Journal of Computer and Systems Sciences **58**, 572–596 (1999)
Article ID jcss.1998.1608, available online at http://www.idealibrary.com on IDEAL®

The Computational Complexity of Some Problems of Linear Algebra*

Jonathan F. Buss†

Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
E-mail: jfbuss@math.uwaterloo.ca

Gudmund S. Frandsen‡

BRICS, Department of Computer Science, University of Aarhus, Ny Munkegade DK-8000, Aarhus, Denmark
E-mail: gudmund@brics.dk

and

Jeffrey O. Shallit†

Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
E-mail: shallit@graceland.uwaterloo.ca

Received October 21, 1996; revised January 5, 1998

We consider the computational complexity of some problems dealing with matrix rank. Let E, S be subsets of a commutative ring R. Let $x_1, x_2, ..., x_t$ be variables. Given a matrix $M = M(x_1, x_2, ..., x_t)$ with entries chosen from $E \cup \{x_1, x_2, ..., x_t\}$, we want to determine $\text{maxrank}_S(M) = \max_{(a_1, a_2, ..., a_t) \in S^t} \text{rank } M(a_1, a_2, ..., a_t)$ and $\text{minrank}_S(M) = \min_{(a_1, a_2, ..., a_t) \in S^t} \text{rank } M(a_1, a_2, ..., a_t)$. There are also variants of these problems that specify more about the structure of M, or instead of asking for the minimum or maximum rank, they ask if there is some substitution of the variables that makes the matrix invertible or noninvertible. Depending on E, S, and which variant is studied,

* An extended abstract of an early version of this paper was presented at STACS'97, Lübeck, Germany, See *Lecture Notes in Comput. Sci.* (R. Reischuk and M. Morvan, Eds.), Vol. 1200, pp. 451–462, Springer-Verlag, Berlin, 1997.

† Supported in part by grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada and by the Information Technology Research Centre (ITRC) of Ontario.

‡ Supported by the ESPRIT Long Term Research Programme of the EU, under Project 20244 (ALCOM-IT), and by Basic Research in Computer Science (BRICS), Centre of the Danish National Research Foundation.

0022-0000/99 $30.00

the complexity of these problems can range from polynomial-time solvable to random polynomial-time solvable to *NP*-complete to *PSPACE*-solvable to unsolvable. An approximation version of the minrank problem is shown to be *MAXSNP*-hard.

1. INTRODUCTION

We consider the computational complexity of some problems of linear algebra —more specifically, problems dealing with matrix rank. Our mathematical framework is as follows. If R is a commutative ring, then $\mathcal{M}_n(R)$ is the ring of $n \times n$ matrices with entries in R. The rows α_i of a matrix are *linearly independent* over R if $\sum_i c_i \alpha_i = 0$ (with $c_i \in R$) implies $c_i = 0$ for all i, and similarly for the columns.

The *determinant* of $M = (a_{ij})_{1 \leqslant i, j \leqslant n}$ is defined as

$$\det M = \sum_{P = (i_1, i_2, \ldots, i_n)} (\operatorname{sgn} P)\, a_{1, i_1} a_{2, i_2} \cdots a_{n, i_n},$$

where

$$P = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

is a permutation of $\{1, 2, \ldots, n\}$. We know that a matrix is invertible over R if and only if its determinant is invertible over R [13].

The *rank* of a matrix M is the maximum number of linearly independent rows. Rank can also be defined as the maximum number of linearly independent columns, and it is well known [13] that these two definitions coincide. We denote the rank of M as rank M. An $n \times n$ matrix is invertible iff its rank is n.

A $k \times k$ *submatrix* of M is the array formed by the elements in k specified rows and columns; the determinant of such a submatrix is called a $k \times k$ *minor*. The rank of M can also be defined as the maximum size of an invertible minor.

The problems we consider are along the following lines: let E, S be two subsets of R. We are given an $n \times n$ matrix $M = M(x_1, x_2, \ldots, x_t)$, with entries chosen from $E \cup \{x_1, x_2, \ldots, x_t\}$, where the x_i are distinct variables. We want to compute

$$\operatorname{maxrank}_S(M) = \max_{(a_1, a_2, \ldots, a_t) \in S^t} \operatorname{rank} M(a_1, a_2, \ldots, a_t) \tag{1}$$

and

$$\operatorname{minrank}_S(M) \min_{(a_1, a_2, \ldots, a_t) \in S^t} \operatorname{rank} M(a_1, a_2, \ldots, a_t). \tag{2}$$

Evidently there is no need to distinguish between column rank and row rank in this definition. Note also that we do not necessarily demand that we be able to exhibit the actual t-tuple that achieves the maximum or minimum rank.

We will show that, depending on the arrangement of the variables in M and on the sets E, S, the complexity of the minrank and maxrank problems ranges from being in P to being unsolvable.

There are several reasons for studying these problems. First, the problems seem —to us, at least—natural questions in linear algebra. Second, a version of the minrank problem is very closely related to determining the minimum rank rational series that approximates a given formal power series to a given order; see [9, 19] and Section 13 of the present paper. Third, the maxrank problem is related to the problem of matrix rigidity which has recently received much attention [20, 6, 14] and may help explain why good bounds on matrix rigidity are hard to obtain.

Fixed: R, a commutative ring.
$E, S \subseteq R$.
Input: M, an $n \times n$ matrix with entries from $E \cup \{x_1, ..., x_t\}$.
k, a non-negative integer.

