On the Approximability of Removing the Smallest Number of Relations from Linear Systems to Achieve Feasibility

Edoardo Amaldi
Department of Mathematics
Swiss Federal Institute of Technology
CH-1015 Lausanne
amaldi@dma.epfl.ch

Viggo Kann
Department of Numerical Analysis
and Computing Science
Royal Institute of Technology
S-100 44 Stockholm
viggo@nada.kth.se

August 11, 1994

Abstract

We investigate the computational complexity of the problem which consists, given a system of linear relations, of finding a solution violating as few relations as possible while satisfying all the others. This general combinatorial problem, referred to as MIN ULR, is considered for the four basic types of relational operators $=$, $\geq$, $>$ and $\neq$. We proved in [3] that MIN ULR with $=$, $\geq$ or $>$ relations is NP-hard even when restricted to homogeneous systems with bipolar coefficients, whereas it is trivial for $\neq$ relations. In this paper we determine strong bounds on the approximability of various intractable variants, including constrained ones where the variables are restricted to take bounded discrete values. The various NP-hard versions of MIN ULR belong to different approximability classes depending on the type of relations and the additional constraints, but none of them can be approximated within any constant factor unless P=NP. In the process of studying MIN ULR we also derive strong bounds on the approximability of a class of closely related problems which consist, given a feasible linear system, of finding a solution with as few nonzero variables as possible. Finally we comment on the approximability of two important special cases that arise in discriminant analysis and machine learning.

Keywords: Linear relations, feasible subsystems, computational complexity, approximability, MIN DOMINATING SET-hard problems, NPO PB-complete problems, designing linear classifiers, training perceptrons.

1 Introduction

We consider the general problem of finding the minimum set of relations that must be removed from a given linear system to make it feasible for the four types of relational operators $=$, $\geq$, $>$ and $\neq$. The basic versions, referred to as Min ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$ for minimum Unsatisfied Linear Relations, are defined as follows: Given a linear system $A\mathbf{x}\mathcal{R}\mathbf{b}$ with a matrix A of size $p \times n$, find a solution $\mathbf{x} \in \mathbb{R}^n$ which violates as few relations as possible while satisfying all the others.

Many variants of these combinatorial problems arise in various fields such as operations research [26, 23, 22], pattern recognition [39, 15, 31] and artificial neural networks [1, 25, 34].

It is well-known that feasible linear systems with equations or inequalities can be solved in polynomial time using an adequate linear programming method [29]. For infeasible systems, least mean squared algorithms, which are suited to linear regression, are not appropriate when the objective is to minimize the number of unsatisfied relations.

A number of algorithms have been proposed for tackling various versions of Min ULR. The weighted variants were also considered in which each relation has an associated weight and the goal is to minimize the total weight of the unsatisfied relations. Johnson and Preparata proved that special cases of Min ULR$^{>}$ and Min ULR$^{\geq}$ with homogeneous systems are NP-hard and devised a complete enumeration method which is also applicable to the weighted and mixed variants [26]. Greer developed a tree algorithm for optimizing functions of systems of linear relations that is more efficient than complete enumeration but still exponential in the worst case [23]. This general procedure can be used to solve Min ULR with any of the four types of relations.

During the last decade many mathematical programming formulations have been studied to design linear discriminant classifiers (see [31, 9] as well as the included references). When the goal is to determine optimal linear classifiers which misclassify the least number of points in the training set, the problem amounts to a special case of Min ULR$^{>}$ and Min ULR$^{\geq}$. Increasingly sophisticated models have been proposed in order to try to avoid unacceptable or trivial solutions [9].

The same type of problem has also attracted a considerable interest in learning theory because it arises when training perceptrons, in particular when minimizing their number of errors. While some heuristic algorithms were devised in [18, 17, 2], Amaldi showed that solving these problems to optimality is NP-hard even when restricted to perceptrons with bipolar inputs -1 or 1 [1]. In [25] minimizing the number of misclassifications was proved at least as hard to approximate as minimum set cover and the special case with binary variables not easier than all minimization problems with polynomially bounded objective functions.

In recent years a growing attention has been paid to infeasible linear programs [22]. When formulating or modifying very large and complex models, it is hard to prevent errors and to guarantee feasibility. Infeasible programs with thousands of constraints frequently occur and cannot be repaired by simple inspection. With the new software and hardware advances, dealing with infeasible programs is becoming a major bottleneck in linear programming. Several methods have been proposed in order to try to locate the source of infeasibility [20, 11] and to remove as few constraints as possible to achieve feasibility [21, 37].

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 [27]). 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 [7, 33, 8]. These results have also important consequences on the approximability of other optimization problems.

In [3, 4] we performed a thorough study of the approximability of the complementary problems, named Max FLS, where one looks for maximum feasible subsystems of linear systems. In particular, we proved that the basic versions with $=$, $\geq$ or $>$ relations are NP-hard even for homogeneous systems with bipolar coefficients. While Max FLS with equations cannot be approximated within p^ε for some $\varepsilon > 0$ where p is the number of relations, the variants with strict or nonstrict inequalities can be approximated within 2 but not within every constant factor. Although complementary pairs of problems such as Min ULR and Max FLS are equivalent to solve optimally their degree of approximability can differ enormously. This is, for instance, the case for the minimum node cover and the maximum independent set problems [19].

By delving into the interactive proof connection, Arora et al. lately established that Min ULR$^=$ cannot be approximated within any constant unless P=NP and within a factor of $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$ unless NP $\subseteq$ DTIME($n^{\text{polylog } n}$) [5]. Moreover, they extended this nonapproximability result to the problem of minimizing the number of misclassifications of a perceptron.

This paper is organized as follows. Section 2 gives a brief account on approximability classes and their hierarchy for minimization problems. In section 3 we recall the known results about the complexity of solving the basic Min ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$ to optimally and we determine alternative bounds on their approximability. In section 4 we study the approximability of Min RVLS, a closely related class of problems which consist, given a feasible linear system, of finding a solution with as few nonzero variables as possible. We also discuss its consequences on Min ULR. While section 5 is devoted to the weighted variants where a different importance may be attached to each relation, section 6 deals with those where the variables are restricted to take a finite number of discrete values. The relation between Min ULR with real and binary variables is analysed in section 7. In section 8 we comment on the approximability of two interesting special cases of Min ULR and Min RVLS with inequalities which arise in discriminant analysis and machine learning. Finally, in section 9 the various nonapproximability results are summarized and some open questions are mentioned.

2 Approximability hierarchy of minimization problems

Definition 1 [14] An NP optimization (NPO) problem over an alphabet Σ is a four-tuple $\Pi = (\mathcal{I}_\Pi, S_\Pi, m_\Pi, opt_\Pi)$ where $\mathcal{I}_\Pi \subseteq \Sigma^*$ is the space of *input instances*, $S_\Pi(I) \subseteq \Sigma^*$ is the space of *feasible solutions* on input $I \in \mathcal{I}_\Pi$, $m_\Pi : \mathcal{I}_\Pi \times \Sigma^* \rightarrow \mathbb{N}$, the *objective function*, is a polynomial time computable function, $opt_\Pi \in \{\max, \min\}$ tells if Π is a *maximization* or a *minimization* problem.

The set $\mathcal{I}_\Pi$ must be recognizable by a Turing machine in polynomial time. The only requirement on S_Π is that there exist a polynomial q and a polynomial time computable predicate π such that for all I in $\mathcal{I}_\Pi$, S_Π can be expressed as $S_\Pi(I) = \{y : |y| \leq q(|I|) \wedge \pi(I, y)\}$ where q and π only depend on Π.

Solving an optimization problem Π given the input $I \in \mathcal{I}_\Pi$ means finding a $y \in S_\Pi(I)$ such that $m_\Pi(I, y)$ is optimum, that is as large as possible if $opt_\Pi = \max$ and as small as possible if $opt_\Pi = \min$. Let $opt_\Pi(I)$ denote this optimal value of m_Π.

Approximating an optimization problem Π given the input $I \in \mathcal{I}_\Pi$ means finding any $y' \in S_\Pi(I)$. How good the approximation is depends on the relation between $m_\Pi(I, y')$ and $opt_\Pi(I)$. The *performance ratio* of a feasible solution with respect to the optimum of a minimization problem Π is defined as $R_\Pi(I, y) = m_\Pi(I, y)/opt_\Pi(I)$ where $I \in \mathcal{I}_\Pi$ and $y \in S_\Pi(I)$.

Definition 2 An optimization problem Π *can be approximated within* c for a constant c if there exists a polynomial time algorithm A such that for all instances $I \in \mathcal{I}_\Pi$, $A(I) \in S_\Pi(I)$ and $R_\Pi(I, A(I)) \leq c$. More generally, an optimization problem Π *can be approximated within* $p(n)$ for a function $p : \mathbb{Z}^+ \rightarrow \mathbb{R}^+$ if there exists a polynomial time algorithm A such that for every $n \in \mathbb{Z}^+$ and for all instances $I \in \mathcal{I}_\Pi$ with $|I| = n$ we have that $A(I) \in S_\Pi(I)$ and $R_\Pi(I, A(I)) \leq p(n)$.

