

Ordinal Optimization Through Multi-objective Reformulation

Kathrin Klamroth^a, Michael Stiglmayr^a, Julia Sudhoff^{a,*}

^a*University of Wuppertal, School of Mathematics and Natural Sciences, IMACM, Gaußstr. 20, 42119 Wuppertal, Germany*

Abstract

We analyze combinatorial optimization problems with ordinal, i.e., non-additive, objective functions that assign categories (like good, medium and bad) rather than cost coefficients to the elements of feasible solutions. We review different optimality concepts for ordinal optimization problems and discuss their similarities and differences. We then focus on two prevalent optimality concepts that are shown to be equivalent. Our main result is a bijective linear transformation that transforms ordinal optimization problems to associated standard multi-objective optimization problems with binary cost coefficients. Since this transformation preserves all properties of the underlying problem, problem-specific solution methods remain applicable. A prominent example is dynamic programming and Bellman's principle of optimality, that can be applied, e.g., to ordinal shortest path and ordinal knapsack problems. We extend our results to multi-objective optimization problems that combine ordinal and real-valued objective functions.

Keywords: multiple objective programming, ordering cones, ordinal objective functions, combinatorial optimization

1. Introduction

Ordinal objective functions occur whenever it is only possible to represent the quality of an element by a ordered category and not by a numerical value. As an example, consider the problem of finding optimized routes for cyclists in a road network: While edges may be associated with different categories like asphalt, gravel or sand—or, when related to safety considerations, very safe (there is a bicycle path), neutral (a quiet road) or unsafe (a main road without bicycle path)—such categories do not immediately translate into monetary or cost values. Bi-objective shortest path problems with route safety criteria

*Corresponding author

Email addresses: klamroth@math.uni-wuppertal.de (Kathrin Klamroth), stiglmayr@math.uni-wuppertal.de (Michael Stiglmayr), sudhoff@math.uni-wuppertal.de (Julia Sudhoff)

are addressed, for example, in the web application *geovelo* and in the associated publications Kergosien et al. (2021); Sauvanet & Néron (2010). In these references, only two categories are considered (safe or unsafe edges), and the safety criterion is translated into a cost function that evaluates the total length of unsafe route segments. In contrast, an ordinal shortest path problem is investigated in Schäfer et al. (2020). A major difficulty when considering ordinal objective functions is that “optimality” may be defined in many different ways. Schäfer et al. (2020) suggests an optimality concept that is based on sorted category vectors. A similar concept is used in Klamroth et al. (2021), where matroid optimization problems with one real-valued and one ordinal objective function are investigated. A different perspective is proposed in Schäfer et al. (2021) who define ordinal optimality for knapsack problems on the basis of numerical representations for the categories.

In this paper, we consider general combinatorial optimization problems and provide a new cone-based interpretation of the optimality concept for ordinal objectives suggested in Schäfer et al. (2021). In particular, we interrelate ordinal optimality with the classical concept of Pareto optimality for an associated multi-objective optimization problem. For a general introduction to multi-objective optimization we refer to Ehrgott (2005). Since the underlying transformation of the objective function is linear and bijective and hence preserves the combinatorial structure of the respective problems, our results immediately lead to efficient solution strategies as, for example, dynamic programming for shortest path problems. The respective transformations are based on a representation of dominance relations by cones.

The paper is organized as follows. In Section 2 we review three optimality concepts for ordinal optimization based on Schäfer et al. (2021). These results and optimality concepts are then extended and re-interpreted as a special case of cone-optimality (see, e.g., Engau, 2007) in Section 3. A detailed analysis of the properties of the corresponding ordering cones then leads to a linear transformation of ordinal optimization problems to associated multi-objective optimization problems with the Pareto cone defining dominance and optimality. This transformation is used in Section 4 to formulate a general algorithm for solving ordinal optimization problems. Furthermore, we investigate the relation between the definition of ordinal optimality and the weight space decomposition of the associated multi-objective problem. In Section 5 we extend our results to more general problem types with additional real-valued objective functions. We conclude in Section 6 with a summary and an outlook on future research.

2. Single-objective Ordinal Optimization

2.1. Problem Definition

We consider combinatorial optimization problems with an ordinal objective function. In general, an *ordinal optimization problem* (OOP) can be formulated as

$$\begin{aligned} \text{“ordinally minimize”} \quad & o(x) \\ \text{s. t.} \quad & x \in X, \end{aligned} \tag{OOP}$$

where X is the set of feasible solutions. We assume that X is a subset of the power set of a discrete set S , i.e., $X \subseteq 2^S$. Every element of S is assigned to one of K ordered categories. This assignment is encoded by a mapping $o : S \rightarrow \mathcal{C}$ with $\mathcal{C} = \{\eta_1, \dots, \eta_K\}$. We assume that category η_i with $i \in \{1, \dots, K-1\}$ is strictly preferred over category η_{i+1} , written as $\eta_i \prec \eta_{i+1}$. The objective function of a feasible solution $x = \{e_1, \dots, e_n\}$ is given by the *ordinal vector* $o(x) = \text{sort}(o(e_1), \dots, o(e_n))$, where the operator $\text{sort}()$ means that the components of $o(x)$ are sorted w.r.t. non-decreasing preferences, i.e., $o_1(x) \preceq o_2(x) \preceq \dots \preceq o_n(x)$. Note that different feasible solutions may have different numbers of elements, and hence the length of the ordinal vector $o(x)$ may vary for different $x \in X$.

Instead of using the un-aggregated, ordered ordinal vector $o(x)$ one can count the number of elements in a feasible solution per category. Accordingly, we use the *counting vector* $c : X \rightarrow \mathbb{Z}_{\geq}^K$ with $\mathbb{Z}_{\geq}^K := \{y \in \mathbb{Z}^K : y_i \geq 0 \text{ for all } i = 1, \dots, K\}$. Thereby, the i -th component of $c(x)$ equals the number of elements in x which are in category η_i , i.e., $c_i(x) = |\{e \in x : o(e) = \eta_i\}|$. Obviously, there is a one to one correspondence between the vectors $o \in \mathbb{R}^n$ and $c \in \mathbb{R}^K$, since the ordinal vector o can be determined from a given counting vector c by $o_i(x) = \eta_j$ with $j = \text{argmin}\{j \in \{1, \dots, K\} : i \leq \sum_{l=1}^j c_l(x)\}$. Note again that the number of elements n of a feasible solution, and hence the length of the ordinal vectors $o \in \mathbb{R}^n$, may vary while the number of categories K and therefore the length of the counting vectors $c \in \mathbb{R}^K$ is fixed. Hence, we get the following formulation of an *ordinal counting optimization problem* (OCOP)

$$\begin{aligned} \text{“ordinally minimize”} \quad & c(x) \\ \text{s. t.} \quad & x \in X, \end{aligned} \tag{OCOP}$$

which will be shown to be equivalent to problem (OOP) for an appropriate definition of “ordinal minimization”.

In the following, we also consider an *incremental tail counting vector* $\tilde{c} \in \mathbb{R}^K$ that counts, in its i -th component, the number of elements of a feasible solution x which are in category η_i or worse, i.e., $\tilde{c}_i(x) = |\{e \in x : \eta_i \preceq o(e)\}| = \sum_{j=i}^K c_j(x)$. In particular, the total number of elements of a solution x is given in the first component of \tilde{c} , i.e., $|x| = \tilde{c}_1(x) = \sum_{i=1}^K c_i(x)$.

As an example, consider the shortest path problem shown in Figure 1 together with the outcome vectors $o(x)$ (for problem (OOP)) and $c(x)$ (for problem (OCOP)) for all feasible solutions $x \in X$. In addition, the incremental tail counting vector $\tilde{c}(x)$ is given for all $x \in X$.

2.2. Optimality Concepts for Ordinal Objective Functions

In the following, we review three different concepts of optimality for ordinal optimization. All of them try to answer the question what minimization could mean for the problems (OOP) and (OCOP). The first concept is to use *numerical representations* that assign a numerical value to every category such that the order of the categories is respected. In this context, a numerical representation respects the order of the categories whenever the numerical value of a

	o	c	\tilde{c}
$x^1 = \{e_1, e_2, e_5\}$	$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$
$x^2 = \{e_4, e_5\}$	$\begin{pmatrix} \eta_1 \\ \eta_3 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$
$x^3 = \{e_1, e_3\}$	$\begin{pmatrix} \eta_2 \\ \eta_2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$
$x^4 = \{e_6, e_8\}$	$\begin{pmatrix} \eta_2 \\ \eta_2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$
$x^5 = \{e_4, e_7, e_8\}$	$\begin{pmatrix} \eta_1 \\ \eta_1 \\ \eta_2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$
$x^6 = \{e_1, e_2, e_7, e_8\}$	$\begin{pmatrix} \eta_1 \\ \eta_1 \\ \eta_2 \\ \eta_2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}$

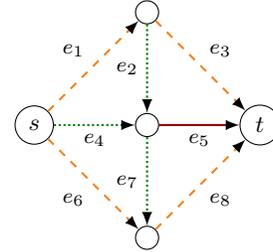


Figure 1: Instance of an ordinal shortest path problem. A dotted-green edge is in the best category η_1 , a dashed-orange edge is in category η_2 and a solid-red edge is in the worst category η_3 . All possible $s - t$ paths x^i , $i = 1, \dots, 6$ and their respective objective function vectors are given.

better category is strictly smaller than the numerical value of a less preferred category. If we take the sum over all numerical values of a vector $o(x')$ for a feasible solution x' , we obtain a unique numerical value that can be compared to the corresponding numerical value of another feasible solution \hat{x} . A feasible solution x' is called *efficient* if there is no other feasible solution \hat{x} which is better w.r.t. *all* numerical representations.

The second concept is to maximize the number of elements in the good categories, and the third concept is to minimize the number of elements in the bad categories. After the formal introduction of these three optimality concepts, we investigate their interrelation.

Optimality by Numerical Representations. We consider the problems (OOP) and (OCOP). The concept of *optimality by numerical representation* for ordinal objectives as introduced in Schäfer et al. (2021) is based on a previous and more general work of Fishburn (1999). It assigns an order preserving numerical value to each category. Following Schäfer et al. (2021), we call a function $\nu : \mathcal{C} \rightarrow \mathbb{Z}_{\geq}$

a numerical representation if

$$\eta_i \prec \eta_j \iff \nu(\eta_i) < \nu(\eta_j) \text{ for all } i, j \in \{1, \dots, K\}.$$

Note that we assume strictly ordered categories, i.e., there are no categories that are indifferent. As a consequence, we do not allow $\nu(\eta_i) = \nu(\eta_j)$ for $i \neq j$ since this would make two different categories indistinguishable in the numerical representation. Let \mathcal{V} denote the set of all numerical representations for a given number of categories K .