2. SUMMARY OF RESULTS

Most of our complexity results for the computation of minrank and maxrank are naturally phrased in terms of the decision problems given in Table 1. We have introduced two special problems, SING(ularity) and NONSING(ularity), which could possibly be easier than the more general minrank/maxrank problems.

Table 2 summarizes our results on the complexity of the four decision problems. We put the problems MAXRANK and NONSING together, since we have not been able to separate their complexities, although we do not know whether they have the same complexity in general. We have good evidence that the MINRANK and SING problems do not, in general, have the same complexity: over $\mathbb{C}$, MINRANK is *NP*-hard (Section 8), but SING has a random polynomial time solution (Section 4).

The exact value of E is not important for our bounds. All our lower bounds are valid for $E = \{0, 1\}$ and all our upper bounds are valid for E being $\mathbb{Q}$ or a finite-dimensional field extension of $\mathbb{Q}$ (respectively, E being $GF(q)$ or a finite-dimensional field extension of $GF(q)$, when the characteristic is finite). For the upper bounds, we assume the input size to be the total number of bits needed to specify the matrix M, when using the standard binary representation of numbers, representing a finite-dimensional algebraic extension by arithmetic modulo an irreducible polynomial, representing polynomials by coefficient vectors and listing the value of each entry in M. The upper bounds are also robust in another sense. We may allow

TABLE 1

Decision Problems

Fixed: R, a commutative ring.
$E, S \subseteq R$.
Input: M, an $n \times n$ matrix with entries from $E \cup \{x_1, \ldots, x_t\}$.
k, a non-negative integer.

Problem	Input	Decide
MINRANK	M, k	$\min_{(a_1,\ldots,a_t)\in S^t} \text{rank } M(a_1, \ldots a_t) \leq k$?
MAXRANK	M, k	$\max_{(a_1,\ldots,a_t)\in S^t} \text{rank } M(a_1, \ldots a_t) \geq k$?
SING	M	$\exists (a_1, \ldots, a_t) \in S^t$ such that $\det M(a_1, \ldots a_t) = 0$?
NONSING	M	$\exists (a_1, \ldots, a_t) \in S^t$ such that $\det M(a_1, \ldots a_t) \neq 0$?

TABLE 2

Complexity Bounds for Decision Problems: The General Case

S	E	MAXRANK NONSING	SING	MINRANK
$GF(q)$	$\{0,1\} \subseteq E \subseteq GF(q)$	*NP*-complete	*NP*-complete	*NP*-complete
$\mathbb{Z}$	$\{0,1\} \subseteq E \subseteq \mathbb{Q}$	*RP*	r.e.; undecidable	r.e.; undecidable
$\mathbb{Q}$	$\{0,1\} \subseteq E \subseteq \mathbb{Q}$	*RP*	r.e.; *NP*-hard	r.e.; *NP*-hard
$\mathbb{Q}_p$	$\{0,1\} \subseteq E \subseteq \mathbb{Q}$	*RP*	*EXPEXPSPACE*; *NP*-hard	*EXPEXPSPACE*; *NP*-hard
$\mathbb{R}$	$\{0,1\} \subseteq E \subseteq \mathbb{Q}$	*RP*	*PSPACE*; *NP*-hard	*PSPACE*; *NP*-hard
$\mathbb{C}$	$\{0,1\} \subseteq E \subseteq \mathbb{Q}$	*RP*	*RP*	*PSPACE*; *NP*-hard

entire multivariate polynomials (with coefficients from E) in a single entry of the matrix M and still preserve our upper bounds, provided such a multivariate polynomial is specified by an arithmetic formula using binary multiplication and binary addition, but no power symbol, so that the representation length of a multivariate polynomial is at least as large as its degree.

S is significant for the complexity, as is apparent from Table 2. However, our upper and lower bounds for $S = \mathbb{C}$ are valid for S being any algebraically closed field (in the case of S having finite characteristic, so must E of course).

The results of Table 2 fall into three groups, according to the proof technique used. The random polynomial time upper bounds use a result due to Schwartz [18]. The undecidability result for $\mathbb{Z}$ uses a combination of Valiant's result that the determinant is universal [21] and Matiyasevich's proof that Hilbert's Tenth Problem is unsolvable [15]. All the remaining problems of the result table (those that are not marked either *RP* or *undecidable*) are equivalent (under polynomial-time transformations) to deciding the existential first-order theory over the field S. The equivalence implies the *NP*-hardness of all these problems and lets us use results by Egidi [4], Ierardi [11], and Canny [3] to obtain the doubly exponential space upper bound for a p-adic field $\mathbb{Q}_p$ and the PSPACE upper bounds for $\mathbb{C}$ and $\mathbb{R}$, respectively. Since it is presently an open problem whether the existential first-order theory over $\mathbb{Q}$ is decidable or not, we suspect it will be difficult to determine the decidability status of MINRANK and SING over $\mathbb{Q}$. Koiran [12] has recently shown that ETh($\mathbb{C}$) is contained in the second level of the polynomial hierarchy if the generalized Riemann hypothesis is true. By our results the same conditional bound is valid for the minrank problem over $\mathbb{C}$.

We also consider the special case when each variable in the matrix occurs exactly once. None of our lower bound proofs are valid under this restriction, and we have improved some of the upper bounds. See Table 3 for a summary. The improved upper bounds all rely on the determinant polynomial being multi-affine when no variable occurs twice. In such a case the *RP*-algorithm for singularity over $\mathbb{C}$ can be generalized to work for singularity over any field.