Although various reductions preserving approximability within constants have been proposed (see [10, 27, 35]), we will use the S-reduction which is suited to relate problems that cannot be approximated within any constant.

Definition 3 [28] Given two NPO problems Π and Π' and a polynomial time transformation $f : \mathcal{I}_\Pi \to \mathcal{I}_{\Pi'}$. $f = (t_1, t_2, a(n), c)$ is an *S-reduction with size amplification* $a(n)$ from Π to Π' if

i) t_1, t_2 are polynomial time computable functions, $a(n)$ is a monotonously increasing positive function and c is a positive constant.

ii) $t_1 : \mathcal{I}_\Pi \to \mathcal{I}_{\Pi'}$ and $\forall I \in \mathcal{I}_\Pi$ and $\forall y \in S_{\Pi'}(t_1(I))$, $t_2(I, y) \in S_\Pi(I)$.

iii) $\forall I \in \mathcal{I}_\Pi$ and $\forall y \in S_{\Pi'}(t_1(I))$, $R_\Pi(I, t_2(I, y)) \leq c \cdot R_{\Pi'}(t_1(I), y)$.

iv) $\forall I \in \mathcal{I}_\Pi, |t_1(I)| \leq a(|I|)$.

The composition of S-reductions is an S-reduction. If Π S-reduces to Π' with size amplification $a(n)$ and Π' can be approximated within some monotonously increasing function $u(n)$ in the size of the input instance, then Π can be approximated within $c \cdot u(a(n))$. For constant and polylogarithmic approximable problems the S-reduction preserves approximability within a constant for any polynomial size amplification. For n^c approximable problems the S-reduction preserves approximability within a constant just for linear size amplification. An S-reduction with $c = 1$ is a *cost preserving transformation.*

Definition 4 [10, 32] An NPO problem Π is *polynomially bounded* if there is a polynomial p such that

$$\forall I \in \mathcal{I}_\Pi \, \forall y \in S_\Pi(I), \, m_\Pi(I, y) \leq p(|I|).$$

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

All unweighted versions of Min ULR are included in NPO PB since their objective function is the number of unsatisfied relations. The weighted versions of Min ULR are included in NPO PB if the weights are polynomially bounded, since the objective function is a sum of a polynomially bounded number of polynomially bounded weights.

Definition 5 Given an NPO problem Π and a class C, Π *is C-hard* if every $\Pi' \in C$ can be S-reduced, with some size amplification, to Π. Π *is C-complete* if $\Pi \in C$ and Π is C-hard.

The range of approximability of NPO minimization problems stretches from problems which can be approximated within every constant in polynomial time, i.e. that have a *polynomial time approximation scheme* like the knapsack problem, to problems that cannot be approximated within $n^{1-\varepsilon}$ for every $\varepsilon > 0$, where n is the size of the input instance, unless P=NP.

Recently Lund and Yannakakis proved, using results from interactive proofs, that Min Graph Colouring cannot be approximated within n^ε for some $\varepsilon > 0$ unless P=NP. Moreover, they derived a strong bound on the approximability of Set Covering and of several closely related minimization problems such as Min Dominating Set [33]. In [8] Bellare et al. improved this result by showing that approximating Set Covering (or some of the related problems) within $c \log n$ for any $c < 1/8$ would imply NP $\subseteq$ DTIME$(n^{\log \log n})$. Since DTIME$(T(n))$ denotes the class of problems which can be solved in time $T(n)$, this inclusion is widely believed to be unlikely. Furthermore, it was established that Min Dominating Set and the related problems cannot be approximated within any constant factor unless P=NP. If there is an approximation preserving reduction from Min Dominating Set to an NPO problem Π we say that Π is Min Dominating Set-hard, which means that it is at least as hard to approximate as the minimum dominating set problem.

If we require the dominating set in MIN DOMINATING SET to be independent we get the *minimum independent dominating set problem* or MIN IND DOM SET. Halldórsson established in [24] that MIN IND DOM SET is very hard to approximate. Assuming P≠NP, the problem cannot be approximated within a factor of $n^{1-\varepsilon}$ for any $\varepsilon > 0$, where n is the number of nodes in the graph. Inspection of Halldórsson's proof shows that the result is still valid if n is the input size, i.e., the sum of the number of nodes and edges in the graph. Furthermore, Kann proved that MIN IND DOM SET is complete for NPO PB in the sense that every polynomially bounded NP optimization problem can be reduced to it using an approximation preserving reduction [28, 13].

The purpose of this paper is to investigate the approximability status of the different MIN ULR variants and to locate them in the hierarchy of approximability classes. We will see that the hardness of approximating apparently similar and closely related versions (for instance a pair of complementary problems) can differ enormously.

3 Approximability of MIN ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$

In this section we focus on the basic versions of MIN ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$.

Unlike the other problems, MIN ULR$^{\neq}$ is trivially solvable because any such system is feasible. Indeed, for any finite set of hyperplanes associated with a set of linear relations there exists a vector $\mathbf{x} \in \mathbb{R}^n$ that does not belong to any of them.

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

For MIN ULR$^{=}$ and MIN ULR$^{\geq}$ with homogeneous systems, we are not interested in the trivial solutions where all variables occurring in the satisfied relations are zero. Even if we forbid the solution $\mathbf{x} = \mathbf{0}$ there might be other trivial solutions where almost all variables are zero except a few that only occur in a few equations. In order to rule out all these trivial solutions for the equality and nonstrict inequality case, we only consider solutions where the variable (variables) occurring in the largest number of satisfied equations is (are) nonzero.

Theorem 1 ([3]) MIN ULR$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq, >\}$ *is* NP*-hard even when restricted to homogeneous systems with bipolar coefficients in* $\{-1, 1\}$.

In other words, MIN ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$ is intractable even when the points corresponding to the rows of A belong to the n-dimensional hypercube.

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. Thus these polynomially bounded problems 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 [19].

It is worth noting that if the number of variables n is constant these three basic versions of MIN ULR can be solved in polynomial time using Greer's algorithm which has an $O(n \cdot p^n / 2^{n-1})$ time-complexity, where p denotes the number of relations and n the number of variables [23]. These problems are trivial when the number of relations p is constant because all subsystems can be checked in $O(n)$ time. Furthermore, they are easy when all maximal feasible subsystems contain a maximum number of relations because a greedy procedure is guaranteed to give a solution that minimizes the number of unsatisfied relations.

Since it is extremely unlikely that the three basic versions of MIN ULR can be solved optimally in polynomial time, we are interested in the existence of efficient algorithms providing solutions that are guaranteed to be a fixed percentage away from the actual optimum. Such approximate algorithms are of great interest in practice.

Before establishing strong bounds on the approximability of MIN ULR$^{\geq}$ and MIN ULR$^{>}$, we point out the relationship between the approximability of MIN ULR$^{\geq}$ and that of MIN ULR$^{=}$.

Proposition 2 MIN ULR$^{\geq}$ *is at least as hard to approximate as* MIN ULR$^{=}$ *but not harder than* MIN ULR$^{=}$ *with nonnegative variables.*

PROOF We first reduce MIN ULR$^{=}$ to MIN ULR$^{\geq}$ and then reduce MIN ULR$^{\geq}$ to MIN ULR$^{=}$ with nonnegative variables.

Let $(A, \mathbf{b})$ be an arbitrary instance of MIN ULR$^{=}$. For each equation $\mathbf{a}^i\mathbf{x} = b_i$ where $\mathbf{a}^i$ denotes the ith row of A, $1 \leq i \leq p$, we consider the two inequalities $\mathbf{a}^i\mathbf{x} \geq b_i$ and $-\mathbf{a}^i\mathbf{x} \geq -b_i$. Since any solution $\mathbf{x}$ satisfies at least one inequality of each such pair, finding a solution that minimizes the number of unsatisfied equations in $A\mathbf{x} = \mathbf{b}$ is equivalent to finding a solution that minimizes the number of unsatisfied inequalities of the corresponding system of size $2p \times n$.

Let $(A, \mathbf{b})$ be an arbitrary instance of MIN ULR$^{\geq}$. By replacing each variable x_i unrestricted in sign by the difference $x_i' - x_i''$ of two nonnegative variables $x_i', x_i'' \geq 0$ and by adding a slack variable for each inequality, we obtain an equivalent system with p equations and $2n + p$ nonnegative variables. □

Proposition 3 *The problems* MIN ULR$^{\geq}$ *and* MIN ULR$^{>}$ *with integer coefficients are equivalent.*

PROOF We simply use the fact that any system $A\mathbf{x} \leq \mathbf{b}$ has a solution if and only if the system $A\mathbf{x} < \mathbf{b} + \varepsilon\mathbf{1}$ has a solution, where $\varepsilon = 2^{-2L}$ and L is the size (in bits) of the binary encoded input instance [36]. □

It has recently been shown in [5] that MIN ULR$^{=}$ cannot be approximated within any constant unless P=NP and within a factor of $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$ unless NP $\subseteq$ DTIME($n^{\text{polylog}\, n}$), where n is, for instance, the number of variables. According to propositions 2 and 3, this is also true for MIN ULR with strict and nonstrict inequalities.