For a given numerical representation ν , we define the numerical value of a feasible solution $x = \{e_1, \dots, e_n\} \in X$ w.r.t. ν (cf. Schäfer et al., 2021) as

$$\nu(x) := \sum_{i=1}^n \nu(o(e_i)) = \sum_{i=1}^K \nu(\eta_i) \cdot c_i(x).$$

The numerical value $\nu(x)$ of a feasible solution $x \in X$ can be evaluated in different ways by re-arranging the terms and using the counting vector c or the incremental tail counting vector \tilde{c} , respectively:

$$\begin{aligned} \nu(x) &= \sum_{i=1}^K \nu(\eta_i) \cdot c_i(x) \\ &= \sum_{i=1}^{K-1} \nu(\eta_i) \left(\sum_{j=i}^K c_j(x) - \sum_{j=i+1}^K c_j(x) \right) + \nu(\eta_K) c_K(x) \\ &= \nu(\eta_1) \cdot \sum_{i=1}^K c_i(x) + \sum_{i=2}^K (\nu(\eta_i) - \nu(\eta_{i-1})) \cdot \sum_{j=i}^K c_j(x) \\ &= \nu(\eta_1) \cdot \tilde{c}_1(x) + \sum_{i=2}^K (\nu(\eta_i) - \nu(\eta_{i-1})) \cdot \tilde{c}_i(x). \end{aligned}$$

An illustration for different ways to evaluate $\nu(x)$ is given in Figure 2.

Example 1. We apply the concept of numerical representations to the shortest path problem given in Figure 1. As a motivation, suppose that there are two decision makers A and B who have to select a most preferred path. They would agree, for example, that path x^1 is worse than path x^2 , because x^1 has a dashed-orange edge more than x^2 and, other than that, their outcome vectors are the same. But they do not agree on the question whether x^2 or x^5 is preferred, because decision maker A chooses the numerical representation $\nu_A(\eta_1) = 1$, $\nu_A(\eta_2) = 2$ and $\nu_A(\eta_3) = 5$, while decision maker B chooses $\nu_B(\eta_1) = 2$, $\nu_B(\eta_2) = 3$ and $\nu_B(\eta_3) = 4$. Therefore, decision maker A would prefer path x^5 because $\nu_A(x^5) = 4 < 6 = \nu_A(x^2)$ while decision maker B would prefer path x^2 because $\nu_B(x^2) = 6 < 7 = \nu_B(x^5)$. Hence, the path x^2 does not ordinally dominate the path x^5 , i.e., x^2 is not better than x^5 for all numerical representations. Similarly, the path x^5 does not ordinally dominate the path x^2 since also x^5 is not better than x^2 for all numerical representations.

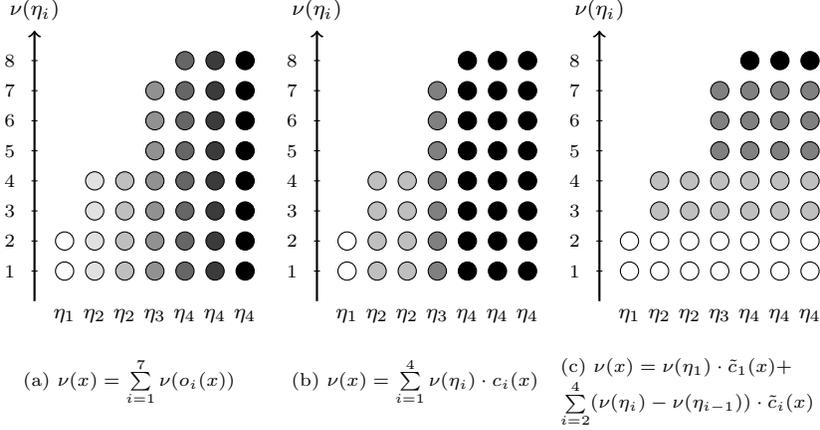


Figure 2: Consider an example with $n = 7$ and $K = 4$. Different ways to compute the numerical value of a feasible solution x with $o(x) = (\eta_1, \eta_2, \eta_2, \eta_3, \eta_4, \eta_4, \eta_4)^\top$, $c(x) = (1, 2, 1, 3)^\top$, $\nu(\eta_1) = 2$, $\nu(\eta_2) = 4$, $\nu(\eta_3) = 7$ and $\nu(\eta_4) = 8$ are illustrated. The different colors represent the summands and visualize the different slicing strategies.

Definition 2 (cf. Schäfer et al., 2021). *Let $x', \hat{x} \in X$ be feasible solutions. Then,*

1. x' weakly ordinally dominates \hat{x} , $o(x')$ weakly ordinally dominates $o(\hat{x})$ and $c(x')$ weakly ordinally dominates $c(\hat{x})$, denoted by $x' \preceq \hat{x}$, $o(x') \preceq o(\hat{x})$, $c(x') \preceq c(\hat{x})$, respectively, if and only if for every $\nu \in \mathcal{V}$, it holds that $\nu(x') \leq \nu(\hat{x})$.
2. x' ordinally dominates \hat{x} , $o(x')$ ordinally dominates $o(\hat{x})$ and $c(x')$ ordinally dominates $c(\hat{x})$, denoted by $x' \preceq \hat{x}$, $o(x') \preceq o(\hat{x})$, $c(x') \preceq c(\hat{x})$, respectively, if and only if x' weakly ordinally dominates \hat{x} and there exists $\nu^* \in \mathcal{V}$ such that $\nu^*(x') < \nu^*(\hat{x})$.
3. $x^* \in X$ is called ordinally efficient, if there does not exist an $x \in X$ such that $x \preceq x^*$.
4. $o(x^*)$ and $c(x^*)$ are called ordinally non-dominated outcome vectors of Problem (OOP) and (OCOP), respectively, if x^* is ordinally efficient.

Optimality by Maximization of Elements in Good Categories. Another optimality concept in ordinal optimization is to maximize the number of elements in good categories. The intuition behind this concept is that solutions with many good elements are to be preferred over solutions with few good elements. The drawback, however, is that this concept rewards solutions with larger numbers of elements as long as these are in (relatively) good categories, which may not be wanted in practice. This optimality concept is defined only for the problem (OCOP), as we need the counting vector c for its definition.

Definition 3. *We say x' weakly head-dominates \hat{x} , denoted by $x' \preceq_h \hat{x}$ or*

$c(x') \geq_h c(\hat{x})$, if and only if

$$\sum_{i=1}^j c_i(x') \geq \sum_{i=1}^j c_i(\hat{x}) \text{ for all } j = 1, \dots, K. \quad (1)$$

Furthermore, x' head-dominates \hat{x} , denoted by $x' \geq_h \hat{x}$ or $c(x') \geq_h c(\hat{x})$, if and only if (1) holds and $c(x') \neq c(\hat{x})$. Moreover, $x^* \in X$ is called head-efficient if there is no $x \in X$ such that $x \geq_h x^*$. The corresponding outcome vector $c(x^*)$ is called head-non-dominated.

Optimality by Minimization of Elements in Bad Categories. The drawback that longer solutions may be preferred over shorter solutions, as long as the elements are in good categories, can be avoided by taking the converse perspective, i.e., when minimizing the number of elements in the bad categories. Again, this optimality concept is defined only for the problem (OCOP).

Definition 4. We say x' weakly tail-dominates \hat{x} , denoted by $x' \leq_t \hat{x}$ or $c(x') \leq_t c(\hat{x})$, if and only if

$$\tilde{c}_j(x') = \sum_{i=j}^K c_i(x') \leq \sum_{i=j}^K c_i(\hat{x}) = \tilde{c}_j(\hat{x}) \text{ for all } j = 1, \dots, K. \quad (2)$$

Again, x' tail-dominates \hat{x} , denoted by $x' \leq_t \hat{x}$ or $c(x') \leq_t c(\hat{x})$, if and only if (2) holds and $c(x') \neq c(\hat{x})$. Moreover, $x^* \in X$ is called tail-efficient if there is no $x \in X$ such that $x \leq_t x^*$. The corresponding outcome vector $c(x^*)$ is called tail-non-dominated.

Remark 5. Note that head-dominance as well as tail-dominance are equivalently defined on the feasible set $X \subseteq 2^S$ and on its image set $c(X) \subseteq \mathbb{R}^K$. The definitions immediately extends to the complete \mathbb{R}^K .

2.3. Properties of and Interrelations between Optimality Concepts for Ordinal Optimization

In addition to the concepts described above, there are further ways to define efficiency. This has been done, for example, in Schäfer et al. (2020) for the ordinal shortest path problem. Their definition has the disadvantage that *Bellman's principle of optimality* (see Bellman, 1957) does not hold in general, i.e., not every subpath of an efficient path is necessarily efficient w.r.t. this optimality concept. The definition of head-optimality has the same disadvantage, see Remark 10 below for more details. In contrast, the definitions of ordinal optimality and tail-optimality can be proven to be equivalent. Moreover, they are compliant with Bellman's principle of optimality. Note that, for the special case of a knapsack problem, this was shown in Schäfer et al. (2021).

Lemma 6. For feasible solutions $\bar{x} = \{\bar{e}_1, \dots, \bar{e}_n\}$, $x' = \{e'_1, \dots, e'_m\} \in X$ with $c_i(\bar{x}) \leq c_i(x')$ for all $i = 1, \dots, K$ and $n < m$ it holds $\bar{x} \preceq x'$.

Proof. First let $\nu \in \mathcal{V}$ be an arbitrary numerical representation. Then $\nu(\bar{x}) = \sum_{i=1}^K \nu(\eta_i) \cdot c_i(\bar{x}) \leq \sum_{i=1}^K \nu(\eta_i) \cdot c_i(x') = \nu(x')$, i.e., $\bar{x} \preceq x'$. Note that since $n < m$, and due to the above assumptions, there must exist a category η_j , $j \in \{1, \dots, K\}$ such that $c_j(\bar{x}) < c_j(x')$. Therefore, there exists a numerical representation ν^* such that $\nu^*(\eta_j) > 0$ and thus $\nu^*(\bar{x}) < \nu^*(x')$, which concludes the proof. \square

Note that the condition of Lemma 6 is always satisfied if $\bar{x}, x' \in X$ with $\bar{x} \subsetneq x'$. Thus, in Example 1 the path x^2 is always preferred over the path x^1 . This is also the case for tail-efficiency, but not for head-efficiency, see Remark 10 below. In many application contexts this is meaningful property, since adding additional elements to a solution (no matter from which category) does generally not improve the solution quality. As an example, we refer again to paths representing bicycle routes as in the app *geovelo*, where we are not interested in routes that are unnecessarily long.

Lemma 7 (Schäfer et al. 2021). *The ordinal dominance relation \preceq defined on the feasible set X is a preorder, i.e., it is reflexive and transitive.*

Proof. Let $x', \hat{x}, \bar{x} \in X$. Obviously, $\nu(x') \leq \nu(x')$ holds for every $\nu \in \mathcal{V}$. Hence, the relation \preceq is reflexive. If $\nu(x') \leq \nu(\hat{x})$ and $\nu(\hat{x}) \leq \nu(\bar{x})$ for every $\nu \in \mathcal{V}$ it follows that $\nu(x') \leq \nu(\bar{x})$ for every $\nu \in \mathcal{V}$ by definition and therefore, we have shown transitivity. \square

Note that the ordinal dominance relation \preceq is in general not antisymmetric on the feasible set X since two different feasible solutions may have the same number of elements in each category like the paths x^3 and x^4 in Example 1.

The following two results show that (weak) ordinal dominance and (weak) tail-dominance are actually equivalent on the feasible set X .

Lemma 8 (Schäfer et al. 2021). *Let $x', \hat{x} \in X$ be two feasible solutions. Then x' weakly ordinally dominates \hat{x} , i.e., $x' \preceq \hat{x}$ if and only if $x' \preceq_t \hat{x}$.*

Proof. The proof is a simplified variant of the proof in Schäfer et al. (2021). Note, that they consider maximization problems while we consider minimization problems.

