ELSEVIER Theoretical Computer Science 147 (1995) 181–210

Theoretical Computer Science

The complexity and approximability of finding maximum feasible subsystems of linear relations

Edoardo Amaldi [a], Viggo Kann [b,*]

[a] Department of Mathematics, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland

[b] Department of Numerical Analysis, and Computing Science, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received September 1993; revised July 1994

Communicated by G. Ausiello

Abstract

We study the combinatorial problem which consists, given a system of linear relations, of finding a maximum feasible subsystem, that is a solution satisfying as many relations as possible. The computational complexity of this general problem, named Max FLS, is investigated for the four types of relations $=$, $\geqslant$, $>$ and $\neq$. Various constrained versions of Max FLS, where a subset of relations must be satisfied or where the variables take bounded discrete values, are also considered. We establish the complexity of solving these problems optimally and, whenever they are intractable, we determine their degree of approximability. Max FLS with $=$, $\geqslant$ or $>$ relations is NP-hard even when restricted to homogeneous systems with bipolar coefficients, whereas it can be solved in polynomial time for $\neq$ relations with real coefficients. The various NP-hard versions of Max FLS belong to different approximability classes depending on the type of relations and the additional constraints. We show that the range of approximability stretches from Apx-complete problems which can be approximated within a constant but not within every constant unless P = NP, to NPO PB-complete ones that are as hard to approximate as all NP optimization problems with polynomially bounded objective functions. While Max FLS with equations and integer coefficients cannot be approximated within p^{ε} for some $\varepsilon > 0$, where p is the number of relations, the same problem over $GF(q)$ for a prime q can be approximated within q but not within q^{ε} for some $\varepsilon > 0$. Max FLS with strict or nonstrict inequalities can be approximated within 2 but not within every constant factor. Our results also provide strong bounds on the approximability of two variants of Max FLS with $\geqslant$ and $>$ relations that arise when training perceptrons, which are the building blocks of artificial neural networks, and when designing linear classifiers.

* Corresponding author. Email: viggo@nada.kth.se.

SSDI 0304-3975(94)00254-1

1. Introduction

We consider the general problem of finding maximum feasible subsystems of linear relations for the four types of relations $=$, $\geqslant$, $>$ and $\neq$. The basic versions, named MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >, \neq\}$, are defined as follows: Given a linear system $\boldsymbol{A}\boldsymbol{x}\mathcal{R}\boldsymbol{b}$ with a matrix A of size $p \times n$, find a solution $\boldsymbol{x} \in \mathbb{R}^n$ which satisfies as many relations as possible.

Different variants of these combinatorial problems occur in various fields such as pattern recognition [38, 13], operations research [23, 18, 17] and artificial neural networks [3, 21, 33, 32].

Whenever a system of linear equations or inequalities is consistent, it can be solved in polynomial time using an appropriate linear programming method [27]. If the system is inconsistent, standard algorithms provide solutions that minimize the least mean squared error. But such solutions, which are appropriate in linear regression, are not satisfactory when the objective is to maximize the number of relations that can be simultaneously satisfied.

Previous works have focused mainly on algorithms for tackling various versions of MAX FLS. Among others, the weighted variants were studied in which each relation has an associated weight and the goal is to maximize the total weight of the satisfied relations. Surprisingly enough, only a few results are known on the complexity of solving some special cases of MAX FLS to optimality and none concerns their approximability.

Johnson and Preparata proved that the OPEN HEMISPHERE and CLOSED HEMISPHERE problems, which are equivalent to MAX FLS$^{>}$ and MAX FLS$^{\geqslant}$, respectively, with homogeneous systems and no pairs of collinear row vectors of A, are NP-hard [23]. Moreover, they devised a complete enumeration algorithm with $O(p^{n-1} \log p)$ time-complexity, where n and p denote the number of variables and relations, that is also applicable to the weighted and mixed variant.

Greer developed a tree method for maximizing functions of systems of linear relations that is more efficient than complete enumeration but still exponential in the worst case [18]. This general procedure can be used to solve MAX FLS with any of the four types of relations.

Recently the problem of training perceptrons, which is closely related to MAX FLS$^{>}$ and MAX FLS$^{\geqslant}$, has attracted a considerable interest in machine learning and discriminant analysis [19]. For nonlinearly separable sets of vectors, the objective is either to maximize the consistency, i.e. the number of vectors that are correctly classified, or to minimize the number of misclassifications. These complementary problems are equivalent to solve optimally but their approximability can differ enormously. While some heuristic algorithms have been proposed in [15, 14, 32], Amaldi extended Johnson's and Preparata's result by showing that solving these problems to optimality is NP-hard even when restricted to perceptrons with bipolar inputs in $\{-1, 1\}$ [3]. In other words, MIXED HEMISPHERE remains NP-hard if the coefficients take bipolar values. Höffgen et al. proved in [21] that minimizing the number of misclassifications is at least as hard to approximate as minimum set cover. Thus, according to [9], it is very hard to approximate. But nothing is known about the approximability of maximizing perceptron consistency.

Variants with mixed types of relations do also occur in practice. A simple example arises for instance in the field of linear numeric editing. Assuming that a database is characterized by all vectors in a given polytope, we try to associate to every given vector a database vector while leaving unchanged as many components as possible [18]. In terms of linear systems, this amounts to finding a solution that satisfies as many $\neq$ relations as possible subject to a set of nonstrict inequality constraints.

There have recently been new substantial progresses in the study of the approximability of NP-hard optimization problems. Various classes have been defined and different reductions preserving approximability have been used to compare the approximability of optimization problems (see [25]). Moreover, the striking results which have recently been obtained in the area of interactive proofs triggered new advances in computational complexity theory. Strong bounds were derived on the approximability of several famous problems like maximum independent set, minimum graph colouring and minimum set cover [6, 31, 9]. These results have also important consequences on the approximability of other optimization problems [8].

A remarkably well-characterized approximation problem is that relative to Max FLS$^=$ over $GF(q)$ where the equations are degree 2 polynomials that do not contain any squares as monomials. Håstad et al. have shown that this problem can be approximated within $q^2/(q-1)$ but not within $q-\varepsilon$ for any $\varepsilon > 0$ unless P = NP [20]. The same problem over the rational numbers or over the real numbers cannot be approximated within $n^{1-\varepsilon}$ for any $\varepsilon > 0$, where n is the number of variables.

The paper is organized as follows. Section 2 provides a brief overview of the important facts about the hierarchy of approximability classes that will be used throughout this work. In Section 3 we prove that solving the basic Max FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ optimally is intractable even for homogeneous systems with bipolar coefficients and we determine their degree of approximability. Various constrained versions of the basic problems are considered in Sections 4 and 5. First we focus on variants where a subset of relations must be satisfied and the objective is to find a solution fulfilling all mandatory relations and as many optional ones as possible. Then we consider the particular cases in which the variables are restricted to take a finite number of discrete values. In Section 6 the overall structure underlying the various results is discussed and open questions are mentioned. The appendix is devoted to three interesting special cases whose last two arise in discriminant analysis and machine learning.

2. Approximability classes

Definition 1 (Crescenzi and Panconesi [12]). An NP optimization (NPO) problem over an alphabet Σ is a four-tuple $F = (\mathscr{I}_F, S_F, m_F, opt_F)$, where

- $\mathscr{I}_F \subseteq \Sigma^*$ is the space of *input instances*. The set $\mathscr{I}_F$ must be recognizable in polynomial time.

- $S_F(x) \subseteq \Sigma^*$ is the space of *feasible solutions* on input $x \in \mathscr{I}_F$. The only requirement on S_F is that there exist a polynomial q and a polynomial time computable predicate π such that for all x in $\mathscr{I}_F$, S_F can be expressed as $S_F(x) = \{y : |y| \leqslant q(|x|) \wedge \pi(x, y)\}$ where q and π only depend on F.
- $m_F : \mathscr{I}_F \times \Sigma^* \to \mathrm{N}$, the *objective function*, is a polynomial time computable function. $m_F(x, y)$ is defined only when $y \in S_F(x)$.
- $opt_F \in \{\max, \min\}$ tells if F is a *maximization* or a *minimization* problem.

Solving an optimization problem F given the input $x \in \mathscr{I}_F$ means finding a $y \in S_F(x)$ such that $m_F(x, y)$ is optimum, that is as large as possible if $opt_F = \max$ and as small as possible if $opt_F = \min$. Let $opt_F(x)$ denote this optimal value of m_F.

Approximating an optimization problem F given the input $x \in \mathscr{I}_F$ means finding any $y' \in S_F(x)$. How good the approximation is depends on the relation between $m_F(x, y')$ and $opt_F(x)$. The *performance ratio* of a feasible solution with respect to the optimum of a maximization problem F is defined as $R_F(x, y) = opt_F(x)/m_F(x, y)$ where $x \in \mathscr{I}_F$ and $y \in S_F(x)$.

Definition 2. An optimization problem F *can be approximated within c* for a constant c if there exists a polynomial-time algorithm A such that for all instances $x \in \mathscr{I}_F$, $A(x) \in S_F(x)$ and $R_F(x, A(x)) \leqslant c$. More generally, an optimization problem F *can be approximated within $p(n)$* for a function $p : \mathrm{Z}^+ \to \mathbb{R}^+$ if there exists a polynomial-time algorithm A such that for every $n \in \mathrm{Z}^+$ and for all instances $x \in \mathscr{I}_F$ with $|x| = n$ we have that $A(x) \in S_F(x)$ and $R_F(x, A(x)) \leqslant p(n)$.

Although various reductions preserving approximability within constants have been proposed (see [7, 10, 26, 35]), the L-reduction is the most easy to use and the most restrictive one [25].

Definition 3 (Papadimitriou and Yannakakis [36]). Given two NPO problems F and G and a polynomial-time transformation $f : \mathscr{I}_F \to \mathscr{I}_G$. f is an *L-reduction* from F to G if there are positive constants α and β such that for every instance $x \in \mathscr{I}_F$

(i) $opt_G(f(x)) \leqslant \alpha \cdot opt_F(x)$,

(ii) for every solution y of $f(x)$ with objective value $m_G(f(x), y) = c_2$ we can in polynomial time find a solution y' of x with $m_F(x, y') = c_1$ such that $|opt_F(x) - c_1| \leqslant \beta\ |opt_G(f(x)) - c_2|$.

If F L-reduces to G we write $F \leqslant_L^p G$.

The composition of L-reductions is an L-reduction. If F L-reduces to G with constants α and β and there is a polynomial-time approximation algorithm for G with worst-case relative error ε, then there is a polynomial-time approximation algorithm for F with worst-case relative error $\alpha\beta\varepsilon$ [36]. Obviously, an L-reduction with $\alpha = \beta = 1$ is a *cost preserving transformation*.

Definition 4 (Berman and Schnitger [10] and Kolaitis and Thakur [29]). An NPO problem F is *polynomially bounded* if there is a polynomial p such that

$$\forall x \in \mathscr{I}_F \, \forall y \in S_F(x), \quad m_F(x, y) \leqslant p(|x|).$$

The class of all polynomially bounded NPO problems is called NPO PB.

All versions of MAX FLS are included in NPO PB since their objective function is the number of satisfied relations or the total weight of the satisfied relations.

Definition 5. Given an NPO problem F and a class C, F *is* C*-hard* if every $G \in C$ can be L-reduced to F. F *is* C*-complete* if $F \in C$ and F is C-hard.