We now prove a nonapproximability result for MIN ULR$^{\geq}$ and MIN ULR$^{>}$ which is more likely to be true but whose bound is not as strong.

Theorem 4 MIN ULR$^{\geq}$ *and* MIN ULR$^{>}$ *are* MIN DOMINATING SET*-hard. They cannot be approximated within any constant unless* P=NP *and within* $c \log n$ *for any* $c < 1/8$ *unless* NP $\subseteq$ DTIME($n^{\log\log n}$)*, even when restricted to homogeneous systems with ternary coefficients in* $\{-1, 0, 1\}$.

PROOF We proceed by cost preserving reduction from the known NP-complete problem MIN DOMINATING SET which is defined as follows [19]. Given an undirected graph $G = (V, E)$, find a minimum cardinality set $V' \subseteq V$ that dominates all nodes of G, i.e. for all $v \in V \setminus V'$ there exists $v' \in V'$ such that $[v, v'] \in E$.

Let $G = (V, E)$ be an arbitrary instance of MIN DOMINATING SET. For each node $v_i \in V$, $1 \leq i \leq n$, we consider the homogeneous inequality

$$x_i \geq 0 \tag{1}$$

and the inhomogeneous inequality

$$x_i + \sum_{j \in N(v_i)} x_j \leq -1. \tag{2}$$

where j is included in $N(v_i)$ if and only if v_j is adjacent to v_i. Thus we have a system with $2n$ inequalities and n variables.

We claim that there exists a dominating set in G of size at most s if and only if there exists a solution $\mathbf{x}$ that violates at most s inequalities of the corresponding system.

Given a dominating set $V' \subseteq V$ of size s, the solution $\mathbf{x}$ defined by

$$x_i = \begin{cases} -1 & \text{if } v_i \in V' \\ 0 & \text{otherwise} \end{cases}$$

satisfies all inhomogeneous inequalities and $n - s$ homogeneous ones.

Conversely, given a solution vector $\mathbf{x}$ that violates s inequalities, we can always satisfy every inhomogeneous inequality that is not already satisfied by making one variable x_i in the inequality negative enough. This operation yields a solution that satisfies at least as many inequalities as $\mathbf{x}$. Consider the set of nodes $V' \subseteq V$ containing all nodes v_i such that $x_i \neq 0$. V' is clearly a dominating set of size s, because $x_i + \sum_{j \in N(v_i)} x_j \leq -1$ only when at least one of the variables is negative, which corresponds to the case where at least one of the nodes is in the dominating set.

Since the above reduction is cost preserving and without amplification, we have exactly the same nonapproximability bounds for Min ULR$^{\geq}$ and Min ULR$^{>}$ as for Min Dominating Set.

This reduction can easily be adapted to Min ULR$^{>}$ restricted to homogeneous systems. It suffices to replace $\geq$ by $>$ and the -1 in the right hand side of the second type of inequalities by 0.

There is a standard way to transform an inhomogeneous Min ULR$^{=}$ or Min ULR$^{\geq}$ system to a homogeneous one. Given a system consisting of p relations, we first multiply all constant right hand sides by a new variable x_0. Then we only have to ensure that this constant is nonzero (for the equality case) or positive (for the inequality case). Therefore we introduce new relations $k \cdot x_0 = k \cdot y_0$ or $k \cdot x_0 \geq 0$, for $k \in [1..p+1]$, where y_0 is a new variable. In every solution that satisfies more than p relations, at least one of the new relations must be satisfied, and since all the new relations are satisfied simultaneously, x_0 will become the variable occurring in the largest number of satisfied equations. Hence x_0 is nonzero. For the inequality case this implies that $x_0 \geq 0$.

Since the same trick can also be applied to all variants of Min ULR$^{=}$ or Min ULR$^{\geq}$, we will not give this reduction every time it is used. □

Clearly, for large n and small $\varepsilon > 0$, a factor of $2^{\log^{1-\varepsilon} n}$ is larger than $c \log n$ with any $c < 1/8$, but NP $\subseteq$ DTIME$(n^{\mathrm{polylog}\, n})$ is more likely to be true than NP $\subseteq$ DTIME$(n^{\log\log n})$. Furthermore, the above proof is much simpler than those given in [5].

Notice that the basic versions of Min ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$ can be approximated within $n+1$, where n is the number of variables. This is a consequence of Helly's theorem for convex sets which states (when applied to linear systems) that every infeasible system with p inequalities or equations and n variables contains an infeasible subsystem with at most $n+1$ relations [12].

The question of whether a logarithmic performance ratio $c \log n$ for some $c \geq 1/8$ can be guaranteed in polynomial time unless P = NP is still open.

4 Complexity of a class of closely related problems

In this section we consider a class of problems pertaining to feasible systems but which are closely related to Min ULR. Given any feasible system of linear relations, one seeks a solution satisfying all relations with as few nonzero variables as possible. Since the purpose is to minimize the number of Relevant Variables in Linear Systems, these problems are referred to as MIN RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$. MIN RVLS$^{=}$ is called *minimum weight solution to linear equations* in the list of NP-complete problems in [19].

There is a simple relation between the approximability of MIN ULR and that of MIN RVLS.

Proposition 5 MIN ULR$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq, >\}$ *is at least as hard to approximate as* MIN RVLS$^{\mathcal{R}}$ *with the same type of relations.*

PROOF It suffices to note that for any instance of MIN RVLS$^{=}$ given by $A\mathbf{x} = \mathbf{b}$ one can construct an equivalent instance of MIN ULR$^{=}$. This is simply achieved by considering, for each variable x_i with $1 \leq i \leq n$, the equation

$$x_i = 0 \tag{3}$$

and by eliminating variables in this system using the set of equations $A\mathbf{x} = \mathbf{b}$ that must be satisfied.

Similarly, by replacing each equation (3) by two nonstrict inequalities one shows that MIN ULR$^{\geq}$ is at least as hard to approximate as MIN RVLS$^{\geq}$.

Since by proposition 3 MIN ULR$^{>}$ is equivalent to MIN ULR$^{\geq}$ for systems with integer coefficients, MIN ULR$^{>}$ is also at least as hard as MIN RVLS$^{\geq}$. $\square$

Thus any nonapproximability bound for MIN RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$ implies an equivalent bound for the corresponding MIN ULR$^{\mathcal{R}}$.

Proposition 6 MIN RVLS$^{=}$ *cannot be approximated within any constant factor unless* P=NP.

PROOF By reduction from a variant of SET COVERING with disjoint sets. In SET COVERING, given a collection $\mathcal{C} = \{C_1, \ldots, C_n\}$ of subsets of a finite set S, one seeks a subcollection $\mathcal{C}' \subseteq \mathcal{C}$ of minimum cardinality such that $\cup_{i=1}^{m} C_{j_i} = S$ with $m \leq n$. Any such $\mathcal{C}'$ is a *cover* of S. If all the sets in $\mathcal{C}'$ are pairwise disjoint, it is an *exact cover*.

The proof is based on the following result by Bellare et al. on exact covers [8]. For every $c > 1$, there exists a polynomial time reduction that transforms any instance ϕ of the satisfiability problem SAT (see [19]) into an instance of SET COVERING with a positive integer K such that

- if ϕ is satisfiable there exists an exact cover $\mathcal{C}'$ of size K,
- if ϕ is unsatisfiable no set cover has size less than $\lfloor c \cdot K \rfloor$.

For any instance $(S, \mathcal{C})$ of set cover, we can construct a system with n variables and $|S|$ equations

$$A\mathbf{x} = \mathbf{1}, \tag{4}$$

where $a_{ij} = 1$ if the ith element of S belongs to C_j and 0 otherwise. $\mathbf{1}$ denotes the $|S|$-dimensional vector with all 1 components. Clearly, the nonzero variables in any solution $\mathbf{x}$ of the above system define a set cover. Conversely, given any exact cover $\mathcal{C}'$ of cardinality K, the vector $\mathbf{x}$ given by

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

satisfies all equations and has K nonzero variables. Thus the minimum number of nonzero variables in a solution of $A\mathbf{x} = \mathbf{1}$ is either K or greater or equal to $\lfloor c \cdot K \rfloor$.