First we show by contradiction that $x' \preceq \hat{x}$ implies $x' \preceq_t \hat{x}$. Let $x', \hat{x} \in X$ and let $j^* \in \{1, \dots, K\}$ with

$$\sum_{i=j^*}^K c_i(x') > \sum_{i=j^*}^K c_i(\hat{x}).$$

The idea of the proof is to make the bad categories $\eta_{j^*}, \dots, \eta_K$ very expensive, such that an element of this category can not be replaced by elements of the lower categories. Hence, we define the numerical representation

$$\nu(\eta_i) = \begin{cases} i, & \text{if } i < j^* \\ i + 2|\hat{x}|K, & \text{if } i \geq j^*. \end{cases}$$

This implies

$$\begin{aligned}
\nu(x') &\geq 2|\hat{x}|K \cdot \sum_{i=j^*}^K c_i(x') \geq 2|\hat{x}|K \cdot \left(1 + \sum_{i=j^*}^K c_i(\hat{x})\right) \\
&> |\hat{x}|K + 2|\hat{x}|K \cdot \sum_{i=j^*}^K c_i(\hat{x}) \geq \sum_{i=1}^K i c_i(\hat{x}) + 2|\hat{x}|K \cdot \sum_{i=j^*}^K c_i(\hat{x}) \\
&= \sum_{i=1}^{j^*-1} i c_i(\hat{x}) + \sum_{i=j^*}^K i c_i(\hat{x}) + \sum_{i=j^*}^K 2|\hat{x}|K \cdot c_i(\hat{x}) = \nu(\hat{x}).
\end{aligned}$$

For the other direction we use the reformulation of $\nu(x)$, which is visualized in Figure 2(c). It follows that for any $\nu \in \mathcal{V}$

$$\begin{aligned}
\nu(x') &= \nu(\eta_1) \tilde{c}_1(x') + \sum_{i=2}^K (\nu(\eta_i) - \nu(\eta_{i-1})) \tilde{c}_i(x') \\
&\leq \nu(\eta_1) \tilde{c}_1(\hat{x}) + \sum_{i=2}^K (\nu(\eta_i) - \nu(\eta_{i-1})) \tilde{c}_i(\hat{x}) = \nu(\hat{x}).
\end{aligned}$$

The inequality holds because of the assumption $\tilde{c}_j(x') \leq \tilde{c}_j(\hat{x})$ for all $j = 1, \dots, K$ and $\nu(\eta_i) - \nu(\eta_{i-1}) > 0$ for all $\nu \in \mathcal{V}$ and $i = 2, \dots, K$. Hence, we have shown that $\sum_{i=j}^K c_i(x') \leq \sum_{i=j}^K c_i(\hat{x})$ for all $j = 1, \dots, K$ implies $x' \preceq \hat{x}$, which concludes the proof. \square

Lemma 9. *Let $x', \hat{x} \in X$ be two feasible solutions. Then x' ordinally dominates \hat{x} , i.e., $x' \preceq \hat{x}$ if and only if $x' \preceq_t \hat{x}$.*

Proof. We first show that $x' \preceq \hat{x}$ implies $x' \preceq_t \hat{x}$. If x' ordinally dominates \hat{x} , then $x' \preceq \hat{x}$ which implies $x' \preceq_t \hat{x}$ due to Lemma 8. It remains to show that $c(x') \neq c(\hat{x})$. As x' ordinally dominates \hat{x} it holds that there is a numerical representation ν^* such that $\nu^*(x') < \nu^*(\hat{x})$. Hence,

$$\begin{aligned}
0 &< \nu^*(\hat{x}) - \nu^*(x') \\
\iff 0 &< \sum_{i=1}^K \nu^*(\eta_i) (c_i(\hat{x}) - c_i(x')).
\end{aligned}$$

Since $\nu(\eta_i) > 0$ for all $i = 1, \dots, K$, it holds $c(x') \neq c(\hat{x})$. Consequently, we have shown that when x' ordinally dominates \hat{x} , then $x' \preceq_t \hat{x}$ holds for all $j = 1, \dots, K$ and $c(x') \neq c(\hat{x})$.

For the other direction it is sufficient to show that $c(x') \neq c(\hat{x})$ implies that there exists a numerical representation $\nu^* \in \mathcal{V}$ such that $\nu^*(x') < \nu^*(\hat{x})$. Let j^* be the largest category such that $c_{j^*}(x') \neq c_{j^*}(\hat{x})$. This implies

$$\sum_{i=j^*}^K c_i(x') < \sum_{i=j^*}^K c_i(\hat{x}).$$

Now the result follows analogously to the proof of Lemma 8 with exchanged roles of x' and \hat{x} . \square

Remark 10. Lemma 8 (and thus also Lemma 9) does not hold in general for the relation \geq_h . As a counter example, consider the paths x^1 and x^2 with $c(x^1) = (1, 1, 1)^\top$ and $c(x^2) = (1, 0, 1)^\top$ from Figure 1. Obviously, x^1 head-dominates x^2 . But for every numerical representation it follows $\nu(x^1) > \nu(x^2)$, which contradicts $x^1 \preceq x^2$.

Note that the crucial point in the counter example given in Remark 10 is the different cardinality of the solutions, $|x^1| \neq |x^2|$. For ordinal optimization problems with fixed cardinality, for which matroid optimization problems as studied in Klamroth et al. (2021) are an example, it can be shown that head- and tail-dominance are equivalent.

Lemma 11. If all feasible solutions have the same cardinality, i.e., if $|x'| = |\hat{x}|$ for all $x', \hat{x} \in X$, then head- and tail-dominance as defined in Definitions 3 and 4, respectively, are equivalent.

Proof. First assume that x' head-dominates \hat{x} , i.e., inequality (1) is satisfied. We show that then x' also tail-dominates \hat{x} , i.e., inequality (2) holds. Towards this end, let $j \in \{2, \dots, K\}$. Then

$$\begin{aligned} \sum_{i=j}^K c_i(x') &= \sum_{i=1}^K c_i(x') - \sum_{i=1}^{j-1} c_i(x') \stackrel{(1)}{\leq} \sum_{i=1}^K c_i(x') - \sum_{i=1}^{j-1} c_i(\hat{x}) \\ &\stackrel{|x'|=|\hat{x}|}{=} \sum_{i=1}^K c_i(\hat{x}) - \sum_{i=1}^{j-1} c_i(\hat{x}) = \sum_{i=j}^K c_i(\hat{x}), \end{aligned}$$

which implies (2).

Now let x' tail-dominate \hat{x} , i.e., (2) is satisfied. We show that then also x' head-dominates \hat{x} , i.e., (1) holds. Hence, let $j \in \{1, \dots, K-1\}$. Then

$$\begin{aligned} \sum_{i=1}^j c_i(x') &= \sum_{i=1}^K c_i(x') - \sum_{i=j+1}^K c_i(x') \stackrel{(2)}{\geq} \sum_{i=1}^K c_i(x') - \sum_{i=j+1}^K c_i(\hat{x}) \\ &\stackrel{|x'|=|\hat{x}|}{=} \sum_{i=1}^K c_i(\hat{x}) - \sum_{i=j+1}^K c_i(\hat{x}) = \sum_{i=1}^j c_i(\hat{x}), \end{aligned}$$

which implies (1). \square

Lemma 12. The relation \leq_t is a partial order on \mathbb{R}^K , i.e., it is reflexive, transitive and antisymmetric. Moreover, the relation \leq_t is a strict partial order on \mathbb{R}^K , i.e., it is irreflexive and transitive.

Proof. Let $u \in \mathbb{R}^K$. Then $u \leq_t u$, i.e., \leq_t is reflexive. Furthermore, for $u, v, w \in \mathbb{R}^K$ such that $u \leq_t v$ and $v \leq_t w$ it follows that $\sum_{i=j}^K u_i \leq \sum_{i=j}^K v_i \leq \sum_{i=j}^K w_i$

for all $j = 1, \dots, K$, i.e., $u \leq_t w$ which means \leq_t is transitive. To show that the relation \leq_t is antisymmetric, consider two vectors $u, v \in \mathbb{R}^K$ with $u \leq_t v$ and $v \leq_t u$. Then $\sum_{i=j}^K u_i = \sum_{i=j}^K v_i$ for all $j = 1, \dots, K$. This implies that $u = v$ and hence \leq_t is antisymmetric. Therefore, \leq_t is a partial order.

Now consider the relation \leq_t . Since $u \not\leq_t u$ for all $u \in \mathbb{R}^K$, it holds that \leq_t is irreflexive. It remains to show that \leq_t is transitive. Towards this end, consider three vectors $u, v, w \in \mathbb{R}^K$ such that $u \leq_t v$ and $v \leq_t w$. This implies $\sum_{i=j}^K u_i \leq \sum_{i=j}^K v_i \leq \sum_{i=j}^K w_i$ for all $j = 1, \dots, K$, and there exist indices $s, t \in \{1, \dots, K\}$ such that $\sum_{i=s}^K u_i < \sum_{i=s}^K v_i$ and $\sum_{i=t}^K v_i < \sum_{i=t}^K w_i$. Hence, we can conclude that $\sum_{i=j}^K u_i \leq \sum_{i=j}^K w_i$ for all $j = 1, \dots, K$ and $u \neq w$, i.e., $u \leq_t w$. Consequently, we have shown that \leq_t is irreflexive and transitive which concludes the proof. \square

As a consequence of the discussion in Sections 2.2 and 2.3, we focus in the following on the ordinal optimization problem (OOP) w.r.t. optimality by numerical representation, or equivalently, on the ordinal counting optimization problem (OCOP) w.r.t. the concept of tail-dominance.

3. Ordinal Optimality versus Pareto Optimality: An Interpretation Based on Ordering Cones

A prominent example of an order relation in the context of multi-objective optimization is the *component-wise order* or *Pareto order*, see, e.g., Ehrgott (2005). For two vectors $u, v \in \mathbb{R}^K$, we write

$$u \leq v :\iff u_i \leq v_i, \quad i = 1, \dots, K \text{ and } u \neq v, \quad (3)$$

and say that u *Pareto dominates* v if and only if $u \leq v$. The associated weak and strict component-wise orders, respectively, are defined by

$$u \leq v :\iff u_i \leq v_i, \quad i = 1, \dots, K, \quad (4)$$

$$u < v :\iff u_i < v_i, \quad i = 1, \dots, K. \quad (5)$$

Note that \leq defines a partial order in \mathbb{R}^K while \leq as well as $<$ define strict partial orders in \mathbb{R}^K .

In the following, basic properties of orders and their relation to cones are discussed and illustrated at the example of the Pareto order. This leads to a novel perspective on ordinal optimality in comparison to Pareto optimality.

3.1. Orders and Cones

Orders and cones are closely related. The following review of basic concepts relevant in our context is on Ehrgott (2005); Engau (2007); Ziegler (1995).