For a very special kind of matrix, viz., row-partitionable matrices where each variable occurs exactly once, we give in Section 13 a polynomial time algorithm for computing the minimum possible rank. The algorithm works in the case where S is any field.

Since minrank is at least *NP*-hard to compute over $\mathbb{Z}$ or a field, one might consider the existence of an efficient approximation algorithm. Suppose, however,

TABLE 3

Upper Bounds When Each Variable Occurs Exactly Once

		MAXRANK		
S	E	NONSING	SING	MINRANK
$GF(q)$	$GF(q)$	*RP*	*RP*	*NP*
$\mathbb{Z}$	$\mathbb{Q}$	*RP*	r.e.	r.e.
$\mathbb{Q}$	$\mathbb{Q}$	*RP*	*RP*	r.e.
$\mathbb{Q}_p$	$\mathbb{Q}$	*RP*	*RP*	*EXPEXPSPACE*
$\mathbb{R}$	$\mathbb{Q}$	*RP*	*RP*	*PSPACE*
$\mathbb{C}$	$\mathbb{Q}$	*RP*	*RP*	*PSPACE*

that for some fixed S (S being $\mathbb{Z}$ or a field) and $E = \{0, 1\}$, there is a polynomial-time algorithm that when given matrix $M = M(x_1, ..., x_t)$ always returns a vector $(a_1, ..., a_t) \in S^t$ satisfying rank $(M(a_1, ..., a_t)) \leqslant (1 + \varepsilon) \cdot \text{minrank}_S(M)$. Then the assumption $P \neq NP$ implies $\varepsilon \geqslant \frac{7}{520} \approx 0.0134615$, as we prove in Section 11. The proof uses reduction from MAXEXACT3SAT; i.e., we use a known nonapproximability result for MAXEXACT3SAT [8], combined with a *MAXSNP*-hardness proof, for the minrank approximation problem.

Note added in proof. We have recently learned that the MAXRANK have previously been studied in the literature. Edmonds [25] posed the general MAXRANK problem, and Lovász [30] stated our random polynomial-time results for MAXRANK.

The special case of MAXRANK, where each variable occurs exactly once, is known in the linear algebra literature under the name *maximum rank matrix completion* [27, 28, 31]. Murota [31] reduced the maximum rank matrix completion problem to matroid intersection, which, when combined with an algorithm of Edmonds [24], shows that all our *RP* results in Table 3 may be improved to *P*. In a recent paper, Geelen [27] gave another deterministic polynomial-time algorithm for this problem. We are very grateful to J. Geelen for providing this information.

Geelen also notes (personal communication) that it can be proved, using a greedy algorithm, that the general problems MAXRANK and NONSING are polynomially related.

The special case of MINRANK, where each variable occurs exactly once, is known in the linear algebra literature under the name *minimum rank matrix completion* [22, 23]. A recent survey of the work on both minimum and maximum rank matrix completion may be found in [29].

In a recent paper [26], Egidi withdrew the *EXPEXPSPACE* upper bound on deciding $\mathbb{Q}_p$ that was claimed in [4]. As a consequence, our *EXPEXSPACE* upper bound in Table 2 should be relaxed to a decidability upper bound [23].

3. COMPUTING MAXRANK OVER INFINITE FIELDS

In this section we show how to compute maxrank with a (Monte-Carlo) random polynomial-time algorithm over any infinite field. We will also show that to solve the problem for $R = S = F$, it suffices to consider the case $R = S = \mathbb{Z}$, when F contains $\mathbb{Z}$. Our main tool is the following lemma, adapted from a paper of Schwartz [18].

LEMMA 1. *Let F be a field and let $p \in F[x_1, ..., x_t]$ be a multivariate polynomial that is not the zero polynomial. Let V be an arbitrary finite subset of F.*

(i) *Let d_i be the maximum degree of x_i occurring in p. If $d_i \leqslant |V|$ for all i then $p(\mathbf{a}) \neq 0$ for at least $\prod_{i=1}^{t} (|V| - d_i)$ tuples $\mathbf{a}$ out of all $\mathbf{a} \in V^t$.*

(ii) *Let d be the total degree of p. If $d \leqslant |V|$ then $p(\mathbf{a}) \neq 0$ for at least $(|V| - d) \cdot |V|^{t-1}$ tuples $\mathbf{a}$ out of all $\boldsymbol{a} \in V^t$.*

We now prove

THEOREM 2. *Let $M = M(x_1, x_2, ..., x_t)$ be a $n \times n$ matrix with entries in $F \cup \{x_1, x_2, ..., x_t\}$. Let $V \subseteq F$ be a set of at least $2n$ distinct elements. Choose a t-tuple $(a_1, a_2, ..., a_t) \in V^t$ at random. Then with probability at least $1/2$, we have* $\operatorname{maxrank}_F(M) = \operatorname{rank} M(a_1, a_2, ..., a_t)$.