Note that while there is a one–to–one correspondence between the exact covers and the solutions of (4) with 0 or 1 components, minimum cardinality covers do not necessarily correspond to a solution. For example, in the set covering instance associated to

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

all subcollections of cardinality 2 are minimum (nonexact) covers but the system has the unique solution $(1/2, 1/2, 1/2)$ with three nonzero variables. For infeasible systems like

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

the minimum number of nonzero variables can be considered as larger than n. □

According to proposition 5, the above result directly implies that MIN ULR$^=$ cannot be approximated within any constant factor unless P=NP, which was established in [5]. Furthermore, propositions 2 and 3 show that MIN ULR$^\geq$ and MIN ULR$^>$ are at least as hard to approximate.

In fact, MIN RVLS$^=$ and MIN RVLS with inequalities are not approximable within a factor of $2^{\log^{0.5-\varepsilon} n}$ for any $\varepsilon > 0$. In order to prove this strong nonapproximability bound we consider a natural variant of the LABEL COVER problem studied in the above reference.

Definition 6 Let $G = (V_1, V_2, E)$ be a bipartite graph, B_1 and B_2 two sets of possible labels for the nodes in V_1 and V_2 respectively, and $\Pi \subseteq E \times B_1 \times B_2$ a relation consisting of admissible pairs of labels for each edge. A *labelling* of G is a function $\mathcal{P} : V_1 \cup V_2 \to 2^{B_1} \cup 2^{B_2}$ that assigns a set of possible labels to each node of G. We say that a labelling $\mathcal{P}$ *covers* an edge $e = [v_1, v_2]$ if the two sets of labels $\mathcal{P}(v_1)$ and $\mathcal{P}(v_2)$ are nonempty and for every $b_2 \in \mathcal{P}(v_2)$ there exists an admissible label $b_1 \in \mathcal{P}(v_1)$ such that $(e, b_1, b_2) \in \Pi$. A labelling that covers all edges of G is called a *total cover.* Furthermore the *cost* of a cover $\mathcal{P}$ is defined as $\sum_{v \in V_1 \cup V_2} |\mathcal{P}(v)|$. The TOTAL LABEL COVER problem is then to find a minimum cost cover. The *size* of an instance of TOTAL LABEL COVER is the total length of a description of the sets G, B_1, B_2 and Π.

The only difference between TOTAL LABEL COVER and LABEL COVER is in the objective function. In TOTAL LABEL COVER the cost is equal to the total number of labels needed to label all the nodes of the bipartite graph (instead of only considering those in V_1).

Lemma 7 *For any fixed $\varepsilon > 0$, there exists a quasi–polynomial reduction*[1] *that reduces an instance ϕ of* SAT *of size s to an instance of* TOTAL LABEL COVER *of size s', with $s' \leq 2^{\text{polylog}\, s}$, such that*

- *if ϕ is satisfiable there is a total cover with cost $|V_1| + |V_2|$,*
- *if ϕ is unsatisfiable, then any labelling covering more than half of the edges has cost greater or equal to $2^{\log^{0.5-\varepsilon} s'} \cdot (|V_1| + |V_2|)$.*

PROOF By construction, the bipartite graphs derived from Feige's and Lováz's 2–prover 1–round interactive proof systems have the following properties (see [16, 33]). The graphs are regular (every node has the same degree), $|V_1| = |V_2|$ and for any e and $b_1 \in B_1$ there is at most one $b_2 \in B_2$ such that $(e, b_1, b_2) \in \Pi$. Furthermore, we may assume that the relation Π is restricted to have only valid labels for any node in V_1, where b_1 is *valid* for $v_1 \in V_1$ if, for every

[1] A reduction is said quasi–polynomial if it has an $O(s^{\text{polylog}\, s})$ time–complexity.

edge e incident to v_1, there is a label b_2 such that $(e, b_1, b_2) \in \Pi$. In fact, the total label covers corresponding to any satisfiable ϕ use only one label per node of the graph. Thus one easily verify that the gap for Label Cover implies an equivalent gap for Total Label Cover. $\square$

Theorem 8 Min RVLS$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq, >\}$ *cannot be approximated within a factor of* $2^{\log^{0.5-\varepsilon} n}$ *for any* $\varepsilon > 0$ *unless* $\mathrm{NP} \subseteq \mathrm{DTIME}(n^{\mathrm{polylog}\, n})$*, where* n *is the number of variables.*

Proof For any Total Label Cover instance (G, B_1, B_2, Π) arising from a Feige-Lováz interactive proof system, one can construct a particular instance of Min RVLS$^=$ with exactly the same objective function. Therefore we can use the first part of the construction described in [5].

The idea is to introduce one variable for each pair (v_i, b_i) where $v_i \in V_i$ and $b_i \in B_i$, $i = 1, 2$, which satisfy the following conditions. While every pair (v_2, b_2) with $v_2 \in V_2$ and $b_2 \in B_2$ is considered, only the pairs (v_1, b_1) where $v_1 \in V_1$, $b_1 \in B_1$ and v_1 is valid are retained. For each pair e, $b_1 \in B_1$ the unique label $b_2 \in B_2$ such that $(e, b_1, b_2) \in \Pi$ is denoted by $b_2[e, b_1]$. We construct an inhomogeneous system

$$A\mathbf{x} = \mathbf{1} \tag{5}$$

with $|E|(|B_2| + 1)$ equations, namely $|B_2| + 1$ for each edge $e \in E$. For each column of A, the components corresponding to a given e are called its *e–projection*. The columns are defined as follows. For each pair $v_2 \in V_2$, $b_2 \in B_2$ the e–projection is $\mathbf{u}_j$ (the $|B_2| + 1$–dimensional vector in which the jth component is 1 and all the others 0) and $\mathbf{0}$ otherwise, where $\mathbf{0}$ denotes the trivial vector with all zero components. For each valid pair $v_1 \in V_1$, $b_1 \in B_1$ the e–projection is $\mathbf{1} - \mathbf{u}_{b_2[e,b_1]}$ if e is incident to v_1 and $\mathbf{0}$ otherwise. Thus we have a linear system with $|E|(|B_2|+1)$ equations and at most $|V_1| \cdot |B_1| + |V_2| \cdot |B_2|$ variables.

It is easily verified that the nonzero variables of any solution $\mathbf{x}$ of (5) define a total cover of the original bipartite graph G. Thus, according to lemma 7, Min RVLS$^=$ cannot be approximated within the given factor unless $\mathrm{NP} \subseteq \mathrm{DTIME}(n^{\mathrm{polylog}\, n})$.

By inspecting the construction, one observes that the same bound holds even for homogeneous systems with binary coefficients. Since any equation can be replaced by two nonstrict inequalities, the result is also valid for Min RVLS$^{\geq}$. Using the equivalence between systems with nonstrict and strict inequalities, we obtain the same bound for Min RVLS$^{>}$. $\square$

This nonapproximability result can be still strengthened.

Theorem 9 Min RVLS$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq, >\}$ *is not approximable in polynomial time within a factor of* $2^{\log^{1-\varepsilon} n}$ *for any* $\varepsilon > 0$ *unless* $\mathrm{NP} \subseteq \mathrm{DTIME}(n^{\mathrm{polylog}\, n})$*, where* n *is the number of variables.*

Proof We proceed by self–improvement as in [6]. The idea is to consider the reduction in proposition 6 from set cover to Min RVLS with a fixed constant gap $c > 1$ between the satisfiable and unsatisfiable cases and to increase it recursively. For any particular set covering instance, we start with the corresponding system (4) whose coefficients are 0 or 1 and we construct the squared system

$$A'\mathbf{x} = \mathbf{1}' \tag{6}$$

obtained by replacing each 1 coefficient of A by the whole matrix A and each 0 coefficient by the $p \times n$ matrix with all 0. Thus A' is a matrix of size $p^2 \times n^2$ and $\mathbf{1}'$ is a vector with p^2 components all equal to 1. Clearly, if the corresponding instance ϕ of the Sat is satisfiable there is a solution $\mathbf{x}$ of (6) with K^2 variables equal to 1 and all the other to 0. On the contrary, if ϕ is unsatisfiable any solution $\mathbf{x}$ that satisfies all equations of (6) has at least $\lfloor c^2 K^2 \rfloor$ nonzero variables.

By applying this construction t times recursively, we obtain a system with p^{2^t} equations and n^{2^t} variables. Let us take $t = \lfloor \log(\log^\beta n) \rfloor$, where n is the number of variables in the Min RVLS$^=$ instance corresponding to the considered Sat instance ϕ, and β is a positive real number. The construction requires $O(n^{\text{polylog}\, n})$ time because the system has $p' = p^{O(\log^\beta n)}$ equations and $n' = n^{O(\log^\beta n)} = 2^{O(\log^{\beta+1} n)}$ variables. Since $\log n' = O(\log^{\beta+1} n)$, we obtain a gap of $c^{2^t} = c^{\log^\beta n} = 2^{O(\log n')^{\beta/(\beta+1)}}$.