A *cone* in \mathbb{R}^K is a subset $C \subseteq \mathbb{R}^K$ such that $\lambda u \in C$ for all $u \in C$ and for all $\lambda \in \mathbb{R}, \lambda > 0$. A cone $C \subseteq \mathbb{R}^K$ is called *pointed* if $u \in C$ implies that $(-u) \notin C$ for all $u \neq 0$. Moreover, a cone $C \subseteq \mathbb{R}^K$ is called a *polyhedral cone* if there exists a matrix $A \in \mathbb{R}^{m \times K} \setminus \{0\}$ such that $C = \text{hcone}(A) := \{y \in \mathbb{R}^K : Ay \geq 0\}$. The

rows of the matrix A are normal vectors of hyperplanes, and thus a polyhedral cone can be seen as the finite intersection of m (closed and linear) halfspaces. Polyhedral cones can also be described by their extreme rays. This property is an immediate consequence of the well-known Weyl-Minkowski Theorem:

Theorem 13 (Weyl-Minkowski-Theorem, cf. Ziegler 1995). *A cone $C \subseteq \mathbb{R}^K$ is finitely generated by n vectors in \mathbb{R}^K , i.e.,*

$$C = \text{vcone}(B) := \{B\lambda : \lambda \in \mathbb{R}^n, \lambda \geq 0\} \text{ for some } B \in \mathbb{R}^{K \times n}$$

if and only if it is a finite intersection of m halfspaces in \mathbb{R}^K , i.e.,

$$C = \text{hccone}(A) = \{y \in \mathbb{R}^K : Ay \geq 0\} \text{ for some } A \in \mathbb{R}^{m \times K}.$$

Now let $C \subseteq \mathbb{R}^K$ be a cone. Then the sets

$$\begin{aligned} C^* &:= \{d \in \mathbb{R}^K : d^\top c \geq 0 \text{ for all } c \in C\} \\ C_s^* &:= \{d \in \mathbb{R}^K : d^\top c > 0 \text{ for all } c \in C \setminus \{0\}\} \end{aligned}$$

are called the *dual cone* and the *strict dual cone* of C , respectively.

Every cone $C \subseteq \mathbb{R}^K$ induces a binary (ordering) relation $\mathcal{R} \subseteq \mathbb{R}^K \times \mathbb{R}^K$ by defining that $(u, v) \in \mathcal{R}$ if and only if $(v - u) \in C$. Depending on the context, we also write $u\mathcal{R}v$ whenever $(u, v) \in \mathcal{R}$ (as, for example, in the case of the binary relations defined in (3), (4) and (5) above). Binary relations that are induced by cones are always compatible with scalar multiplication, i.e., $(u, v) \in \mathcal{R}$ implies $(\lambda u, \lambda v) \in \mathcal{R}$ for all $u, v \in \mathbb{R}^K$ and $\lambda > 0$. Moreover, they are compatible with addition, i.e., $(u, v) \in \mathcal{R}$ implies $(u + w, v + w) \in \mathcal{R}$ for all $u, v, w \in \mathbb{R}^K$.

Conversely, binary (ordering) relations that are compatible with scalar multiplication induce cones that represent the respective relation. The following result will be particularly useful in our context.

Lemma 14 (see, e.g., Ehrgott 2005). *Let $\mathcal{R} \subseteq \mathbb{R}^K \times \mathbb{R}^K$ be a binary relation on \mathbb{R}^K which is compatible with scalar multiplication. Then $C_{\mathcal{R}} := \{(v - u) \in \mathbb{R}^K : (u, v) \in \mathcal{R}\}$ is a cone, and $C_{\mathcal{R}}$ induces the binary relation \mathcal{R} . If \mathcal{R} is additionally compatible with addition, then the following statements hold:*

1. $0 \in C_{\mathcal{R}}$ if and only if \mathcal{R} is reflexive.
2. $C_{\mathcal{R}}$ is pointed if and only if \mathcal{R} is antisymmetric.
3. $C_{\mathcal{R}}$ is convex if and only if \mathcal{R} is transitive.

It is easy to see that all three orders (3), (4) and (5) are compatible with scalar multiplication and addition. The component-wise order (3) induces the *Pareto cone* given by $P := \mathbb{R}_{\geq}^K = \{(v - u) \in \mathbb{R}^K : u \leq v\} = \{y \in \mathbb{R}^K : y \geq 0\}$. Similarly, the weak component-wise order (4) induces the *closure of the Pareto cone* given by $\mathbb{R}_{\geq}^K = \{y \in \mathbb{R}^K : y \geq 0\} = P \cup \{0\}$, which is a proper, pointed and convex cone (see Lemma 14). Moreover, it is a polyhedral cone that is defined by the identity matrix, i.e., $P \cup \{0\} = \text{hccone}(I) = \text{vccone}(I)$ (where I is

the $K \times K$ identity matrix). Note also that the Pareto cone is self dual, i.e., $P^* = P$.

Lemma 14 implies that binary relations \mathcal{R} that are compatible with scalar multiplications can be equivalently represented by associated cones $C_{\mathcal{R}}$. This interrelation was used, among others, in Engau (2007) to define the concept of *cone-efficiency* (or $C_{\mathcal{R}}$ -*efficiency*). For a general introduction to ordering cones in the context of vector optimization see, e.g., Tammer & Göpfert (2003); Jahn (2011).

Definition 15 (c.f. Engau 2007). *Let $Y \subset \mathbb{R}^K$ be a nonempty set and let $C_{\mathcal{R}} \subset \mathbb{R}^K$ be a cone induced by a strict partial order $\mathcal{R} \subset \mathbb{R}^K \times \mathbb{R}^K$ (i.e., \mathcal{R} is irreflexive and transitive). Then the sets*

$$\begin{aligned} N(Y, C_{\mathcal{R}}) &:= \{y \in Y : (y - C_{\mathcal{R}}) \cap Y = \emptyset\} \\ N_w(Y, C_{\mathcal{R}}) &:= \{y \in Y : (y - \text{int}(C_{\mathcal{R}})) \cap Y = \emptyset\} \end{aligned}$$

are called the $C_{\mathcal{R}}$ -non-dominated and the weakly $C_{\mathcal{R}}$ -non-dominated set of Y , respectively. The corresponding pre-images $x \in X$ are called $C_{\mathcal{R}}$ -efficient and weakly $C_{\mathcal{R}}$ -efficient, respectively. Thereby, $\text{int}(C_{\mathcal{R}})$ denotes the interior of $C_{\mathcal{R}}$.

Furthermore, we say that u $C_{\mathcal{R}}$ -dominates v if $u \in (v - C_{\mathcal{R}})$, and that u weakly $C_{\mathcal{R}}$ -dominates v if $u \in (v - \text{int}(C_{\mathcal{R}}))$.

Note that when \mathcal{R} is the component-wise order (or Pareto order) defined in (3), then $C_{\mathcal{R}} = P$ is the Pareto cone in \mathbb{R}^K , $N(Y, P)$ is the non-dominated set (or *Pareto set*) of Y , and $N_w(Y, P)$ is the weakly non-dominated set of Y , see, e.g., Ehrgott (2005).

3.2. The Ordinal Cone

In the following we show that tail-dominance, represented by the binary relation \leq_t , is induced by a polyhedral cone in \mathbb{R}^K . We will refer to this cone as the *ordinal cone*. As a first step towards this goal, we prove that \leq_t is compatible with scalar multiplication and addition. As a second step, the associated ordinal cone is constructed and analyzed. Afterwards, we can reinterpret problem (OCOP) based on cone optimality.

Lemma 16. *The relation \leq_t is compatible with scalar multiplication and with addition.*

Proof. To show that \leq_t is compatible with scalar multiplication, let $\lambda > 0$ and $u, v \in \mathbb{R}^K$ with $u \leq_t v$. It follows that $\lambda \sum_{i=j}^K u_i \leq \lambda \sum_{i=j}^K v_i$ which implies $\sum_{i=j}^K \lambda u_i \leq \sum_{i=j}^K \lambda v_i$ for all $j = 1, \dots, K$. Furthermore, it holds that $\lambda u \neq \lambda v$, and hence $\lambda u \leq_t \lambda v$, which implies that \leq_t is compatible with scalar multiplication.

It remains to show that \leq_t is also compatible with addition. Let $u, v, w \in \mathbb{R}^K$ with $u \leq_t v$, i.e., $\sum_{i=j}^K u_i \leq \sum_{i=j}^K v_i$ for all $j = 1, \dots, K$ and $u \neq v$. This implies that $\sum_{i=j}^K (u_i + w_i) \leq \sum_{i=j}^K (v_i + w_i)$ for all $j = 1, \dots, K$ and $(u + w) \neq (v + w)$, i.e., $u + w \leq_t v + w$. Hence we have proven the compatibility with addition. \square

It can be proven analogously that \leq_t is also compatible with scalar multiplication and addition.

From Lemma 14 and Lemma 16 we can conclude that the cone $C_{\leq_t} := \{(v-u) \in \mathbb{R}^K : u \leq_t v\}$ induced by the strict partial order \leq_t is pointed, convex and it does not contain 0. We call this cone the *ordinal cone* to emphasize that C_{\leq_t} equivalently represents ordinal dominance and show that its closure, the cone $C_{\leq_t} \cup \{0\}$, is a polyhedral cone that can be described as the intersection of K halfspaces.

Theorem 17. *The closure of the ordinal cone is a polyhedral cone. In particular, it holds that $C_{\leq_t} \cup \{0\} = \text{hcone}(A_{\leq_t})$ with $A_{\leq_t} \in \mathbb{R}^{K \times K}$ given by*

$$A_{\leq_t} = (a_{ij})_{i,j=1,\dots,K} \text{ with } a_{ij} = \begin{cases} 1, & \text{if } i \leq j \\ 0, & \text{otherwise} \end{cases}, \quad \text{i.e., } A_{\leq_t} = \begin{pmatrix} 1 & \cdots & \cdots & 1 \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Proof. First note that $0 \in (C_{\leq_t} \cup \{0\}) \cap \text{hcone}(A_{\leq_t})$. It thus remains to show that for all $\tilde{u} \in \mathbb{R}^K \setminus \{0\}$, it holds that $\tilde{u} \in C_{\leq_t}$ if and only if $A_{\leq_t} \tilde{u} \geq 0$.

Now let $\tilde{u} \in \mathbb{R}^K \setminus \{0\}$ with $A_{\leq_t} \tilde{u} \geq 0$. We define $u := 0 \in \mathbb{R}^K$ and $v := \tilde{u}$. Hence, it holds $v - u = \tilde{u}$ and

$$\begin{aligned} & A_{\leq_t} \tilde{u} \geq 0 \text{ and } \tilde{u} \neq 0 \\ \iff & \sum_{i=j}^K \tilde{u}_i \geq 0 \quad \text{for all } j = 1, \dots, K, \text{ and } \tilde{u} \neq 0 \\ \iff & \sum_{i=j}^K v_i \geq \sum_{i=j}^K u_i \quad \text{for all } j = 1, \dots, K, \text{ and } v \neq u \\ \iff & u \leq_t v \\ \iff & \tilde{u} = v - u \in C_{\leq_t}. \end{aligned}$$

Thus, we obtain $\text{hcone}(A_{\leq_t}) \setminus \{0\} = C_{\leq_t}$, which concludes the proof. \square

Theorem 13 implies that the closure of the ordinal cone $C_{\leq_t} \cup \{0\}$, which is a polyhedral cone by Theorem 17, must also have a description based on a finite number of extreme rays. Indeed, the following result provides such a description based on exactly K extreme rays.