Proof. Suppose $\operatorname{maxrank}_F(M) = k$. Then there exists some t-tuple $(a_1, a_2, ..., a_t) \in F^t$ such that $\operatorname{rank} M(a_1, a_2, ..., a_t) = k$. Hence, in particular, there must be some $k \times k$ minor of $M(a_1, a_2, ..., a_t)$ with nonzero determinant. Consider the corresponding $k \times k$ submatrix M' of $M(x_1, x_2, ..., x_t)$. Then the determinant of M', considered as a multivariate polynomial p in the indeterminates $x_1, x_2, ..., x_t$, cannot be identically zero (since it is nonzero when $x_1 = a_1, ..., x_t = a_t$). It now follows from Lemma 1(ii) that p is nonzero for at least half of all elements of V^t. Thus for at least half of all these t-tuples $(a_1, a_2, ..., a_t)$, the corresponding $k \times k$ minor of M must be nonzero, and hence, $M(a_1, a_2, ..., a_t)$ has rank at least k. Since $\operatorname{maxrank}_F(M) = k$, it follows that $\operatorname{rank} M(a_1, a_2, ..., a_t) = k$ for at least half of the choices $(a_1, a_2, ..., a_t) \in V^t$. ∎

The theorem justifies the following random polynomial-time algorithm to compute $\operatorname{maxrank}_F(M)$ over an infinite field F: choose r t-tuples of the form $(a_1, a_2, ..., a_t)$ independently at random and compute $\operatorname{rank} M(a_1, a_2, ..., a_t)$ for each of them, obtaining ranks $b_1, b_2, ..., b_r$. Then with probability at least $1 - 2^{-r}$, we have $\operatorname{maxrank}_F(M) = \max_{1 \leqslant i \leqslant r} b_i$.

It also follows from Theorem 2 that over an infinite field F, the quantity $\operatorname{maxrank}(M)$ cannot change when we consider an extension field F' with $F \subseteq F'$, or when we consider an infinite subset $S \subseteq F$. In particular, the decision problem MAXRANK is in the complexity class RP for $E = \mathbb{Q}$ and $\mathbb{Z} \subseteq S$.

4. THE SINGULARITY PROBLEM OVER AN ALGEBRAICALLY CLOSED FIELD

In this section we consider the complexity of the decision problem SING in the case $R = S = F$, where F is an algebraically closed field. We will show that in this case, SING $\in RP$. The proof uses the following lemmas.

LEMMA 3. *Let $P(x_1, x_2, ..., x_t)$ be a multivariate polynomial over an infinite field F. Then p is identically zero iff p is the zero polynomial.*

Proof. This is implied by Lemma 1. ∎

LEMMA 4. *Let $p(x_1, x_2, ..., x_t)$ be a nonconstant multivariate polynomial over a field F. If F is algebraically closed, then p takes on all values in F.*

Proof. This may be proved by induction on t, the number of variables. We leave the details to the reader. ∎

THEOREM 5. *If $R = S = F$, and F is algebraically closed, then* SING $\in$ *RP.*

Proof. Consider the following algorithm: Let $V \subseteq F$ be a set of at least $2n$ distinct elements. Choose r t-tuples $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_r$ at random from V^t, and evaluate the determinant $\det M(\mathbf{a}_i)$ for $1 \leqslant i \leqslant r$. If at least two different values are obtained, return "yes." If all the values obtained are the same and all are nonzero, return "no." If all the values are the same and all are zero, return "yes."

We claim that if there exists a t-tuple $\mathbf{a}$ such that $\det M(\mathbf{a}) = 0$, then this algorithm returns the correct result with probability at least $1 - 1/2^{r-1}$, while if there is no such t-tuple, the algorithm always returns the correct result.

To prove the claim, define $p(x_1, x_2, ..., x_t) = \det M(x_1, x_2, ..., x_t)$, a multivariate polynomial. If p is nonconstant, then by Lemma 4 it takes on all values in F, including 0. If p is constant and nonzero, then it cannot take on the value 0. Finally, if p is constant and zero, then it clearly takes on the value 0.

It now follows that our algorithm always returns the correct result, except possibly when all the values obtained are the same and nonzero. In this case we return "no," whereas if we are unlucky the answer could possibly be "yes." However, if the polynomial p is not the constant polynomial, then the polynomial $p - p(\mathbf{a}_1)$ is nonzero, and by Lemma 1(ii) we know $p(\mathbf{a}_i) \neq p(\mathbf{a}_1)$ with probability at least $1/2$ for $2 \leqslant i \leqslant r$. It follows that the probability of making an error in this case is bounded by $1/2^{r-1}$. ∎

5. UNIVERSALITY OF THE DETERMINANT

In this section, we prove a result that underlies all our lower bounds for the singularity and minrank problems: any multivariate polynomial is the determinant of a fairly small matrix. The result was first proven by Valiant [21], but since we need a slightly modified construction and the result is fundamental to our lower bound proofs, we make this paper self-contained and give the details of the construction.

To state the result, we need a few definitions. Let an *arithmetic formula* F be a well-formed formula using constants, variables, the unary operator $\{-\}$ and the binary operators $\{+, \cdot\}$. The *length* of a formula F (denoted by $|F|$) is defined as the total number of occurrences of constants, variables, and operators. For example,

$$|3x(y-z)-3| = |3 \cdot x \cdot (y + (-(z))) + (-(3))| = 11.$$

(Note that our definition of formula length is not the same as Valiant's.)