The theorem follows by contradiction. Suppose there exists a polynomial time algorithm that approximates Min RVLS$^=$ instances with n variables within a factor $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 1/(\beta+1)$. By applying it to the resulting instance of Min RVLS$^=$, one could decide in $O(n^{\text{polylog}\, n})$ time whether any given instance ϕ of Sat is satisfiable. But this would imply $\text{NP} \subseteq \text{DTIME}(n^{\text{polylog}\, n})$.

The same bound is also valid for strict and nonstrict inequalities. □

This result for Min RVLS$^=$ immediately implies the nonapproximability of Min ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$ within the same factor.

In fact, there is a reduction from homogeneous Min ULR$^{\mathcal{R}}$ to Min RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq\}$.

Proposition 10 Min ULR$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq\}$ *is equally hard to approximate as* Min RVLS$^{\mathcal{R}}$ *with the same type of relations.*

Proof Min RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq\}$ can be reduced to Min ULR$^{\mathcal{R}}$ by proposition 5, and Min ULR$^{\mathcal{R}}$ can be transformed to homogeneous Min ULR$^{\mathcal{R}}$ as shown in the proof of theorem 4. Thus it suffices to show that homogeneous Min ULR$^{\mathcal{R}}$ can be reduced to Min RVLS$^{\mathcal{R}}$ for $\mathcal{R} \in \{=, \geq\}$.

Let $\sum_{j=1}^{n} a_{ij} x_j = 0$ with $1 \leq i \leq p$ be an instance of homogeneous Min ULR$^=$. We construct an instance of Min RVLS$^=$ consisting of $p(n+1)$ equations of the following type:

$$\sum_{j=1}^{n} a_{ij} x_j = y_{ik}$$

where $1 \leq i \leq p$, $1 \leq k \leq n+1$ and y_{ik} are $p(n+1)$ new variables.

If s equations of $A\mathbf{x} = \mathbf{0}$ are satisfied, then we will get a solution of Min RVLS$^=$ with between $s(n+1)$ and $s(n+1)+n$ variables equal to zero. Conversely, a solution of Min RVLS$^=$ with between $s(n+1)$ and $s(n+1)+n$ variables equal to zero, will give us a solution of Min ULR$^=$ with s satisfied equations. The reduction is an S-reduction with size amplification $O(pn)$.

The same construction can be used for Min RVLS$^{\geq}$, just by substituting $=$ with $\geq$. □

Note that due to the size amplification the nonapproximability bound for Min ULR$^=$ established in [5] leads to a smaller bound for Min RVLS$^=$ than that of theorems 8 and 9.

Using the same reduction, the complementary maximization problem Max IVLS$^=$, Maximum number of Irrelevant Variables in Linear Systems, can be shown in its homogeneous version to be equally hard to approximate as homogeneous Max FLS$^=$, i.e. not approximable within p^ε for some $\varepsilon > 0$ unless P=NP [3].

Interestingly, Max IVLS$^{\geq}$ and Max IVLS$^{>}$ are much harder to approximate than Max FLS$^{\geq}$ and Max FLS$^{>}$, respectively. It is easily proved that the former problems are harder than the maximum independent set problem (which is not approximable within n^ε for some $\varepsilon > 0$ unless P=NP, where n is the number of nodes), while the latter ones can be approximated within 2 [3]. It suffices to construct, for each edge $e = [v_i, v_j]$, the inequality $x_i + x_j \geq 1$ or $x_i + x_j > 0$ and to observe that there is a correspondence between the independent sets of cardinality at least s and the solutions with at least s zero components.

5 Approximability of weighted and constrained Min ULR

In many practical situations different relations may have different importances. This can be modeled by assigning a weight to each relation and by seeking a solution that minimizes the total weight of the unsatisfied relations.

Interesting special cases of weighted Min ULR include the constrained versions where some relations are mandatory while the others are optional. C Min ULR$^{\mathcal{R}_1;\mathcal{R}_2}$ with $\mathcal{R}_1, \mathcal{R}_2 \in \{=, \geq, >, \neq\}$ denotes the variant where the mandatory relations are of type $\mathcal{R}_1$ and the optional ones of type $\mathcal{R}_2$. When $\mathcal{R}_1 = \mathcal{R}_2$ the problem can be seen as a weighted Min ULR$^{\mathcal{R}_1}$ problem in which the weight of every mandatory relation is larger than the total weight of all optional ones.

Proposition 11 *Weighted* Min ULR$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq, >\}$ *and positive integer weights, that are bounded in size by a polynomial in the size of the whole input, are equally hard to approximate as the basic versions.*

Proof Basic Min ULR$^{\mathcal{R}}$ is clearly a special case of weighted Min ULR$^{\mathcal{R}}$ where all weights are equal to one. For proving the other direction, it suffices to note that any weighted instance can be associated to an equivalent unweighted one. This is simply achieved by making for each relation a number of copies equal to the corresponding weight. The number of relations will still be polynomial since the weights are polynomially bounded. □

This equivalence between weighted and unweighted variants holds also for the constrained versions where the optional relations and mandatory relations are of the same type. To make sure that the mandatory relations are satisfied, a large enough number of equivalent relations is required, namely more than the number of optional ones. Thus the approximability of constrained Min ULR can be very well characterized using the unconstrained versions.

It is worth noting that no such relation exists between weighted and unweighted versions of the complementary problems Max FLS. As we proved in [3], just enforcing some mandatory relations makes Max FLS with inequalities harder to approximate. While Max FLS$^{>}$ and Max FLS$^{\geq}$ can be approximated within a factor 2, the constrained variants are at least as hard as the maximum independent set problem and hence cannot be approximated within a factor n^{ε} for some $\varepsilon > 0$, where n is the instance size.

Any instance of a constrained problem C Min ULR$^{=;\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$ can be transformed into an equivalent instance of Min ULR$^{\mathcal{R}}$ by eliminating variables in the set of optional relations using the set of mandatory ones.

All the problems C Min ULR$^{\mathcal{R};\neq}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$ are trivial for the same reason as Min ULR$^{\neq}$, see section 3.

Proposition 12 C Min ULR$^{\geq;=}$ *is equally hard to approximate as* Min ULR$^{\geq}$.

Proof Proposition 2 tells us that Min ULR$^{\geq}$ can be reduced to Min ULR$^{=}$ with nonnegative variables. This can be expressed as a C Min ULR$^{\geq;=}$ problem where the mandatory inequalities ensure that all variables are nonnegative.

Conversely, C Min ULR$^{\geq;=}$ can be reduced to C Min ULR$^{\geq;\geq}$ by substituting each optional equation by two inequalities as in the proof of proposition 2. We can then use the standard reduction from C Min ULR$^{\geq;\geq}$ to Min ULR$^{\geq}$. □

Similarly we can show that C Min ULR$^{>;=}$ is at least as hard to approximate as Min ULR$^{>}$.

Proposition 13 C Min ULR$^{\neq;=}$ *is* Min Dominating Set*-hard. Therefore it cannot be approximated within any constant unless* P=NP *and not within* $c \log n$ *for any* $c < 1/8$ *unless* NP $\subseteq$ DTIME$(n^{\log\log n})$*, even when restricted to homogeneous systems with binary coefficients in* $\{0, 1\}$.

PROOF The proof is by reduction from MIN DOMINATING SET. Let $G = (V, E)$ be an arbitrary instance of MIN DOMINATING SET. For each node $v_i \in V$, $1 \le i \le n$, we consider the optional equation

$$x_i = 0 \tag{7}$$

and the mandatory inhomogeneous relation

$$x_i + \sum_{j \in N(v_i)} x_j \neq 0 \tag{8}$$

where j is included in $N(v_i)$ if and only if v_j is adjacent to v_i.

We claim that there exists a dominating set in G of size at most s if and only if there exists a solution that violates at most s relations of the corresponding linear system.

Given a dominating set $V' \subseteq V$ of size s, the solution $\mathbf{x}$ defined by

$$x_i = \begin{cases} 1 & \text{if } v_i \in V' \\ 0 & \text{otherwise} \end{cases}$$

satisfies all mandatory relations and all but s of the optional equations.

Conversely, given a solution vector $\mathbf{x}$ that satisfies all mandatory relations and violates s of the optional relations, we consider the set of nodes $V' \subseteq V$ containing all nodes v_i such that $x_i \neq 0$. V' is clearly a dominating set of size s, because $x_i + \sum_{j \in N(v_i)} x_j \neq 0$ only when at least one of the variables is nonzero, which corresponds to the case where at least one of the nodes is in the dominating set. $\Box$

6 Hardness of approximation for bounded discrete variables

In this section we determine the approximability of MIN ULR$^{\mathcal{R}}$ and MIN RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \ge, >, \neq\}$ when the variables are restricted to take a finite number of discrete values. In particular, we consider the cases with binary variables in $\{0, 1\}$ and bipolar ones in $\{-1, 1\}$. The corresponding variants of MIN ULR and MIN RVLS are referred to as BIN MIN ULR, BIP MIN ULR, BIN MIN RVLS and BIP MIN RVLS respectively.