Theorem 18. *It holds that $\text{hcone}(A_{\leq_t}) = \text{vcone}(B_{\leq_t})$ for A_{\leq_t} defined according to Theorem 17 and $B_{\leq_t} \in \mathbb{R}^{K \times K}$ given by $B_{\leq_t} = (b_{ij})_{i,j=1,\dots,K}$ with*

$$b_{ij} = \begin{cases} 1, & \text{if } i = j \\ -1, & \text{if } i = j - 1, \\ 0, & \text{otherwise} \end{cases}, \quad \text{i.e., } B_{\leq_t} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -1 & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Proof. We first show that $\text{hcone}(A_{\leq t}) \subseteq \text{vcone}(B_{\leq t})$. Let $d \in \text{hcone}(A_{\leq t})$. Hence, it holds $A_{\leq t} d \geq 0$ which is equivalent to $\sum_{i=j}^K d_i \geq 0$ for all $j = 1, \dots, K$. Set $\lambda_j := \sum_{i=j}^K d_i \geq 0$ and let $B_{j\bullet}$ denote the j -th row of $B_{\leq t}$, for $j = 1, \dots, K$. Then $B_{j\bullet} \lambda = \lambda_j - \lambda_{j+1} = d_j$ for $j = 1, \dots, K-1$ and $B_{K\bullet} \lambda = \lambda_K = d_K$. Consequently, we have shown that $d \in \text{vcone}(B_{\leq t})$.

For the other direction, let $d \in \text{vcone}(B_{\leq t})$, i.e., $d = B_{\leq t} \lambda$ for some $\lambda \geq 0$. The definition of $B_{\leq t}$ implies that $\sum_{i=j}^K B_{i\bullet} \lambda = \sum_{i=j}^{K-1} (\lambda_i - \lambda_{i+1}) + \lambda_K = \lambda_j$ for all $j = 1, \dots, K-1$ and $B_{K\bullet} \lambda = \lambda_K$. Hence, it follows that $A_{\leq t} \cdot d = A_{\leq t} \cdot (B_{\leq t} \lambda) = \lambda \geq 0$ and thus $d \in \text{hcone}(A_{\leq t})$, which concludes the proof. \square

Note that these descriptions of the closure of the ordinal cone $C_{\leq t} \cup \{0\}$ are not unique. Indeed, both the normal vectors in $A_{\leq t}$ as well as the extreme rays in $B_{\leq t}$ could be reordered, and they could be multiplied by arbitrary positive scalars without changing the cone that they define. In the particular description given in Theorems 17 and 18, however, we observe that the matrix $B_{\leq t}$ is the inverse of the matrix $A_{\leq t}$, i.e., $(A_{\leq t})^{-1} = B_{\leq t}$.

Remark 19. *It holds $(\text{hcone}(A_{\leq t}))^* = \text{hcone}((B_{\leq t})^\top)$ and $(\text{vcone}(B_{\leq t}))^* = \text{vcone}((A_{\leq t})^\top)$ since the normal vectors of the halfspaces given in $A_{\leq t}$ are orthogonal to the extreme rays contained in $B_{\leq t}$.*

It is easy to see that the Pareto cone, P , is a subset of the ordinal cone, $C_{\leq t}$. Moreover, the dual cone of the ordinal cone is a subset of the Pareto cone, i.e., $(C_{\leq t})^* \subseteq P \subseteq C_{\leq t}$. This holds since $z \in P$ implies that $z \geq 0$ and hence $A_{\leq t} \cdot z \geq 0$, i.e., $z \in C_{\leq t}$. Moreover, $z^* \in (C_{\leq t})^*$ is equivalent to $(z^*)^\top c \geq 0$ for all $c \in C_{\leq t}$ which implies $z_i^* \geq 0$, because the i -th unit vector $e_i \in \mathbb{R}^K$ is contained in $C_{\leq t}$ for all $i = 1, \dots, K$. These cones and their duals are visualized in Figure 3.

Remark 20. *Note that head-dominance (c.f. equation (1)), represented by the binary relation \geq_h , also induces a polyhedral cone. Indeed, we have that $C_{\geq_h} \cup \{0\} = \text{hcone}(A_{\geq_h}) = \text{vcone}(B_{\geq_h})$, where $A_{\geq_h} = A_{\leq t}^\top$ and $B_{\geq_h} = B_{\leq t}$.*

Now we can use the ordinal cone $C_{\leq t}$ as described in Definition 15 to reformulate the optimization problem (OCOP) as follows:

$$\begin{aligned} \min_{C_{\leq t}} \quad & c(x) \\ \text{s. t.} \quad & x \in X. \end{aligned} \tag{OCOP}$$

Here, $\min_{C_{\leq t}}$ denotes the minimization in the sense of Definition 15 for the ordinal cone $C_{\leq t} = \text{hcone}(A_{\leq t}) \setminus \{0\}$. In other words, the $C_{\leq t}$ -non-dominated set of problem (OCOP) is given by $N(Y, C_{\leq t})$, where $Y = c(X)$. In the following, we use this notation to clearly distinguish between the optimization w.r.t. different ordering cones.

3.3. Bijective Linear Transformation Between Ordinal and Pareto Optimization

In the previous subsection we showed that tail-dominance, and hence also ordinal dominance due to Lemma 9, can be equivalently described by the ordinal

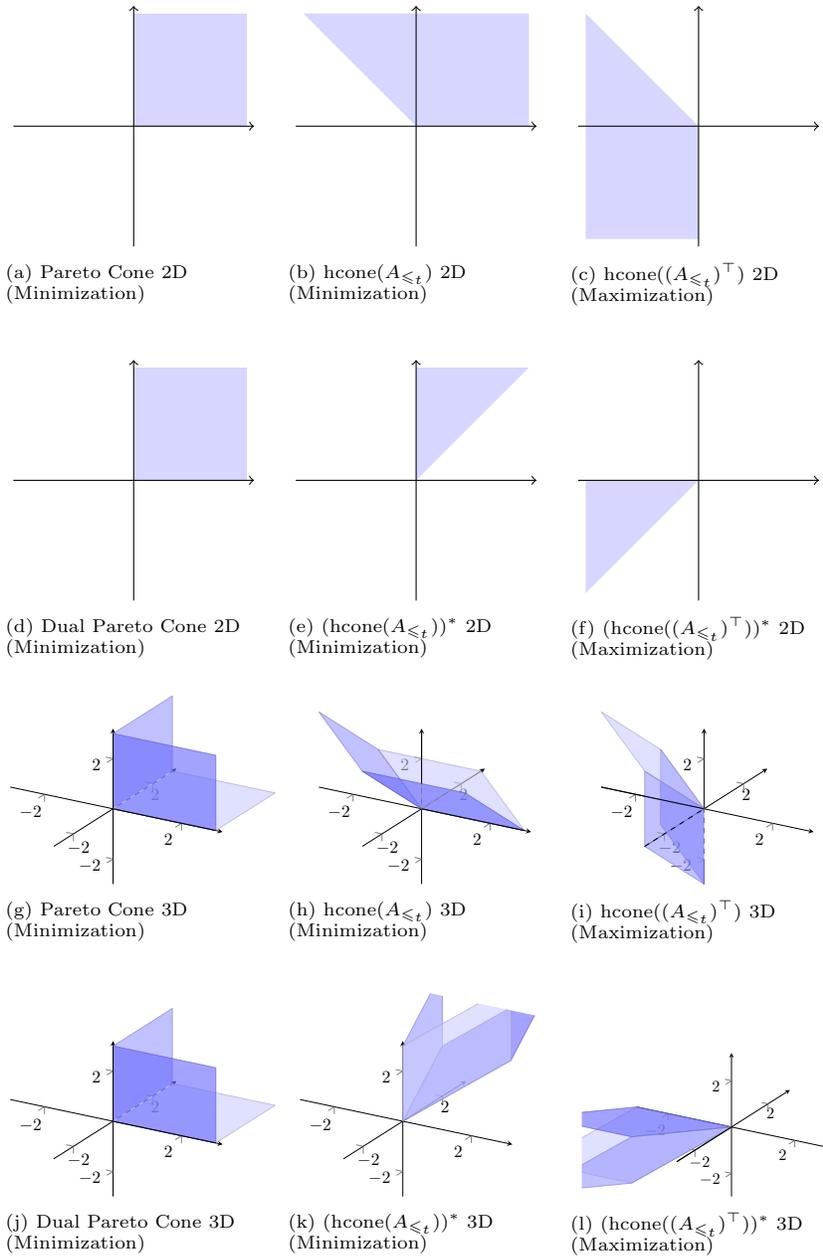


Figure 3: Cones and their dual cones

cone $C_{\leq t}$. Moreover, the closure $C_{\leq t} \cup \{0\}$ of the ordinal cone is the polyhedral cone $\text{hcone}(A_{\leq t}) = \text{vcone}(B_{\leq t})$ that is spanned by K linearly independent extreme rays in \mathbb{R}^K , c.f. Theorems 17 and 18. Since the closure of the Pareto cone $P \cup \{0\}$ is also a polyhedral cone that is spanned by K linearly independent extreme rays in \mathbb{R}^K (namely the K unit vectors in \mathbb{R}^K), there exists a bijective linear transformation that maps the (closure of the) ordinal cone onto the (closure of the) Pareto cone.

We thus define the following *transformed Pareto cone optimization problem* (TOP)

$$\begin{aligned} \min_P \quad & A_{\leq t} \cdot c(x) \\ \text{s. t.} \quad & x \in X, \end{aligned} \tag{TOP}$$

where \min_P denotes the optimization w.r.t. the Pareto cone P according to Definition 15. Note that the objective vector of problem (TOP) corresponds to the incremental tail counting vector $\tilde{c}(x) = A_{\leq t} \cdot c(x) \in \mathbb{R}^K$ introduced in Section 2.1, that counts in its j th component the number of elements of x that are in category η_j or worse. Indeed, for a feasible solution $x = \{e_1, \dots, e_n\} \in X$ we get $\tilde{c}(x) = \sum_{i=1}^n \tilde{c}(e_i)$, where

$$\tilde{c}_j(e_i) = \begin{cases} 1, & \text{if } \eta_j \preceq o(e_i) \\ 0, & \text{otherwise} \end{cases} \quad \text{for all } j = 1, \dots, K.$$

Thus, problem (TOP) is actually a multi-objective optimization problem with K binary objective functions $\tilde{c}_1, \dots, \tilde{c}_K$ defined on the ground set S , and with feasible set $X \subseteq 2^S$. Recall from Section 2.1 that $\tilde{c}_1(e) = 1$ for all $e \in S$ and hence $\tilde{c}_1(x)$ simply counts the number of elements in a solution $x \in X$. Moreover, the vector $\tilde{c}(e)$ has the consecutive ones property in the sense that whenever a component of $\tilde{c}(e)$ is zero, then all subsequent components of $\tilde{c}(e)$ are also zero.

As an example, consider the ordinal shortest path problem introduced in Example 1. The path x^1 consists of the green-dotted edge e_2 with $\tilde{c}(e_2) = (1, 0, 0)^\top$, the orange-dashed edge e_1 with $\tilde{c}(e_1) = (1, 1, 0)^\top$, and the red-solid edge e_5 with $\tilde{c}(e_5) = (1, 1, 1)^\top$. Hence, we compute $\tilde{c}(x^1) = \tilde{c}(e_2) + \tilde{c}(e_1) + \tilde{c}(e_5) = (3, 2, 1)^\top$, see also Figure 1.

