A GENERALIZED SIMPLEX METHOD FOR INTEGER PROBLEMS GIVEN BY VERIFICATION ORACLES

SERGEI CHUBANOV

Bosch Center for Artificial Intelligence, Renningen, Germany, and University of Siegen, Germany

Abstract. We consider a linear problem over a finite set of integer vectors and assume that there is a verification oracle, which is an algorithm being able to verify whether a given vector optimizes a given linear function over the feasible set. Given an initial solution, the algorithm proposed in this paper finds an optimal solution of the problem together with a path, in the 1-skeleton of the convex hull of the feasible set, from the initial solution to the optimal solution found. The length of this path is bounded by the sum of distinct values which can be taken by the components of feasible solutions, minus the dimension of the problem. In particular, in the case when the feasible set is a set of binary vectors, the length of the constructed path is bounded by the number of variables, independently of the objective function.

Key words: linear programming, duality, simplex method, normal fan.

1. Introduction and formulation of the main result

Let S be a finite subset of $\mathbb{Z}^n$. We consider the following linear optimization problem over S :

$$\min\{c^T x : x \in S\}, \tag{1.1}$$

where $c \in \mathbb{Z}^n$.

Some examples of (1.1):

- S is the set of the integer points of a polytope. That is, (1.1) is an integer linear problem.
- S is the set of vertices of an integral polytope. In this case we have a usual linear program if the polytope is given by a system of linear constraints.

Further, we assume that we have a feasibility oracle which is able to verify whether a given solution belongs to S and a *verification oracle* which is able to verify whether a given solution in S minimizes a given objective function. Calls to the verification oracle will be called *verifications* for short. More precisely, given $y \in \mathbb{Q}^n$ and $x^0 \in S$, a *verification oracle* checks whether x^0 is optimal for

$$\min\{y^T x : x \in S\}. \tag{1.2}$$

E-mail address: `sergei.chubanov@de.bosch.com`.

The presence of an *augmentation oracle* is a much stronger assumption. Given x^0 and y, an augmentation oracle either decides that x^0 is optimal for (1.2) or returns x such that $y^T x < y^T x^0$.

For instance, in the case of minimum-cost network flow problems, a verification oracle can be based on checking whether there are negative-cost cycles in the residual graph for a given flow. Another example of a combinatorial problem where we have an efficient verification oracle is the minimum-cost spanning tree problem. In both examples, all known verification oracles are actually augmentation oracles, because they not only verify whether the current solution is optimal but also provide an improvement direction in the case of non-optimality.

If S is the vertex set of the polytope defined by a system $Ax \leq b$, then a verification oracle can be represented by an algorithm which is able to check whether a given vector y belongs to the conic hull of the left-hand sides of the inequalities which are active at a given vertex. Of course, such a situation may be worth considering only if the maximum number of inequalities active at a vertex is much smaller than the entire number of constraints.

Our algorithm can be viewed as the most abstract form of the simplex method which is only allowed to use arithmetic operations and verifications. At each iteration, by means of verifications, the algorithm finds a suitable adjacent vertex in the convex hull of S.

More formally, by a *simplex method* we understand any algorithm constructing a path in the 1-skeleton of the convex hull P of S from a given vertex of P to another vertex being an optimal solution. The length of a path in the 1-skeleton of P is the number of edges in this path.

Let an integer Δ be known such that

$$\forall x', x'' \in S : \|x' - x''\|_\infty \leq \Delta. \tag{1.3}$$

We will use the bound Δ in the sense of (1.3) throughout the paper.

Further, by *elementary operations* we understand arithmetic operations, comparisons, and computations of fractional parts, least common multiples, and greatest common divisors. The main contribution of the present paper is the following theorem and the subsequent corollary:

Theorem 1.1. *Let x^0 be a vertex of the convex hull of S. Let c^0 define a linear function for which x^0 is a unique minimizer. Let $c^0, \ldots, c^k$ be given such that $c^t - c^{t-1}$ is an integer vector for all $t = 1, \ldots, k$. There exists a simplex method which solves (1.1) by visiting at most*

$$\sum_{t=1}^{k} |\{(c^t - c^{t-1})^T x : x \in S\}| - k \tag{1.4}$$

vertices of the convex hull of S. If S is the set of integer vectors of a polytope, then the number of verifications, feasibility tests, and elementary operations is bounded by a polynomial in (1.4), n, $\log \Delta$, and in the binary sizes of c^t.

Corollary 1.1. *Let u_i be the number of distinct values the ith component of a vertex of the convex hull of S may take. There is a simplex method solving the problem (1.1) by visiting at most $\sum_{i=1}^n u_i - n$ feasible solutions.*

Proof. Let $k = n$ and $c^t = c^{t-1} + c_t \cdot \mathbf{e}^t$, where $\mathbf{e}^t$ is the unit vector whose tth component is 1. Now apply Theorem 1.1. ■

The above corollary implies that the number of feasible solutions visited by the algorithm does not depend on c provided that c^t are chosen as in the proof.

It seems to be an open question whether it is possible to achieve an oracle complexity bound which would be polynomial in the binary size of u, even in the case when S is the set of integer vectors of some polyhedron given by a list of constraints. If there is an evaluation oracle, i.e., an oracle which returns the optimal value for any given linear objective function, the problem is known to be solvable in polynomial time; see Orlin, Punnen, and Schulz [14]. However, the assumption that there is an evaluation oracle is too strong, even compared to the assumption that an augmentation oracle is given.

The polynomial equivalence of verification and optimization in 0-1 integer linear programming was shown by Schulz [18], who gave an algorithm using only polynomially many verifications. Also, in [18] it was proved that there is a variant of the simplex method which solves linear programs over 0-1 polytopes by visiting at most $O(n^4)$ vertices (see Theorem 26 in [18]). So in the present paper we improve this bound to $O(n)$. A model of computation with an augmentation oracle for arbitrary objective functions over polytopes was considered in Chubanov [3]. As a corollary of the analysis of greedy algorithms proposed in [3] it was shown that the diameter of a 0-1 polytope is bounded by $O(m \log n)$ where m is an upper bound on the number of nonzero components of a vertex.

