economics and Mathematical Systems*, volume 491, chapter Multiobjective Combinatorial Optimization. Springer-Verlag, 2000.

[8] K. Klamroth and M. M. Wiecek. Dynamic programming approaches to the multiple criteria knapsack problem. *Naval Research Logistics*, 47(1):5776, 2000.

[9] A. Moitra and R. O'Donnel. Pareto optimal solutions for smoothed analysts. *SIAM J. Comput*, 41, 5:1266–1284, 2012.

[10] M. Müller-Hannemann and K. Weihe. Pareto shortest paths is often feasible in practice. *In Proceedings of the 5th International Workshop on Algorithm Engineering (WAE)*, page 185198, 2001.

[11] A. Mustafa and M. Goh. Finding integer efficient solutions for bicriteria and tricriteria network flow problems using dinas. *Computers & OR*, 25(2):139157, 1998.

[12] G. L. Nemhauser and Z. Ullmann. Discrete dynamic programming and capital allocation. *Management Science*, 15:494505, 1969.

[13] K. Röglin and S.-H. Teng. Smoothed analysis of multiobjective optimization. *In Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 681–690, 2009.

[14] A. J. V. Skriver and K. A. Andersen. A label correcting approach for solving bicriterion shortest-path problems. *Computers & OR*, 27(6):507524, 2000.

[15] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385463, 2004.