In order to show that the ordinal counting optimization problem (OCOP) (and hence the ordinal optimization problem (OOP)) can be solved by using the above transformation to the “standard” multi-objective optimization problem (TOP), we use a classical non-dominance mapping result for polyhedral cones. This result can be found in Engau (2007) and the references therein among several others. We include a proof, which is similar to the more general proof in Hunt & Wiecek (2003), for the sake of completeness.

Theorem 21 (see, e.g., Engau, 2007). *Let $Y \subset \mathbb{R}^K$ be a nonempty set and let $\text{hcone}(A)$ be a cone induced by a matrix $A \in \mathbb{R}^{m \times K}$. Then it holds*

$$A \cdot N(Y, \text{hcone}(A) \setminus \{0\}) \subseteq N(A \cdot Y, P).$$

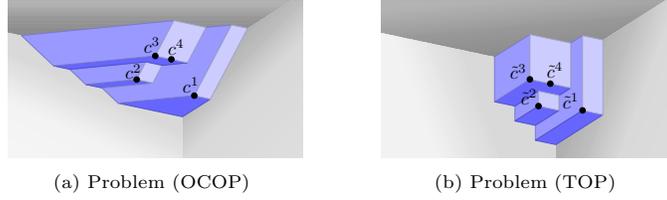


Figure 4: Illustration of the tail-efficient solutions of the instance of (OCOP) introduced in Example 23 (left) and of the respective Pareto-efficient solutions of the transformed problem (TOP) (right). The dominated areas are shown up to the reference points $(4, 4, 4)^\top$ (left) and $(5, 5, 5)^\top$ (right).

If $\text{rank}(A) = K$, then equality holds, i.e., then we have that $A \cdot N(Y, \text{hcone}(A) \setminus \{0\}) = N(A \cdot Y, P)$.

Proof. Suppose that $\bar{y} \in Y$ such that $\bar{y} \in N(Y, \text{hcone}(A) \setminus \{0\})$ and $A \cdot \bar{y} \notin N(A \cdot Y, P)$. Then, by Definition 15, there exists a $\hat{y} \in Y \setminus \{\bar{y}\}$ such that $A \cdot \hat{y} \in (A \cdot \bar{y} - P)$, i.e., there exists $d \in P$ such that $A \cdot \hat{y} = A \cdot \bar{y} - d$. Hence, it follows that $d = A \cdot \bar{y} - A \cdot \hat{y} = A \cdot (\bar{y} - \hat{y}) \geq 0$ and thus $\bar{d} := \bar{y} - \hat{y} \in \text{hcone}(A) \setminus \{0\}$. Finally, we can deduce that $\hat{y} \in (\bar{y} - \text{hcone}(A) \setminus \{0\})$, with $\hat{y} \in Y$. But then $\bar{y} \notin N(Y, \text{hcone}(A) \setminus \{0\})$, which contradicts the assumption.

It remains to show that $A \cdot N(Y, \text{hcone}(A) \setminus \{0\}) \supseteq N(A \cdot Y, P)$ if $\text{rank}(A) = K$. Towards this end, suppose that $\bar{y} \in Y$ such that $A \cdot \bar{y} \in N(A \cdot Y, P)$ and $\bar{y} \notin N(Y, \text{hcone}(A) \setminus \{0\})$. Hence, there exists a $d \in \text{hcone}(A) \setminus \{0\}$ such that $\hat{y} = \bar{y} - d \in Y$. This implies $A \cdot \hat{y} = A \cdot \bar{y} - A \cdot d \in A \cdot Y$. From $\text{rank}(A) = K$ and $d \neq 0$ we deduce that $A \cdot d \neq 0$ and thus $A \cdot d \geq 0$. Consequently, $(A \cdot \bar{y} - P) \cap A \cdot Y \neq \emptyset$ which contradicts the assumption that $A \cdot \bar{y} \in N(A \cdot Y, P)$. \square

Theorem 22. *The set of ordinally efficient solutions for problem (OOP), the set of tail-efficient (ordinally efficient) solutions of problem (OCOP) and the set of Pareto-efficient solutions of problem (TOP) are equal.*

Proof. This follows immediately from Lemma 9, the relation between orders and cones and Theorem 21. \square

Example 23. *Consider an instance of problem (OCOP) with $K = 3$ categories that has four feasible counting vectors $c^1 = (3, 1, 0)^\top$, $c^2 = (0, 2, 1)^\top$, $c^3 = (0, 0, 2)^\top$ and $c^4 = (1, 0, 2)^\top$. The transformation to problem (TOP) yields the corresponding incremental tail counting vectors as $\tilde{c}^1 = (4, 1, 0)^\top$, $\tilde{c}^2 = (3, 3, 1)^\top$, $\tilde{c}^3 = (2, 2, 2)^\top$ and $\tilde{c}^4 = (3, 2, 2)^\top$. The outcome spaces of both formulations are depicted in Figure 4 together with the dominance cones C_{\leq_t} and P , respectively.*

4. Solution Strategies

In this section, we discuss a generic algorithmic framework for ordinal optimization problems that takes advantage of the close relationship to multi-objective optimization problems. Since the weighted sum scalarization is a

popular approach in multi-objective optimization, the interpretation of weights in the context of ordinal optimization and Pareto optimization is analyzed in more detail, and their relation to numerical representations is discussed.

4.1. Ordinal Optimization by Pareto Transformation

From the theory above it follows that we can solve the problems (OCOP) and (OOP) by solving the transformed problem (TOP), which is a standard multi-objective combinatorial problem w.r.t. Pareto optimality. After the computation of the Pareto-efficient set of problem (TOP), or of a minimal complete Pareto-efficient set, respectively, it is necessary to re-compute the corresponding outcome vectors of either problem (OCOP) or (OOP). In this context, a minimal complete Pareto-efficient set of (TOP) is a subset of the Pareto-efficient set that contains one Pareto-efficient solution for each Pareto-non-dominated outcome vector. We refer to Serafini (1987) for different solution concepts in multi-objective optimization. The efficient sets of the problems (TOP), (OCOP) and (OOP) are equal and will be denoted by X_{eff} in the following. The respective non-dominated sets are denoted by $Y_{\text{nd}}^{\text{TOP}} := N(\tilde{c}(X), P)$, $Y_{\text{nd}}^{\text{OCOP}} := N(c(X), C_{\leq t})$ and $Y_{\text{nd}}^{\text{OOP}}$, respectively. A procedure for the computation of the efficient set and the respective non-dominated sets based on this Pareto transformation is outlined in Algorithm 1.

Algorithm 1: Ordinal optimization by Pareto transformation (OOPT)

Input: feasible set $X \subseteq 2^S$ and ordinal function $o : S \rightarrow \mathcal{C}$
Output: efficient set X_{eff} and non-dominated sets $Y_{\text{nd}}^{\text{OCOP}}$ and $Y_{\text{nd}}^{\text{OOP}}$

- 1 Compute $c(x)$ for all $x \in X$ // compute counting objective c
- 2 $X_{\text{eff}} := \min_P \{A_{\leq t} \cdot c(x) : x \in X\}$ // solve lin. transf. (TOP)
- 3 $Y_{\text{nd}}^{\text{OCOP}} := c(X_{\text{eff}})$ // map efficient set to ...
- 4 $Y_{\text{nd}}^{\text{OOP}} := o(X_{\text{eff}})$ // ... resp. obj. spaces
- 5 **return** efficient set X_{eff} and non-dominated sets $Y_{\text{nd}}^{\text{OCOP}}$ and $Y_{\text{nd}}^{\text{OOP}}$

Note that the structural properties of the problems (OCOP) and (OOP) are preserved by the transformation as we do not change the feasible set and as the transformation of the objective function is linear and bijective. In particular, combinatorial solution strategies, like e.g. Bellman’s principle of optimality for knapsack problems, can be applied in step 2 of Algorithm 1 to efficiently compute X_{eff} . Ordinal optimization is thus in general no more complex than standard multi-objective optimization.

4.2. Weighted Sum Scalarization and Ordinal Weight Space Decomposition

In the following we investigate the interrelation between weighted sum scalarizations for (TOP) and (OCOP) and numerical representations for (OOP). Thereby we rely on the concept of weight space decompositions, which were introduced by Benson & Sun (2000) for multi-objective linear programming and extended to integer linear problems in Przybylski et al. (2010).

The *weighted sum scalarization for (TOP)* is

$$\begin{aligned} \min \quad & \sum_{i=1}^K \lambda_i \tilde{c}_i(x) && \text{(WSTOP}(\lambda)) \\ \text{s. t.} \quad & x \in X \end{aligned}$$

with $\lambda_i > 0$ for $i = 1, \dots, K$ and $\sum_{i=1}^K \lambda_i = 1$. Analogously, the *weighted sum scalarization for (OCOP)* can be formulated as

$$\begin{aligned} \min \quad & \sum_{i=1}^K \mu_i c_i(x) && \text{(WSOCOP}(\mu)) \\ \text{s. t.} \quad & x \in X \end{aligned}$$

with $\mu \in (C_{\leq t})_s^*$, where $(C_{\leq t})_s^*$ is the strict dual cone of the ordinal cone $C_{\leq t}$, and $\sum_{i=1}^K \mu_i = 1$. Recall that the strict dual cone $(C_{\leq t})_s^*$ is the interior of the dual cone $(C_{\leq t})^*$, which is visualized in Figures 3(e) and 3(k).

It is a well-known fact that when considering a multi-objective optimization problem, then optimal solutions of weighted sum scalarizations with weighting vectors $\lambda \in \mathbb{R}_{>}^K$ are always Pareto-efficient (see, e.g., Ehrgott, 2005). Such solutions are called *supported* efficient solutions. Thus, problem (WSTOP(λ)) always yields Pareto-efficient solutions for problem (TOP). Since (OCOP) can be interpreted as a multi-objective optimization problems w.r.t. the ordering cone $C_{\leq t}$, every optimal solution of the associated weighted sum problem (WSOCOP(μ)) with weights in the strict dual cone $(C_{\leq t})_s^*$ of $C_{\leq t}$ is ordinally efficient for (OCOP) (see, e.g., Engau, 2007).

The supported efficient solutions of problems (TOP) and (OCOP) are the same, hence there is a one-to-one correspondence between appropriate weighting vectors λ and μ . For a given $\lambda \in \mathbb{R}_{>}^K$ and $x \in X$ we define $\mu_i = \sum_{j=1}^i \lambda_j$ for all $i = 1, \dots, K$. Then it holds that

$$\sum_{i=1}^K \lambda_i \tilde{c}_i(x) = \sum_{i=1}^K \lambda_i \sum_{j=i}^K c_j(x) = \sum_{i=1}^K c_i(x) \sum_{j=1}^i \lambda_j = \sum_{i=1}^K c_i(x) \mu_i,$$

which shows that problems (WSOCOP(μ)) and (WSTOP(λ)) have the same objective functions in this case.