PROPOSITION 6. *Let R be a commutative ring. Let F be an arithmetic formula using constants from $E \subseteq R$ and variables from $\{x_1, ..., x_t\}$. For some $n \leqslant |F| + 2$, we may in time $n^{O(1)}$ construct an $n \times n$ matrix M with entries from $E \cup \{0, 1\} \cup \{x_1, ..., x_t\}$ such that $p_F = \det M$ and* $\operatorname{minrank}_R(M) \geqslant n - 1$, *where p_F denotes the polynomial described by formula F.*

Proof. We use a modified version of Valiant's construction [21]. The main difference is that we insist that the rank of the constructed $n \times n$ matrix cannot be less than $n - 1$ under any substitution for the variables. We also consider the negation operation explicitly, which allows us to avoid the use of negative constants

in the formula, when wanted. Our construction is essentially a modification of Valiant's construction to take care of these extra requirements, combined with a simplification that leads to matrices of somewhat larger size than Valiant's original construction.

Let a formula F be given. The construction falls into two parts. In the first part, we construct a series–parallel $s-t$-graph G_F with edge weights from $E \cup \{1\} \cup \{x_1, ..., x_t\}$ by induction on the structure of F as sketched in Fig. 1. To such a series–parallel $s-t$-graph G_F, we associate the polynomial

$$p(G_F) = \sum_{\pi \text{ is } s-t\text{-path in } G_F} (-1)^{\text{length}(\pi)} \cdot \prod_{e \text{ an edge of } \pi} \text{weight}(e).$$

By structural induction on F, one may verify that $p_F = p(G_F)$.

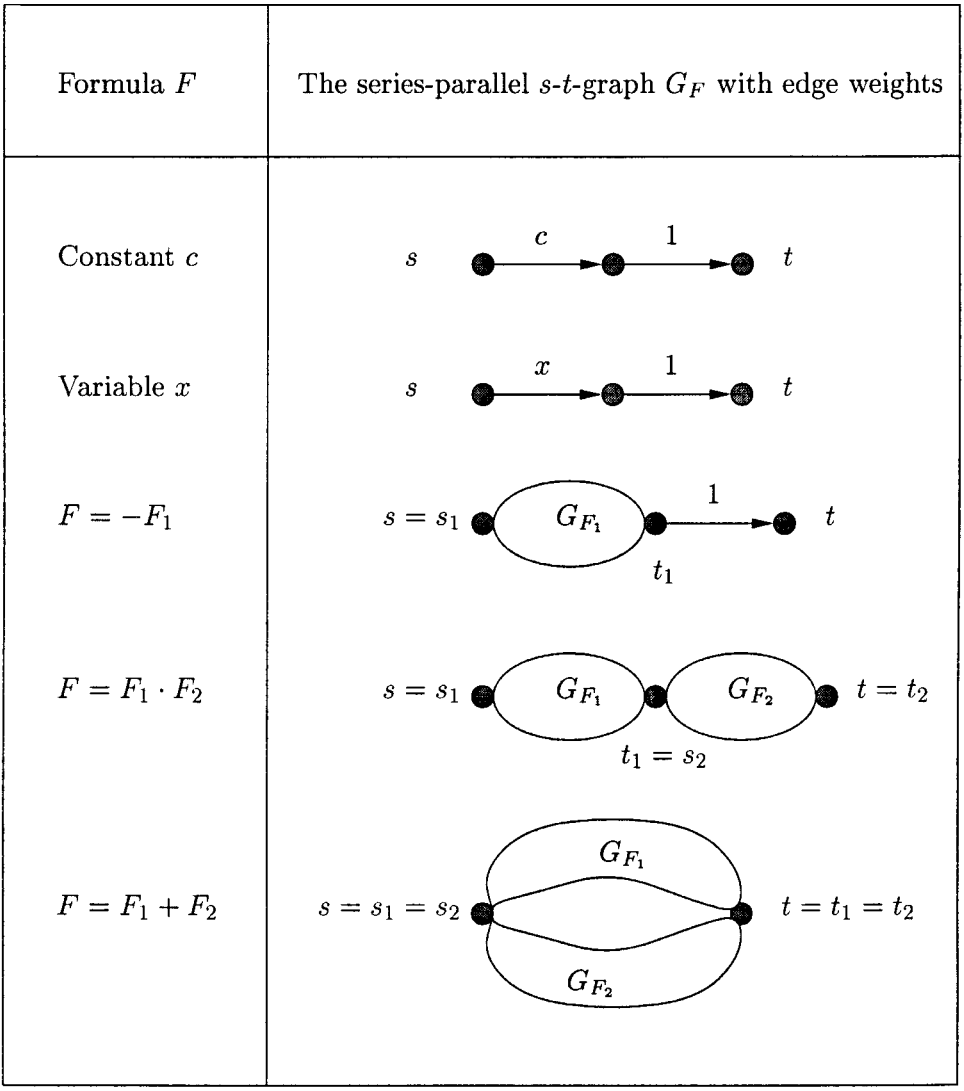

FIG. 1. Inductive construction of G_F.

In the second part of the construction, we change G_F into a cyclic graph G'_F by adding an edge from t to s of weight 1 and adding self-loops with weight 1 to all vertices different from s. The matrix $M = \{m_{ij}\}$ is simply the weight matrix for G'_F; i.e., m_{ij} is the weight of the edge from vertex i to vertex j if it exists, and $m_{ij} = 0$ otherwise. The determinant of M is a sum of monomials, where each monomial is the product of the weights in a specific cycle cover of G'_F (with sign ± 1, depending on the length of the cycles). But because of the special form of G'_F each cycle cover will consist of a number of self-loops (possibly zero) and a single cycle arising from an $s-t$-path in G_F, combined with the added edge from t to s. Hence, each $s-t$-path in G_F gives rise to one monomial in $\det M$, and the sign of the monomial will be -1 if and only if the path has odd length. Thus $\det M = p(G_F) = p_F$.