The range of approximability of NPO problems stretches from problems that can be approximated within every constant, i.e. that have a *polynomial-time approximation scheme*, to problems that cannot be approximated within n^{ε} for some $\varepsilon > 0$, where n is the size of the input instance, unless P = NP.

In the middle of this range we find the important class APX, which consists of problems that can be approximated within some constant, and the subclass MAX SNP, which is syntactically defined [36]. Several maximization problems have been shown to be MAX SNP-complete, and recently it was shown that these problems are also APX-complete [28, 7].

Provided that P $\neq$ NP it is impossible to find a polynomial-time algorithm that approximates a MAX SNP-hard (or APX-hard) problem within every constant [6]. Thus showing a problem to be APX-complete describes the approximability of the problem quite well: it can be approximated within a constant but not within every constant.

The maximum independent set problem cannot be approximated within n^{ε} for some $\varepsilon > 0$, where n is the number of nodes in the input graph [6]. If there is an approximation preserving reduction from MAX IND SET to an NPO problem F we say that F is MAX IND SET-hard, which means that it is at least as hard to approximate as the maximum independent set problem.

There exist natural problems that are complete in NPO PB, for example MAX PB 0–1 PROGRAMMING [10]. These are the hardest problems to approximate in this class since every NPO PB problem can be reduced to them using an approximation preserving reduction [11].

The purpose of this paper is to show where the different versions of MAX FLS are placed in this hierarchy of approximability classes. We will see that the approximability of apparently similar variants can differ enormously.

3. Complexity of MAX FLS$^{\mathscr{R}}$

In this section we focus on the basic MAX FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$. We first prove that these problems are hard to solve optimally and then determine their degree of approximability. Several special cases that can be solved in polynomial time are also mentioned. Note that MAX FLS with $\neq$ relations is trivial because any such system is feasible. Indeed, for any finite set of hyperplanes associated with a set of linear relations there exists a vector $x \in \mathbb{R}^n$ that does not belong to any of them.

3.1. Optimal solution

In order to determine the complexity of solving MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >\}$ to optimality, we consider the corresponding decision versions that are no harder than the original optimization problems. Given a linear system $A\boldsymbol{x}\mathcal{R}\boldsymbol{b}$ where A is of size $p \times n$ and an integer K with $1 \leqslant K \leqslant p$, does there exist a solution $\boldsymbol{x} \in \mathbb{R}^n$ satisfying at least K relations of the system?

In the homogeneous versions of MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant\}$ we are not interested in the trivial solutions where all variables occurring in the satisfied relations are zero.

Geometrically, homogeneous MAX FLS$^=$ can be viewed as follows. Given a set of p points in $\mathbb{R}^n$, find a hyperplane passing through the origin and containing the largest number of points.

Theorem 1. MAX FLS$^=$ *is* NP-*hard even when restricted to homogeneous systems with ternary coefficients in* $\{-1, 0, 1\}$.

Proof. We proceed by polynomial-time reduction from the known NP-complete problem EXACT 3-SETS COVER that is defined as follows [16]. Given a set S with $|S| = 3q$ elements and a collection $C = \{C_1, \ldots, C_m\}$ of subsets $C_j \subseteq S$ with $|C_j| = 3$ for $1 \leqslant j \leqslant m$, does C contain an exact cover, i.e. $C' \subseteq C$ such that each element s_i of S belongs to exactly one element of C' ?

Let (S, C) be an arbitrary instance of EXACT 3-SETS COVER. We will construct a particular instance of MAX FLS$^=$ denoted by $(A, \boldsymbol{b})$ such that the answer to the former one is affirmative if and only if the answer to the latter one is also affirmative.

The idea is to construct a system containing one variable x_j for each subset $C_j \in C$, $1 \leqslant j \leqslant m$, and at least one equation for each element s_i of S, $1 \leqslant i \leqslant 3q$. We consider the following set of equations:

$$\sum_{j=1}^{|C|} a_{ij} x_j = 1 \quad \text{for } i = 1, \ldots, 3q, \tag{1}$$

where $a_{ij} = 1$ if the element $s_i \in C_j$ and $a_{ij} = 0$ otherwise, as well as the additional ones

$$x_j = 1 \quad \text{for } j = 1, \ldots, m, \tag{2}$$

$$x_j = 0 \quad \text{for } j = 1, \ldots, m. \tag{3}$$

Moreover, we set $K = 3q + m$. Clearly K is equal to the largest number of equations that can be simultaneously satisfied.

Given any exact cover $C' \subseteq C$ of (S, C), the vector $\boldsymbol{x}$ defined by

$$x_j = \begin{cases} 1 & \text{if } C_j \in C', \\ 0 & \text{otherwise} \end{cases}$$

satisfies all equations of type (1) and exactly m equations of types (2)–(3). Hence $\boldsymbol{x}$ fulfils $K = 3q + m$ equations.

Conversely, suppose that we have a solution $\boldsymbol{x}$ that satisfies at least $K = 3q + m$ equations of $(A, \boldsymbol{b})$. By construction, this implies that $\boldsymbol{x}$ fulfils all equations of type (1) and m equations of types (2)–(3). Thus the subset $C' \subseteq C$ defined by $C_j \in C'$ if and only if $x_j = 1$ is an exact cover of (S, C).

The reduction can easily be extended to homogeneous MAX FLS$^=$. We just need to add a new variable x_{m+1} with $a_{i,m+1} = -b_i$ for all i, $1 \leqslant i \leqslant p$, and to observe that in any nontrivial solution $\boldsymbol{x}$ we must have $x_{m+1} \neq 0$. Indeed, $x_{m+1} = 0$ would necessarily imply $x_j = 0$ for all j, $1 \leqslant j \leqslant m$. □

The question arises as to whether the problem is still intractable for systems with bipolar coefficients in $\{-1, 1\}$.

Corollary 2. MAX FLS$^=$ *remains* NP-*hard for homogeneous systems with bipolar coefficients.*

Proof. We extend the above reduction by using a duplication technique that allows to reduce systems with ternary coefficients in $\{-1, 0, 1\}$ to systems with bipolar coefficients in $\{-1, 1\}$. The idea is to replace each variable x_j, $1 \leqslant j \leqslant n$, by two variables that are forced to be equal and that have only bipolar coefficients.

Consider an arbitrary instance of homogeneous MAX FLS$^=$ with ternary coefficients arising from an instance of EXACT 3-SETS COVER. Without loss of generality, we can assume that $a_{ij} \in \{-2, 0, 2\}$. This simple multiplication by a factor 2 does not affect the set of solutions but makes all coefficients even. Since the absolute value of a_{ij} is either 0 or 2, we can construct a system with bipolar coefficients that is equivalent to (1)–(3) by duplicating the variables and adding new equations.

Suppose that n different variables occur in (1)–(3). We associate with each equation $\boldsymbol{a}\boldsymbol{x} = 0$ with n variables an equation $\tilde{\boldsymbol{a}}\boldsymbol{y} = 0$ with $2n$ variables. The coefficient vector $\tilde{\boldsymbol{a}}$ is determined as follows:

$$\tilde{a}_j = \begin{cases} 1 & \text{if } a_{\lceil j/2 \rceil} = 2 \text{ or } a_{\lceil j/2 \rceil} = 0 \text{ and } j \text{ is odd}, \\ -1 & \text{if } a_{\lceil j/2 \rceil} = -2 \text{ or } a_{\lceil j/2 \rceil} = 0 \text{ and } j \text{ is even}, \end{cases}$$

where $1 \leqslant j \leqslant 2n$. This defines a mapping from $\{-2, 0, 2\}^n$ into $\{-1, 1\}^{2n}$ that associates to each component a_i of $\boldsymbol{a}$ the two components $\tilde{a}_{2i-1}$ and $\tilde{a}_{2i}$ of $\tilde{\boldsymbol{a}}$ such that $a_i = \tilde{a}_{2i-1} + \tilde{a}_{2i}$ for all $i, 1 \leqslant i \leqslant n$. For any $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\boldsymbol{a}\boldsymbol{x} = 0$, the vector $\boldsymbol{y} \in \mathbb{R}^{2n}$ given by $y_j = x_{\lceil j/2 \rceil}$ satisfies the corresponding equation $\tilde{\boldsymbol{a}}\boldsymbol{y} = 0$. Furthermore, if $y_{2i} = y_{2i-1}$ for $1 \leqslant i \leqslant n$ and $\tilde{\boldsymbol{a}}\boldsymbol{y} = 0$ then the vector $\boldsymbol{x}$ given by $x_i = y_{2i-1}$, $1 \leqslant i \leqslant n$, is a solution of $\boldsymbol{a}\boldsymbol{x} = 0$. Thus, in order to construct an equivalent system with only $\{-1,1\}$ coefficients we must add new equations that eliminate the n additional degrees of freedom that have been introduced by mapping the original n-dimensional problem into the $2n$-dimensional one. In particular, we should ensure that $y_{2i} = y_{2i-1}$ for all $1 \leqslant i \leqslant n$. This can be achieved by satisfying simultaneously the 2 following homogeneous equations with bipolar coefficients:

$$\begin{aligned} y_{2i} - y_{2i-1} + \sum_{1 \leqslant l \leqslant n,\ l \neq i} (y_{2l} - y_{2l-1}) = 0, \\ y_{2i} - y_{2i-1} - \sum_{1 \leqslant l \leqslant n,\ l \neq i} (y_{2l} - y_{2l-1}) = 0. \end{aligned} \tag{4}$$

Indeed, $v + w = 0$ and $v - w = 0$ necessarily imply $v = w = 0$. Many equivalent equations of type (4) are needed in order to guarantee that $y_{2i} = y_{2i-1}$ for each i, $1 \leqslant i \leqslant n$. For each constraint $y_{2i} = y_{2i-1}$, we include a number of pairs of Eqs. (4) that is larger than the number of equations of type (1)–(3). This can always be done by selecting the coefficients $\tilde{a}_{2l}, \tilde{a}_{2l-1} \in \{-1,1\}$ occurring in

$$\sum_{1 \leqslant l \leqslant n,\ l \neq i} \tilde{a}_{2l} y_{2l} + \tilde{a}_{2l-1} y_{2l-1}$$

in different ways (any choice with $\tilde{a}_{2l} = -\tilde{a}_{2l-1}$ is adequate).

It is worth noting that this general technique can be adapted to reduce any system of equations whose coefficients are restricted to take a finite number of values to an equivalent system with only bipolar coefficients. □

This result has an immediate consequence on the complexity of MAX FLS$^{\geqslant}$ that is stronger than that established in [23].

Corollary 3. MAX FLS$^{\geqslant}$ *is* NP-*hard for homogeneous systems with bipolar coefficients.*

Proof. By simple polynomial-time reduction from MAX FLS$^{=}$. Let $(A, \boldsymbol{b})$ be an arbitrary instance of MAX FLS$^{=}$. For each equation $\boldsymbol{a}^i \boldsymbol{x} = 0$ where $\boldsymbol{a}^i$ denotes the ith row of A, $1 \leqslant i \leqslant p$, we consider the two inequalities $\boldsymbol{a}^i \boldsymbol{x} \geqslant 0$ and $-\boldsymbol{a}^i \boldsymbol{x} \geqslant 0$. Clearly, there exists a vector satisfying at least K equations of $(A, \boldsymbol{b})$ if and only if there exists a solution satisfying at least $p+K$ inequalities of the corresponding system with $2p$ inequalities. □

This is also true for systems with strict inequalities.