Note that $\mu_i = \sum_{j=1}^i \lambda_j$ and $\lambda_i > 0$ for all $i = 1, \dots, K$ implies that $\mu_i < \mu_j$ for all $i < j$, as required. Conversely, weighting vectors $\mu \in (C_{\leq t})_s^*$ satisfy $\mu_i < \mu_j$ for all $i < j$ and hence yield associated weighting vectors $\lambda \in \mathbb{R}_{>}^K$ by setting $\lambda_1 := \mu_1 > 0$ and $\lambda_i := \mu_i - \mu_{i-1} > 0$ for all $i = 2, \dots, K$. Note also that while the values of $\mu_i = \sum_{j=1}^i \lambda_j$, $i = 1, \dots, K$ (for given $\lambda \in \mathbb{R}_{>}^K$) are in general not normalized to satisfy $\sum_{i=1}^K \mu_i = 1$, such weighting vectors μ can be easily normalized by setting

$$\mu_i := \frac{\sum_{j=1}^i \lambda_j}{\sum_{\ell=1}^K \sum_{j=1}^{\ell} \lambda_j} = \frac{\sum_{j=1}^i \lambda_j}{\sum_{j=1}^K (K - j + 1) \lambda_j}.$$

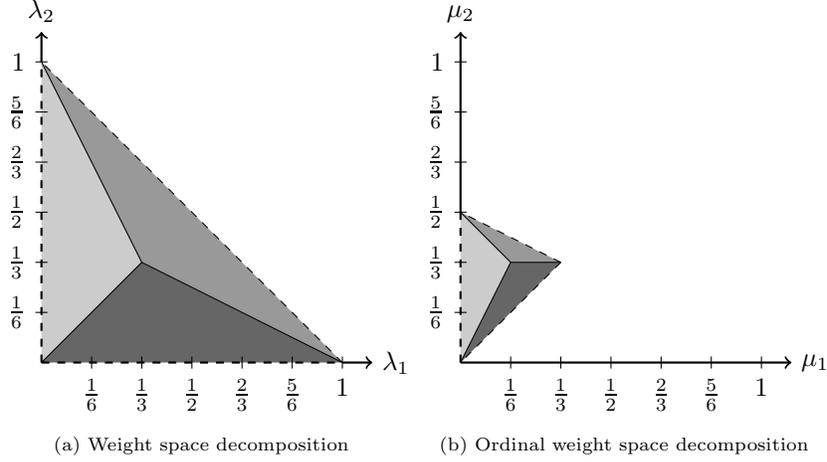


Figure 5: Weight space decomposition and corresponding ordinal weight space decomposition for the shortest path problem given in Example 1. The efficient solution x^2 corresponds to the light grey triangle, both x^3 and x^4 correspond to the middle grey triangle and x^5 corresponds to the dark grey triangle. The values on dashed lines may not be chosen for λ and μ , because for the weight space decomposition we assume that $\lambda \in \mathbb{R}_>^K$ and $\sum_{i=1}^3 \lambda_i = 1$, and for the ordinal weight space decomposition we require $0 < \mu_1 < \mu_2 < \mu_3$ and $\sum_{i=1}^3 \mu_i = 1$.

Note that this normalization is applicable since $(C_{\leq t})_s^* \subset \mathbb{R}_>^K$. As a consequence, a weight space decomposition for the multi-objective problem (TOP) can be translated into an associated *ordinal weight space decomposition* for the ordinal counting optimization problem (OCOP). In this context, a weight space decomposition subdivides the space of relevant weighting vectors $\lambda \in \mathbb{R}_>^K$ with $\sum_{i=1}^K \lambda_i = 1$ into polyhedral cells such that all weighting vectors from the same cell generate the same efficient solution(s).

Example 24. *In the shortest path problem of Example 1 the solutions x^2 , x^3 , x^4 and x^5 are efficient. In Figure 5 the corresponding weight space decomposition is depicted showing the values of λ and μ for which the respective efficient solution is obtained.*

From yet another perspective, weighting vectors $\mu \in (C_{\leq t})_s^*$, i.e., weighting vectors $\mu \in \mathbb{R}^K$ satisfying $0 < \mu_i < \mu_j$ for all $i < j$, are related to numerical representations as introduced in Section 2.2. Indeed, numerical representations assign a numerical value $\nu(\eta_i)$ to every ordinal category η_i , $i = 1, \dots, K$, such that $\nu(\eta_i) < \nu(\eta_j)$ whenever $i < j$. Hence we can choose the values μ_i , $i = 1, \dots, K$, equal to the values $\nu(\eta_i)$ of any numerical representation that satisfies $\nu(\eta_1) > 0$. These values can again be normalized without changing the optimal solutions of (WSOCOP(μ)) by setting

$$\mu_i := \frac{\nu(\eta_i)}{\sum_{j=1}^K \nu(\eta_j)}, \quad i = 1, \dots, K.$$

It is important to note that this does not imply that numerical representations and weighted sum scalarizations are equivalent. Similarly, it is in general not possible to compute all ordinally efficient solutions of problem (OOP) by solving (WSOCOP(μ)) for an appropriate μ . To see this, recall that a solution $x' \in X$ is called ordinally efficient for problem (OOP) if and only if there is no $\hat{x} \in X$ that ordinally dominates x' , i.e., if for every $\hat{x} \in X$ there exists a numerical representation $\nu^{\hat{x}} \in \mathcal{V}$ such that $\nu^{\hat{x}}(x') \leq \nu^{\hat{x}}(\hat{x})$. In contrast, a solution $x' \in X$ is optimal for problem (WSOCOP(μ)) with appropriate μ if and only if there exists a numerical representation $\nu^* \in \mathcal{V}$ such that $\nu^*(x') \leq \nu^*(\hat{x})$ for all $\hat{x} \in X$.

Remark 25. *Ordinal optimization problems may have non-supported efficient solutions. This is illustrated in Example 26. Hence, we can not expect to determine all efficient solutions with the weighted sum method.*

Example 26. *Consider an instance with two categories and three efficient solutions x', \hat{x}, \bar{x} with counting vectors $c(x') = (3, 1)^\top$, $c(\hat{x}) = (5, 0)^\top$ and $c(\bar{x}) = (0, 2)^\top$. The incremental tail counting vectors in the transformed problem (TOP) are $\tilde{c}(x') = (4, 1)^\top$, $\tilde{c}(\hat{x}) = (5, 0)^\top$ and $\tilde{c}(\bar{x}) = (2, 2)^\top$, respectively. Obviously, $\tilde{c}(x')$ is non-dominated in (TOP) but unsupported, and thus x' is not optimal for (WSTOP(λ)), irrespective of the choice of $\lambda \in \mathbb{R}_{>}^K$. Similarly, there is no numerical representation such that x' is simultaneously better than \hat{x} and \bar{x} , i.e., there is no numerical representation ν such that $\nu(x') \leq \nu(\hat{x})$ and $\nu(x') \leq \nu(\bar{x})$. Indeed, the numerical values of the points are $\nu(x') = 3\nu(\eta_1) + \nu(\eta_2)$, $\nu(\hat{x}) = 5\nu(\eta_1)$ and $\nu(\bar{x}) = 2\nu(\eta_2)$. $\nu(x') \leq \nu(\hat{x})$ implies $\nu(\eta_2) \leq 2\nu(\eta_1)$ and $\nu(x') \leq \nu(\bar{x})$ implies $3\nu(\eta_1) \leq \nu(\eta_2)$, which is a contradiction to $\nu(\eta_1) < \nu(\eta_2)$ and $\nu(\eta_i) \geq 0$ for $i = 1, 2$. However, neither \hat{x} nor \bar{x} ordinally dominate x' , i.e., neither \hat{x} nor \bar{x} yield a better objective value for every numerical representation.*

5. Multi-objective Ordinal Optimization

5.1. Conflicting Real-valued Objectives and Ordinal Objectives

The results of Section 3 can be extended to multi-objective optimization problems that combine a finite number of p “standard” real-valued objective functions $w^j : X \rightarrow \mathbb{R}$ with a finite number of r ordinal objective functions $o^l : X \rightarrow \mathcal{C}^l$, $l = 1, \dots, r$, that are in mutual conflict. The number of categories in the l -th ordinal objective function is denoted by K_l , i.e., $\mathcal{C}^l = \{\eta_1^l, \dots, \eta_{K_l}^l\}$ for $l = 1, \dots, r$.

For a feasible solution $x = \{e_1, \dots, e_n\} \in X$, we assume that $w^j(x) := \sum_{i=1}^n w^j(e_i)$, $j = 1, \dots, p$. Moreover, $o^l(x) = \text{sort}(o^l(e_1), \dots, o^l(e_n))$ for $l = 1, \dots, r$. This leads to the *multi-objective ordinal optimization problem with*

additional cost functions (MOOP):

$$\begin{aligned}
& \min_{\mathcal{P}} (w^1(x), \dots, w^p(x))^\top \\
& \min_{\preceq} o^1(x) \\
& \quad \vdots \\
& \min_{\preceq} o^r(x) \\
& \text{s. t. } x \in X.
\end{aligned} \tag{MOOP}$$

By replacing the ordered vectors $o^l(x)$ for $l = 1, \dots, r$ by the counting vectors $c^l(x)$ for $l = 1, \dots, r$ we get a corresponding *multi-objective ordinal counting optimization problem with additional cost functions* (MCOP):

$$\begin{aligned}
& \min_{\mathcal{P}} (w^1(x), \dots, w^p(x))^\top \\
& \min_{C_{\leq t}} c^1(x) \\
& \quad \vdots \\
& \min_{C_{\leq t}} c^r(x) \\
& \text{s. t. } x \in X.
\end{aligned} \tag{MCOP}$$

We denote the concatenated outcome vectors of (MCOP) as

$$v(x) := (w^1(x), \dots, w^p(x), (c^1(x))^\top, \dots, (c^r(x))^\top)^\top \in \mathbb{R}^{p+\tilde{r}},$$

where $\tilde{r} := \sum_{l=1}^r K_l$. Then problem (MCOP) can be transformed into an equivalent standard multi-objective optimization problem w.r.t. Pareto dominance using a linear transformation that is defined by the block diagonal matrix

$$\tilde{A} := \begin{pmatrix} I_{p \times p} & & & \\ & A_{\leq t}^1 & & \\ & & \ddots & \\ & & & A_{\leq t}^r \end{pmatrix}.$$

Here, $I_{p \times p} \in \mathbb{R}^{p \times p}$ denotes the identity matrix and $A_{\leq t}^l$ is the transformation matrix corresponding to the objective c^l for $l = 1, \dots, r$, c.f. Theorem 17. Thus, we get the *multi-objective transformed Pareto cone optimization problem* (MTOP)

$$\begin{aligned}
& \min_{\mathcal{P}} \tilde{A} \cdot v(x) \\
& \text{s. t. } x \in X.
\end{aligned} \tag{MTOP}$$

Now, problem (MOOP) or, equivalently, problem (MCOP) can be solved by using a simple adaptation of Algorithm 1, c.f. Section 4.

Example 27. We consider a problem of type (MCOP) with one real-valued objective w and one counting objective c with $K = 2$ categories (i.e., $p = r = 1$). Consider an instance with four feasible outcome vectors $v = (w, c_1, c_2)^\top$ given by $v^1 = (4, 1, 0)^\top$, $v^2 = (3, 2, 1)^\top$, $v^3 = (2, 0, 2)^\top$ and $v^4 = (3, 0, 2)^\top$.