To see the lower bound on minrank, consider the $(n-1) \times (n-1)$ submatrix M' of M arising from erasing the column and row corresponding to the vertex s. The determinant of M' has one monomial for each cycle cover of $G'_F - \{s\}$. However, removing the vertex s breaks all cycles corresponding to paths from s to t in G_F, but with s removed all the remaining vertices have a self loop, so there is precisely one cycle cover and it consists of all the self-loops. Since all the self-loops have weight 1, we find that $\det M' = 1$, so $\text{minrank}_R(M) \geqslant n-1$.

The bound $2 + |p_F|$ on the size of G_F arises because the graph G_F has in addition to the vertices s and t at most one vertex for each application of a rewrite rule from Fig. 1.

6. THE SINGULARITY PROBLEM OVER THE INTEGERS

In this section we prove that the decision problem SING is unsolvable over $\mathbb{Z}$.

THEOREM 7. *Given a matrix $M = M(x_1, ..., x_t)$ with entries from $\{0, 1\} \cup \{x_1, ..., x_t\}$, it is undecidable whether there exist $a_1, ..., a_t \in \mathbb{Z}$ such that* $\det M(a_1, ..., a_t) = 0$.

Proof. We reduce from Hilbert's Tenth Problem [15]. An instance of Hilbert's Tenth Problem is a Diophantine equation $p(x_1, ..., x_t) = 0$, where p is a multivariate polynomial with integer coefficients. We construct a formula for p using only $+, -, \cdot, 0, 1$ in addition to the indeterminates by replacing each integer constant $c \geqslant 2$ having binary representation $c = \sum_{i=0}^{l} b_i 2^i$ with the formula

$$b_0 + (1+1)[b_1 + (1+1)[b_2 + (1+1)[b_3 + \cdots + (1+1)[b_l]\cdots]]].$$

By the construction of Proposition 6, the resulting formula f_p for the polynomial $p(x_1, ..., x_t)$ is turned into a matrix $M = M(x_1, ..., x_t)$ such that $\det M(x_1, ..., x_t) = p(x_1, ..., x_t)$. The assertion of the theorem follows from the undecidability of Hilbert's Tenth Problem. ∎

7. EXISTENTIAL FIRST-ORDER THEORIES

In this section, we describe the syntax of existential first-order theories over fields and state some complexity results for the corresponding decision problems. We will apply this later to our rank problems.

For any field F, we have arithmetic operations $+$, $\cdot$, constants 0, 1, and equality relation $=$. Adding the Boolean operations $\wedge$, $\vee$, $\neg$ and the existential quantifier $\exists$, we get the first-order language specified by the following grammar. (Note that we require all quantifiers to be collected in a prefix to the formula, thereby avoiding implicit universal quantification and alternation of quantifiers.)

$$\begin{aligned}
\mathtt{V} &::= x_1 \mid x_2 \mid x_3 \mid \cdots \mid x_n \mid \cdots \\
\mathtt{C} &::= 0 \mid 1 \\
\mathtt{AT} &::= \mathtt{V} \mid \mathtt{C} \\
\mathtt{T} &::= \mathtt{AT} \mid (\mathtt{T}+\mathtt{T}) \mid (\mathtt{T}\cdot\mathtt{T}) \\
\mathtt{AF} &::= \mathtt{T} = \mathtt{T} \\
\mathtt{BF} &::= \mathtt{AF} \mid (\neg \mathtt{BF}) \mid (\mathtt{BF} \wedge \mathtt{BF}) \mid (\mathtt{BF} \vee \mathtt{BF}) \\
\mathtt{F} &::= \mathtt{BF} \mid \exists \mathtt{V}\ \mathtt{F}
\end{aligned}$$

A *sentence* is a formula with no free variables (all variables are bound by quantifiers). We say that sentence φ is true in the field F (the field F is a *model* of the sentence φ), if the sentence evaluates to true, when quantifications are interpreted over elements in F and arithmetic operations and constants are given the natural interpretations, and we write

$$F \models \varphi.$$

For a more formal definition of the semantics see, for example, Enderton [5].

Note that we may allow the use of subtraction without increasing the descriptive power, since the sentence $\exists x\, \exists y.\ (1-x)\ y = 1$ is merely a shorthand for $\exists x\, \exists y\, \exists x'.\ (1+x')\ y = 1 \wedge x + x' = 0$.

For a field F, we define the existential theory of F:

$$\mathrm{ETh}(F) = \{\varphi : F \models \varphi\}.$$

The decision problem for $\mathrm{ETh}(F)$ is: on input φ, decide whether $F \models \varphi$.

PROPOSITION 8. *For F being any fixed field, $\mathrm{ETh}(F)$ is NP-hard.*

Proof. We reduce from 3SAT. Let C be an instance of 3SAT; i.e.,

$$C \equiv C_1 \wedge C_2 \wedge \cdots \wedge C_k,$$

where $C_i \equiv (l_{i1} \vee l_{i2} \vee l_{i3})$ and $l_{ij} \in \{y_1, y_2, ..., y_t\} \cup \{\overline{y_1}, \overline{y_2}, ..., \overline{y_t}\}$. We modify C to be an arithmetic formula f_C by replacing each y_i with the atomic formula $x_i = 1$ and replacing each $\overline{y_i}$ with the atomic formula $x_i = 0$. Clearly,

$$C \text{ is satisfiable iff } \quad F \models \exists x_1, \exists x_2 \cdots \exists x_t \cdot f_C.$$

The *NP*-hardness follows from the *NP*-hardness of 3SAT. ∎

TABLE 4

Upper Bounds on Deciding ETh(F)

F	Upper bound on ETh(F)	Reference
$GF(q)$	*NP*	
$\mathbb{Q}$	Recursively enumerable	
$\mathbb{Q}_p$	*EXPEXPSPACE*	Egidi, 1993 [4]
$\mathbb{R}$	*PSPACE*	Canny, 1988 [3]; Renegar, 1992 [17]
$\mathbb{C}$	*PSPACE*	Ierardi, 1989 [11]

The complexity of deciding ETh(F) seems to depend on the field F. Table 4 summarizes the upper bounds that we are aware of.