Theorem 4. MAX FLS$^{>}$ *is* NP-*hard for homogeneous systems with bipolar coefficients.*

Proof. We proceed by polynomial-time reduction from the known NP-complete problem MAX IND SET that is defined as follows [16]. Given an undirected graph $G = (V, E)$, find a largest independent set $V' \subseteq V$, i.e. a largest subset of nonadjacent nodes.

Let $G = (V, E)$ be an arbitrary instance of MAX IND SET. For each edge $(v_i, v_j) \in E$ we construct the inequality

$$x_i + x_j < 0 \tag{5}$$

and for each node $v_i \in V$ the inequality

$$x_i > 0. \tag{6}$$

Thus we have a system with $|V|$ variables and $|E| + |V|$ strict inequalities.

We claim that the given graph G contains an independent set I of size s if and only if there exists a solution $\boldsymbol{x}$ satisfying all inequalities of type (5) and s inequalities of type (6).

Given an independent set $I \subseteq V$ of size s, the solution obtained by setting

$$x_i = \begin{cases} 1 & \text{if } v_i \in I, \\ -2 & \text{otherwise} \end{cases}$$

satisfies all edge-inequalities (5) and all the node-inequalities (6) corresponding to a node $v_i \in I$. Moreover, the $|V| - s$ inequalities associated with $v_i \notin I$ are not fulfilled because $x_i < 0$.

Conversely, given an appropriate solution $\boldsymbol{x}$ we consider the set $I \subseteq V$ containing all nodes whose second type inequality is satisfied. The size of I is obviously equal to s. We verify by contradiction that I is an independent set. Suppose that $\boldsymbol{x}$ fulfils all edge-inequalities and s node-inequalities. If I contains two adjacent nodes v_i and v_j, then we must have, on one hand, $x_i > 0$ and $x_j > 0$ and, on the other hand, $x_i + x_j < 0$, which is impossible. Hence I is an independent set of cardinality s.

In order to complete the proof we must make sure that all edge-inequalities are satisfied. This can be achieved by adding $|V|$ equivalent copies of each one of them, and in particular by multiplying each edge-inequality by different integer factors $f \in \{2, \ldots, |V| + 1\}$. Thus we have a system with $(|V| + 1)|E|$ inequalities of the first type and $|V|$ of the second one. Clearly, the given graph G contains an independent set I of size s if and only if there exists a solution $\boldsymbol{x}$ satisfying $(|V| + 1)|E| + s$ strict inequalities.

This polynomial-time reduction can be extended to MAX FLS$^>$ with bipolar coefficients by applying Carver's transposition theorem [37]. According to this result, a homogeneous system $A\boldsymbol{x} < \boldsymbol{0}$ is feasible if and only if $\boldsymbol{y} = \boldsymbol{0}$ is the unique solution of

$$A^{\mathrm{t}}\boldsymbol{y} = \boldsymbol{0}, \qquad \boldsymbol{y} \geqslant \boldsymbol{0}. \tag{7}$$

Thus any instance of MAX FLS$^>$ can be associated with such a system. Using the technique described in the proof of Corollary 2, it is then possible to construct, for each system (7) with integer coefficients taking their values in $\{-(|V|+1), \ldots, 0, \ldots, |V|+1\}$, an equivalent system with only bipolar coefficients. It suffices to add a large enough number of appropriate equations forcing the new variables associated with any original variable to be equal. $\square$

Consequently, MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >\}$ is intractable not only when the points corresponding to the rows of A lie on the n-dimensional hypersphere but also when they belong to the n-dimensional hypercube. In such instances no pairs of relations differ by a multiplicative factor.

Since these problems are NP-hard for bipolar coefficients, they turn out to be strongly NP-hard, i.e. intractable even with respect to unary coding of the data. According to a well-known result concerning polynomially bounded problems [16], they do not have a fully polynomial-time approximation scheme (an ε-approximation scheme where the running time is bounded by a polynomial in both the size of the instance and $1/\varepsilon$) unless P = NP.

Before turning to the approximability of MAX FLS, it is worth noting that some simple special cases are polynomially solvable. If the number of variables n is constant, MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >\}$ can be solved in polynomial time using Greer's algorithm that has an $O(n \cdot p^n/2^{n-1})$ time-complexity, where p and n denote, respectively, the number of relations and variables. For a constant number of relations, these problems are trivial since all subsystems can be checked in time $O(n)$. Moreover, they are easy when all maximal feasible subsystems (with respect to inclusion) have a maximum number of relations because a greedy procedure is guaranteed to find a maximum feasible subsystem.

3.2. Approximate solution

The previous NP-hardness results make extremely unlikely the existence of polynomial time methods for solving the three basic versions of MAX FLS to optimality. But in practice optimal solutions are not always required and approximate algorithms providing solutions that are guaranteed to be a fixed percentage away from the actual optimum are often satisfactory.

We will now show that MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >\}$ and integer coefficients cannot be approximated within every constant unless P = NP. The proofs are by L-reductions from the known APX-complete problem MAX 2SAT that is defined as follows [16]. Given a finite set X of variables and a set $C = \{C_1, \ldots, C_m\}$ of disjunctive clauses with at most 2 literals in each clause, find a truth assignment for X that satisfies as many clauses of C as possible.

In homogeneous MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant\}$ we are only interested in solutions where the variable(s) occurring in the largest number of satisfied equations are nonzero. This rules out the trivial solutions as well as those obtained by setting all variables to zero except one of the variables occurring in the smallest number of equations.

Theorem 5. MAX FLS$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geqslant\}$ *is* APX-*hard even when restricted to homogeneous systems with discrete coefficients in* $\{-4,-3,-2,-1,0,1,2,3,4\}$ *and no pairs of identical relations.*

Proof. We first consider MAX FLS$^{=}$. Let (X, C) with $C = \{C_1, \ldots, C_m\}$ be an arbitrary instance of MAX 2SAT. For each clause C_i, $1 \leqslant i \leqslant m$, containing two variables x_{j_1} and x_{j_2}

we construct the following equations:

$$a_{ij_1} x_{j_1} + a_{ij_2} x_{j_2} = 2, \tag{8}$$

$$a_{ij_1} x_{j_1} + a_{ij_2} x_{j_2} = 0, \tag{9}$$

$$x_{j_1}, x_{j_2} = 1, \tag{10}$$

$$x_{j_1}, x_{j_2} = -1, \tag{11}$$

where $a_{ij} = 1$ if x_j occurs positively in C_i and $a_{ij} = -1$ if x_j occurs negatively. Thus we have a system with $6m$ equations.

Given a truth assignment that satisfies s clauses of the MAX 2SAT instance, we immediately get a solution $\boldsymbol{x}$ that satisfies $2m+s$ equations of the MAX FLS$^=$ instance. This is simply achieved by setting the variables x_j to 1 or -1 depending on whether the corresponding boolean variable is TRUE or FALSE in the assignment.

Consider any solution $\boldsymbol{x}$ of the MAX FLS$^=$ instance. For each i, $1 \leqslant i \leqslant m$, at most 3 equations can be simultaneously satisfied: at most one of (8)–(9) and at most one of (10)–(11) for each of the two variables. If any component of $\boldsymbol{x}$ is neither 1 nor -1, we can set it to 1 without decreasing the number of satisfied equations. In other words, we can suppose that any solution $\boldsymbol{x}$ has bipolar components.

Consequently, we have a correspondence between solutions of the MAX 2SAT instance satisfying s clauses and solutions of the MAX FLS$^=$ instance fulfilling $2m+s$ equations. Since opt[MAX FLS$^=$]$\leqslant 3m$ and since there exists an algorithm providing a truth assignment that satisfies at least $\lceil m/2 \rceil$ of the clauses in any MAX 2SAT instance (see for example [22]), we have opt[MAX FLS$^=$]$\leqslant 6 \cdot$ opt[MAX 2SAT]. Thus all conditions for an L-reduction are fulfilled.

This L-reduction can be extended to homogeneous systems with coefficients in $\{-4,-3,-2,-1,0,1,2,3,4\}$ and no pairs of identical equations. For each clause C_i, $1 \leqslant i \leqslant m$, we add a new variable $x_{|X|+i}$ and we consider the following equations:

$$a_{ij_1} x_{j_1} + a_{ij_2} x_{j_2} - 2x_{|X|+i} = 0, \tag{12}$$

$$a_{ij_1} x_{j_1} + a_{ij_2} x_{j_2} = 0, \tag{13}$$

$$x_{j_1} - x_{|X|+i} = 0, \tag{14}$$

$$x_{j_1} + x_{|X|+i} = 0, \tag{15}$$

$$x_{j_2} - x_{|X|+i} = 0, \tag{16}$$

$$x_{j_2} + x_{|X|+i} = 0, \tag{17}$$

$$f x_{|X|+i} - f x_{|X|+m+1} = 0 \qquad \text{for all } f \in \{1,2,3,4\}, \tag{18}$$

$$-f x_{|X|+i} + f x_{|X|+m+1} = 0 \qquad \text{for all } f \in \{1,2,3,4\}, \tag{19}$$

where the coefficients a_{ij} are defined as above. Thus we have a system with $14m$ equations and $|X| + m + 1$ variables. Here the correspondence is between solutions of the MAX 2SAT instance satisfying s clauses and feasible solutions $\boldsymbol{x}$ fulfilling $10m+s$ equations of this homogeneous system.

Given a truth assignment that satisfies s clauses, the solution obtained by setting $x_j = \pm 1$ for $1 \leqslant j \leqslant |X|$ depending on whether the corresponding boolean variable is TRUE or FALSE and $x_{|X|+i} = 1$ for $1 \leqslant i \leqslant m+1$ satisfies at least $10m+s$ equations. Conversely, consider an arbitrary solution $\boldsymbol{x}$ satisfying $10m+s$ equations. By definition of homogeneous MAX FLS$^=$, we know that $x_{|X|+m+1}$ is nonzero for $m \geqslant 3$ because it occurs in at least $4m+s$ satisfied equations while any x_j with $1 \leqslant j \leqslant |X|$ occurs in at most $4m$ equations and any $x_{|X|+i}$ with $1 \leqslant i \leqslant m$ in at most 13. If $x_{|X|+i} \neq x_{|X|+m+1}$ for any i, $1 \leqslant i \leqslant m$, we can set it to $x_{|X|+m+1}$ without decreasing the number of satisfied equations. According to the same argument, any x_j with $1 \leqslant j \leqslant |X|$ that is neither $x_{|X|+m+1}$ nor $-x_{|X|+m+1}$ can be set to $x_{|X|+m+1}$ since $x_{|X|+i} = x_{|X|+m+1} \neq 0$ for $1 \leqslant i \leqslant m$. Thus we can assume that all equations of types (18)–(19) are satisfied, that $x_j = \pm x_{|X|+m+1}$ for $1 \leqslant j \leqslant |X|$ and therefore that s equations of types (12)–(13) are fulfilled.

Now $x_{|X|+m+1}$ is either positive or negative. If $x_{|X|+m+1} > 0$ it is equivalent to satisfy $10m+s$ equations of the above system and to satisfy $2m+s$ equations of the system (8)–(11). If $x_{|X|+m+1} < 0$ the truth assignment given by

$$y_j = \begin{cases} \text{TRUE} & \text{if } x_j = -x_{|X|+m+1}, \\ \text{FALSE} & \text{otherwise} \end{cases}$$

fulfils at least s clauses of the MAX 2SAT instance.