In the presence of the so-called directed augmentation oracle, which is a somewhat stronger assumption than requiring an augmentation oracle, Schulz and Weismantel [17] show that an integer linear problem of the form

$$\max\{c^T x : Ax = b, \mathbf{0} \leq x \leq u, x \in \mathbb{Z}^n\}$$

can be solved in $O(n \log(\|c\|_\infty \|u\|_\infty))$ calls to the directed augmentation oracle. De Loera, Hemmecke, and Lee [5] studied integer linear problems of the same form in the presence of a special class of augmentation oracles where it is additionally required that they deliver augmentation directions satisfying some conditions. In particular, they even obtain a polynomial oracle complexity when the augmentation oracle is able to deliver the best augmentation direction in a subset of potential augmentation directions (see [5] for more details). Under some less restrictive conditions on the quality of the directions returned by the augmentation oracle, the algorithm proposed in [5] needs at most $(4n-4)\|u\|_\infty \log(c^T x^0 - OPT)$ augmentations starting at an initial solution x^0. Here, OPT denotes the optimal value.

At first glance, the problem in question may seem to be trivially reducible to the binary case in polynomial time where the feasible solutions would be binary encodings of the solutions of the original problem. However, such a reduction ignores our oracle model. Indeed, let S be a subset of $\{-1, 0, 1\}^n$. An obvious reduction to a binary problem would be to represent a solution x as $x = -w^1 + w^2$ where w^1 and w^2 are binary vectors such that $-\mathbf{1} \leq -w^1 + w^2 \leq \mathbf{1}$. Now assume that we wish to verify if a linear function given by vectors c^1 and c^2 is minimized at a given solution (w^1, w^2) of the obtained binary problem, where $-w^1 + w^2 \in S$. If $c^1 = c^2$, this can be easily done by applying the verification oracle to $c = c^1$ and $x = -w^1 + w^2$. However, if $c^1 \neq c^2$, we cannot apply the verification oracle in such a straightforward way. On the other hand, the algorithms for the binary case mentioned above are based on an analysis of how optimal solutions change when a single component of the objective function is modified, i.e., these algorithms need to consider cases where $c^1 \neq c^2$. So the mentioned complexity results for binary problems do not apply to the more general problem (1.1) even if S is a subset of $\{-1, 0, 1\}^n$.

Our algorithm implies some known bounds on the diameters of polytopes. For instance, the diameters of 0-1 polytopes are bounded by the dimension of the solution space according to Theorem 1.1. Naddef [13] proposed a non-constructive and very simple proof for this bound. Recently, Kitahara and Mizuno [11] observed that Naddef's bound can be obtained constructively in the following way. Note that any vertex x of a 0-1 polytope is the unique minimizer of the linear function defined by the coefficient vector $\mathbf{1} - 2x$, where $\mathbf{1}$ is the all-one vector of the respective dimension. Then, when minimizing this linear function, any suitable variant of the simplex method with the property that the current objective value is improved whenever a new feasible solution is found visits at most n distinct vertices before it comes to the optimal solution x. This follows from the fact that the number of distinct values which can be taken by the linear function defined by $\mathbf{1} - 2x$ over a set of binary vectors is bounded by $n + 1$. However, this approach does not imply a polynomial number of vertices considered by the algorithm for an arbitrary linear function with integer coefficients.

The bound of Corollary 1.1 belongs to the family of bounds on diameters of polytopes presented in Kleinschmidt and Onn [12] and generalizes the bound obtained by Naddef [13]. On the other hand, for the case when the ith coordinate of vertices ranges from 0 to k_i, Del Pia and Michini [7] improved these bounds to $2 + k_2 + k_3 + ... + k_n - \lceil n/2 \rceil$, where k_i are sorted in nondecreasing order and $k_1 \geq 2$. Under the same setting, this bound was in turn improved by Deza and Pournin [9] to $3 + k_2 + k_3 + ... + k_n - \lceil 2n/3 \rceil$ for $k_1 \geq 3$. These bounds are somewhat stronger, but only within a constant factor compared to those following from our algorithm. Moreover, these bounds are non-constructive. At the same time, Del Pia and Michini [8] propose a variant of the simplex method which traces a path of length $O(n^4 k \log(nk))$ for lattice polytopes contained in $[0, k]^n$. In this case, our variant of the simplex method visits at most nk vertices. On the other hand, one should note that Deza et al. [10] derived a lower bound of the order $k^{n/(n+1)}$, for any fixed n, on the largest possible diameter of a lattice polytope contained in $[0, k]^n$.

The earlier works mentioned above either propose algorithmic results for special cases of (1.1) under stronger assumptions than the presence of a verification oracle or provide nonconstructive bounds on diameters. The main contribution of the present paper is a simplex algorithm for the most general form of a bounded integer linear problem (1.1) which constructs a path in the 1-skeleton of the convex hull of the feasible set of length similar to the bounds obtained in the mentioned earlier works and has an access to the feasible set of the problem via exclusively feasibility and verification oracles.

Our algorithm can be viewed as a variant of the shadow vertex algorithm for linear programming, with the reservation that our algorithm does not assume a polyhedral description of the convex hull of S as a part of the input, whereas all traditional variants of the simplex method including the shadow vertex algorithm are applied to linear programs given by lists of linear constraints. Let us recall that the shadow vertex algorithm for linear programming (see Borgwardt [1]) has the following geometric interpretation. Consider an initial vertex and a linear function which is optimized by the initial vertex. In the space of linear functions, the segment connecting the initial function with the objective function intersects a family of normal cones of the feasible polytope. The vertices whose normal cones share a facet form and edge of the polytope. That is, that family of normal cones corresponds to a family of paths in the 1-skeleton of the polytope (if the initial linear function is in a general position, in a certain sense, then the mentioned family of normal cones corresponds to a single path). The shadow vertex algorithm traverses one of these paths. Some variations of the shadow vertex algorithm were proposed in attempt to improve average and worst-case estimates of the number of vertices traversed by the algorithm. One of the most recent results is that of Dadush and Hähnle [4] who propose a variant of the shadow vertex algorithm. They obtain complexity estimates which are polynomial in the dimension of the problem and in the reciprocal of the so-called curvature of the polytope.

The algorithm proposed in the present paper is in part a further development of the algorithm proposed in Chubanov [2] (this is an unpublished manuscript), where a non-cycling variant of the shadow vertex algorithm was considered. Similarly to the shadow vertex algorithm, our algorithm constructs a path in the space of linear functions, from a function for which a unique optimizer is known to the objective function. The algorithm traverses the normal cones intersected by this path and finally comes to a cone containing the objective function. Note that since a polyhedral description of the convex hull of S is not available, the traditional variants of the simplex method, including those of the shadow vertex algorithm mentioned above, are not applicable, whereas our algorithm is able to discover the necessary information about the local polyhedral structure of the convex hull of S exclusively by means of verifications.