Theorem 14 *The following results are valid for every combination of $\mathcal{R}_1, \mathcal{R}_2 \in \{=, \ge, >, \neq\}$. BIN MIN ULR$^{\mathcal{R}_1}$ and C BIN MIN ULR$^{\mathcal{R}_1;\mathcal{R}_2}$ are NPO PB-complete. Assuming P$\neq$NP, C BIN MIN ULR$^{\mathcal{R}_1;\mathcal{R}_2}$ cannot be approximated within $s^{1-\varepsilon}$ for any $\varepsilon > 0$ and BIN MIN ULR$^{\mathcal{R}_1}$ cannot be approximated within $s^{0.5-\varepsilon}$ for any $\varepsilon > 0$, where $s = \max\{n, p\}$ i.e. the maximum of the number of variables and the number of relations.*

PROOF We first show the result for the problem C BIN MIN ULR$^{\ge;\ge}$ and then extend the result to the other variants.

We proceed by reduction from MIN IND DOM SET which is defined as follows [19]. Given an undirected graph $G = (V, E)$, find a minimum cardinality independent set $V' \subseteq V$ that dominates all nodes of G.

Let $G = (V, E)$ be an arbitrary instance of MIN IND DOM SET. For each node $v_i \in V$, $1 \le i \le n$, we consider the optional inequality

$$x_i \le 0 \tag{9}$$

and the inhomogeneous mandatory inequality

$$x_i + \sum_{j \in N(v_i)} x_j \ge 1 \tag{10}$$

where j is included in $N(v_i)$ if and only if v_j is adjacent to v_i. Furthermore, we construct for each edge $[v_i, v_j] \in E$ the inhomogeneous mandatory inequality

$$x_i + x_j \leq 1. \tag{11}$$

Thus we have a system with n variables, n optional inequalities and $n + |E|$ mandatory ones.

We claim that there exists an independent dominating set in G of size at most s if and only if there exists a solution $\mathbf{x} \in \{0,1\}^n$ that violates s relations of the corresponding linear system. Given an independent dominating set $V' \subseteq V$ of size s, the solution $\mathbf{x}$ defined by

$$x_i = \begin{cases} 1 & \text{if } v_i \in V' \\ 0 & \text{otherwise} \end{cases}$$

satisfies all mandatory inequalities and $n - s$ of the optional ones. Conversely, given a solution vector $\mathbf{x}$ that violates s of the relations, we consider the set of nodes $V' \subseteq V$ containing all nodes v_i such that $x_i = 1$. V' is clearly of size s, independent (because of the mandatory relations (11)) and dominating (because of the mandatory relations (10)).

The theorem follows because Min Ind Dom Set is NPO PB-complete and cannot be approximated within $n^{1-\varepsilon}$ for any $\varepsilon > 0$, where n is the sum of the number of nodes and edges in the graph.

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

	type $\geq$	type $>$	type $\neq$	type $=$
type (9)	≤ 0	< 1	$\neq 1$	$= 0$
type (10)	≥ 1	> 0	$\neq 0$	$= \lvert N(v_i)\rvert - \sum_{j=1}^{\lvert N(v_i)\rvert-1} y_{ij}$
type (11)	≤ 1	< 2	$\neq 2$	$= 1 - z_{ij}$

In the case of mandatory equations we need to introduce $2|E| - n$ additional slack variables y_{ij} and $|E|$ additional slack variables z_{ij}. Thus the total number of variables will be $3|E|$, that is, still a linear number in n and $|E|$.

To get the hardness results for the unconstrained problems, we add $|V| + 1$ copies of each mandatory relation so that there are more valuable than the optional ones. Since such a reduction has a quadratic size amplification, we get a weaker nonapproximability bound than for Min Ind Dom Set. $\square$

It is worth noting that Bin Min ULR$^{\mathcal{R}_1}$ and C Bin Min ULR$^{\mathcal{R}_1;\mathcal{R}_2}$ with $\mathcal{R}_1, \mathcal{R}_2 \in \{=, \geq, >, \neq\}$ remain NPO PB-complete for homogeneous systems. In the above reduction from Min Ind Dom Set, we multiply each nonzero constant in the right hand side of a relation by a new variable x_0 which should be viewed as the constant 1. In order to prevent x_0 from being zero we introduce new mandatory relations $x_0 \geq 0$, $x_0 > 0$, $x_0 \neq 0$, or $x_0 = y_0$ with a new variable y_0, depending on the type of relations. In the case of nonstrict inequalities and equalities we introduce (as in the proof of theorem 4) a lot of copies of the relations in order to make x_0 the variable occurring the most frequently in the satisfied relations.

Corollary 15 *The following results are valid for every combination of $\mathcal{R}_1, \mathcal{R}_2 \in \{=, \geq, >, \neq\}$.* Bip Min ULR$^{\mathcal{R}_1}$ *and* C Bip Min ULR$^{\mathcal{R}_1;\mathcal{R}_2}$ *are* NPO PB-*complete. Assuming* P$\neq$NP, C Bip Min ULR$^{\mathcal{R}_1;\mathcal{R}_2}$ *cannot be approximated within $s^{1-\varepsilon}$ for any $\varepsilon > 0$ and* Bip Min ULR$^{\mathcal{R}_1}$ *cannot be approximated within $s^{1-\varepsilon}$ for any $\varepsilon > 0$, where $s = \max\{n, p\}$.*

PROOF By cost preserving transformation from BIN MIN ULR$^{\mathcal{R}_1}$ and C BIN MIN ULR$^{\mathcal{R}_1;\mathcal{R}_2}$. For any relation

$$\sum_{j=1}^{n} a_{ij} x_j \ \mathcal{R}\ b_i$$

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

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

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

MIN RVLS restricted to binary or bipolar variables turns out to be also very hard to approximate.

Proposition 16 BIN MIN RVLS$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=,\geq,>\}$ *is* NPO PB*-complete. Assuming* P$\neq$NP *neither* BIN MIN RVLS$^{\geq}$ *nor* BIN MIN RVLS$^{>}$ *can be approximated within* $n^{1-\varepsilon}$ *for any* $\varepsilon > 0$*, and* BIN MIN RVLS$^{=}$ *cannot be approximated within* $n^{0.5-\varepsilon}$ *for any* $\varepsilon > 0$*, where* n *is the number of variables.*

PROOF As before, we proceed by reduction from MIN IND DOM SET. The reduction is very similar to the one used in theorem 14 for proving the NPO PB-hardness of C BIN MIN ULR$^{\mathcal{R};\mathcal{R}}$ with $\mathcal{R} \in \{=,\geq,>\}$. The BIN MIN RVLS$^{\mathcal{R}}$ instance is simply composed of the mandatory relations constructed in the C BIN MIN ULR$^{\mathcal{R};\mathcal{R}}$ instance. The number of violated optional relations in the proof exactly corresponds to the number of nonzero variables. This implies that BIN MIN RVLS$^{\geq}$ and BIN MIN RVLS$^{>}$ are NPO PB-hard.

For BIN MIN RVLS$^{=}$, we still have to deal with the slack variables y_{ij} and z_{ij} that have been added. Suppose there is a total number of N slack variables.

In order to make the value of the x variables more important than the values of the N slack variables altogether, we introduce, for each variable x_i, N new variables $x_{i1},\ldots,x_{iN}$ and the N additional equations $x_i - x_{ij} = 0$ for $j \in [1..N]$. In any solution $\mathbf{x}$ of the resulting instance we will have, for each variable x_i, that $x_i = x_{i1} = \ldots = x_{iN}$. Consider the set of nodes $V' \subseteq V$ containing all nodes v_i such that $x_i = 1$. V' is clearly independent and dominating. If t variables in $\mathbf{x}$ are equal to 1, the size of the independent set will be $\lfloor t/(N+1) \rfloor$.

Conversely, an independent dominating set containing s nodes corresponds to a solution of the BIN MIN RVLS$^{=}$ instance with between $s(N+1)$ and $s(N+1)+N$ variables equal to 1. Thus the reduction is an S-reduction with size amplification $O(nN)$ and we get the nonapproximability bound $n^{0.5-\varepsilon}$ for any $\varepsilon > 0$, where n is the number of variables. $\square$

Using the cost preserving transformation in the proof of corollary 15 one can show the same nonapproximability results for BIP MIN RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=,\geq,>\}$.

Note that the BIN MIN RVLS$^{\geq}$ problem is exactly the same problem as MIN POLYNOMIALLY BOUNDED 0-1 PROGRAMMING, which was shown to be NPO PB-complete in [28]. Moreover, the corresponding maximization problem BIN MAX IVLS$^{\geq}$ is the same problem as MAX POLYNOMIALLY BOUNDED 0-1 PROGRAMMING, which is NPO PB-complete [10] and cannot be approximated within $n^{0.5-\varepsilon}$ for any $\varepsilon > 0$, where n is the number of variables, unless P=NP [13].

