On the number of non-dominated points of a multicriteria optimization problem*

Cristina Bazgan[1,2] Florian Jamain[1] Daniel Vanderpooten[1]

1. PSL, Université Paris-Dauphine, LAMSADE UMR 7243
Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France
2. Institut Universitaire de France
{bazgan,florian.jamain,vdp}@lamsade.dauphine.fr

Abstract

This work proposes an upper bound on the maximal number of non-dominated points of a multicriteria optimization problem. Assuming that the number of values taken on each criterion is known, the criterion space corresponds to a comparability graph or a product of chains. Thus, the upper bound can be interpreted as the stability number of a comparability graph or, equivalently, as the width of a product of chains. Standard approaches or formulas for computing these numbers are impractical. We develop a practical formula which only depends on the number of criteria. We also investigate the tightness of this upper bound and the reduction of this bound when feasible, possibly efficient, solutions are known.

Keywords: multicriteria optimization, non-dominated points, comparability graph, stability number, product of chains, Sperner property.

1 Introduction

In multicriteria optimization, in opposition to single criterion optimization, there is typically no optimal solution i.e. one that is best for all the criteria. Therefore, the standard situation is that any solution can always be improved on at least one criterion. The solutions of interest, called *efficient* solutions, are those such that any other solution which is better on one criterion is necessarily worse on at least one other criterion. In other words, a solution is efficient if its corresponding vector of criterion values is not dominated by any other vector of criterion values corresponding to a feasible solution. These vectors, associated to efficient solutions, are called *non-dominated points*. For many multicriteria optimization problems, one of the main difficulties is the large cardinality of the set of non-dominated points, and the even larger cardinality of the set of efficient solutions (considering that several solutions can have the same image in the criterion space). However, similarly to single criterion optimization where we usually look for one among all optimal solutions, we usually look for all non-dominated points and a corresponding efficient solution for each such point. Thus, we can restrict our

*This work is supported by the project ANR-09-BLAN-0361 "GUaranteed Efficiency for PAReto optimal solutions Determination (GUEPARD)"

study to the set of non-dominated points. Even with this restriction, it is well-known, that most multicriteria combinatorial optimization problems are *intractable*, in the sense that they admit families of instances for which the number of non-dominated points is exponential in the size of the instance [4]. This situation arises when the number of values taken on each criterion is itself exponential in the size of the instance. It is thus interesting to investigate the number of non-dominated points when we know (or have an upper bound on) the number of values taken on each criterion. This problem can be stated within different theoretical frameworks. Using graph theory, the maximal cardinality of a set of non-dominated points corresponds to the *stability number* of a given graph. Using ordered set theory, this maximal cardinality corresponds to the *width* of a *product of chains*. These two frameworks provide different insights on our problem.

Up to our knowledge, this problem has not been dealt with, except very recently by Stanojević et al. in [9]. The best bound they give is obtained by a recursion formula which is well-known in ordered set theory [8] and that we recall in our Proposition 1. Unfortunately, this formula becomes quickly impractical when the number of values on each criterion increases. One of our purposes is to provide an alternative formula which does not depend on these numbers.

In the following section, we define the basic concepts and formalize the problem both in the context of graphs and ordered sets. Then, in section 3, we deal with simple cases and provide, in the general case, a formula using a combinatorial version of the inclusion-exclusion principle [3]. The time for computing this formula is only exponential in the number of criteria. We also make comparisons with other bounds which are easier to compute. In section 4, we show that the proposed bound is tight for many classical multicriteria optimization problems. In section 5, we try to reduce the maximal number of non-dominated points using known feasible solutions, possibly efficient. We conclude with some possible extensions to this work.

2 Basic concepts and problem statements

2.1 Basic concepts

In this paper, we consider multicriteria optimization problems formulated as:

$$\min_{x \in S}\{f_1(x), \ldots, f_p(x)\}, \tag{1}$$

where $f_1, \ldots, f_p$ are $p \geq 2$ criteria functions to be minimized and S is the set of feasible solutions.

We distinguish the decision space X which contains the set S of feasible solutions from the criterion space $Y \subseteq R^p$ which contains the criterion vectors associated to these solutions. We denote by $f(x) = (f_1(x), \ldots, f_p(x))$ the feasible point associated to a feasible solution $x \in S$, and by $Z = f(S)$ the set of images of the feasible solutions. We define in the criterion space Y, the following partial strict order, denoted by $\leq$, such that for any $y, y' \in Y$, $y \leq y'$ if $y_i \leq y'_i$ for all $i \in \{1, \ldots, p\}$ and $y \neq y'$. Relation $\leq$ corresponds to the standard dominance relation used in multicriteria optimization.

Then we define *efficient* solutions and *non-dominated* points, respectively, in the decision space X and in the criterion space Y, as follows:

Definition 1. *A feasible solution $x \in S$ is called* efficient *if there is no other feasible solution $x' \in S$ such that $f(x') \leq f(x)$. We denote by S_{Eff} the set of efficient solutions. If x is efficient, $f(x)$ is a* non-dominated point *in the criterion space, and let $Z_{ND} = f(S_{Eff})$.*

In this context formulation (1) means that we aim at generating the set of all non-dominated points and a corresponding efficient solution for each such point.

In this paper, we assume that f_i can take up to $c_i + 1$ values, where c_i is a nonnegative integer. Thus, we consider, without loss of generality, that each f_i can take integer values between 0 and c_i, $i = 1, \ldots, p$.