2. Main notation and notions used

The linear hull and the convex hull are denoted by $lin.hull(\cdot)$ and $conv.hull(\cdot)$, respectively. The interior of a set of vectors is denoted by $int(\cdot)$ and the relative interior of a set is denoted by $rel.int(\cdot)$.

The boundary of a set K in a Euclidean space is denoted by ∂K.

The (line) segment connecting vectors a and b is denoted by $[a, b]$. If we wish to exclude an endpoint, say a, we replace the respective square bracket by a round bracket, i.e., we write $(a, b]$.

Two segments of nonzero length are called parallel if they are contained in parallel lines.

The (binary) size is denoted by $size(\cdot)$. In the case of a rational vector, its size is the sum of the sizes of the pairs of coprime numbers representing the components of the vector. If a component is equal to zero, it is assumed to be given by the fraction 0/1.

For a rational vector v whose components are represented as fractions of coprimes where the denominators are positive integers, let $\alpha(v)$ denote the least common multiple of the denominators of the components of v. A vector v can be uniquely represented as

$$v = \frac{r}{\alpha(v)},$$

where r is an integer vector.

Let x be in S. Denote

$$C(x) = \{y \in \mathbb{R}^n : y^T x = \min_{x' \in S} y^T x'\}.$$

We call $C(x)$ *the normal cone* of x. The normal cone of x can be viewed as the set of all linear functions minimized at x.

The family of the normal cones $C(x)$ where x is an element in S being at the same time a vertex of $conv.hull(S)$ is called *the normal fan* and denoted by $\mathcal{F}$. The normal fan covers the whole space $\mathbb{R}^n$. It is clear that for any rational vector y and for any $x \in S$ it can be verified whether $y \in C(x)$ in a single call to the verification oracle.

We will often refer to a linear function by means of its coefficient vector, i.e, function y means the respective linear function whose coefficients are $y_1, \ldots, y_n$. In fact, the set of normal cones is a polyhedral cell complex over the respective space of linear functions.

The normal cone of x can be equivalently defined as

$$C(x) = \{y \in \mathbb{R}^n : (x - x')^T y \leq 0, \forall x' \in S\}.$$

Since S is finite, the normal cones are *polyhedral*. In most cases, we will implicitly use this definition.

An example of a normal fan is given in Figure 1.

The following properties of normal fans are well known. For a reference, their proofs are given in the appendix; see [6] and [19] for more information about normal fans.

Proposition 2.1. *For any $C(x^1)$ and $C(x^2)$ in $\mathcal{F}$, $C(x^1) \cap C(x^2)$ is a common face of $C(x^1)$ and $C(x^2)$.*

Denote

$$H(x', x'') = \{y : (x' - x'')^T y = 0\}.$$

Here and in the subsequent sections we will consider the following family of hyperplanes:

$$\mathcal{H} := \{H(x', x'') : x', x'' \in S, x' \neq x''\}. \qquad (2.1)$$

Each face of a normal cone is contained in some hyperplane of this family.

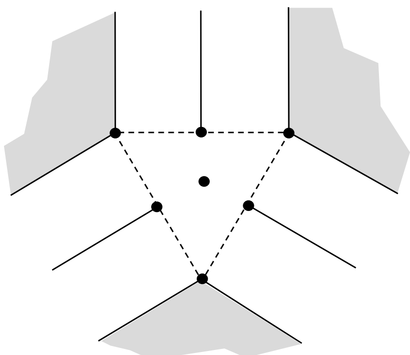

FIGURE 1. In this example, $n = 2$ and $|S| = 7$. The black circles are the elements of S. The dashed triangle is *conv.hull*(S). The picture represents the Minkowski sums of the elements of S with their normal cones. The full-dimensional cones, which are three, belong to the normal fan.

.

Lemma 2.1. *If $C(x^0)$ belongs to $\mathcal{F}$, then each facet of $C(x^0)$ is a facet of some other cone $C(x) \in \mathcal{F}$.*

It is clear that all cones $C(x) \in \mathcal{F}$ have full dimension.

We call two normal cones of $\mathcal{F}$ *adjacent* if and only if they share a facet.

Lemma 2.2. *Two vertices x^1 and x^2 of conv.hull(S) are adjacent if and only if $C(x^1)$ and $C(x^2)$ are adjacent.*

3. Procedure for finding normal cones intersected by a line segment

Further, we say that a segment $[a, b]$ is in a general position if neither a nor b are in a hyperplane of the family $\mathcal{H}$ and $[a, b] \cap (H_1 \cap H_2) = \emptyset$ for any H_1 and H_2 in $\mathcal{H}$ such that $H_1 \neq H_2$. That is, if $[a, b]$ is in the general position, then a and b are not contained in any of the facets of the normal cones of the family $\mathcal{F}$ and $[a, b] \cap F = \emptyset$ for all faces F, of the normal cones in $\mathcal{F}$, whose dimensions are less than $n - 1$.

Note that if $[a, b]$ is in the general position, then both a and b belong to the interiors of some normal cones, which follows from the definition.

Let $w \in S$ such that $C(w) \in \mathcal{F}$ and z in the interior of $C(w)$ be given. Now we will develop a procedure which, given a linear function z^*, will allow us to find $w^* \in S$ such that $z^* \in C(w^*)$, provided that $[z, z^*]$ is in the general position. Along with $C(w^*)$, the procedure finds all the normal cones intersected by $[z, z^*]$.

First of all, we need to understand how to find the intersection of a half-open segment $(a, b]$ with the boundary of a normal cone $C(x^0) \in \mathcal{F}$ containing a. So let us consider $a \in C(x^0)$ and $b \notin C(x^0)$ such that

$$e := b - a$$

is an integer vector and $(a, b]$ intersects $\partial C(x^0)$ at a single point. The intersection point of $(a, b]$ with $\partial C(x^0)$ has the form

$$a + \lambda(b - a), \quad \lambda = \frac{-(x - x^0)^T a}{(x - x^0)^T e}, \tag{3.1}$$

for some $x \in S$, $x \neq x^0$.