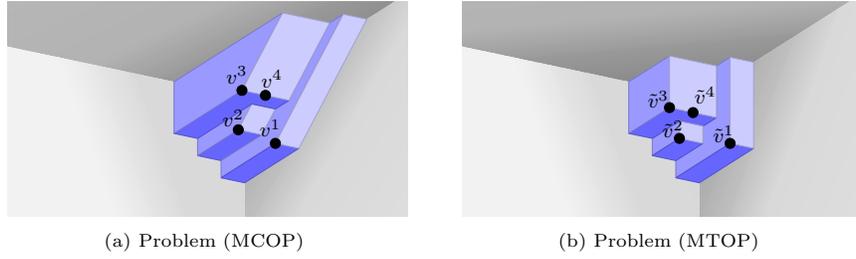


Figure 6: Original and transformed outcome space for the multi-objective problem with one real-valued and one ordinal objective function introduced in Example 27. In both figures the upper corner of the bounding box is located at the point $(5, 5, 5)^\top$.

Then the corresponding outcome vectors of problem (MTOp), $\tilde{v}^i = \tilde{A} v^i$ for $i = 1, \dots, 4$, are obtained as $\tilde{v}^1 = (4, 1, 0)^\top$, $\tilde{v}^2 = (3, 3, 1)^\top$, $\tilde{v}^3 = (2, 2, 2)^\top$ and $\tilde{v}^4 = (3, 2, 2)^\top$. In this case, the transformation matrix is given by

$$\tilde{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The feasible points and the dominated volumes in the respective outcome spaces are depicted in Figure 6 for both problems, (MCOP) and (MTOp).

5.2. Coherent Real-valued Objectives and Ordinal Objectives

In some practical applications, the elements of S have a real-valued cost (e.g., the length of an edge) and an associated category (e.g., the safety of the corresponding road segment for a cyclist) such that the real-valued cost is, rather than in conflict, coherent with the respective category. This situation is illustrated at the following example:

Example 28. Consider the shortest path problem shown in Figure 7. Let $w(e)$ denote the length of an edge e and let $o(e)$ denote its safety: dotted green edges are safe and in category η_1 , while solid red edges are insecure and in category η_2 . Then, irrespective of the number of edges contained in the respective paths, the path $x^1 = \{e_1, e_2, e_3\}$ should be preferred over the path $x^2 = \{e_4, e_5, e_6\}$ since the total weights are equal $w(x^1) = w(x^2) = 10$, and the red sub-path in x^1 has a smaller weight than that of x^2 . In this sense, the weight or length of an edge can be interpreted as an attribute of its respective category. However, x^1 is dominated by x^2 w.r.t. problem (MOOP).

To model the situation where a real-valued objective function $w : S \rightarrow \mathbb{R}$ is in accordance with an ordinal objective function $o : S \rightarrow \mathcal{C}$ with K categories, i.e., the situation where the weight $w(e)$ reflects the multiplicity with which the category $o(e)$ of the element e is to be counted, we introduce a *weighted counting*

vector

$$c_i^w(e) := \begin{cases} w(e) & \text{if } o(e) = \eta_i \\ 0 & \text{otherwise.} \end{cases}$$

The basic idea of this concept is also used in the risk-aware bicycle routing application *geovelo*, which takes, besides the route length, also the total length of unsafe route segments into account. For example, when $K = 4$, $w(e) = 7$ and $o(e) = \eta_2$, then $c^w(e) = (0, 7, 0, 0)^\top$. The weighted counting objective of a feasible solution $x = \{e_1, \dots, e_n\} \in X$ equals the sum of the weighted counting vectors of all elements in x , i.e., $c^w(x) = \sum_{i=1}^n c^w(e_i)$. Thereby, the i -th component of $c^w(x)$ corresponds to the total weight of the elements in x that are in category η_i , $i = 1, \dots, K$. Now the weighted counting vector can be handled analogously to the counting vector. Indeed, as in the previous chapter we consider the transformation $\tilde{c}_i^w(x) := \sum_{j=i}^K c_j^w(x) = A_{\leq i} \cdot c^w(x)$ to obtain the *weighted transformed Pareto cone optimization problem*

$$\begin{aligned} \min_P \quad & \tilde{c}^w(x) \\ \text{s. t.} \quad & x \in X \end{aligned} \tag{WTOP}$$

w.r.t. the concept of Pareto optimality, that can be solved with the methods developed in the preceding sections.

5.3. Modelling Aspects

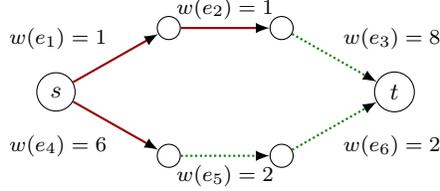
We emphasize that it depends on the context of the respective application whether the multi-objective model (MTOP) or the aggregated model (WTOP) is more suitable. The following example illustrates that the aggregated model (WTOP) is meaningful whenever w and c are interrelated and coherent objectives, while the multi-objective model (MTOP) is particularly useful for unrelated or incompatible objectives.

Example 29. Consider again the shortest path problem depicted in Figure 7. Obviously, the path $x^1 = \{e_1, e_2, e_3\}$ is the unique efficient solution for problem (WTOP), while the path $x^2 = \{e_4, e_5, e_6\}$ is the unique efficient solution for problem (MOOP).

Whether x^1 or x^2 are actually preferred thus depends on the interpretation of the weights and of the ordinal categories. Towards this end, suppose that, as in Example 28, the category of an edge corresponds to its security level.

First, consider the case that the real-valued objective $w(e)$ represents the length of the edge e as in Example 28. Then both objectives are interrelated and hence model (WTOP) is appropriate.

If, on the other hand, $w(e)$ represents the toll of the edge or road e , then this real-valued objective is not an attribute of the corresponding category. In other words, both objectives are potentially conflicting and not coherent. Hence, in this case the path x^2 is preferred since the total amount of toll is the same for both paths, but the second path has more green and fewer red edges.



	w	o	c^w	\tilde{c}^w
$x^1 = \{e_1, e_2, e_3\}$	10	$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_2 \end{pmatrix}$	$\begin{pmatrix} 8 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 10 \\ 2 \end{pmatrix}$
$x^2 = \{e_4, e_5, e_6\}$	10	$\begin{pmatrix} \eta_1 \\ \eta_1 \\ \eta_2 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 6 \end{pmatrix}$	$\begin{pmatrix} 10 \\ 6 \end{pmatrix}$

Figure 7: Instance of a shortest path problem. A dotted-green edge is in the best category η_1 and the solid-red edges are in the worst category η_2 . The possible s - t -paths x^1 and x^2 and their different objective function vectors are given.

6. Conclusion

In this paper we investigate ordinal combinatorial optimization problems. We describe different optimality concepts for ordinal objective functions, namely ordinal optimality, which was introduced first by Schäfer et al. (2021), as well as tail- and head-optimality. We prove that all three concepts are equivalent if all feasible solutions have the same length. In general, only ordinal optimality and tail-optimality are equivalent.

We provide alternative descriptions of these three optimality concepts based on associated ordering cones. Using the fact that ordinal optimality and tail-optimality are equivalent, and that tail-optimality can be represented by a polyhedral cone with K extreme rays in \mathbb{R}^K , we show that ordinal optimization problems can be transformed into equivalent multi-objective optimization problems with binary cost coefficients. The transformation is realized by a bijective linear mapping. The resulting problem can be solved with standard methods from multi-objective optimization, and hence ordinal optimization is as easy or hard as the associated, “standard” multi-objective problems. For example, ordinal knapsack problems and ordinal shortest path problems can be solved by multi-objective dynamic programming, using Bellman’s principle of optimality.

The results can be extended to problems with more than one objective function. We suggest two modelling approaches to combine an ordinal objective function with a real-valued objective function. While in the first approach all objectives are considered in a standard multi-objective setting, the second approach allows to model interrelated and coherent objective functions, where the real-valued objective is interpreted as an attribute of the respective category in

the ordinal objective.

Future work should focus on the development of tailored optimization algorithms for the associated multi-objective optimization problems that exploit the fact that these problems have binary cost coefficients. Moreover, specific combinatorial problems like, for example, shortest path, knapsack, assignment and general routing and network flow problems should be analyzed both with coherent and with conflicting ordinal and real-valued objective functions.

Acknowledgment

The authors thankfully acknowledge financial support by Deutsche Forschungsgemeinschaft, project number KL 1076/11-1.

References

- Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.
- Benson, H. P., & Sun, E. (2000). Outcome space partition of the weight set in multiobjective linear programming. *Journal of Optimization Theory and Applications*, *105*, 17–36.
- Ehrgott, M. (2005). *Multicriteria Optimization*. Berlin, Heidelberg: Springer. doi:10.1007/3-540-27659-9.
- Engau, A. (2007). *Domination and decomposition in multiobjective programming*. Ph.D. thesis Clemson University.
- Fishburn, P. (1999). Preference structures and their numerical representations. *Theoretical Computer Science*, *217*, 359–383. doi:10.1016/S0304-3975(98)00277-1.
- geovelo (2022). Geovelo, an application for short and safe bicycle route computations. <https://geovelo.fr>. Accessed: 2022-01-28.
- Hunt, B. J., & Wiecek, M. M. (2003). Cones to aid decision making in multicriteria programming. In *Multi-Objective Programming and Goal Programming* (p. 153–158). Springer Berlin Heidelberg. doi:10.1007/978-3-540-36510-5.
- Jahn, J. (2011). *Vector optimization: Theory, Applications, and Extensions*. (2nd ed.). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-17005-8.
- Kergosien, Y., Giret, A., Néron, E., & Sauvanet, G. (2021). An efficient label-correcting algorithm for the multiobjective shortest path problem. *INFORMS Journal on Computing*, . doi:10.1287/ijoc.2021.1081.
- Klamroth, K., Stiglmayr, M., & Sudhoff, J. (2021). Multi-objective matroid optimization with ordinal weights, . arXiv:2109.13804. Submitted to European Journal of Operational Research.

- Przybylski, A., Gandibleux, X., & Ehrgott, M. (2010). A recursive algorithm for finding all nondominated extreme points in the outcome set of a multiobjective integer programme. *INFORMS Journal on Computing*, 22, 371–86. doi:10.1287/ijoc.1090.0342.
- Sauvanet, G., & Néron, E. (2010). Search for the best compromise solution on multiobjective shortest path problem. *Electron. Notes Discrete Math.*, 36, 615–22. doi:10.1016/j.endm.2010.05.078.
- Schäfer, L. E., Dietz, T., Barbati, M., Figueira, J. R., Greco, S., & Ruzika, S. (2021). The binary knapsack problem with qualitative levels. *European Journal of Operational Research*, 289, 508–514. doi:10.1016/j.ejor.2020.07.040.
- Schäfer, L. E., Dietz, T., Fröhlich, N., Ruzika, S., & Figueira, J. R. (2020). Shortest paths with ordinal weights. *European Journal of Operational Research*, 280, 1160–1170. doi:10.1016/j.ejor.2019.08.008.
- Serafini, P. (1987). Some considerations about computational complexity for multi objective combinatorial problems. In J. Jahn, & W. Krabs (Eds.), *Recent Advances and Historical Development of Vector Optimization* (pp. 222–32). Springer volume 294 of *Lecture Notes in Economics and Mathematical Systems*. doi:10.1007/978-3-642-46618-2.
- Tammer, C., & Göpfert, A. (2003). Theory of vector optimization. In *Multiple Criteria Optimization: State of the Art Annotated Bibliographic Surveys* (pp. 1–70). Springer US. doi:10.1007/0-306-48107-3_1.
- Ziegler, G. M. (1995). *Lectures on Polytopes*. Springer New York. doi:10.1007/978-1-4613-8431-1.