7 Relation between MIN ULR with real and binary variables

The question naturally arises as to whether the strong nonapproximability results for the binary versions of MIN ULR have any consequences for the basic versions. The answer is positive if we slightly modify the definition of the objective function by introducing a *threshold*. Indeed, we obtain a set of problems that are at least as hard as the binary versions.

The threshold variant of MIN ULR, named THRESHOLD MIN ULR, with any type of relations is defined as follows. The input consists of a linear system of relations together with a threshold T, which is a lower bound on the minimum number of unsatisfiable relations of the system. Given a solution vector $\mathbf{x}$ that violates s relations, the objective function is defined as $s - T$. Since at least T relations are violated, this value is always nonnegative.

Theorem 17 THRESHOLD MIN ULR$^{\mathcal{R}}$ *with* $\mathcal{R} \in \{=, \geq, >\}$ *is* NPO PB-*complete. Assuming* P$\neq$NP, *threshold* MIN ULR$^{\mathcal{R}}$ *cannot be approximated within* $n^{1-\varepsilon}$ *for any* $\varepsilon > 0$, *where* n *is the number of variables, and not within* $p^{0.5-\varepsilon}$ *for any* $\varepsilon > 0$, *where* p *is the number of relations.*

PROOF We proceed by reduction from BIN MIN ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$.

Given an input instance of BIN MIN ULR$^{\mathcal{R}}$ with n binary variables $x_1, \ldots, x_n$ and p relations, we construct an instance of THRESHOLD MIN ULR$^{\mathcal{R}}$ by extending the system with relations which ensure that the variables just take zero or one values.

In the equality case we add, for every variable x_i, $p+1$ copies of the equations $x_i = 0$ and $x_i = 1$. If all variables in a solution vector take values in $\{0, 1\}$, exactly $n(p+1)$ of the added equations will be satisfied and $n(p+1)$ violated. For each variable that is not either zero or one, $p+1$ additional equations will be violated, which are more than the number of equations in the original system. If we choose $n(p+1)$ as the threshold we will get an S-reduction from BIN MIN ULR$^{=}$ to threshold MIN ULR$^{=}$ without variable amplification and with a relation amplification of $O(np)$. In the proof of the nonapproximability bound of BIN MIN ULR$^{\mathcal{R}}$ in theorem 14, the number of relations is about the same as the number of variables, and consequently THRESHOLD MIN ULR$^{=}$ cannot be approximated within $p^{0.5-\varepsilon}$ for any $\varepsilon > 0$, where p is the number of relations.

For threshold MIN ULR$^{\geq}$ we use the same construction, but instead of the two equations $x_i = 0$ and $x_i = 1$ we include the four inequalities $x_i \leq 0$, $x_i \geq 0$, $x_i \leq 1$, and $x_i \geq 1$. If all components of a solution vector are either zero or one, just one of these four inequalities will be violated. Otherwise two of the inequalities will be violated. The threshold and the rest of the proof is the same as for the equality case.

Finally, for threshold MIN ULR$^{>}$ we just use the equivalence between MIN ULR$^{\geq}$ and MIN ULR$^{>}$ described in proposition 3. □

8 Two special cases of MIN ULR and MIN RVLS

In this section we discuss two special cases of MIN ULR and MIN RVLS with inequalities which arise when designing two-class linear classifiers and when training perceptrons.

Given a set of vectors $S = \{a^k\}_{1 \leq k \leq p} \subset \mathbb{R}^n$ labelled as positive or negative examples, we look for a hyperplane H, specified by a normal vector $\mathbf{x} \in \mathbb{R}^n$ and a bias x_0, such that all the positive vectors lie on the positive side of H while all the negative ones on the negative side. In other words, we seek a discriminant hyperplane H which separates the examples in the first class from those in the second class. This problem is referred to as MIN H-MISCLASSIFICATIONS.

A hyperplane H is said to be *consistent* with an example $\mathbf{a}^k$ if $\mathbf{a}^k\mathbf{x} > x_0$ or $\mathbf{a}^k\mathbf{x} \leq x_0$ depending on whether $\mathbf{a}^k$ is positive or negative. In the general case where S is nonlinearly separable, a natural objective is to minimize the number of input vectors that are misclassified[2] (see [31, 17] and the included references).

[2]The approximability of the complementary problem where one looks for a hyperplane which is consistent with as many $\mathbf{a}^k \in S$ as possible has been studied in [3].

Note that the symmetric variant of MIN H-MISCLASSIFICATIONS is also frequently considered, where we ask for $\mathbf{a}^k\mathbf{x} > x_0$ for positive $\mathbf{a}^k$ and $\mathbf{a}^k\mathbf{x} < x_0$ for negative ones.

In [5] a way of extending the nonapproximability bounds for MIN ULR$^=$ to the symmetric version of MIN H-MISCLASSIFICATIONS is suggested. Although the argument used does not suffice to complete the proof, it can easily be fixed.

The problem is related to the fact that starting with any instance of MIN ULR$^=$ we must construct a system with strict inequalities with a particular variable playing the role of the threshold x_0. As mentioned in [5], one can easily associate to any considered instance of MIN ULR$^=$ an equivalent inhomogeneous instance of MIN ULR$^>$. Therefore each equation is replaced by two nonstrict inequalities, and then a new slack variable δ is added to change each nonstrict inequality into a strict one. More precisely, every relation $\mathbf{ax} \geq 0$ is replaced by $\mathbf{ax}+\delta > 0$. Now, in order to make sure that the two systems are equivalent we must have $\delta < 1/L$, where L is a given positive integer. This can of course be guaranteed by introducing a large enough number of copies of this strict inequality, but then the resulting system is not an instance of symmetric MIN H-MISCLASSIFICATIONS. Indeed, if δ is considered as the threshold the inequalities ensuring $\delta < 1/L$ are not homogeneous.

Fortunately, there exists a simple and general technique to construct, for any instance of inhomogeneous MIN ULR$^>$, an equivalent instance of symmetric MIN H-MISCLASSIFICATIONS.

Observation 18 Suppose we have a system $\mathbf{a}^k\mathbf{x} > b^k$ with $1 \leq k \leq p$ where all b^k are nonzero. Multiply each inequality by an appropriate constant so that all right hand sides are equal to 1. By replacing all right hand sides constants 1 by a variable x_0, we get a system with either $\mathbf{a}^k\mathbf{x} > x_0$ type or $\mathbf{a}^k\mathbf{x} < x_0$ type inequalities. Clearly, any solution of this new system such that $x_0 > 0$ gives a solution of the original system. Thus by adding a large enough number of copies of $x_0 > 0$ the two problems are guaranteed to be equivalent.

In order to complete the reduction in [5], we just apply this technique to the system consisting of $\mathbf{ax} + \delta > 1/(2L)$ inequalities and enough copies of $\delta < 1/L$.

It is worth noting that the same argument can be used to show that (nonsymmetric) MIN H-MISCLASSIFICATIONS cannot be approximated within $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$ unless NP $\subseteq$ DTIME($n^{\text{polylog}\, n}$).

A special case of MIN RVLS$^>$ and MIN RVLS$^\geq$ is also of particular interest in discriminant analysis and machine learning. The problem occurs when, given a linearly separable set S of positive and negative examples, one wants to minimize the number of parameters that are required to correctly classify all examples in S [30, 38]. In other words, we look for a solution vector $\mathbf{x}$ with as few nonzero components as possible.

This objective, which is related to the concept of parsimony, is crucial because the number of nonzero parameters has a strong impact on the performances of the classifier for unseen data. In [38] the symmetric variant with strict inequalities was proved to be at least as hard to approximate as MIN DOMINATING SET. Moreover, it was shown that an approximate algorithm minimizing the number of nonzero parameters within a factor of $O(\log p)$, where p is the number of examples, would require far fewer examples to achieve a given level of accuracy than any algorithm which does not minimize this quantity. Finally, it was left as an open question whether this number could be approximated within a factor of $O(\log p)$.

We will show elsewhere that the nonapproximability bounds for MIN RVLS$^\geq$ established in section 4 lead to a negative answer to this question unless NP $\subseteq$ DTIME($p^{\text{polylog}\, p}$). The consequences of this result on the hardness of designing compact feedforward neural networks will be also discussed.

9 Conclusions

The various versions of Min ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >, \neq\}$ that we have considered are obtained by placing constraints on the coefficients, on the variables and by assigning a possibly different importance to each relation.