Now our goal is to develop a procedure for constructing this intersection point using the verification oracle.

Lemma 3.1. *The intersection point of $(a, b]$ with $\partial C(x^0)$ can be found in*

$$O\left(\log(n\Delta \cdot \|e\|_\infty)\right)$$

calls to the verification oracle.

Proof. The denominator in the expression for λ in (3.1) is not greater than

$$M = n\Delta\|e\|_\infty.$$

- Using binary search, find λ_1 and $\lambda_2 > \lambda_1$ (starting with $\lambda_1 = 0$ and $\lambda_2 = 1$) such that
$$\lambda \in [\lambda_1, \lambda_2], \quad |\lambda_1 - \lambda_2| \leq 1/(2M^2).$$
This can be done in $O(\log M)$ calls to the verification oracle. Note that λ is the unique rational number in $[\lambda_1, \lambda_2]$ whose denominator is not greater than M.
- Use the continued fraction method to find the best approximation of $\frac{1}{2}(\lambda_1 + \lambda_2)$ by a rational number whose denominator is not greater than M. This is exactly the λ in question. The method of continued fractions runs in time $O(\log M)$ provided that a computation of the fractional part of a rational number is considered to be an elementary operation; see [15] for a reference.

■

Let $C(x^0)$ and $C(x)$ be normal cones in $\mathcal{F}$. If $C(x^0) \cap C(x)$ contains $n-1$ linearly independent vectors $y^i \in \mathbb{R}^n$, $i = 1, \ldots, n-1$, then $C(x^0)$ and $C(x)$ are adjacent. Indeed, $F = C(x^0) \cap C(x)$ is a common face of $C(x^0)$ and $C(x)$. At the same time F has dimension $n-1$ because it contains $n-1$ linearly independent vectors. That is, $C(x^0)$ and $C(x)$ share the facet F.

If both $C(x)$ and $C(x^0)$ contain y^i, then

$$(y^i)^T x = (y^i)^T x^0. \tag{3.2}$$

Since vectors y^i are linearly independent, the coefficient matrix of the above system is of rank $n-1$. That is, if we know y^i, then $x - x^0$ can be found up to a scalar multiplier.

The above observations lead us to the following algorithm for finding all the cones intersected by $[z, z^*]$ in the general position.

Procedure 3.1.

Input: $w \in S$ such that $C(w) \in \mathcal{F}$, $z \in int(C(w))$, and $z^* \in \mathbb{R}^n$ such that $[z, z^*]$ is in the general position.
Output: $w^* \in S$ such that $z^* \in C(w^*)$.

$x^0 := w$;
$a := z$;
$\varepsilon := 1$;
while $z^* \notin C(x^0)$.

1. Choose $n-1$ vectors b_i with $\|z^* - b^i\| \leq \varepsilon$ such that $z^* - b^i$, $i = 1, \ldots, n-1$, are linearly independent.
2. Find the intersection points y^i of $(a, b_i]$ with $\partial C(x^0)$, $i = 1, \ldots, n-1$, and the intersection point y of $(a, z^*]$ with $\partial C(x^0)$.
3. Solve the system (3.2). Let v be a nontrivial solution of the system.
4. Let $x := x^0 + \gamma_{\max} v$, where
$$\gamma_{\max} = \arg\max_{\gamma \geq 0} \{\gamma : w = x^0 + \gamma v \in S\}.$$
5. **if** $x = x^0$ or there exists i such that $y^i \notin C(x)$, then set $\varepsilon := \varepsilon/2$;
else $x^0 := x$ and $a := y$;

end
Return $w^* := x^0$.

The above procedure is explained in more detail in the proof of Lemma 3.3.

To analyze the progress for every transition from x^0 to x in the course of Procedure 3.1, we need the following lemma, which can be viewed as a generalization of some observations used by Schulz [18] and earlier by Schulz, Weismantel, and Ziegler [16] for 0-1 problems:

Lemma 3.2. *Consider $x^0 \in S$. Let*

$$y \in C(x^0), \quad y + \lambda \cdot e \notin C(x^0),$$

where $\lambda > 0$ and e is a nonzero vector. Let $x \in S$ be such that

$$y + \lambda \cdot e \in C(x).$$

Then

$$e^T x < e^T x^0.$$

Proof. From the condition of the lemma it follows that

$$y^T x + \lambda \cdot e^T x = (y + \lambda \cdot e)^T x < (y + \lambda \cdot e)^T x^0 = y^T x^0 + \lambda \cdot e^T x^0 \leq y^T x + \lambda \cdot e^T x^0.$$

This implies the required inequality. ∎

Let $\delta([z, z^*])$ be the distance between $[z, z^*]$ and $\cup_{H_1, H_2 \in \mathcal{H}: H_1 \neq H_2} (H_1 \cap H_2)$. Consider

$$[z, z^*]^\# := [z, z^*] + B(\mathbf{0}, \delta([z, z^*]),$$

i.e., the Minkowski sum of $[z, z^*]$ and the ball centered at $\mathbf{0}$ with radius $\delta([z, z^*])$.

If $[z, z^*]$ is in the general position, then $[z, z^*]^\#$ has the property that any of its intersections with a facet of a normal cone in $\mathcal{F}$ is a subset of that facet. If a positive value smaller than $\delta([z, z^*])$ was known to us, we could initialize ε with that value in Procedure 3.1, in which case no divisions of ε would occur because all $(a, b^i]$ would be subsets of $[z, z^*]^\#$ and, therefore, at each iteration of the while-loop, all y^i would belong to the relative interior of some facet of the current cone $C(x^0)$. So since a positive value smaller than $\delta([z, z^*])$ is in general not known in advance, ε is initialized with 1 and divided by 2 whenever Procedure 3.1 recognizes that ε is not sufficiently small to guarantee that all y^i belong to the relative interior of the same facet of $C(x^0)$.

Lemma 3.3. *The number of transitions from x^0 to x in the course of of Procedure 3.1 is bounded by*

$$|(z^* - z)^T x : x \in S| - 1. \tag{3.3}$$

The overall number of iterations of the while-loop in the course of Procedure 3.1 is bounded by

$$O\left(\log \frac{1}{\delta([z, z^*])} + \left|(z^* - z)^T x : x \in S\right|\right), \tag{3.4}$$

each iteration taking

$$O(n(\log \Delta + \log(\|z - z^*\| + 1))) \tag{3.5}$$

calls to the verification oracle,

$$O(\Delta) \tag{3.6}$$

calls to the feasibility oracle, and

$$O(n^3 + n\Delta) \tag{3.7}$$

elementary operations, in the sense of the definition given before Theorem 1.1, where computations of greatest common divisors and least common multiplies are considered to be elementary operations.