ETh($GF(q)$) is in *NP* for any fixed finite field ($GF(q)$), since one may replace the variables with nondeterministically chosen field elements and evaluate the resulting variable-free formula in polynomial time.

Similarly, ETh($\mathbb{Q}$) is recursively enumerable, but to the best of our knowledge it is still an open problem whether ETh($\mathbb{Q}$) is in fact decidable.

The *EXPEXPSPACE* bound for the field of p-adic numbers, $\mathbb{Q}_p$ (for some fixed prime p) is proven for a more general theory than the one considered here. It is quite conceivable that a better bound can be found for our existential sentences.

One may get a *PSPACE* bound for $\mathbb{C}$ as a corollary to the *PSPACE* bound for $\mathbb{R}$, since arithmetic in $\mathbb{C}$ can be represented by arithmetic on pairs of numbers in $\mathbb{R}$. However, the proof of Ierardi [11] uses a different technique and holds for any algebraically closed field. Recently, Koiran [12] showed that in fact ETh(C) is contained in the second level of the polynomial hierarchy—conditional upon the generalized Riemann hypothesis being true.

8. LOWER BOUND FOR MINRANK OVER A FIELD

In this section, we prove that over a field that is not algebraically closed, the decision problem SING is as hard as deciding the corresponding existential first-order theory (although for some fields we use extra constants in addition to 0, 1 to establish the correspondence). Only one step in the proof does not seem to generalize to an arbitrary field—namely the reduction of a system (conjunction) of equations to a single equation, which is necessary for encoding a general existential sentence as a singularity problem. However, we observe that a *system* of equations can be encoded as a single minrank problem. This allows us to show that over any field the more general decision problem MINRANK is indeed as hard as the corresponding existential first-order theory.

LEMMA 9. *Let F be a field. Given an existential sentence $\exists x_1 \cdots \exists x_t \cdot \varphi(x_1, ..., x_t)$, of length m, we can in time $m^{O(1)}$ construct an equivalent existential sentence $\exists x_1 \cdots \exists x_{t'} \cdot \psi(x_1, ..., x_{t'})$ such that ψ contains no negations.*

TABLE 5

Construction for Elimination of $\neg$

	Rewrite rules
Step 1	$\neg(F_1 \wedge F_2) \to (\neg F_1) \vee (\neg F_2)$ $\neg(F_1 \vee F_2) \to (\neg F_1) \wedge (\neg F_2)$
Step 2	$\neg t(\mathbf{x}) = 0 \to 1 - z \cdot t(\mathbf{x}) = 0$

Proof. The formula ψ is constructed from φ using the rewrite rules of Table 5.

In step 1, we use de Morgan's laws to move all negations down so that they are applied directly to the atomic formulas.

In step 2, we replace each negated atomic formula by an unnegated formula. We introduce a new variable z for each such atomic formula, which represents the inverse of the term $t(\mathbf{x})$. These new variables must be existentially quantified. ▮

LEMMA 10. *Let F be a field. Given an existential sentence $\exists x_1 \cdots \exists x_t \cdot \varphi(x_1, ..., x_t)$ of length m, that contains no negation, we can in time $m^{O(1)}$ construct an equivalent existential sentence $\exists x_1 \cdots \exists x_{t'} \cdot \psi(x_1, ..., x_{t'})$ such that ψ contains neither negation nor disjunction.*

Proof. Let φ have s subformulas $f_1, ..., f_s$, each of which may be atomic or composite. For each such subformula f_i, we introduce a new (existentially quantified) variable z_i, and we construct a new formula f'_i that is either atomic or the conjunction of two atomic formulas. The f'_is will be constructed such that

$$\begin{array}{l} \exists x_1 \cdots \exists x_{t'} \cdot \text{“}f_i \text{ is satisfied”} \\ \qquad\qquad \Updownarrow \\ \exists x_1 \cdots \exists x_t \, \exists z_1 \cdots \exists z_s \text{ “}z_i = 0 \text{ and } f'_j \text{ is satisfied} \\ \qquad \text{for all subformulas } f_j \text{ of } f_i \text{ (including } f_i\text{).”} \end{array} \tag{3}$$

If the subformula f_1 corresponds to the entire formula φ, this implies that

$$\begin{array}{l} \exists \mathbf{x} \cdot \varphi(\mathbf{x}) \\ \qquad\qquad \Updownarrow \\ \exists \mathbf{x}, \mathbf{z}. \, z_1 = 0 \wedge f'_1(\mathbf{x}, \mathbf{z}) \wedge \cdots \wedge f'_s(\mathbf{x}, \mathbf{z}). \end{array}$$