Table 1 summarizes the nonapproximability results that hold for Min ULR variants unless P=NP. As seen in the table, the approximability of similar variants of Min ULR differs

	real variables	binary variables
Min ULR$^{=}$	not within any constant [5]	NPO PB-complete
Min ULR$^{\geq}$	Min Dominating Set-hard	
Min ULR$^{>}$		
Min ULR$^{\neq}$	trivial	
C Min ULR$^{\geq;=}$	as hard as Min ULR$^{\geq}$	
C Min ULR$^{>;=}$	harder than Min ULR$^{>}$	
C Min ULR$^{\neq;=}$	Min Dominating Set-hard	
C Min ULR$^{\geq;\geq}$	as hard as Min ULR$^{\geq}$	
C Min ULR$^{>;\geq}$	harder than Min ULR$^{\geq}$	
C Min ULR$^{\neq;\geq}$	harder than Min ULR$^{\geq}$	
C Min ULR$^{\geq;>}$	harder than Min ULR$^{>}$	
C Min ULR$^{>;>}$	as hard as Min ULR$^{>}$	
C Min ULR$^{\neq;>}$	harder than Min ULR$^{>}$	
C Min ULR$^{\geq;\neq}$	trivial	
C Min ULR$^{>;\neq}$		
C Min ULR$^{\neq;\neq}$		

Table 1: Main approximability results for Min ULR variants which hold provided P$\neq$NP. The constrained versions C Min ULR$^{=;\mathcal{R}}$ with mandatory equations and $\mathcal{R} \in \{=, \geq, >, \neq\}$ have the same approximability as Min ULR$^{\mathcal{R}}$.

depending on the type of relations. The results are valid for inhomogeneous systems with integer coefficients and no pairs of identical relations, and some of them are still valid for homogeneous systems with ternary, and even binary, coefficients. In order to avoid trivial solutions for relations that are equalities or nonstrict inequalities, we required the variables occurring most frequently in the satisfied relations to be nonzero.

Min ULR$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, >\}$ have recently been shown not to be approximable within any constant unless P=NP and not within $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$ unless NP $\subseteq$ DTIME$(n^{\text{polylog}\, n})$. Using a simple reduction from Min Dominating Set, we obtained a weaker but more likely logarithmic bound for Min ULR$^{\geq}$ as well as Min ULR$^{>}$. Moreover, we got for free the constant factors results.

In the process of studying Min ULR, we proved that Min RVLS$^{\mathcal{R}}$ with $\mathcal{R} \in \{=, \geq, >\}$ cannot be approximated within a constant factor and within $2^{\log^{1-\varepsilon} n}$ under the usual assumptions. These nonapproximability results, which are of interest by their own, directly implies the bounds established in [5] for the basic versions of Min ULR. Interestingly both problems with equations or nonstrict inequalities are equivalent to approximate.

The weighted variants of Min ULR are not harder to approximate than the unweighted ones if the weights are polynomially bounded. Restricting the variables to binary (or bipolar) values makes these problems harder to approximate, namely NPO PB-complete. As to Min ULR$^{\neq}$, while the basic version is trivial, some of the constrained variants are hard to approximate, and all binary versions even turn out to be NPO PB-complete. The nonapproximability bounds such as $n^{1-\varepsilon}$ for any $\varepsilon > 0$ makes the existence of any nontrivial approximation algorithm extremely unlikely.

It is worth noting that the overall situation for MIN ULR differs considerably from that for the complementary MAX FLS problems [3]. Here all basic versions can be approximated within a factor $n+1$, where n is the number of variables, while MAX FLS$^=$ cannot be approximated within p^ε for some $\varepsilon > 0$, where p is the number of equations.

Some interesting questions related to the approximability of MIN ULR$^\mathcal{R}$ with $\mathcal{R} \in \{=, \geq, >\}$ are still open. Assuming P$\neq$NP, can we guarantee a logarithmic performance ratio $c \log n$ for some $c \geq 1/8$? Do these problems admit polynomial time algorithms providing approximate solutions within $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$?

References

[1] E. Amaldi. On the complexity of training perceptrons. In T. Kohonen et al., editor, *Artificial Neural Networks*, pages 55–60, Amsterdam, 1991. Elsevier science publishing company.

[2] E. Amaldi and C. G. Diderich. On the probabilistic and thermal perceptron training algorithms. Technical Report ORWP-15-93, Department of Mathematics, Swiss Federal Institute of Technology, Lausanne, 1993.

[3] E. Amaldi and V. Kann. The complexity and approximability of finding maximum feasible subsystems of linear relations. Technical Report ORWP-11-93, Department of Mathematics, Swiss Federal Institute of Technology, Lausanne and Technical Report TRITA-NA-9313, ISSN 0348-2952, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1993. To be published in Theoretical Computer Science.

[4] E. Amaldi and V. Kann. On the approximability of finding maximum feasible subsystems of linear systems. In *Proc. 11th Ann. Symp. on Theoretical Aspects of Comput. Sci.*, pages 521–532. Springer-Verlag, 1994. Lecture Notes in Comput. Sci. 775.

[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 34rd Ann. IEEE Symp. on Foundations of Comput. Sci.*, pages 724–733, 1993.

[6] S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. Manuscript, 1994.

[7] 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.*, pages 14–23, 1992.

[8] 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.*, pages 294–304. ACM, 1993.

[9] K. P. Bennett and O. L. Mangasarian. Robust linear programming discrimination of two linearly inseparable sets. *Optimization Methods and Software*, 1:23–34, 1992.

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

[11] J. Chinneck and E. Dravnieks. Locating minimal infeasible constraint sets in linear programs. *ORSA Journal on Computing*, 3:157–168, 1991.

[12] V. Chvatal. *Linear programming.* Freeman and Co., New York, 1983.

[13] P. Crescenzi, V. Kann, and L. Trevisan. On the structure of the classes NPO and APX. Manuscript in preparation, 1994.

[14] P. Crescenzi and A. Panconesi. Completeness in approximation classes. *Inform. and Comput.*, 93(2):241–262, 1991.

[15] R. O. Duda and P. E. Hart. *Pattern classification and scene analysis.* Wiley, New York, 1973.

[16] U. Feige and L. Lovász. Two–prover one round proof systems: Their power and their problems. In *Proc. Twenty fourth Ann. ACM Symp. on Theory of Comp.*, pages 733–744, 1992.

[17] M. R. Frean. A thermal perceptron learning rule. *Neural Computation*, 4:946–957, 1992.

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

[19] 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.

[20] J. Gleeson and J. Ryan. Identifying minimally infeasible subsystems of inequalities. *ORSA Journal on Computing*, 3:61–63, 1990.

[21] H. J. Greenberg. Consistency, redundancy, and implied equalities in linear systems. Manuscript, November 1993.

[22] H. J. Greenberg and F. H. Murphy. Approaches to diagnosing infeasible linear programs. *ORSA Journal on Computing*, 3:253–261, 1991.

[23] R. Greer. *Trees and Hills: Methodology for Maximizing Functions of Systems of Linear Relations*, volume 22 of *Annals of Discrete Mathematics.* Elsevier science publishing company, Amsterdam, 1984.

[24] M. M. Halldórsson. Approximating the minimum maximal independence number. *Inform. Process. Lett.*, 46:169–172, 1993.

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

[26] D. S. Johnson and F. P. Preparata. The densest hemisphere problem. *Theoretical Computer Science*, 6:93–107, 1978.

[27] V. Kann. *On the Approximability of NP-complete Optimization Problems.* PhD thesis, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1992.

[28] V. Kann. Polynomially bounded minimization problems which are hard to approximate. In *Proc. of 20th International Colloquium on Automata, Languages and Programming*, pages 52–63. Springer-Verlag, 1993. Lecture Notes in Comput. Sci. 700.

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

[30] G. J. Koehler. Linear discriminant functions determined by genetic search. *ORSA Journal on Computing*, 3:345–357, 1991.

[31] G. J. Koehler and S. S. Erenguc. Minimizing misclassifications in linear discriminant analysis. *Decision Sciences*, 21:63–85, 1990.

[32] P. G. Kolaitis and M. N. Thakur. Logical definability of NP optimization problems. Technical Report UCSC-CRL-93-10, Board of Studies in Computer and Information Sciences, University of California at Santa Cruz, 1993. To be published in Information and Computation.

[33] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. In *Proc. Twenty fifth Ann. ACM Symp. on Theory of Comp.*, pages 286–293, 1993.

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

[35] A. Panconesi and D. Ranjan. Quantifiers and approximation. *Theoretical Computer Science*, 107:145–163, 1993.

[36] C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization, Algorithms and Complexity.* Prentice Hall, 1982.

[37] M. Parker and J. Ryan. Covering IIS's. in preparation.

[38] K. S. van Horn and T. Martinez. The minimum feature set problem. Technical Report CS-92-8, Computer Science Department, Brigham Young University, Provo, 1992.

[39] R. E. Warmack and R. C. Gonzalez. An algorithm for optimal solution of linear inequalities and its application to pattern recognition. *IEEE Trans. on Computers*, 22:1065–1075, 1973.