A similar construction can be used to show that MAX FLS$^{\geqslant}$ is APX-hard. For each clause C_i, $1 \leqslant i \leqslant m$, containing two variables x_{j_1} and x_{j_2} we consider the following equations:

$$a_{ij_1}x_{j_1} + a_{ij_2}x_{j_2} \geqslant -1, \tag{20}$$

$$x_{j_1}, x_{j_2} \geqslant 1, \tag{21}$$

$$-x_{j_1}, -x_{j_2} \geqslant -1, \tag{22}$$

$$x_{j_1}, x_{j_2} \geqslant -1, \tag{23}$$

$$-x_{j_1}, -x_{j_2} \geqslant 1, \tag{24}$$

where $a_{ij} = 1$ if x_j occurs positively in C_i and $a_{ij} = -1$ if x_j occurs negatively. The overall system has $9m$ inequalities.

Clearly, any solution $\boldsymbol{x}$ satisfies at least two and at most three of Eqs. (21)–(24) for each variable and when three of them are simultaneously satisfied then it is either equal to 1 or to -1. If in any MAX FLS$^{\geqslant}$ solution a variable is neither 1 nor -1, we can modify it without decreasing the number of satisfied equations. Thus we have a correspondence between solutions of the MAX 2SAT instance satisfying s clauses and solutions of the MAX FLS$^{\geqslant}$ instance fulfilling $6m+s$ inequalities. Moreover, opt[MAX FLS$^{\geqslant}$] $\leqslant 14 \cdot$ opt[MAX 2SAT] since opt[MAX FLS$^{\geqslant}$] $\leqslant 7m$.

As for MAX FLS$^=$, this L-reduction can be extended to homogeneous systems with discrete coefficients in $\{-4,-3,-2,-1,0,1,2,3,4\}$ and with no pairs of identical relations. □

The question of whether this result holds for homogeneous systems with bipolar coefficients is still open. The duplication technique described in Corollary 2 leads to a polynomial-time reduction but not to an L-reduction because the new systems have $O(m^2)$ equations, where m is the number of clauses in the MAX 2SAT instance.

By taking the right-hand side terms with 0.1 or 0.9 absolute values, the L-reduction for MAX FLS$^{\geqslant}$ can be adapted to show that MAX FLS$^{>}$ with no pairs of identical relations is APX-hard. This holds for homogeneous systems with no identical inequalities and integer coefficients.

If identical relations are allowed, MAX FLS$^{\geqslant}$ is APX-hard for systems with ternary coefficients while the L-reduction for MAX FLS$^{>}$ can be extended to bipolar coefficients using the duplication technique of Theorem 4.

The following results give a better characterization of the approximability of MAX FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ in terms of the various classes mentioned in Section 2.

MAX FLS$^{=}$ can obviously be approximated within $p/\min\{n-1, p\}$, where p is the number of equations and n the number of variables occurring in the system. The next proposition shows that it cannot be approximated within a constant factor.

Proposition 6. MAX FLS$^{=}$ *restricted to homogeneous systems with integer coefficients is not in* APX *unless* P = NP.

Proof. Suppose that MAX FLS$^{=}$ can be approximated within a constant $c > 1$ and consider an arbitrary instance with p homogeneous equations $e_1 = 0, \ldots, e_p = 0$. Let s be the number of equations contained in a maximum feasible subsystem.

Construct a new problem with the equations $e_{i,j,k} = 0$ where $e_{i,j,k} = e_i + k \cdot e_j$, $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant p$, $1 \leqslant k \leqslant T$ for an integer T. Since $e_i = 0$ and $e_j = 0$ imply $e_{i,j,k} = 0$ for every value of k, the s satisfied equations of the original problem give $T \cdot s^2$ satisfied equations of the new problem. However, some additional equations may be satisfied when $e_i = -k \cdot e_j$ and $e_i \neq 0$. But no more than p^2 equations are fulfilled in such a way because there is at most one such equation for each pair (i,j).

Since the optimal solution contains at least $T \cdot s^2$ satisfied equations, the approximation algorithm provides a solution that fulfils at least $T \cdot s^2/c$ equations. We examine the satisfied equations and throw away every equation $e_i + k \cdot e_j$ where $e_i \neq 0$. This leaves us with at least $T \cdot s^2/c - p^2$ equations. Since there are at most T equations for every pair (i,j), we obtain at least

$$\sqrt{\frac{T \cdot s^2/c - p^2}{T}} = \sqrt{\frac{s^2}{c} - \frac{p^2}{T}}$$

satisfied equations of the original problem. If we run the approximation algorithm directly on the original problem we are guaranteed to find s/c satisfied equations.

By choosing

$$T \geqslant \left\lceil \frac{p^2 c^2}{s^2(c-1)} \right\rceil + 1,$$

more equations are satisfied by applying the approximation algorithm to the $e_{i,j,k}$ problem than by applying it to the original problem. This can be done over and over again to get better constants in the approximation. But Theorem 5 states that MAX FLS$^=$ is APX-hard and thus there exists a constant δ between 0 and 1 such that it cannot be approximated within a smaller constant than $1/(1-\delta)$. Hence MAX FLS$^=$ is not in APX. □

By using tuples of $\log p$ equations instead of pairs of equations and by using walks on expander graphs in order to choose a polynomial number of these tuples it is possible to show a yet stronger result.

Theorem 7. *Unless P = NP, there is a positive constant ε such that homogeneous* MAX FLS$^=$ *cannot be approximated within p^ε, where p is the number of equations.*

Proof. We start as in the proof of Proposition 6 with an arbitrary instance of MAX FLS$^=$ with p homogeneous equations $e_1 = 0, \ldots, e_p = 0$ and let s be the number of equations contained in a maximum feasible subsystem. In this proof we assume that a fixed percentage of the equations can be satisfied, i.e. that $s/p \geqslant \beta$, where β is a constant. The APX-hard problem instances constructed in the first reduction in Theorem 5 have $\beta > 10/14$, but we can increase β by adding a lot of trivially satisfiable equations in a new variable.

Instead of using pairs of equations when constructing the new problem we take m-tuples of equations, where m is about $\log p$. Let us consider

$$e_{i,k} = \sum_{j=1}^{m} e_{i_j} k^j \quad \text{with} \quad 1 \leqslant i \leqslant N, 1 \leqslant k \leqslant T$$

for some integers N and T to be determined. For each i with not all $e_{i_j} = 0$ the polynomial equation $\sum e_{i_j} x^j = 0$ can have at most m solutions of which one is $x = 0$. Thus at most $m-1$ of the T equations $e_{i,k} = 0, 1 \leqslant k \leqslant T$, can be satisfied unless $e_{i_j} = 0$ for all j in $[1 \ldots m]$. We call an m-tuple for which every $e_{i_j} = 0$ a *good tuple.*

The problem here is that we cannot form new equations from every m-tuple of old equations, since there are $p^m \sim p^{\log p}$ tuples, that is more than a polynomial number. We would like to find a polynomial subset of all possible m-tuples such that the proportion of good tuples in the subset is about the same as the proportion of good tuples in the set of all possible m-tuples, which is s^m/p^m.

We need the following lemma, proved by Alon et al. [1] using [2, 24].

Lemma 8. *Let G be a d-regular, p-node graph in which the absolute value of every nontrivial eigenvalue (of the adjacency matrix) is at most λd. Let S be a set of s nodes of G and $\mu = s/p$. Let $P = P(S,m)$ be the total number of walks of m nodes that stay in S (i.e. the number of m-paths where every node is in S). Then for every odd integer m we have*

$$\mu(\mu - \lambda(1-\mu))^{m-1} \leqslant \frac{P}{p \cdot d^{m-1}} \leqslant \mu(\mu + \lambda(1-\mu))^{m-1}.$$

We will consider Ramanujan graphs as G. By definition, a Ramanujan graph is a connected d-regular graph whose eigenvalues are either $\pm d$ or at most $2\sqrt{d-1}$ in absolute value. Thus we can take $\lambda = 2/\sqrt{d}$. Infinite families of Ramanujan graphs can be constructed when $d-1$ is prime congruent to 1 modulo 4 [24, 30].

Choose m as the least odd integer greater than $\log p$ (where log is the logarithm in base 2). We identify each node in G with an equation $e_i = 0$, $1 \leqslant i \leqslant p$. As m-tuples in the constructed problem we use every possible m-path in G. There are $p \cdot d^{m-1}$ m-paths in G so $N = p \cdot d^{m-1}$.

If S is the set of nodes corresponding to the s equations contained in some maximum feasible subsystem, $P(S,m)$ will be exactly the number of good m-tuples.

By assigning variables as in the optimal solution of the original problem we will satisfy at least $T \cdot P(S,m)$ equations, so we know that any optimal solution of the new problem fulfils at least

$$NT\frac{s}{p}\left(\frac{s}{p} - \lambda\left(1 - \frac{s}{p}\right)\right)^{m-1} \tag{25}$$

(rounded to an integer) equations. On the other hand, we know that if the optimal solution of the original problem only had satisfied $\lfloor s(1-\delta) \rfloor$ equations, for some δ between 0 and 1, then any optimal solution of the new problem would fulfil at most

$$NT\frac{s(1-\delta)}{p}\left(\frac{s(1-\delta)}{p} + \lambda\left(1 - \frac{s(1-\delta)}{p}\right)\right)^{m-1} + N(m-1) \tag{26}$$

(rounded to an integer) equations.

We will show that the quotient between (25) and (26) can be bounded from below by $(NT)^{\varepsilon}$ for some $\varepsilon > 0$. If there exists an algorithm approximating MAX FLS$^=$ within p^{ε} (where p is the number of equations) we can apply it to the constructed problem with NT equations and be sure that it gives us a solution containing more satisfied equations than (26). Therefore the assignment given by the algorithm will approximate the original problem within $1/(1-\delta)$, but this is NP-hard since MAX FLS$^=$ is APX-hard (by Theorem 5).

Thus we just have to bound the quotient between (25) and (26) appropriately.