If it is known that S is the set of integer vectors of a convex polyhedron, then (3.6) reduces to $O(\log \Delta)$ and (3.7) reduces to $O(n^3 + n \log \Delta)$

Proof. Since $[z, z^*]$ is in the general position, $\delta([z, z^*]) > 0$. According to the observations preceding the lemma, it follows that a division of ε can occur only if $\varepsilon \geq \delta([z, z^*])$ because otherwise all y^i are in the same facet of the current cone $C(x^0)$. Then, taking the step of maximum length from x^0 in direction v so that the resulting point is in S, we find another vertex x of the convex hull of S (because

$C(x)$ is adjacent to $C(x^0)$ by the facet containing the points y^i); finding $\gamma_{\max}$ at Step 4 serves for exactly this purpose.

Solving the maximization problem of Step 4:

Let v be the vector found at Step 3. Let $\gamma_{\min} > 0$ be the smallest positive value such that $\gamma_{\min} v$ is an integer vector ($\gamma_{\min} = \mathrm{lcm}(v'')/\gcd(v')$ where v'' is the vector of denominators and v' is the vector of numerators in a rational representation of v; lcm is the least common multiple and gcd is the greatest common divisor). Find the maximum integer s such that $x^0 + s\lambda_{\min} v \in S$. Then $\gamma_{\max} = s\gamma_{\min}$. Since the $x \in S$ we are looking for should satisfy the condition $\|x^0 - x\|_\infty \le \Delta$, it follows that $s \in \{1, \dots, \Delta\}$. Therefore, if it is known that S is the set of integer points of a polytope, then s can be found by a binary search in $O(\log \Delta)$ calls to the feasibility oracle (observe that in this case all integer vectors in $[x^0, x]$ belong to S). In the general case, we check all integers s from 1 to Δ, which gives (3.6). That is, the number of elementary operations is bounded by (3.7), where we take into account that the linear system at Step 3 can be solved in time $O(n^3)$ and the multiplication by $\gamma_{\max}$ at Step 4 can be performed in time $O(n)$. Recall that, according to our definition given before Theorem 1.1, for the sake of convenience we consider the computation of fractional parts, least common multiplies, and greatest common divisors as elementary operations in our model of computation. The computation of these values can be performed in a number of arithmetic operations which is polynomial in the encoding length of the input data.

Whenever a transition from x^0 to x takes place, x^0 and x are adjacent vertices of the convex hull of S. Thus, the while-loop constructs a sequence of vertices forming a path in the 1-skeleton of *conv.hull*(S). Consider $C(x^0) \cap [a, y]$ and $C(x) \cap [y, z^*]$. Since $\delta([z, z^*]) > 0$, both intersections are segments of nonzero length whose centers z' and z'' lie in the interiors of $C(x^0)$ and $C(x)$, respectively. Since $z'' - z' = \lambda e$ where $\lambda > 0$, from Lemma 3.2 it follows that $e^T x < e^T x^0$. Therefore, the number of transitions from x^0 to x is bounded by the number of different values that can be taken by $e^T x$ for $x \in S$ minus one (we subtract one because the initial $e^T x^0$ is one of those values). Thus, we get the bound (3.3) on the number of transitions from x^0 to x and the bound (3.4) on the number of iterations of the while-loop, taking into account that the number of divisions of ε is bounded by $O(\log(1/\delta([z, z^*])))$ according to the observations made at the beginning of the proof. ■

4. A generalized simplex method

Now consider the original optimization problem (1.1) where we are looking for $x^* \in S$ being a vertex of *conv.hull*(S) such that $c \in C(x^*)$.

Assume that we are given c^0 and a vertex w^0 of *conv.hull*(S) such that $c^0 \in$ *int*$(C(w^0))$. That is, w^0 is a unique minimizer of the linear function defined by c^0.

Let c be represented as

$$c = c^0 + (c^1 - c^0) + \dots (c^k - c^{k-1}),$$

where $c^k = c$.

If all segments $[c^t, c^{t-1}]$ were in the general position, then we could immediately apply Procedure 3.1 k times; see Algorithm 4.1 below. (In fact, under certain probabilistic assumptions, we can guarantee the general position with probability 1.)

Since the general position is not guaranteed by default, later we replace c^t by z^t such that the segments $[z^{t-1}, z^t]$ are in the general position and all relevant properties are preserved so that we obtain an optimal solution of the original problem.

So the following algorithm finds an optimal solution for an auxiliary problem

$$\{(z^k)^T x : x \in S\}. \tag{4.1}$$

Algorithm 4.1.

Input: A vertex x^0 of the convex hull of S and $[z^{t-1}, z^t]$, $t = 1, \ldots, k$, in the general position such that $z^0 \in int(C(x^0))$.
Output: An optimal solution x^* for (4.1) and a path P in the 1-skeleton of the convex hull of S connecting x^0 and x^*.

$P = \emptyset$;
$w^* := x^0$;
for $t = 1, \ldots, k$:
  Run Procedure 3.1 with $w = w^*$, $z = z^{t-1}$ and $z^* = z^t$;
  Add the solutions found in the course of Procedure 3.1 to the path P;
  Let w^* be the solution returned by Procedure 3.1;
end
Return $x^* := w^*$.

Lemma 4.1. *Algorithm 4.1 works correctly. The length of the path P is bounded by*

$$\sum_{t=1}^{k} |(z^t - z^{t-1})^T x : x \in S| - k.$$

Proof. Since all segments $[z^{t-1}, z^t]$ are in the general position, it follows that the calls to Procedure 3.1 are correct in the course of the algorithm. Lemma 3.3 implies the required bound on the length of P. ∎

Let $Z(y)$ denote the union of all normal cones which contain $y \in \mathbb{R}^n$:

$$Z(y) = \bigcup_{x \in S : y \in C(x)} C(x).$$

So in place of c^t we will consider z^t, $t = 1, \ldots, k$, with the following properties:

(i) $z^t - z^{t-1} = c^t - c^{t-1}$
(ii) $z^t \in Z(c^t)$, for all $t = 0, \ldots, k$.
(iii) $[z^{t-1}, z^t]$, $t = 0, \ldots, k$, are in the general position.