In some cases, the c_i values are known precisely, e.g. for qualitative criteria which take values on a scale whose grades correspond to predefined judgements. In other cases, these values can only be approximated. For instance, assuming that criterion functions are integer-valued, we can find an upper bound on c_i by computing the coordinates of the *ideal* and *anti-ideal* points, corresponding, respectively, to the best and the worst possible values on each criterion. Better bounds can be given if we can compute the coordinates of the *nadir* point, which corresponds to the worst possible values over the set of non-dominated points. Unfortunately, this is not easy in general, especially when the number of criteria is at least 3 [5].

The problem of determining the maximum cardinality of the non-dominated set can be stated as follows.

Max SizeND
Input: an integer p and p integers c_i, $i = 1, \ldots, p$.
Output: maximum cardinality of the non-dominated set Z_{ND} associated to a set Z of p-dimensional points such that at most $c_i + 1$ values are taken on the i^{th} dimension, $i = 1, \ldots, p$.

Let $(\overline{c_i + 1}) = \{0, \ldots, c_i\}$, $i = 1, \ldots, p$ and $P = (\overline{c_1 + 1}) \times \ldots \times (\overline{c_p + 1})$. Any relevant set Z, and in particular any of those leading to a non-dominated set of maximum cardinality, is included in P.

2.2 Statement as a graph theory problem

Consider the graph $G = (P, E)$ whose set of vertices is $P = (\overline{c_1 + 1}) \times \ldots \times (\overline{c_p + 1})$ and set of edges is $E = \{(u, v) \in P \times P : u \leq v\}$. By construction, G is a *comparability* graph (i.e. a graph that admits a transitive orientation), since relation $\leq$ is transitive.

In this context, determining the maximum number of non-dominated points amounts to determining the maximum cardinality of a stable set in G, i.e. computing $\alpha(G)$, the stability number of G. It is well-known that $\alpha(G)$ can be determined in polynomial time when G is a comparability graph [6]. In our case, this is achieved by computing a minimum flow in the digraph $G' = (P, \leq)$ from $(0, \ldots, 0)$ to $(c_1, \ldots, c_p)$ where each vertex has a lower bound of 1. Then $\alpha(G)$ corresponds to the value of this minimum flow in G'.

Computing a minimum flow in G' can be performed in a time polynomial in the number of vertices $P = \prod_{i=1}^{p}(c_i + 1)$. Since the input of Max SizeND is not G' but only values $c_1, \ldots, c_p$, which are encoded in binary, this approach only gives us a pseudo-polynomial time algorithm to solve Max SizeND.

2.3 Statement as an ordered set theory problem

Given a partially ordered set (S, R), we recall that a *chain* is a totally ordered subset and an *antichain* is a subset whose elements are pairwise incomparable. Moreover, the *height* of

(S,R), denoted by $h(S)$, is the maximal cardinality of a chain in S, and the *width* of (S,R), denoted by $\alpha(S)$, is the maximal cardinality of an antichain in S. (S,R) is said to be *ranked* if we can define a function r such that for any $x,y \in S$, whenever xRy and there is no element $z \in S$ such that $xRzRy$, we have $r(y) = r(x)+1$. Calling L_k the level of rank k in S, i.e. the subset of elements of S with rank k, we define $n_k = |L_k|$ and $\sigma(S) = \max n_k$. Since the levels are antichains, we have $\alpha(S) \geq \sigma(S)$. Finally, a partially ordered ranked set S is said to satisfy the Sperner property if $\alpha(S) = \sigma(S)$ [2].

In our case, $P = (\overline{c_1+1}) \times \ldots \times (\overline{c_p+1})$, which is a product of chains, is partially ordered by the dominance relation $\leq$. The resulting partially ordered set $(P, \leq)$ has height $h(P) = \sum_{k=1}^{p} c_k$ (denoted for short by h in the following). Moreover, $(P, \leq)$ can be ranked using rank function r which associates to each element $(y_1, \ldots, y_p) \in P$ its rank $r(y_1, \ldots, y_p) = \sum_{k=1}^{p} y_k$.

In this context, solving Max SizeND is equivalent to determining the width $\alpha(P)$. We recall the following result.

Theorem 1. *(De Bruijn et al. [1]) A product of chains satisfies the Sperner property.*

Therefore, since P is a product of chains, we have $\alpha(P) = \sigma(P)$. Thus, we are interested in determining the cardinality of a level of P which has the largest number of elements. It is well-known that the levels of maximum cardinality are all centered around the level $L_{h/2}$ if h is even, and the levels $L_{(h-1)/2}$ and $L_{(h+1)/2}$ if h is odd [2, 8]. Thus, determining $\alpha(P)$ amounts to computing $n_{\lfloor \frac{h}{2} \rfloor}$.

Leclerc [8] and Caspar et al. [2] proposed induction formulas to compute $n_{\lfloor \frac{h}{2} \rfloor}$. Nevertheless, these induction formulas depend on p but also on the values c_i, $i = 1, \ldots, p$. Since the values c_i may often be large, these formulas are not really usable in practice, as acknowledged by the previous authors. This is, however, another pseudo-polynomial time method to solve Max SizeND. The motivation is to obtain a new more practical formula, the complexity of which does not depend on the values c_i, that is a strongly polynomial time algorithm.