Let $\lambda = \alpha\beta$, where α is a positive constant to be defined later. Then $\alpha\beta = 2/\sqrt{d}$ and $\log d = 2 + 2\log(1/\alpha\beta)$. Hence $N = p \cdot d^{m-1} \approx p \cdot p^{\log d} = p^{3+2\log(1/\alpha\beta)}$ and

$$\begin{aligned}(25) &> NT\frac{s}{p}\left(\frac{s}{p} - \alpha\beta\right)^{m-1} \geqslant NT\frac{s}{p}\left(\frac{s}{p}(1-\alpha)\right)^{m-1} \\ &\geqslant NT\beta(\beta(1-\alpha))^{m-1} \approx NT\beta p^{-\log[1/\beta(1-\alpha)]}.\end{aligned} \tag{27}$$

First we consider the first term of (26):

$$\begin{aligned}&NT\frac{s(1-\delta)}{p}\left(\frac{s(1-\delta)}{p} + \lambda\left(1 - \frac{s(1-\delta)}{p}\right)\right)^{m-1} \\ &\leqslant NT\frac{s(1-\delta)}{p}\left(\frac{s(1-\delta)}{p} + \lambda\right)^{m-1}\end{aligned}$$

$$= NT\frac{s(1-\delta)}{p}\left(\frac{s(1-\delta)}{p}+\alpha\beta\right)^{m-1}$$

$$\leqslant NT\frac{s(1-\delta)}{p}\left(\frac{s}{p}(1-\delta+\alpha)\right)^{m-1}. \tag{28}$$

Using (27) and (28) we can bound the quotient between (25) and the first term of (26) from below by

$$\frac{NT(s/p)\left(\frac{s}{p}(1-\alpha)\right)^{m-1}}{NT[s(1-\delta)/p]\left(\frac{s}{p}(1-\delta+\alpha)\right)^{m-1}}$$

$$= \frac{1}{1-\delta}\left(\frac{1-\alpha}{1-\delta+\alpha}\right)^{m-1} \approx \frac{1}{1-\delta}p^{\log[(1-\alpha)/(1-\delta+\alpha)]} > p^{\log[(1-\alpha)/(1-\delta+\alpha)]}. \tag{29}$$

We now consider $N(m-1)$, the second term of (26). In this case, we have

$$\frac{(25)}{N(m-1)} \geqslant \frac{NT\beta p^{-\log[1/\beta(1-\alpha)]}}{N\log p} \geqslant Tp^{-\log[1/\beta(1-\alpha)]-\delta'} \quad \text{for every } \delta' > 0.$$

We would like both $p^{\log[(1-\alpha)/(1-\delta+\alpha)]}$ and $Tp^{-\log[1/\beta(1-\alpha)]-\delta'}$ to be greater than $2(NT)^{\varepsilon}$ for some constant $\varepsilon > 0$. If we choose $T = \lceil p^{\log[1/\beta(1-\delta+\alpha)]}\rceil$ the second expression is greater than

$$p^{\log[1/\beta(1-\delta+\alpha)]-\log[1/\beta(1-\alpha)]-\delta'} = p^{\log[(1-\alpha)/(1-\delta+\alpha)]-\delta'},$$

which is less than the first quotient $p^{\log[(1-\alpha)/(1-\delta+\alpha)]}$ for every $\delta' > 0$. Thus we only have to bound $p^{\log[(1-\alpha)/(1-\delta+\alpha)]-\delta'}$ from below by

$$2(NT)^{\varepsilon} \approx p^{\varepsilon(3+2\log(1/\alpha\beta)+\log[1/\beta(1-\delta+\alpha)])} = p^{\varepsilon(3+3\log(1/\beta)+2\log(1/\alpha)+\log[1/(1-\delta+\alpha)])}.$$

Let $\alpha = \varepsilon'\delta$ for some $\varepsilon' > 0$. Then we must satisfy

$$\varepsilon < \frac{\log[(1-\varepsilon'\delta)/(1-\delta+\varepsilon'\delta)]-\delta'}{3+3\log(1/\beta)+2\log(1/\varepsilon'\delta)+\log[1/(1-\delta+\varepsilon'\delta)]}$$

$$\approx \frac{\delta(1-2\varepsilon')-\delta' \ \ln\ 2}{3\ln 2+3\ln(1/\beta)+2\ln(1/\varepsilon')+2\ln(1/\delta)+\delta(1-\varepsilon')},$$

which, given β and choosing δ' and ε' small enough, is a positive constant, slightly smaller than $\delta/\ln(1/\delta^2)$. □

While completing this paper we discovered that Arora, Babai, Stern and Sweedyk simultaneously addressed the complexity of one variant of this broad class of problems, namely MAX FLS$^=$ [5]. They independently proved that the problem cannot be approximated within any constant factor unless P = NP, and not within a factor $2^{\log^{0.5-\varepsilon} n}$ for any $\varepsilon > 0$ unless NP $\subseteq$ DTIME($n^{\text{polylog}\, n}$), where n is the number of variables or equations. But Theorem 7 is stronger because, on one hand, for large n the factor $2^{\log^{0.5-\varepsilon} n}$ is smaller than n^{δ} for any fixed $\delta > 0$ and, on the other hand, P $\neq$ NP is more likely to be true than NP $\not\subseteq$ DTIME($n^{\text{polylog}\, n}$).

Max FLS$^{\geqslant}$ and Max FLS$^{>}$ turn out to be much easier to approximate than Max FLS$^{=}$.

Proposition 9. Max FLS$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{\geqslant, >\}$ *is* Apx-*complete and can be approximated within* 2.

Proof. Both problems can be approximated within 2 using the following simple algorithm.

Algorithm:

```
Input: An instance (A,b) of MAX FLS^R with R ∈ {≥,>}
Init: X := {variables occurring in (A,b)} and
                                    E := {inequalities in (A,b)}
  WHILE E ≠ ∅ DO
  IF there are inequalities in E that contain a single variable
  THEN
   U := {x ∈ X | x occurs as a single variable in at least
   one inequality of E}
   Pick at random y ∈ U
   F(y) := {e ∈ E | e contains only the variable y}
   Assign a value to y that satisfies as many inequalities
                                              in F(y) as possible
   E := E − F(y)
  ELSE
   Pick at random a variable y and assign a random value to it
   Reevaluate the inequalities in E that contain y
  END IF
  X := X − {y}
  END WHILE
```

This algorithm is guaranteed to provide a 2-approximation because we can always assign to y a value that satisfies at least half of the inequalities in $F(y)$. Moreover, it runs in polynomial time since each variable and each inequality are considered only once.

Since Max FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{\geqslant, >\}$ is Apx-hard and can be approximated within 2, both problems are Apx-complete. □

Notice that this greedy-like method is similar to the 2-approximation algorithm that has been proposed for Max Sat [22]. As for Max Sat [39], there could exist a better polynomial-time algorithm that guarantees a smaller performance ratio.

Provided that P $\neq$ NP, the previous results describe the approximability of MAX FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ quite well. While MAX FLS$^{=}$ cannot be approximated within p^{ε} for some $\varepsilon > 0$ where p is the number of equations, MAX FLS$^{\geqslant}$ and MAX FLS$^{>}$ can be approximated within a factor 2 but not within every constant.

One can observe that MAX FLS$^{\geqslant}$ and MAX FLS$^{>}$ share a common property: a constant fraction of the relations can always be simultaneously satisfied. The above 2-approximation algorithm is optimal in the sense that no constant fraction larger than $\frac{1}{2}$ can be guaranteed. Of course no such a property holds for MAX FLS$^{=}$.

In the appendix we deal with two interesting special cases of MAX FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ related, on one hand, to finite field computation and, on the other hand, to discriminant analysis and machine learning.

Finally we should point out that in many practical situations different relations may have different importances. This can be modeled by assigning a weight to each relation and looking for a solution that maximizes the total weight of the satisfied relations.

Such weighted versions of MAX FLS turn out to be equally hard to approximate as the corresponding unweighted versions. If the weights are polynomially bounded integers, we just need to make for each relation a number of copies equal to the associated weight. Otherwise we reduce to the polynomially bounded case by dividing the weights by $\tilde{w}/p$, where $\tilde{w}$ is the largest weight, and rounding them to the nearest integer. It is easily verified that the absolute error due to scaling and rounding is bounded by a constant times the optimum value.

4. Hardness of constrained MAX FLS

An interesting and important special case of weighted MAX FLS is the constrained version, denoted by C MAX FLS, where some relations are mandatory while the others are optional. The objective is then to find a solution that satisfies all mandatory relations and as many optional ones as possible. Any instance of C MAX FLS is equivalent to the particular instance of weighted MAX FLS where each optional relation is assigned a unit weight and each mandatory relation a weight larger than the total number of optional ones.

However, while the weighted versions of MAX FLS are equally hard to approximate as the unweighted versions, most of the constrained versions turn out to be at least as hard to approximate as MAX IND SET. Thus, unless P = NP, they cannot be approximated within m^{ε} for some $\varepsilon > 0$, where m is the size of the instance.

When considering mixed variants of C MAX FLS with different types of mandatory and optional relations, C MAX FLS$^{\mathscr{R}_1;\mathscr{R}_2}$ with $\mathscr{R}_1, \mathscr{R}_2 \in \{=, \geqslant, >, \neq\}$ denotes the variant where the mandatory relations are of type $\mathscr{R}_1$ and the optional ones of type $\mathscr{R}_2$.

Theorem 10. C MAX FLS$^{\mathscr{R}_1;\mathscr{R}_2}$ *with* $\mathscr{R}_1, \mathscr{R}_2 \in \{\geqslant, >\}$ *is* MAX IND SET-*hard even for homogeneous systems.*

Proof. The proof is by cost preserving polynomial-time transformations from MAX IND SET. We proceed like in the first part of the reduction of Theorem 4 and start with C MAX FLS$^{>;>}$.

Let $G=(V,E)$ be an arbitrary instance of MAX IND SET. For each edge $(v_i,v_j)\in E$ we construct the mandatory inequality

$$x_i+x_j<0 \tag{30}$$

and for each $v_i\in V$ the optional inequality

$$x_i>0. \tag{31}$$

Thus we have a system with $|V|$ variables and $|E|+|V|$ strict inequalities. As shown in the proof of Theorem 4, the given graph G contains an independent set I of size s if and only if there exists a solution $\boldsymbol{x}$ satisfying all mandatory inequalities and s optional ones.

This cost preserving transformation can easily be adapted to show that the other problems are MAX IND SET-hard. For C MAX FLS$^{\geqslant;>}$ we just change the mandatory inequality of type (30) to $x_i+x_j\leqslant 0$. The proof is the same.

For C MAX FLS$^{\geqslant;\geqslant}$ an additional variable x_{n+1} needed. We include the mandatory inequality

$$x_{n+1}\geqslant 0 \tag{32}$$

and for each edge $(v_i,v_j)\in E$ the mandatory inequality

$$x_i+x_j\leqslant 0. \tag{33}$$

For each $v_i\in V$ we consider the optional inequality

$$x_i-x_{n+1}\geqslant 0. \tag{34}$$

Thus we have a system with $|V|+1$ variables and $|E|+|V|+1$ inequalities.

Now the given graph G contains an independent set I of size s if and only if there exists a solution $\boldsymbol{x}$ satisfying all inequalities of types (32)–(33) and s inequalities of type (34). Given an independent set $I\subseteq V$ of size s, the solution obtained by setting

$$x_i=\begin{cases} 1 & \text{if } v_i\in I \text{ or } i=n+1,\\ -2 & \text{if } 1\leqslant i\leqslant |V| \text{ and } v_i\notin I\end{cases}$$

satisfies all mandatory inequalities and all the optional ones that correspond to a node $v_i\in I$. The $|V|-s$ optional inequalities associated with $v_i\notin I$ are not fulfilled because $x_i<x_{n+1}$. Conversely, given an appropriate solution $\boldsymbol{x}$ that fulfils all mandatory inequalities and s optional ones. Since the variable x_{n+1} is included in every optional inequality and hence is the most common one, it cannot be zero and must be positive because of relation (32). The set $I\subseteq V$ containing all nodes associated to variables with values larger than x_{n+1} is then an independent set of cardinality s.

Finally, by simply changing the mandatory inequalities in this reduction to strict inequalities we get a reduction to C Max FLS$^{>;\geqslant}$. □

Thus forcing a subset of relations makes Max FLS$^{\mathcal{R}}$ harder for $\mathcal{R} \in \{\geqslant, >\}$: Max FLS$^{\mathcal{R}}$ is Apx-complete while C Max FLS$^{\mathcal{R};\mathcal{R}}$ is Max Ind Set-hard. This is not true for Max FLS$^{=}$ since any instance of C Max FLS$^{=;\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >\}$ can be transformed into an equivalent instance of Max FLS$^{\mathcal{R}}$ by eliminating variables in the set of optional equations using the set of mandatory ones.