Property (i) will guarantee the bound (1.4) on the length of the path and (ii) will ensure that the obtained solution is optimal for the original problem.

We will look for z^t of the form

$$z^t = c^t + \mu p, \tag{4.2}$$

where $\mu > 0$ is sufficiently small and p is an integer vector such that $a^T p \neq 0$ for all vectors $a \in A$, where A is defined as

$$A = \{((x - x^2)^T (c^t - c^{t-1}))(x - x^1) - ((x - x^1)^T (c^t - c^{t-1}))(x - x^2)\} \setminus \{\mathbf{0}\} \tag{4.3}$$

where $x, x^1, x^2 \in S$, and $t = 1, \ldots, k$. Note that we don't exclude $x = x^1$ or $x = x^2$. The following proposition gives us a tool to find a suitable vector p with respect to any finite set of nonzero vectors, provided that an upper bound ρ for their maximum norms is known:

Proposition 4.1. *Consider a set A of nonzero vectors in $\mathbb{Z}^n$ such that, for some integer ρ, $\|a\|_\infty \le \rho$ for all $a \in A$. Then $p \in \mathbb{Z}^n$ such that $a^T p \ne 0$ for all $a \in A$ and whose size is polynomially bounded in n and in the size of ρ can be found in linear time.*

Proof. Let

$$p := (p_1, \dots, p_n)^T, \quad p_j := (2\rho)\sum_{i=1}^{j-1} p_i, \quad p_1 := 2\rho. \tag{4.4}$$

Consider $a \in A$ and let a_s be the last nonzero component of a, under the enumeration of the components in the ascending order of their indices. Let $a_s \ge 1$. Then, since $a_j \ge -\rho$ (because $\|a\|_\infty \le \rho$)

$$a^T p \overset{[a_s \ge 1, a_j \ge -\rho]}{\ge} p_s - \rho \sum_{i=1}^{s-1} p_i = 2\rho \sum_{i=1}^{s-1} p_i - \rho \sum_{i=1}^{s-1} p_i > 0.$$

Thus, $a^T p > 0$. The case $a_s \le -1$ is analogous. In this case, we would obtain $a^T p < 0$.

To estimate the size of p, note that

$$p_j \le (2\rho)(j-1)p_{j-1} \le (2\rho)^j (j-1)! \le (2\rho)^n n!.$$

Since p_j and ρ are positive integers, the size of p_j is bounded by the size of the respective upper bound. This implies that the size of p is polynomially bounded in n and in the size of ρ. ∎

The following lemma shows how to modify the endpoints a given segment $[a, a+e]$ so that the resulting segment is in the general position:

Lemma 4.2. *Consider a segment $R = [a+\mu p, a+\mu p+e]$ where $e \in \mathbb{R}^n$ is a rational vector and p is chosen as in the proof of Proposition 4.1 for*

$$A = \{((x - x^2)^T e)(x - x^1) - ((x - x^1)^T e)(x - x^2)\} \setminus \{\mathbf{0}\},$$

where $x, x^1, x^2 \in S$. If μ is such that

$$0 < \mu < \frac{1}{\alpha(a) \cdot 2n\Delta^2(\|e\|_\infty \cdot \|p\|_\infty)}, \tag{4.5}$$

then R is in the general position.

Proof. The segment R is not in the general position when at least one of the following two situations **(a)** and **(b)** takes place:

(a) One of the endpoints y is a solution of the equation

$$(x^0 - x)^T y = 0$$

for some $x, x^0 \in S$, $x \ne x^0$, or

(b) there exist x, x^1, and x^2 in S such that $x - x^1$ and $x - x^2$ are linearly independent and

$$(x - x^1)^T y = 0, \quad (x - x^2)^T y = 0, \tag{4.6}$$

where $y \in R$.

In **(a)**, from the equation it follows that

$$\mu = -\frac{(x^0 - x)^T a}{(x^0 - x)^T p} \quad \text{or} \quad \mu = -\frac{(x^0 - x)^T (a + e)}{(x^0 - x)^T p},$$

where the denominators are nonzero because $x - x^0 \in A$ and $a^T p \ne 0$ for all $a \in A$. This contradicts (4.5), because if μ is positive in any of the above two cases, it must be larger than the upper bound in (4.5). Thus, we have proved that both endpoints of R are in the interiors of some normal cones of $\mathcal{F}$.

In **(b)**, there exists λ such that

$$(x - x^1)^T(a + \mu p + \lambda e) = 0, \quad (x - x^2)^T(a + \mu p + \lambda e) = 0.$$

If both values $(x - x^1)^T e$ and $(x - x^2)^T e$ are zero, then the analysis reduces to one of the cases we have already considered above when analyzing **(a)**. So assume that at least one of those two values is nonzero. Since $x - x^1$ and $x - x^2$ are linearly independent, it follows that their linear combinations with at least one nonzero coefficient are nonzero vectors. Then, according to our choice of p, the determinant of the coefficient matrix is nonzero:

$$((x - x^1)^T p) \cdot ((x - x^2)^T e) - ((x - x^1)^T e) \cdot ((x - x^2)^T p) \neq 0,$$

because the left-hand side can be written as $a^T p$ where $a \in A$. Then (4.6) has a unique solution (μ, λ) where

$$\mu = \frac{-((x - x^1)^T a) \cdot ((x - x^2)^T e) + ((x - x^1)^T e) \cdot ((x - x^2)^T a)}{((x - x^1)^T p) \cdot ((x - x^2)^T e) - ((x - x^1)^T e) \cdot ((x - x^2)^T p)}.$$

Again, this contradicts the condition (4.5) because, since $\mu > 0$, the above equation leads to a value not less than the upper bound in (4.5). ∎

Finally, we need a tool to guarantee (ii):