3 Computation of the width of a product of chains

We first provide an upper bound on the width of a product of chains P, showing that this bound is tight in a special case, which includes the bicriteria case. Then, we propose and compare two formulas for computing exactly $\alpha(P)$.

We assume w.l.o.g. that the criteria are numbered by non-increasing order of values c_i, that is $c_1 \geq ... \geq c_p$.

3.1 A simple upper bound on $\alpha(P)$

A first simple upper bound on $\alpha(P)$ is given by the following result.

Lemma 1. $\alpha(P) \leq \prod_{i=2}^{p}(c_i+1)$.

Proof: By contradiction, if $\alpha(P) > \prod_{i=2}^{p}(c_i+1)$ there exist at least two non-dominated points with the same values on criteria f_i, $i = 2, \ldots, p$. Then, among these two points, the point with a worse value on f_1 is dominated by the other one. □

This upper bound is tight in a particular case, as shown in the following lemma:

Lemma 2. $\alpha(P) = \prod_{i=2}^{p}(c_i + 1)$ *if and only if* $c_1 \geq \sum_{i=2}^{p} c_i$.

Proof:

$\Leftarrow$ If $c_1 \geq \sum_{i=2}^{p} c_i$ then all possible $\prod_{i=2}^{p}(c_i + 1)$ configurations on the last $p-1$ criteria can be completed on criterion f_1 so as to define non-dominated points. Indeed, any point with value v_j on criterion f_j, $j = 2, \dots, p$ is non-dominated if it is assigned the (nonnegative) value $\sum_{i=2}^{p} c_i - \sum_{i=2}^{p} v_j$ on criterion f_1.

$\Rightarrow$ If $\alpha(P) = \prod_{i=2}^{p}(c_i + 1)$, all possible configurations on the last $p-1$ criteria must correspond to non-dominated points. In particular, the $\sum_{i=2}^{p} c_i + 1$ following configurations, which constitute a chain on the last $p-1$ criteria, must correspond to non-dominated points:

$(*, 0, \dots, 0), (*, 1, 0, \dots, 0), \dots, (*, c_2, 0, \dots, 0),$
$(*, c_2, 1, 0, \dots, 0), \dots, (*, c_2, c_3, 0, \dots, 0),$
$\dots$
$(*, c_2, c_3, \dots, c_{p-1}, 1), \dots, (*, c_2, c_3, \dots, c_{p-1}, c_p)$

For this chain on the last $p-1$ criteria to become an antichain on the p criteria, we need $\sum_{i=2}^{p} c_i + 1$ different values on criterion f_1, and thus $c_1 \geq \sum_{i=2}^{p} c_i$. $\square$

In the particular case where $p = 2$, we obtain the following corollary, since $c_1 \geq c_2$.

Corollary 1. *If* $p = 2$, *we have* $\alpha(P) = c_2 + 1$.

3.2 Exact computation of $\alpha(P)$

Since P satisfies the Sperner property, we noticed at the end of section 2.3 that $\alpha(P) = n_{\lfloor \frac{h}{2} \rfloor}$. We first review a well-known recursion formula for computing $n_{\lfloor \frac{h}{2} \rfloor}$, which is not practicable as values c_i grow. Then, we propose an alternative analytical formula, which is shown to be much easier to implement.

3.2.1 A recursion formula

As indicated in [8], the following result is known from "folklore".

Proposition 1. *Let* $P' = P \times (\overline{c+1})$ *where* P *is a product of chains. The values* n'_k, *the size of level of rank* k *in* P', *can be obtained from values* n_k *by the following recursion:*

$$n'_k = \sum_{i=0}^{c} n_{k-i}$$

which can be rewritten as

$$n'_k = n'_{k-1} + n_k - n_{k-c-1} \tag{2}$$

where $n_k = 1$, *for all* $k \geq 0$ *when* P *is a chain and* $n_k = 0$ *for* $k < 0$.

As outlined in [8], this recursion is relevant in practice only for a small number of criteria and small values c_i. More precisely, the complexity of this induction formula is given by the following result.

Lemma 3. *The computation of the width of the product of chains* $P = (\overline{c_1 + 1}) \times \dots \times (\overline{c_p + 1})$ *using formula (2) of Proposition 1 is done in* $\Theta(p^2 c_{max})$ *operations, where* $c_{max} = \max\{c_1, \dots, c_p\}$.

Proof: At each step i of the recursion for $i = 1, \ldots, p$, the computation of the cardinality of $((\sum_{j=p-(i-1)}^{p} c_j) + 1)/2$ levels is needed. Since each of these cardinalities is computed in constant time, the computation of $\alpha(P)$ is performed in $((\sum_{i=1}^{p} ic_i) + p)/2 = \Theta(p^2 c_{max})$ operations. $\square$

Observe additionally that these recursions require to keep in memory all the sizes of the levels of the previous step, which requires a space $\Theta(c_{max})$. In most multiple criteria problems, the number of criteria is rather small and can thus assumed to be constant. On the other hand, values c_i may be rather large. This makes this recursion quickly useless. This is the motivation to obtain a formula computing the width of P whose complexity does not depend on the values c_i.

3.2.2 An analytical formula

We need to compute the number of points on a level of maximum cardinality, which amounts to computing the number of integer solutions of the equation

$$x_1 + \ldots + x_p = k \tag{3}$$

with $k = \lfloor h/2 \rfloor$, under the constraints $0 \leq x_i \leq c_i$.