Nevertheless, two simple variants of C Max FLS$^{=;=}$ and Max FLS$^{=}$ turn out to be Max Ind Set-hard. Replacing the strict inequality of type (30) by $x_i + x_j$**6**1 and that of type (31) by $x_i = 1$, one verifies that C Max FLS$^{\geqslant;=}$ is Max Ind Set-hard even for systems with ternary coefficients and bipolar right-hand-side components. This has an immediate implication on the hardness of Max FLS$^{=}$ with the natural nonnegativeness constraint.

Corollary 11. Max FLS$^{=}$ *restricted to systems with ternary coefficients and nonnegative variables is* Max Ind Set-*hard.*

Proof. By adding a nonnegative slack variable for each mandatory inequality of any C Max FLS$^{\geqslant;=}$ instance, we obtain a particular instance of C Max FLS$^{=;=}$ that can be transformed into an equivalent instance of Max FLS$^{=}$. Any variable x_i unrestricted in sign can then be replaced by two nonnegative variables x_i^1 and x_i^2. The coefficients of these two auxiliary variables have the same absolute value as the coefficient of x_i but opposite signs. □

The question of whether Max FLS$^{=}$ becomes harder to approximate when the variables are restricted to be nonnegative or whether the basic version is already Max Ind Set-hard, is still open. Note that the positiveness constraints do not affect C Max FLS$^{\geqslant;\geqslant}$ since they can be viewed as additional mandatory inequalities.

The approximability of mixed variants involving $\neq$ mandatory relations is somewhat different.

Proposition 12. *While* C Max FLS$^{\neq;=}$ *is* Max Ind Set-*hard,* C Max FLS$^{\neq;>}$ *and* C Max FLS$^{\neq;\geqslant}$ *are* Apx-*complete and can be approximated within 2.*

Proof. For C Max FLS$^{\neq;=}$, we proceed by cost preserving transformation from Max Ind Set. For each edge $(v_i, v_j) \in E$ we consider the mandatory relation $x_i + x_j \neq 2$ and for each node $v_i \in V$ we consider the optional relation $x_i = 1$. Clearly, there exists an independent set of size s if and only if there exists a solution $\boldsymbol{x}$ satisfying all mandatory relations and s optional ones.

According to Theorem 5, Max FLS$^{\geqslant}$ and Max FLS$^{>}$ are Apx-hard and the constrained versions C Max FLS$^{\neq;\geqslant}$ and C Max FLS$^{\neq;>}$ must be at least as hard. To approximate these problems within 2 we modify the greedy algorithm in Proposition 9 so that it also takes into account the mandatory relations. When a variable value is chosen it should not contradict any of the mandatory relations that have a single unassigned variable. This is always possible since there is only a finite number of such relations while the number of possible values satisfying the largest number of optional relations is infinite. □

5. Approximability of Max FLS with bounded discrete variables

In this section we assess the approximability of Max FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ when the variables are restricted to take a finite number of discrete values. Both extreme cases with binary variables in $\{0,1\}$ and bipolar variables in $\{-1,1\}$ are considered. The corresponding variants of Max FLS are named Bin Max FLS$^{\mathscr{R}}$ and Bip Max FLS$^{\mathscr{R}}$, respectively.

Theorem 13. Bin Max FLS$^{\mathscr{R}}$ *with* $\mathscr{R} \in \{=, \geqslant, >\}$ *is* Max Ind Set-*hard even for systems with ternary coefficients.*

Proof. The proof is by cost preserving polynomial transformation from Max Ind Set. We first consider Max FLS$^{>}$.

Let $G = (V, E)$ be the graph of an arbitrary instance of Max Ind Set. For each node $v_i \in V$ we construct the strict inequality

$$x_i - \sum_{j \in N(v_i)} x_j > 0,$$

where j is included in $N(v_i)$ if and only if v_j is adjacent to v_i. Thus we have a system of $|V|$ homogeneous inequalities with ternary coefficients. By construction, the ith inequality is satisfied if and only if $x_i = 1$ and $x_j = 0$ for all j, $1 \leqslant j \leqslant |V|$, such that $a_j = -1$.

It is easy to verify that given an independent set $I \subseteq V$ of size s we get a binary solution satisfying the s corresponding inequalities by setting $x_i = 1$ if $v_i \in I$ and $x_i = 0$ otherwise. Conversely, given any binary solution $\boldsymbol{x}$ satisfying s inequalities we obtain an independent set of size s by including in I all nodes v_i, $1 \leqslant i \leqslant |V|$, such that $x_i = 1$.

Notice that this cost preserving polynomial transformation works also for $A\boldsymbol{x} \geqslant 1$ or $A\boldsymbol{x} = 1$. This construction can be adapted to show that Bin Max FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ is hard to approximate even when restricted to homogeneous systems. However, we must allow the coefficients to take their values in $\{-2, 0, 1\}$ instead of in $\{-1, 0, 1\}$. □

Corollary 14. Bip Max FLS$^{\mathscr{R}}$ *with* $\mathscr{R} \in \{=, \geqslant, >\}$ *is* Max Ind Set-*hard even for systems with ternary coefficients and integer right-hand-side components.*

Proof. By simple cost preserving transformation from BIN MAX FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$. For any relation

$$\sum_{j=1}^{n} a_{ij}x_j \;\mathscr{R}\; b_i,$$

with binary variables $x_j \in \{0,1\}$ and $1 \leqslant i \leqslant p$, we can construct an equivalent relation

$$\sum_{j=1}^{n} a_{ij}y_j \;\mathscr{R}\; 2b_i + \sum_{j=1}^{n} a_{ij},$$

with bipolar variables $y_j \in \{-1,1\}$ using the variable substitution $y_j = 2x_j - 1$. □

Although the above transformation does not preserve homogeneity, we know from the L-reductions used to prove Theorem 5 that homogeneous BIP MAX FLS$^{\mathscr{R}}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ is APX-hard. In fact, homogeneous BIP MAX FLS$^{\geqslant}$ and BIP MAX FLS$^{>}$ are APX-complete.

Proposition 15. *Homogeneous* BIP MAX FLS$^{\geqslant}$ *can be approximated within 2 and homogeneous* BIP MAX FLS$^{>}$ *can be approximated within 4.*

Proof. We first deal with homogeneous BIP MAX FLS$^{\geqslant}$. Take an arbitrary bipolar vector $\boldsymbol{x}$ and consider the number of satisfied relations for $\boldsymbol{x}$ and $-\boldsymbol{x}$. If the left-hand side of a relation is positive for $\boldsymbol{x}$ it will be negative for $-\boldsymbol{x}$ and vice versa. Thus one of these antipodal vectors satisfies at least half of the inequalities.

This trivial algorithm does not work for homogeneous BIP MAX FLS$^{>}$ because many relations may be zero for both antipodal vectors. Therefore we first look for a solution with many nonzero relations. A greedy approximation algorithm similar to the one in Proposition 9 provides a solution $\boldsymbol{x}$ for which at least half of the relations are nonzero. Now one of $\boldsymbol{x}$ and $-\boldsymbol{x}$ makes at least half of these relations, and therefore a quarter of all relations, positive. □

Thus restricting the systems to be homogeneous makes BIP MAX FLS$^{\geqslant}$ and BIP MAX FLS$^{>}$ much easier to approximate. The situation is quite different for homogeneous BIP MAX FLS$^{=}$ with integer coefficients. According to the same arguments as in the proof of Theorem 7, this problem cannot be approximated within p^{ε} for some $\varepsilon > 0$.

Proposition 16. BIN MAX FLS$^{\neq}$ *and* BIP MAX FLS$^{\neq}$ *are* APX-*complete and can be approximated within* 2.

Proof. For both problems the proof is by cost preserving transformation from MAX 2SAT. We first consider BIP MAX FLS$^{\neq}$. Let (X, C) be an arbitrary instance of MAX 2SAT with $C = \{C_1, \ldots, C_m\}$. For each clause C_i, $1 \leqslant i \leqslant m$, containing two variables x_{j_1} and x_{j_2} we consider the relation

$$a_{j_1}x_{j_1} + a_{j_2}x_{j_2} \neq -2,$$

where $a_j = 1$ if the boolean variable x_j occurs positively in C_i, $a_j = -1$ if x_j occurs negatively in C_i, and $a_j = 0$ otherwise. Thus we have a system with m relations. Clearly, there exists a truth assignment satisfying s clauses of (X, C) if and only if there exists a solution $\boldsymbol{x}$ satisfying s relations.

A similar reduction is used for BIN MAX FLS$^{\neq}$. Both problems are in APX since they can be approximated within 2 using a greedy algorithm similar to the one in Proposition 9. □

The following result shows that the constrained variants of BIN MAX FLS$^{\mathscr{R}}$ with mandatory relations, named C BIN MAX FLS$^{\mathscr{R}_1;\mathscr{R}_2}$, are NPO PB-complete, that is, at least as hard to approximate as every NP optimization problem with polynomially bounded objective function.

Proposition 17. C BIN MAX FLS$^{\mathscr{R}_1;\mathscr{R}_2}$ *is* NPO PB-*complete for* $\mathscr{R}_1, \mathscr{R}_2 \in \{=, \geqslant, >, \neq\}$, *even for systems with ternary coefficients.*

Proof. We first show the result for the problem C BIN MAX FLS$^{>;>}$ and then extend the result to the other variants.

We proceed by cost preserving transformations from MAX DONES that is known to be NPO PB-complete [25] and is defined as follows [35]. Given two disjoint sets X, Z of variables and a collection $C = \{C_1, \ldots, C_m\}$ of disjunctive clauses of at most 3 literals, find a truth assignment for X and Z that satisfies every clause in C so that the number of Z variables that are set to TRUE in the assignment is maximized.

Suppose we are given an arbitrary instance of MAX DONES with the boolean variables $y_1, \ldots, y_n$ where $y_j \in Z$ if $1 \leqslant j \leqslant |Z|$ and $y_j \in X$ if $|Z| < j \leqslant n$. For each clause $l_{j_1} \vee l_{j_2} \vee l_{j_3} \in C$ we consider the mandatory inequality

$$t_{j_1} + t_{j_2} + t_{j_3} > 0, \tag{35}$$

where, for $1 \leqslant k \leqslant 3$, $t_{j_k} = x_j$ if $l_{j_k} = y_{j_k}$, $t_{j_k} = 1 - x_{j_k}$ if $l_{j_k} = \bar{y}_{j_k}$, and $t_{j_k} = 0$ if there is no l_{j_k} (i.e. if the clause contains less than three literals). For each variable $y_j \in Z$ with $1 \leqslant j \leqslant |Z|$ we consider the optional inequality

$$x_j > 0. \tag{36}$$

Thus we have a system with $|X| + |Z|$ variables, $|Z|$ optional relations and m mandatory ones.

We claim that there is an assignment to X and Z with s variables from Z set to TRUE if and only if there exists a solution $\boldsymbol{x} \in \{0, 1\}^n$ that satisfies s optional relations of the corresponding linear system.

Given an assignment, the solution $\boldsymbol{x}$ defined by

$$x_j = \begin{cases} 1 & \text{if } y_j \text{ is set to TRUE,} \\ 0 & \text{otherwise} \end{cases}$$

satisfies all mandatory relations and s of the optional ones. Conversely, given a solution vector $\boldsymbol{x}$ that satisfies s of the optional relations, the corresponding assignment (y_j is TRUE if and only if $x_j = 1$) satisfies all clauses (because of the mandatory relations), and s of the Z variables are true (since s of the optional relations are satisfied).