Lemma 4.3. *Consider a rational vector $y \in \mathbb{R}^n$ and $v \in [-\delta, \delta]^n$, where*

$$\delta \leq \frac{1}{2n\Delta\alpha(y)}, \tag{4.7}$$

then

$$y + v \in Z(y).$$

Proof. Consider $x \in S$ such that $y \in C(x)$ and $\bar{x} \in S$ such that $y \notin C(\bar{x})$. Note that

$$(x - \bar{x})^T y < 0.$$

Let us prove that

$$(x - \bar{x})^T (y + v) < 0.$$

Note that

$$|(x - \bar{x})^T y| > 1/\alpha(y). \tag{4.8}$$

Therefore,

$$(x - \bar{x})^T (y + v) < -\frac{1}{\alpha(y)} + (x - \bar{x})^T v \leq -\frac{1}{\alpha(y)} + n\Delta\delta < 0,$$

which takes place for all x and $\bar{x}$ chosen as above. This completes the proof. ∎

Now we are ready to prove our main theorem:

Proof of Theorem 1.1. Let A have the form (4.3) and μ and δ be chosen so that condition (4.5) is satisfied for all $a = c^{t-1}$ and $e = c^t - c^{t-1}$, $t = 1, \ldots, k$, and (4.7) is satisfied for $v = \mu p$ and all $y = c^t$, $t = 0, \ldots, n$. To be more precise, we choose δ as the minimum of the upper bounds in (4.7) where y ranges over c^t.

Let z^t have the form (4.2). To compute p, we need $\rho \geq \|a\|_\infty$, for all $a \in A$. We can e.g. choose $\rho = \max_{t=1,\ldots,k} 2n\Delta^2 \|c^t - c^{t-1}\|_{\max}$. Then, Lemma 4.2 implies (iii), i.e., that all segments $[z^{t-1}, z^t]$ are in the general position and Lemma 4.3 implies (ii). From (ii), it follows that $z^0 \in int(C(x^0))$. Now we run Algorithm 4.1 to solve (4.1). Because of (ii), the obtained solution x^* is optimal for the original problem.

Condition (i) follows immediately from the fact that all z^t are obtained from c^t be translating c^t by the same vector.

Let μ be chosen as 1/2 of the minimum upper bound in (4.5) where the minimum is taken over $a = c^{t-1}$, and $e = c^t - c^{t-1}$, $t = 1, \ldots, k$. Now it only remains to note that $\delta([z^{t-1}, z^t])$ are bounded from below by a value whose binary size is polynomial in the sizes of c^t and Δ, which follows from standard results of linear algebra; see the appendix. ■

5. Conclusions.

In this paper, we have presented a simplex method of a general form which only uses a verification oracle at each iteration to find an adjacent vertex of the convex hull of the set of feasible solutions. The algorithm runs in a pseudopolynomial oracle time, which means that verification and optimization are pseudopolynomially equivalent in the case of linear optimization problems over finite sets. The question whether a polynomial bound can be obtained remains open.

Appendix

Proof of Proposition 2.1. Denote $h := x^1 - x^2$ and $H := \{y : h^T y = 0\}$. Note that $H \cap C(x^1)$ is a face of $C(x^1)$. This follows from the fact that $h^T y \leq 0$ is valid for $C(x^1)$. Since we have $h^T y = 0$ for all $y \in C(x^1) \cap C(x^2)$, it follows that

$$C(x^1) \cap C(x^2) \subseteq H \cap C(x^1). \tag{5.1}$$

Note that

$$\forall y \in H \cap C(x^1) : \min_{x \in S} x^T y = (x^1)^T y \overset{y \in H}{=} (x^2)^T y \geq \min_{x \in S} x^T y.$$

This implies

$$\forall y \in H \cap C(x^1) : (x^2)^T y = \min_{x \in S} x^T y.$$

Therefore,

$$H \cap C(x^1) \subseteq C(x^1) \cap C(x^2).$$

Then, taking into account (5.1), we conclude that

$$C(x^1) \cap C(x^2) = H \cap C(x^1).$$

Now we interchange the roles of x^1 and x^2 noting that hyperplane H remains the same, which yields

$$C(x^1) \cap C(x^2) = H \cap C(x^2).$$

It follows that

$$H \cap C(x^1) = H \cap C(x^2) = C(x^1) \cap C(x^2).$$

Thus, the intersection of the two cones is a face for each of them. ■

Proof of Lemma 2.1. Consider a facet Y of $C(x^0)$ and affinely independent vectors $y^1, \ldots, y^n$ in $rel.int(Y)$ such that $conv.hull(y^1, \ldots, y^n)$ is not intersected by any hyperplane of $\mathcal{H} \setminus lin.hull(Y)$. Then, since $conv.hull(y^1, \ldots, y^n) \subset Y$, there is $y^{n+1} \notin C(x^0)$ such that the whole simplex $conv.hull(y^1, \ldots, y^{n+1})$ is not intersected by any of the hyperplanes in $\mathcal{H} \setminus lin.hull(Y)$. This simplex is contained in some cone $C(x)$, $x \in S \setminus \{x^0\}$. It follows that $C(x^0) \cap C(x)$ contains $conv.hull(y^1, \ldots, y^n)$. Therefore, $\dim C(x^0) \cap C(x) \geq n - 1$. At the same time, $\dim C(x^0) \cap C(x) \leq n - 1$ because $C(x^0) \neq C(x)$ and $C(x^0) \cap C(x)$ is a common face of $C(x^0)$ and $C(x)$ by Proposition 2.1. Hence, the dimension of the common face $C(x^0) \cap C(x)$ is exactly $n - 1$. The cone $C(x)$ belongs to the normal fan because $conv.hull(y^1, \ldots, y^{n+1})$ has dimension n and is contained in $C(x)$. Since the vectors $y^1, \ldots, y^n$ are affinely independent, they uniquely identify Y. At the same time they belong to $C(x^0) \cap C(x)$, which is a common facet of $C(x^0)$ and $C(x)$, as proved above. Therefore, $C(x^0) \cap C(x)$ is exactly Y. This implies the lemma. ∎

Proof of Lemma 2.2. That x^1 and x^2 are vertices follows from the full-dimensionality of the respective normal cones.

Consider z in the relative interior of the common facet $Y = C(x^1) \cap C(x^2)$. Note that Y is the minimal face of $C(x^1)$ containing z. Assume that x^1 and x^2 are not adjacent. Then there is $x^3 \in S$ such that x^3 is a vertex of *conv.hull*(S) and

$$z^T x^3 = z^T x^2 = z^T x^1.$$

Since x^3 is a vertex of the convex hull, $x^1 - x^3$ and $x^1 - x^2$ are linearly independent. Therefore, the face

$$C(x^1) \cap \{y : (x^1 - x^3)^T y = 0, (x^1 - x^2)^T y = 0\}$$

of $C(x^1)$ has dimension less than $n - 1$. This face should contain z. We have a contradiction because Y is the minimal face containing z and Y has dimension $n - 1$. Therefore, x^1 and x^2 are adjacent.