We recall the following result, presented in standard textbooks on combinatorics such as [3], which is a combinatorial version of the inclusion-exclusion principle.

Lemma 4. *The number of integer solutions of equation (3) with the restrictions*

$$s_i \leq x_i \leq m_i, \quad i = 1, \ldots, p$$

where s_i and m_i are given for $i = 1, \ldots, p$ with $s \leq k \leq m$, $s = s_1 + \ldots + s_p$ and $m = m_1 + \ldots + m_p$, with $u_i = m_i - s_i \geq 0$, $i = 1, \ldots, p$ is given by

$$\binom{p+k-s-1}{p-1} + \sum_{r=1}^{p} (-1)^r \sum_{I \subseteq \{1,\ldots,p\}:|I|=r} \binom{p+k-s-\sum_{i\in I} u_i - r - 1}{p-1}$$

Applied in our context, the previous lemma gives the following result.

Theorem 2. *The width $\alpha(P)$ of a product of chains $P = (\overline{c_1+1}) \times \cdots \times (\overline{c_p+1})$ is given by the following formula:*

$$\alpha(P) = \sum_{I \subseteq \{1,\ldots,p\}:|I| \leq \lfloor \frac{h}{2} \rfloor - c_I} (-1)^{|I|} \prod_{k=1}^{p-1} \left(1 + \frac{\lfloor h/2 \rfloor - c_I - |I|}{k}\right) \tag{4}$$

where $c_I = \sum_{i\in I} c_i$ and $c_\varnothing = 0$.

Proof: Using the formula of Lemma 4 with $s_i = 0$ and $m_i = c_i$ for $i = 1, \ldots, p$ we obtain the following formula:

$$\alpha(P) = \binom{p+\lfloor h/2 \rfloor - 1}{p-1} + \sum_{r=1}^{p} (-1)^r \sum_{I \subseteq \{1,\ldots,p\}:|I|=r} \binom{p+\lfloor h/2 \rfloor - \sum_{i\in I} c_i - r - 1}{p-1}$$

Combining the two members of this formula we have:

$$\alpha(P) = \sum_{I \subseteq \{1,\ldots,p\}: |I| \leq \lfloor \frac{h}{2} \rfloor - c_I} (-1)^{|I|} \binom{p + \lfloor h/2 \rfloor - c_I - |I| - 1}{p-1}$$

where $c_I = \sum_{i \in I} c_i$ and $c_\varnothing = 0$, which can be rewritten as (4) using $\binom{n}{t} = \frac{1}{t!}(n-t+1)\ldots n$. □

In the particular case where $p = 3$, the formula can be simplified as follows.

Corollary 2. *If $p = 3$, we have*

$$\alpha(P) = \begin{cases} (c_2+1)(c_3+1) & \text{if } c_1 \geq c_2 + c_3 \\ 1 + \left(\frac{h}{2}\right)^2 + \frac{h}{2} - \frac{c_1^2+c_2^2+c_3^2}{2} & \text{if } c_1 < c_2 + c_3 \text{ and } h = c_1 + c_2 + c_3 \text{ is even} \\ \frac{1}{2} + \left(\frac{h+1}{2}\right)^2 - \frac{c_1^2+c_2^2+c_3^2}{2} & \text{if } c_1 < c_2 + c_3 \text{ and } h = c_1 + c_2 + c_3 \text{ is odd} \end{cases}$$

Moreover, if $c_i = q, i = 1, 2, 3$, we have

$$\alpha(P) = \begin{cases} \frac{3}{4}(q+1)^2 + \frac{1}{4} & \text{if } q \text{ is even} \\ \frac{3}{4}(q+1)^2 & \text{if } q \text{ is odd} \end{cases}$$

Proof: The first case is a consequence of Lemma 2. The second and third cases are obtained from formula (4), observing that the only subsets $I \subseteq \{1, 2, 3\}$ such that $|I| \leq \lfloor h/2 \rfloor - c_I$ are $\varnothing, \{1\}, \{2\}$, and $\{3\}$ when $c_1 < c_2 + c_3$. □

The next lemma gives the complexity for computing $\alpha(P)$, using (4).

Lemma 5. *The computation of the width of a product of chains $P = (\overline{c_1+1}) \times \cdots \times (\overline{c_p+1})$ using formula (4) is performed in $O(p2^p)$ operations.*

Proof: The product $\prod_{k=1}^{p-1}(1 + \frac{\lfloor h/2 \rfloor - c_I - |I|}{k})$ requires $O(p)$ operations and the sum is over $O(2^p)$ subsets, so the computation of $\alpha(P)$ needs $O(p2^p)$ operations. □

Thus, this complexity is exponential in the number of criteria p, but does not depend on the values c_i. Actually, since p is usually small in practice and thus considered constant in theory, the previous discussion can be summarized through the following result.

Theorem 3. Max SizeND *is solvable in constant time when p is constant.*

3.3 Comparison of the different bounds

We propose to compare $\alpha(P)$ to simpler bounds on the number of non-dominated points. Let us first illustrate this comparison on a large instance of the tri-objective Spanning Tree problem. Let $G = (V, E)$ be a complete graph with $n = 101$ vertices, where each edge cost is randomly chosen between 0 and 10 on each criterion. We wish to compute, in the worst case, the number of non-dominated points.