For the other constrained problems C BIN MAX FLS$^{\mathcal{R}_1;\mathcal{R}_2}$, we use the same reduction as above but the right hand side of the relations must be substituted according to the following table.

	type $\geqslant$	type $>$	type $\neq$	type $=$
type (35)	$\geqslant 1$	> 0	$\neq 0$	$= 1 + x' + x''$
type (36)	$\geqslant 1$	> 0	$\neq 0$	$= 1$

In the case of mandatory equations we need to introduce two additional slack variables x' and x'' in each equation (a total of $2m$ new variables). □

These results imply, using the same argument as in corollary 14, that the corresponding bipolar versions C BIN MAX FLS$^{\mathcal{R}_1;\mathcal{R}_2}$ with $\mathcal{R}_1, \mathcal{R}_2 \in \{=, \geqslant, >, \neq\}$ are NPO PB-complete for systems with ternary coefficients and integer right-hand-side components.

Since MAX DONES cannot be approximated within $|Z|^{1-\varepsilon}$, for any $\varepsilon > 0$, unless P = NP [11], the same nonapproximability bound $\hat{p}^{1-\varepsilon}$ is valid for all versions of C BIN MAX FLS$^{\mathcal{R}_1;\mathcal{R}_2}$ and C BIN MAX FLS$^{\mathcal{R}_1;\mathcal{R}_2}$, where $\hat{p}$ is the number of optional relations.

6. Conclusions

The various versions of MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >, \neq\}$ that we have considered are obtained by placing constraints on the coefficients (left- and right-hand sides), on the variables and on the relations that must be satisfied.

Table 1 summarizes our main approximability results. All these results hold for inhomogeneous systems with integer coefficients and no pairs of identical relations, but most of them are still valid for homogeneous systems with ternary coefficients.

Thus the approximability of similar variants of MAX FLS can differ enormously depending on the type of relations. Nevertheless, there is some structure: all basic versions of MAX FLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geqslant, >\}$ are APX-hard, restricting the variables to binary (bipolar) values or introducing a set of relations that must be satisfied makes them harder to approximate, and if both restrictions are considered simultaneously all problems become NPO PB-complete. The case of MAX FLS$^{\neq}$ is considerably different. Its constrained variants are intrinsically easier than the corresponding problems with the other types of relations except when constraints are imposed both on the relations and the variables.

As shown in the appendix, MAX FLS$^{=}$ over $GF(q)$ and therefore C MAX FLS$^{=;=}$ over $GF(q)$ are approximable within q but not within q^{ε} for some $\varepsilon > 0$, while MAX FLS$^{=}$ as well as C MAX FLS$^{=;=}$ restricted to nonnegative variables are MAX IND SET-hard.

Table 1
Main approximability results for MAX FLS variants. $\mathscr{R}$ denotes any relational operator in $\{=, \geqslant, >, \neq\}$. Nonstrict inequality ($\geqslant$) can be substituted by strict inequality ($>$) in every place in the table.

	Real variables	Binary variables
MAX FLS$^=$	Not within p^ε for some $\varepsilon > 0$	MAX IND SET-hard
MAX FLS$^\geqslant$	APX-complete (within 2)	MAX IND SET-hard
MAX FLS$^\neq$	Trivial	APX-complete (within 2)
C MAX FLS$^{=;=}$	Not within p^ε for some $\varepsilon > 0$	NPO PB-complete
C MAX FLS$^{=;\geqslant}$	APX-complete (within 2)	NPO PB-complete
C MAX FLS$^{\geqslant;=}$	MAX IND SET-hard	NPO PB-complete
C MAX FLS$^{\geqslant;\geqslant}$	MAX IND SET-hard	NPO PB-complete
C MAX FLS$^{\neq;=}$	MAX IND SET-hard	NPO PB-complete
C MAX FLS$^{\neq;\geqslant}$	APX-complete (within 2)	NPO PB-complete
C MAX FLS$^{\mathscr{R};\neq}$	Trivial	NPO PB-complete

Moreover, our nonapproximability bounds for basic MAX FLS do also hold for the weighted versions. In the appendix we determine the approximability of two important variants of mixed MAX FLS$^>$ and MAX FLS$^\geqslant$ related to discriminant analysis and machine learning.

Whenever possible we studied the complexity results for homogeneous systems whose coefficients can take as few values as possible. In order to avoid trivial solutions we required the variables occurring most frequently in the satisfied relations to be nonzero. It is worth noting that some problems, like BIP MAX FLS$^>$ and BIP MAX FLS$^\geqslant$, become harder to approximate when inhomogeneous systems are considered.

Several interesting questions are still open. Are there better approximation algorithms for MAX FLS$^>$ and MAX FLS$^\geqslant$? Does MAX FLS$^=$ become harder when the variables are constrained to be nonnegative or is it already MAX IND SET-hard? One could also wonder whether the problems we have shown MAX IND SET-hard are in fact NPO PB-complete.

The approximability of the complementary minimization problems where the objective is to minimize the number of unsatisfied relations instead of maximizing the number of satisfied ones have been studied elsewhere [4, 5].

Acknowledgements

The authors are grateful to Oded Goldreich, Mike Luby and most of all to Johan Håstad for their valuable suggestions concerning the proofs of proposition 6 and theorems 7 and A.2. Edoardo Amaldi thanks Claude Diderich for helpful discussions.

Appendix: Three particular cases

Three interesting special cases of unconstrained and constrained MAX FLS$^\mathscr{R}$ with $\mathscr{R} \in \{=, \geqslant, >\}$ are considered. The last two arise in the important fields of discriminant analysis and machine learning.

The first problem, named MAX FLS$^=$ over $GF(q)$, is obtained by restricting MAX FLS$^=$ to systems where the equations are in $GF(q)$, that is modulo a prime q.

Proposition A.1. *For any prime q,* Max FLS$^=$ over $GF(q)$ *is* Apx*-complete and can be approximated within q.*

Proof. The L-reduction from Max 2Sat to Max FLS$^=$ can be used to show that Max FLS$^=$ over $GF(q)$ is Apx-hard for $q \geqslant 3$. However, it breaks down for $q = 2$ because in that case an equation of the type $x_i + x_j = 0$ has both solutions $x_i = x_j = 0$ and $x_i = x_j = 1$.

For $q = 2$ we proceed by reduction from Max Cut that is defined as follows [16]. Given an undirected graph $G = (V, E)$, find a set of nodes $V' \subseteq V$ such that the number of edges $(v_i, v_j) \in E$ with $v_i \in V'$ and $v_j \in V - V'$ is as large as possible.

For every edge (v_i, v_j) we introduce the equation $x_i + x_j = 1 \pmod 2$. An equation is satisfied if and only if one of the variables is odd and the other is even. The oddness and evenness partition the graph and the size of the cut is exactly the number of satisfied equations. This construction is clearly an L-reduction.

When the equations are in $GF(q)$, it is easy to find a solution that satisfies at least $1/q$ of the equations. This can be achieved using a simple greedy algorithm similar to that presented in the proof of Proposition 9. □

For example, if all coefficients take their values in $\{0, 1\}$ and if all computations are performed modulo 2, Max FLS$^=$ can be approximated within 2. This result is clearly not applicable when standard computations are used.

Using proposition A.1 and proof techniques from Theorem 7 we can show a better lower bound on the approximability of Max FLS$^=$ over $GF(q)$.

Theorem A.2. *For any prime q there is a constant $\varepsilon > 0$ such that* Max FLS$^=$ over *$GF(q)$ cannot be approximated within q^{ε}.*

Proof. We use the same construction as in the proof of theorem 7 but we choose m as the least odd number greater than $\log q$ (instead of $\log p$). Then $N = p \cdot d^{m-1} \approx p \cdot q^{\log d}$. We let $T = q - 1$, which means that we consider every number in $GF(q)$ except 0 as constants in the constructed equations.

Since Max FLS$^=$ over $GF(q)$ is Apx-complete there is a constant $\delta > 0$ such that it cannot be approximated within $1/(1-\delta)$. The quotient between (25) and the second term of (26) now is

$$\frac{(25)}{N(m-1)} \geqslant \frac{NT\beta q^{-\log[1/\beta(1-\alpha)]}}{N \log q} \geqslant q^{1-\log[1/\beta(1-\alpha)]-\delta'} \quad \text{for every } \delta' > 0,$$

and the quotient between (25) and the first term of (26) is, using (27) and (28), bounded by

$$\frac{NT(s/p)((s/p)(1-\alpha))^{m-1}}{NT[s(1-\delta)/p]((s/p)(1-\delta+\alpha))^{m-1}} \approx \frac{1}{1-\delta} q^{\log[(1-\alpha)/(1-\delta+\alpha)]}$$
$$\geqslant q^{\log[(1-\alpha)/[(1-\delta+\alpha)]}.$$

If we choose $\alpha < \delta^2$, $\beta > 1/(2-2\delta)$ and $\delta' < \delta^2$ both quotients will be bounded by q^ε where ε is approximately δ. □

The second problem, MAX H-CONSISTENCY, is a special case of mixed MAX FLS$^>$ and MAX FLS$^\geqslant$. It arises when training perceptrons or designing linear classifiers [19]. Given a set of vectors $S = \{a^k\}_{1 \leqslant k \leqslant p} \subset \mathbb{R}^n$ labeled as positive or negative examples, we look for a hyperplane H, specified by a normal vector $\boldsymbol{w} \in \mathbb{R}^n$ and a bias w_0, such that all the positive vectors lie on the positive side of H while all the negative ones lie on the negative side. A halfspace H is said to be *consistent* with an example $\boldsymbol{a}^k$ if $\boldsymbol{w}\boldsymbol{a}^k > w_0$ or $\boldsymbol{w}\boldsymbol{a}^k \leqslant w_0$ depending on whether $\boldsymbol{a}^k$ is positive or negative. In the general case where S is nonlinearly separable, a natural objective is to maximize the consistency, i.e. to find a hyperplane that is consistent with as many $\boldsymbol{a}^k \in S$ as possible [3, 15, 32].

Proposition A.3. MAX H-CONSISTENCY *is* APX-*complete and can be approximated within* 2.

Proof. MAX H-CONSISTENCY can clearly be approximated within 2 using the greedy algorithm of proposition 9.

In order to show that it is APX-hard we adapt the L-reduction from MAX 2SAT to MAX FLS$^\geqslant$ given in the proof of Theorem 5. Starting with the system of $9m$ inequalities (20)–(24), we add two extra variables $w_{|X|+1}$ and w_0 (the bias) as well as a large enough number of nonstrict inequalities forcing $w_0, w_{|X|+1} \geqslant 0$. For each clause C_i, $1 \leqslant i \leqslant m$, containing two variables x_{j_1} and x_{j_2} we consider the 10 positive examples in $\mathbb{R}^n$ with $n = |X| + 1$ associated to the following inequalities:

$$10a_{ij_1}w_{j_1} + 10a_{ij_2}w_{j_2} + 2w_{|X|+1} > w_0, \tag{A.1}$$

$$10w_{j_1} - 8w_{|X|+1} > w_0, \tag{A.2}$$

$$-10w_{j_1} + 12w_{|X|+1} > w_0, \tag{A.3}$$

$$10w_{j_1} + 12w_{|X|+1} > w_0, \tag{A.4}$$

$$-10w_{j_1} - 8w_{|X|+1} > w_0, \tag{A.5}$$

$$10w_{j_2} - 8w_{|X|+1} > w_0, \tag{A.6}$$

$$-10w_{j_2} + 12w_{|X|+1} > w_0, \tag{A.7}$$

$$10w_{j_2} + 12w_{|X|+1} > w_0, \tag{A.8}$$

$$-10w_{j_2} - 8w_{|X|+1} > w_0, \tag{A.9}$$

$$w_{|X|+1} > w_0. \tag{A.10}$$

Moreover, we include $3m$ identical negative examples $\mathbf{0}$ implying the inequality $w_0 \geqslant 0$. Thus we have a MAX H-CONSISTENCY instance with $10m$ positive examples and $3m$ negative ones.