Let x^1 and x^2 be adjacent. Let $Y = C(x^1) \cap C(x^2)$. The set Y is a common face of $C(x^1)$ and $C(x^2)$. On the other hand, since x^1 and x^2 are adjacent, Y has dimension $n - 1$. ∎

The observation below follows from standard results of the theory of polyhedra:

Observation 5.1. *The distance between two disjoint polyhedra P_1 and P_2 defined by systems of linear inequalities with rational coefficients is lower bounded by a positive value whose binary size is polynomial in the sizes of binary encodings of the linear systems defining P_1 and P_2.*

Proof. Note that the $\| \cdot \|_2$-norm is lower bounded by the $\| \cdot \|_\infty$-norm. The problem of minimizing $\|y_1 - y_2\|_\infty$ subject to $y_1 \in P_1$ and $y_2 \in P_2$ can be written as a linear program whose binary size is polynomial in the binary sizes of the systems describing P_1 and P_2. From the theory of linear programming, it is well known (see e.g. [15]) that if a linear program with rational coefficients has an optimal solution then it has one of polynomial size. Therefore, our linear program has an optimal solution consisting of a pair (y_1, y_2) where y_1 and y_2 are of polynomial size. The respective value $\|y_1 - y_2\|_\infty$ is also of polynomial size. This value is positive because P_1 and P_2 are disjoint. ∎

Remark to the proof of Theorem 1.1: Indeed, $\delta([z^{t-1}, z^t])$ is the distance from $[z^{t-1}, z^t]$ to one of the intersections $H_1 \cap H_2$, where H_1 and H_2 are hyperplanes in $\mathcal{H}$. The respective intersection $H_1 \cap H_2$ is the solution set of a linear system of the form

$$(x_1' - x_1'')^T y = 0, \quad (x_2' - x_2'')^T y = 0,$$

where $x_i', x_i'' \in S$, $i = 1, 2$. Let P_1 be the solution set of the above system. Let $P_2 = [z^{t-1}, z^t]$. The constraint $y \in [z^{t-1}, z^t]$ defining P_2 can be equivalently written as a system of $n + 1$ linear constraints with coefficients whose binary sizes are polynomially bounded in the sizes of z^{t-1} and z^t. According to our choice of μ and p, the binary sizes of z^{t-1} and z^t are polynomially bounded in the sizes of c^t, $t = 0, \ldots, k$, and Δ. Now, taking into account Observation 5.1, we see that $\delta([z^{t-1}, z^t])$, which is the distance between disjoint polyhedra P_1 and P_2, is lower bounded by a positive value whose binary size is polynomially bounded in the sizes of c^t, $t = 0, \ldots, k$, and Δ. ∎

References

[1] Borgwardt, K.-H. The Simplex Method: A Probabilistic Analysis. Algorithms and Combinatorics: Study and Research Texts. Springer, Berlin (1987)

[2] Chubanov, S. A dual polyhedral complex and pivot rules for simplex algorithms. An unpublished manuscript. This paper was presented at GOR 2007.

[3] Chubanov, S. A scaling algorithm for optimizing arbitrary functions over vertices of polytopes. Math. Program. (2020). https://doi.org/10.1007/s10107-020-01522-0

[4] Dadush, D., and Hähnle, N. On the shadow simplex method for curved polyhedra. *Discrete Comput. Geom.* DOI 10.1007/s00454-016-9793-3 (2016)

[5] De Loera, J.A., Hemmecke, R., and Lee, J. On augmentation algorithms for linear and integer-linear programming: from Edmonds-Karp to Bland and beyond. *SIAM J. Optim.* **25**, 2494-2511 (2015)

[6] De Loera, J., Rambau, J., Santos, F. Triangulations: Structures for Algorithms and Applications. Springer-Verlag Berling Heidelberg (2010)

[7] Del Pia, A., Michini, C. On the Diameter of Lattice Polytopes. *Discrete Comput. Geom.* **55**, 681-687 (2016)

[8] Del Pia, A., Michini, C. Short simplex paths in lattice polytopes. *Optimization Online* (2020)

[9] Deza, A., and Pournin, L. Improved bounds on the diameter of lattice polytopes. *Acta Math. Hungar.*, **154** (**2**), 457-469 (2018)

[10] Deza, A., Pournin, L., and Sukegawa, N. The diameter of lattice zonotopes. *Proc. Amer. Math. Soc.*, **148**, 3507-3516 (2020)

[11] Kitahara, T., and Mizuno, S. A proof by the simplex method for the diameter of a (0,1)-polytope. *Optimization Online* (2011)

[12] Kleinschmidt, P., and Onn, S. On the diameter of convex polytopes. *Discrete Mathematics* **102**, 75-77 (1992)

[13] Naddef, D.: The Hirsch conjecture is true for (0,1)-polytopes. Math. Program., Ser. A **45**, 109-110 (1989)

[14] Orlin J.B., Punnen A.P., Schulz A.S. (2009) Integer Programming: Optimization and Evaluation Are Equivalent. In: Dehne F., Gavrilova M., Sack JR., Tóth C.D. (eds) Algorithms and Data Structures. WADS 2009. Lecture Notes in Computer Science, vol 5664. Springer, Berlin, Heidelberg

[15] Schrijver, A. Theory of linear and integer programming. Wiley (1986)

[16] Schulz, A.S., Weismantel R., and Ziegler G.M. 0/1-integer programming: Optimization and augmentation are equivalent. *Lecture Notes in Computer Science* **979** 473-483 (1995)

[17] Schulz, A.S., Weismantel R. The Complexity of Generic Primal Algorithms for Solving General Integer Programs. *Mathematics of Operations Research* **27**, 681-692 (2002)

[18] Schulz, A.S. On the relative complexity of 15 problems related to 0/1-integer programming. Chapter 19 in W.J. Cook, L. Lovász, J. Vygen (eds.): Research Trends in Combinatorial Optimization, Springer, Berlin, 399-428 (2009)

[19] Ziegler, G.M.: Lectures on Polytopes. Graduate Texts in Mathematics. Springer-Verlag, New York, USA (2007)