Considering that for some instances all feasible solutions can give rise to different non-dominated points [7], a first bound is the total number of spanning trees in a complete graph, i.e. $n^{n-2} = 101^{99}$. This huge bound, which can be achieved only when edge costs are exponential, does not take account of values c_i.

A second bound corresponds to the product $\prod_{i=2}^{p}(c_i + 1)$, where $c_1 = c_2 = c_3 = 1000$ and $p = 3$ which gives $1001^2 = 1.002.001$.

Finally, our proposed bound, computed from Corollary 2 with $c_1 = c_2 = c_3 = 1000$, gives $\frac{3}{4}(1001)^2 + \frac{1}{4} = 750.751$.

It is interesting to quantify the ratio between $\alpha(P)$ and $\prod_{i=2}^{p}(c_i + 1)$. The smallest ratio is reached, as in the previous example, when all c_i are equal. Let $\alpha_{p,q}(P)$ be the result of the formula which computes the maximal number of non-dominated points in the worst case when there are p criteria and for all i, $c_i = q - 1$. Thus, we determine $\lim_{q\to\infty} \alpha_{p,q}(P)/q^{p-1}$, where q^{p-1} corresponds to the product $\prod_{i=2}^{p}(c_i + 1)$.

Proposition 2. *For p criteria, we have* $\lim_{q\to\infty} \frac{\alpha_{p,q}(P)}{q^{p-1}} = \sum_{l=1}^{\lceil \frac{p}{2} \rceil - 1} -(1)^l \frac{p}{l!(p-l)!}(\frac{p}{2} - l)^{p-1}$.

Proof: Using formula (4) and keeping only the coefficients of the terms of degree $p - 1$.

□

This way, we can compute all these limits when p is fixed. For instance $\lim_{q\to\infty} \frac{\alpha_{3,q}(P)}{q^2} = \frac{3}{4}$ and $\lim_{q\to\infty} \frac{\alpha_{4,q}(P)}{q^3} = \frac{2}{3}$. When the number of criteria increases, we note that the proposed bound is more and more interesting, as compared with the bound q^{p-1}.

4 Tightness of the bound for multicriteria combinatorial optimization problems

The determination of the maximum number of non-dominated points is particularly relevant for multicriteria combinatorial optimization problems, for which it is well-known that this number can be exponential in the size of the instance [4]. Considering such a problem Π, the problem of determining the maximum cardinality of the non-dominated set associated to Π, knowing values c_i, $i = 1, \ldots, p$, is denoted by Max SizeND Π in the following.

We show in this part that our bound $\alpha(P)$ is tight for the multicriteria version of some classical optimization problems such as Selection, Knapsack, Shortest Path, Spanning Tree, TSP, s-t Cut. We propose some relatively simple families of instances of these problems where the number of non-dominated points is exactly $\alpha(P)$.

We first introduce some notations used in the definitions of these problems. Selection and Knapsack require to define a set O of objects, a capacity b and a nonnegative integer t. Each object $o \in O$ has a criterion vector $v(o) = (v_1(o), \ldots, v_p(o))$ and a weight $w(o)$. We define the criterion functions on a set $O' \subseteq O$ as $v_i(O') = \sum_{o\in O'} v_i(o)$ for all $i \in \{1, \ldots, p\}$.

Selection consists in selecting a subset $O' \subseteq O$ of t objects maximizing $v_i(O')$, $i = 1, \ldots, p$. Knapsack consists in selecting a subset $O' \subseteq O$ satisfying the constraint $\sum_{o\in O'} w(o) \leq b$ maximizing $v_i(O')$, $i = 1, \ldots, p$.

The other problems are defined on a graph. Consider $G = (V, E)$ a graph where $V = \{1, \ldots, n\}$ is the set of vertices and $E \subseteq V \times V$ is the set of edges. Each edge $e \in E$ has a criterion vector $v(e) = (v_1(e), \ldots, v_p(e))$. We define the value function v on a subset E' of edges as follows: $v(E') = (v_1(E'), \ldots, v_p(E'))$ where $v_i(E') = \sum_{e\in E'} v_i(e)$ for all $i \in \{1, \ldots, p\}$.

Proposition 3. *The bound $\alpha(P)$ is tight for* Max SizeND Selection *and* Max SizeND Knapsack.

Proof: Consider p integers $c_1, \ldots, c_p$ and p subsets O_j, $j = 1, \ldots, p$, where each subset O_j contains c_j identical objects $o_j^i, i = 1, \ldots, c_j$ with $v_j(o_j^i) = 1$ and $v_k(o_j^i) = 0$ for $k \neq j$. Let $O = \cup_{j=1}^{p} O_j$ with $|O| = \sum_{j=1}^{p} c_j = n$ and $t = \lfloor \frac{n}{2} \rfloor$.

Selecting $t = \lfloor \frac{n}{2} \rfloor = \lfloor \frac{h}{2} \rfloor$ objects can be seen as selecting x_j objects in subset O_j, $j = 1, \ldots, p$ such that $\sum_{j=1}^{p} x_j = \lfloor \frac{h}{2} \rfloor$ and $0 \leq x_j \leq c_j$, with a resulting non-dominated criterion vector $(x_1, x_2, \ldots, x_p)$. The number of such vectors is the number of integer solutions of equation (3) and thus corresponds to $\alpha(P)$.