It is easy to verify that there is a correspondence between solutions of the MAX 2SAT instance satisfying s clauses and hyperplanes classifying correctly $10m+s$ vectors of S. In fact, in any solution $(\boldsymbol{w}, w_0)$ that fulfils $10m+s$ inequalities we have $w_{|X|+1} > w_0 \geqslant 0$ because the corresponding hyperplane must correctly classify at least a negative example and a positive one of type (A.10). As far as the L-reduction is concerned, the inequalities (A.1)–(A.10) are therefore equivalent to the inequalities (20)–(24).

This L-reduction can be extended to the case where the negative examples are not all identical by considering $(n+3m)$-dimensional examples instead of n-dimensional ones. We construct the same $10m$ positive examples (all additional components are 0) and $3m$ pairs of negative examples implying the following inequalities:

$$w_{|X|+i+1} \leqslant w_0,$$
$$-w_{|X|+i+1} \leqslant w_0,$$

where $1 \leqslant i \leqslant 3m$ and $w_{|X|+i+1}$ are (free) variables occurring in a single inequality. Clearly, any solution $(\boldsymbol{w}, w_0)$ satisfying at least one pair of inequalities is such that $w_0 \geqslant 0$. Since the absolute values of all $w_{|X|+i+1}$ can be taken as small as needed (these variables are unconstrained) and since at most $7m$ inequalities of types (A.2)–(A.10) can be simultaneously satisfied, we have a correspondence between solutions of the MAX 2SAT instance satisfying s clauses and hyperplanes classifying correctly $13m+s$ vectors of S. □

The same nonapproximability result holds in the symmetric case where we look for a hyperplane containing no examples, i.e. such that $\boldsymbol{w}\boldsymbol{a}^k > w_0$ for all positive examples $\boldsymbol{a}^k$ while $\boldsymbol{w}\boldsymbol{a}^k < w_0$ for all negative ones. Note that MAX H-CONSISTENCY is easier to approximate than the complementary problem that consists of minimizing the number of misclassifications [4, 5, 21].

A variant of MAX H-CONSISTENCY, which is a special case of C MAX FLS$^{\geqslant;>}$, occurs as a subproblem in various constructive methods for building multilayer networks [34, 33]. In this problem, named MAX H-COVERING, only a single type of misclassifications is allowed. Given a set of examples, we look for a hyperplane that correctly classifies all negative examples and as many positive ones as possible.

Corollary A.4. MAX H-COVERING *is* MAX IND SET-*hard.*

Proof. By L-reduction from MAX 2 ONES NEG, which is known to be MAX IND SET-hard [35] and is defined as follows [16]. Given a finite set X of variables and a set $C = \{C_1, \ldots, C_m\}$ of 2-literal clauses with only negated variables, find a truth assignment for X that satisfies every clause and that contains as many TRUE variables as possible.

For each clause $\bar{x}_{j_1} \vee \bar{x}_{j_2}$ we construct a negative example $\boldsymbol{a} \in \mathbb{R}^n$ with $n = |X|$, $a_{j_1} = a_{j_2} = 1$ and $a_j = 0$ for $1 \leqslant j \leqslant n$ with $j \neq j_1$ and $j \neq j_2$. The trivial vector $\boldsymbol{0}$ is also included as a negative example. Finally, we construct for each boolean variable x_j the positive example $\boldsymbol{a}$ where $a_j = 1$ and $a_l = 0$ for $1 \leqslant l \leqslant n$ with $l \neq j$.

We claim that there exists a truth assignment with s TRUE variables satisfying all clauses if and only if there exists a hyperplane H classifying correctly all negative examples and s positive ones. Given an appropriate truth assignment, the hyperplane H specified by the bias $w_0 = 1$ and the normal vector $\boldsymbol{w}$ defined by

$$w_j = \begin{cases} 2 & \text{if } x_j \text{ is TRUE,} \\ -2 & \text{otherwise} \end{cases}$$

correctly classifies all negative examples and the positive ones corresponding to a TRUE variable. Moreover, the $n - s$ positive examples associated with a FALSE variable are misclassified. Conversely, given an appropriate hyperplane H we consider the set of boolean variables $Y \subseteq X$ associated to a positive example that is correctly classified. It is easily verified that the assignment where the s variables in Y are TRUE and all the other ones are FALSE satisfies all the clauses of (X, C). □

The symmetric variant of MAX H-COVERING where the objective is to correctly classify all positive examples and as many negative examples as possible is also MAX IND SET-hard.

References

[1] N. Alon, U. Feige, A. Wigderson and D. Zuckerman, Derandomized graph products, *Comput. Complexity*, to appear.

[2] N. Alon and F.R.K. Chung, Explicit construction of linear sized tolerant networks, *Discrete Math.* **72** (1988) 15–19.

[3] E. Amaldi, On the complexity of training perceptrons, in T. Kohonen et al., eds. *Artificial Neural Networks* (Elsevier, Amsterdam, 1991) 55–60.

[4] E. Amaldi and V. Kann, On the approximability of removing the smallest number of relations from linear systems to achieve feasibility, Tech. Report ORWP-6-94, Department of Mathematics, Swiss Federal Institute of Technology, Lausanne, 1994.

[5] S. Arora, L. Babai, J. Stern and Z. Sweedyk, The hardness of approximate optima in lattices, codes, and systems of linear equations, in: *Proc. of 34th Ann. IEEE Symp. on Foundations of Comput. Sci.* (1993) 724–733.

[6] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy, Proof verification and hardness of approximation problems. In *Proc. of 33rd Ann. IEEE Symp. on Foundations of Comput. Sci.*, 1992 14–23.

[7] G. Ausiello, P. Crescenzi and M. Protasi, Approximate solutions of NP optimization problems. Tech. Report SI/RR-94/03, Università di Roma "La Sapienza", 1994.

[8] L. Babai, Transparent proofs and limits to approximation, Manuscript, 1993.

[9] M. Bellare, S. Goldwasser, C. Lund and A. Russell, Efficient probabilistically checkable proofs and applications to approximation, in: *Proc. Twenty fifth Ann. ACM Symp. on Theory of Comp.*, ACM (1993) 294–304.

[10] P. Berman and G. Schnitger, On the complexity of approximating the independent set problem, *Inform. and Comput.* **96** (1992) 77–94.

[11] P. Crescenzi, V. Kann and L. Trevisan, Natural complete and intermediate problems in approximation classes, Manuscript, 1994.

[12] P. Crescenzi and A. Panconesi, Completeness in approximation classes, *Inform. and Comput.* **93** (1991) 241–262.

[13] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis* (Wiley, New York, 1973).

[14] M.R. Frean, The upstart algorithm: a method for constructing and training feedforward neural networks, *Neural Comput.* **2** (1990) 198–209.

[15] S.I. Gallant, Perceptron-based learning algorithms, *IEEE Trans. on Neural Networks* **1** (1990) 179–191.

[16] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness* (W. H. Freeman and Company, San Francisco, 1979).

[17] H.J. Greenberg and F.H. Murphy, Approaches to diagnosing infeasible linear programs, *ORSA J. Comput.* **3** (1991) 253–261.

[18] R. Greer, *Trees and Hills: Methodology for Maximizing Functions of Systems of Linear Relations*, Annals of Discrete Mathematics, Vol. 22 (Elsevier, Amsterdam, 1984).

[19] D.J. Hand, *Discrimination and classification* (Wiley, New York, 1981).

[20] J. Håstad, S. Phillips and S. Safra, A well-characterized approximation problem, *Inform. Process. Lett.* **47** (1993) 301–305.

[21] K-U. Höffgen, H-U. Simon and K. van Horn, Robust trainability of single neurons, Tech. Report CS-92-9, Computer Science Department, Brigham Young University, Provo, UT, 1992.

[22] D.S. Johnson, Approximation algorithms for combinatorial problems, *J. Comput. System Sci.* **9** (1974) 256–278.

[23] D.S. Johnson and F.P. Preparata, The densest hemisphere problem, *Theoret. Comput. Sci.* **6** (1978) 93–107.

[24] N. Kahale, On the second eigenvalue and linear expansion of regular graphs, in: *Proc. of 33rd Ann. IEEE Symp. on Foundations of Comput. Sci.* (1992) 296–303

[25] V. Kann, On the approximability of NP-complete optimization problems, Ph.D. Thesis, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1992.

[26] V. Kann, Polynomially bounded minimization problems that are hard to approximate, in: *Proc. of 20th Internat. Colloq. on Automata, Languages and Programming*, Lecture Notes in Computer Science Vol. 700 (Springer, Berlin 1993) 52–63; *Nordic J. Comput.*, to appear.

[27] N. Karmarkar, A new polynomial time algorithm for linear programming, *Combinatorica* **4** (1984) 373–395.

[28] S. Khanna, R. Motwani, M. Sudan and U. Vazirani, On syntactic versus computational views of approximability, in: *Proc. 35th Ann. IEEE Symp. on Foundations of Comput. Sci.*, 1994.

[29] P.G. Kolaitis and M.N. Thakur, Logical definability of NP optimization problems, *Inform. and Comput.* **115** (1994) 321–353.

[30] A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan graphs, *Combinatorica* **8** (1988) 261–277.

[31] C. Lund and M. Yannakakis, On the hardness of approximating minimization problems, *J. ACM* **41** (1994) 960–981.

[32] M. Marchand and M. Golea, An approximate algorithm to find the largest linearly separable subset of training examples, in: *Proc. of 1993 Ann. Meeting of the International Neural Network Society*, (INNS Press, 1993) 556–559.

[33] M. Marchand and M. Golea, On learning simple neural concepts: from halfspace intersections to neural decision lists, *Network: Comput. Neural Systems* **4** (1993) 67–85.

[34] M. Marchand, M. Golea and P. Ruj
án, A convergence theorem for sequential learning in two-layer perceptrons, *Europhys. Lett.* **11** (1990) 487–492.

[35] A. Panconesi and D. Ranjan, Quantifiers and approximation, *Theoret. Comput. Sci.* **107** (1993) 145–163.

[36] C.H. Papadimitriou and M. Yannakakis, Optimization, approximation, and complexity classes, *J. Comput. System Sci.* **43** (1991) 425–440.

[37] A. Schrijver, *Theory of Linear and Integer Programming*, Interscience Series in Discrete Mathematics and Optimization (Wiley, New York, 1986).

[38] R.E. Warmack and R.C. Gonzalez, An algorithm for optimal solution of linear inequalities and its application to pattern recognition, *IEEE Trans. Comput.* **22** (1973) 1065–1075.

[39] M. Yannakakis, On the approximation of maximum satisfiability, in: *J. Algorithms* **17** (1994) 475–502.