Since SELECTION is a particular case of KNAPSACK, the result also holds for MAX SIZEND KNAPSACK. □

Proposition 4. *The bound $\alpha(P)$ is tight for* MAX SIZEND SHORTEST PATH*,* MAX SIZEND SPANNING TREE*, and* MAX SIZEND TSP.

Proof: Assume that p is even and let q be a nonnegative integer. We consider the following gadget consisting of a graph with two vertices, which are connected by edges corresponding to all the p-tuples containing $p/2$ values 0 and $p/2$ values 1, with the corresponding values on these edges (see Figure 1).

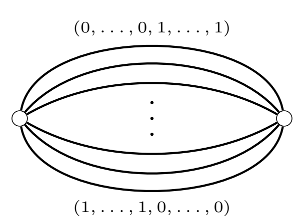

Figure 1: Gadget

Let G be the concatenation of q times this gadget (see Figure 2).

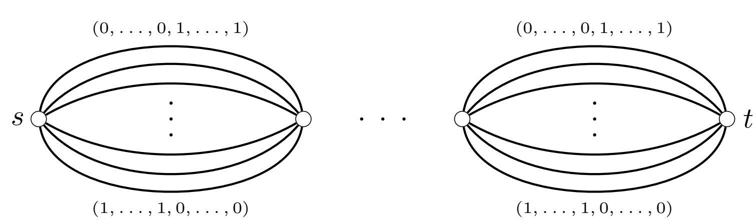

Figure 2: Graph G

Any path between s and t in G uses exactly one edge of each gadget and corresponds to a non-dominated point $(v_1, \ldots, v_p)$ with $0 \leq v_i \leq q$ and $\sum_{i=1}^{p} v_i = \frac{pq}{2}$. The number of such points is the number of integer solutions of equation (3), with $c_i = q$, for $i = 1, \ldots, p$, and thus corresponds to $\alpha(P)$.

Since in the previous construction paths and spanning trees are equivalent, the proof holds for MAX SIZEND SPANNING TREE. Adding edge (s, t) to the above construction with criterion value $(0, \ldots, 0)$, the proof holds also for MAX SIZEND TSP. □

Proposition 5. *The bound $\alpha(P)$ is tight for* MAX SIZEND s-t CUT.

Proof: The proof is essentially the same as in Proposition 4 but using the following gadget consisting of a path whose edges correspond to all the p-tuples containing $p/2$ values 0 and $p/2$ values 1 (see Figure 3) and the following graph G (see Figure 4), where this gadget is duplicated q times, each of these being connected at each end.

Figure 3: Gadget

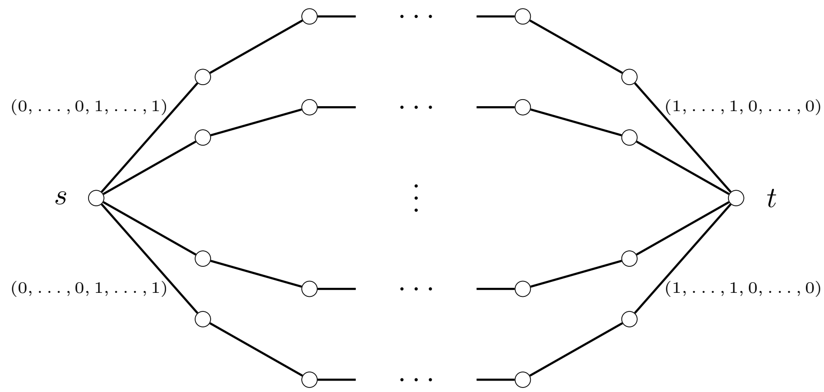

Figure 4: Graph G

In the same way, we have $c_i = q$ for $i = 1, \dots, p$, and the number of non-dominated points is exactly $\alpha(P)$. □

5 Reduction of the maximal number of non-dominated points using known feasible solutions

We investigate now if it is possible to improve the upper bound on the number of non-dominated points when a subset of feasible solutions or a subset of efficient solutions is known. Indeed, feasible solutions can often be easily computed. Moreover, *supported* efficient solutions, which are obtained by optimizing a weighted sum of the criteria, are easily computable, when the corresponding single criterion problem is polynomially solvable.

The knowledge of feasible criterion vectors, possibly known to be non-dominated, involves the elimination of some points in P. More precisely, if a feasible point z is known, all the points dominated by z cannot be part of the non-dominated set and can thus be removed from P. Moreover, if z is known to be non-dominated, we can also remove from P all the points which dominate z. In the graph theory setting, this leads to subgraphs which are still comparability graphs. Therefore, the computation of the maximal number of non-dominated points in this context is still achievable in pseudo-polynomial time. We investigate the problem under the ordered set theory setting.

5.1 When feasible solutions are known

Given $P = (\overline{c_1 + 1}) \times \dots \times (\overline{c_p + 1})$ and k points $z^1, \dots, z^k$ in the criterion space, representing feasible solutions, let D be the subset of P dominated by at least one point from $\{z^1, \dots, z^k\}$, that is the set of points y of P such that there is $j \in \{1, \dots, k\}$ with $z^j \leq y$. We want to study if the set $Q = P - D$ still satisfies the Sperner property and we want to compute $\alpha(Q)$.

5.1.1 Case $p = 2$

In the bicriteria case we have the following result.

Proposition 6. *When $p = 2$, Q satisfies the Sperner property and we have $\alpha(Q) = min(c_2, min_{j=1}^{k} r(z^j)) + 1$ where $r(z^j)$ is the rank of point z^j.*

Proof: When there is no point in $\{z^1, \ldots, z^k\}$ located below the first level of maximum cardinality of P we have $\alpha(Q) = \alpha(P) = c_2 + 1$. Otherwise, let L be the lowest level of P containing an element of the set $\{z^1, \ldots, z^k\}$ and $z^m = (z_1^m, z_2^m)$ such a point. Since points z^j, $j \neq m$, located above level L do not eliminate any point on L, we have $\alpha(Q) \geq |L| = r(z^m)+1$.

Consider now the set $W \subset P$ of points belonging either to the chains containing all the points with a first constant coordinate v_1, for each $v_1 \in \{0, \ldots, z_1^m\}$ or to the chains containing all the points with a second constant coordinate v_2, for each $v_2 \in \{0, \ldots, z_2^m - 1\}$. We have $Q \subset W$ and we use $|L|$ chains to cover W. Therefore, any antichain of Q contains at most $|L|$ points, i.e. we have $\alpha(Q) \leq |L| = r(z^m) + 1$.

In any case, $\alpha(Q)$ corresponds to the cardinality of a level of Q, meaning that Q satisfies the Sperner property. □

5.1.2 Case $p \geq 3$

When $p \geq 3$, the observed structure does not satisfy the Sperner property as will be shown in the next result. We observed in the bicriteria case that $\alpha(Q)$ is determined either from the first level of maximum cardinality or from the level of one of the points z^j. We could expect that, for $p \geq 3$, only these levels are relevant when computing $\alpha(Q)$. Unfortunately, we also show that other levels may contribute to $\alpha(Q)$. This suggests that the determination of $\alpha(Q)$ is difficult.

Proposition 7. *For any $p \geq 3$, Q does not satisfy the Sperner property. Moreover, other levels than the first level of maximum cardinality of P and levels of the points z^j may contribute to $\alpha(Q)$.*

Proof: We first construct a simple example with three criteria. Let $P = \overline{3} \times \overline{3} \times \overline{3}$ be the product of chains and $z = (0, 1, 0)$ a known feasible solution (see Figure 5).

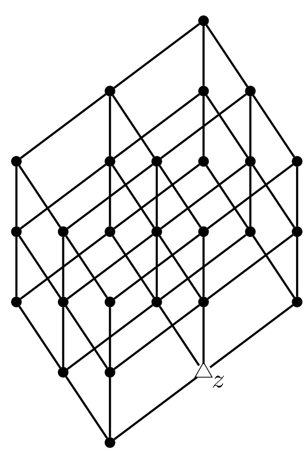

Figure 5: $P = \overline{3} \times \overline{3} \times \overline{3}$

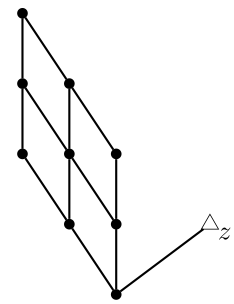

Figure 6: The set $Q = P - D$

The set $Q = P - D$, represented in Figure 6, does not satisfy the Sperner property. Indeed, we have $\sigma(Q) = |L_1| = |L_2| = 3$, while $\alpha(Q) = |L_2| + 1 = 4$ since point z, which belongs to level L_1, is incomparable to the 3 points belonging to level L_2. Observe that L_2 is neither the first level of maximum cardinality of P (L_3) nor the level of z (L_1).

This example can be extended easily to $p \geq 4$ criteria. We just need to extend z with values 0 on the $p-3$ other criteria and add $p-3$ new points z^i, $i = 1, \ldots, p-3$, where z^i has coordinate 1 on criterion $i+3$ and 0 on the other criteria. Doing so, we obtain the same set Q as for $p = 3$ (except that points in Q have now all their $p-3$ last coordinates equal to 0). $\Box$

5.2 When efficient solutions are known

We consider now the same problem when the feasible solutions are known to be efficient.

Given $P = (\overline{c_1+1}) \times \ldots \times (\overline{c_p+1})$ and k non-dominated points $z^1, \ldots, z^k$ in the criterion space, representing efficient solutions, let D be the subset of P corresponding to the set of points y of P such that there is $j \in \{1, \ldots, k\}$ with $z^j \leq y$ or $y \leq z^j$. We are interested in computing $\alpha(Q)$, where $Q = P - D$.

5.2.1 Case $p = 2$

In this case, the set Q does not satisfy the Sperner property. We illustrate this on an instance where $P = \overline{8} \times \overline{6}$, and two known non-dominated points $z^1 = (5,1)$ and $z^2 = (1,4)$ (see Figure 7). Here Q consists of the points represented by squares and the two points z^1 and z^2. A largest antichain in Q is $\{z^1, z^2, y^1, \ldots, y^4\}$ and thus we have $\alpha(Q) = 6$, whereas $\sigma(Q) = 4$.

We show, however, that $Q \backslash \{z^1, \ldots, z^k\}$ is a disjoint union of products of two chains, which allows the computation of $\alpha(Q)$. We assume in this part that the k non-dominated points z^j, $j = 1, \ldots, k$ are ranked by non increasing value on the first criterion, i.e. $z_1^1 \geq \ldots \geq z_1^k$.

Proposition 8. *When $p = 2$, we have $\alpha(Q) = k + \min(c_1 - z_1^1, z_2^1) + \min(z_1^k, c_2 - z_2^k) + \sum_{j=1}^{k-1} \min(z_1^j - z_1^{j+1} - 1, z_2^{j+1} - z_2^j - 1)$.*

Proof: The first term k in the proposed formula corresponds to the k given non-dominated points. These k points delimit exactly $k+1$ disjoint products of two chains, some of them being possibly empty. The first product of chains is of size $c_1 - z_1^1$ on the first criterion and z_2^1 on the second one, the $(k+1)^{\text{th}}$ product of chains is of size z_1^k on the first criterion and $c_2 - z_2^k$ on the second one, whereas the products of chains located between two points z^j and z^{j+1} are of size $z_1^j - z_1^{j+1} - 1$ on the first criterion and $z_2^{j+1} - z_2^j - 1$ on the second one. Each point of any of these $k+1$ products of chains is incomparable with any point of any other

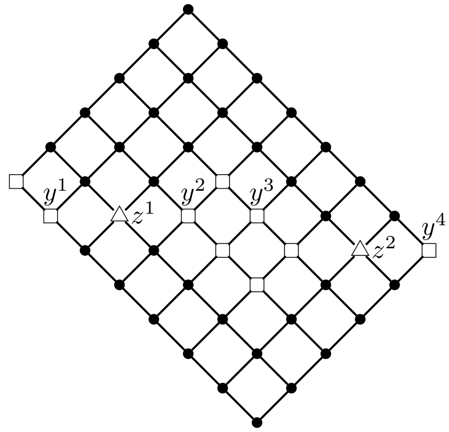

Figure 7: $P = \overline{8} \times \overline{6}$

product and incomparable with each z^j. Since the width of a product of two chains $\overline{c_1} \times \overline{c_2}$ is $\min(c_1, c_2)$, the formula is proved. □

We remark that, to determine $\alpha(Q)$, we can consider only non-dominated points located on the levels which contain the known non-dominated points z^j. Referring again to the instance presented in Figure 7, we illustrate this remark with the largest antichain $\{z^1, z^2, y^1, \ldots, y^4\}$.

5.2.2 Case $p \geq 3$

We observed in the bicriteria case that $\alpha(Q)$ is determined by considering points on the levels of points z^j. Unfortunately, for $p \geq 3$, other levels may contribute to $\alpha(Q)$, as shown in the next result. This suggests that the determination of $\alpha(Q)$ is difficult.

Proposition 9. *For any $p \geq 3$, other levels than the first level of maximum cardinality of P and levels of the points z^j may contribute to $\alpha(Q)$.*

Proof: We consider the same counter-example as in the proof of Proposition 7 and Figure 5. The set $Q = P - D$ is represented in Figure 8.

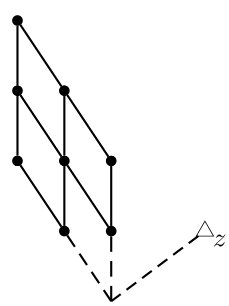

Figure 8: The set $Q = P - D$

We have $\alpha(Q) = |L_2|+1$. Observe that L_2 is neither the first level of maximum cardinality of P (L_3) nor the level of z (L_1). □

6 Conclusions

The purpose of this work was to develop tight and easily computable bounds on the cardinality of the set of non-dominated points. Graph theory and ordered set theory provided complementary insights on this topic. Two main questions require further investigation.

A basic assumption in our work is the a priori knowledge on the number of values taken on each criterion. Obviously, obtaining a good upper bound on these values is itself a difficult question which depends on the problem at hand as well as on the specific instances.

Knowing feasible, possibly efficient, solutions may improve our bound on the number of non-dominated points. The impact is clear in the bicriteria case. For $p \geq 3$, nice properties (the Sperner property, the fact that only the levels of known points are relevant) are no longer valid. Even if we know, from graph theory, that this upper bound can be computed in pseudo-polynomial time, further structural insights are still required.

References

[1] N.G. De Bruijn, C. Tengbergen, and D. Kruyswijk. On the set of divisors of a number. *New Archive for Mathematics*, 23:191–193, 1951.

[2] N. Caspar, B. Leclerc, and B. Monjardet. *Finite ordered sets: concepts, results and uses.* Cambridge University Press, 2012.

[3] C. A. Charalambides. *Enumerative combinatorics.* Chapman, 2002.

[4] M. Ehrgott. *Multicriteria Optimization.* Springer (2nd edition), 2005.

[5] M. Ehrgott and D. Tenfelde-Podehl. Computation of ideal and nadir values and implications for their use in MCDM methods. *European Journal of Operational Research*, 151:119–139, 2002.

[6] M.C. Golumbic. *Algorithmic graph theory and perfect graphs.* Annals of Discrete Mathematics Series, vol 57, North Holland, 2004.

[7] H.W. Hamacher and G. Ruhe. On spanning tree problems with multiple objectives. *Annals of Operations Research*, 52:209–230, 1994.

[8] B. Leclerc. Sur le nombre d'éléments des niveaux des produits de chaînes et des treillis permutoèdres. *Mathématiques et sciences humaines*, 112:37–48, 1990.

[9] M. Stanojević, M. Vujošević, and B. Stanojević. On the cardinality of the nondominated set of multi-objective combinatorial optimization problems. *Operations Research Letters*, 41(2):197–200, 2013.