For each original subformula f_i the new formula f'_i is constructed as described in Table 6. By induction in the structure of f_i, one may verify that this construction does satisfy (3), from which the theorem follows. ▮

TABLE 6

Construction for Elimination of $\vee$

f_i	f'_i
$p_i(\mathbf{x}) = 0$	$p_i(\mathbf{x}) = z_i$
$f_j \vee f_k$	$z_j \cdot z_k = z_i$
$f_j \wedge f_k$	$z_j \cdot z_k = z_i \wedge z_j + z_k = z_i$

LEMMA 11. *Let F be a field. Given an existential sentence φ of length m, we can in time $n^{O(1)}$ construct an integer k and an $n \times n$ matrix with entries from $\{0, 1\} \cup \{x_1, x_2, ..., x_t\}$, where $n = O(m)$ such that* $\text{minrank}_F(M) \leqslant k$ *if and only if* $F \models \varphi$.

Proof. Let an existential sentence be given. First we remove all negations and disjunctions using the constructions of Lemmas 9 and 10. Without loss of generality, we may therefore assume that we are given the existential sentence

$$\exists \mathbf{x}\, p_1(\mathbf{x}) = 0 \wedge \cdots \wedge p_r(\mathbf{x}) = 0$$

for some arithmetic formulas p_i, $i = 1, ..., r$.

By Proposition 6, we may for each $p_i(x_i, ..., x_t)$ find an $n_i \times n_i$ matrix M_i with entries from $\{0, 1\} \cup \{x_1, x_2, ..., x_t\}$ such that $\det M_i = p_i(x_1, ..., x_t)$ and $\text{minrank}_F(M_i) \geqslant n_i - 1$.

Let $n = \sum_{i=1}^{r} n_i$, let $k = \sum_{i=1}^{r} (n_i - 1)$, and construct the $n \times n$ matrix M by placing $M_1, ..., M_r$ consecutively on the main diagonal and zeroes elsewhere. Clearly, $\text{minrank}_F(M) \geqslant k$ and $\text{rank}\, M = k$ only when all the polynomials p_i are simultaneously zero; therefore, $\text{minrank}_F(M) \leqslant k$ iff $F \models \varphi$. ∎

COROLLARY 12. *Let F be a field. The decision problem* MINRANK *for $S = F$ and $E = \{0, 1\}$ is NP-hard.*

Proof. Immediate from Lemma 11 and Proposition 8. ∎

It is possible to eliminate conjunction from formulas when the field is not algebraically closed, allowing us to prove a SING-version of Lemma 11 for such fields.

LEMMA 13. *Let F be a fixed field that is not algebraically closed. Then there exists a finite set of constants $E \subseteq F$ such that given arithmetic formulas $p_1(\mathbf{x}), ..., p_r(\mathbf{x})$ of combined length m, we can in time $m^{O(1)}$ construct a single arithmetic formula $p(\mathbf{x})$ (using constants from E) such that*

$$F \models \exists \mathbf{x}.p_1(\mathbf{x}) = 0 \wedge \cdots \wedge p_r(\mathbf{x}) = 0$$

$$\Updownarrow$$

$$F \models \exists \mathbf{x}.p(\mathbf{x}) = 0.$$

The set of constants $E = \{0, 1\}$ suffices for any of the fields $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}_p$, $GF(q)$.

Proof. Since F is not algebraically closed, there exists a univariate polynomial $f(x) = \sum_{i=0}^{d} a_i x^i$ of degree $d \geqslant 1$ with $a_0, ..., a_d \in F$ such that f has no root in F. Define a new polynomial in two variables by $g(y, z) = z^d \cdot f(y/z) = \sum_{i=0}^{d} a_i y^i z^{d-i}$. The polynomial $g(y, z)$ is nonzero, except when $y = z = 0$. To see this, observe that $g(y, 0) = a_d y^d$ which is nonzero when $y \neq 0$; for $z \neq 0$ we have that $g(y, z) = 0$ implies that $f(y/z) = 0$, which cannot occur since f has no roots in F. From g one may construct a new polynomial of degree d^2 in four variables, namely $h(x, y, z, w) = g(g(x, y), g(z, w))$. Clearly, $h(x, y, z, w) = 0$ only if $x = y = z = w = 0$. By induction one may construct a polynomial $k(x_1, ..., x_r)$ over any specified number of variables such that $k(x_1, ..., x_r)$ has no nontrivial zeros. This is the idea behind our construction of $p(\mathbf{x})$. Using an arithmetic formula for g (of size $O(d^2)$) construct the formula $p(\mathbf{x})$ from $p_1(\mathbf{x}), ..., p_r(\mathbf{x})$ using the rewrite rule of Table 7 repeatedly $\log r$ times. The size of $p(\mathbf{x})$ will be $O(d^{2 \log r} m) = O(r^{2 \log d} m) = m^{O(1)}$.

To see that $E = \{0, 1\}$ suffices for some special fields as claimed in the lemma, choose $f(x) = x^2 + 1$ for F being $\mathbb{Q}$ or $\mathbb{R}$. A monic polynomial in $\mathbb{Z}[x]$ whose reduction modulo p is irreducible over the finite field $GF(p)$ will also be irreducible over the p-adic field $\mathbb{Q}_p$ (see [7, Corollary 5.3.8, p. 139]). Therefore, choose $f(x) = x^2 + x + 1$ for F being $\mathbb{Q}_2$ or $GF(2)$; choose $f(x) = x^2 + (p - a)$ for some quadratic nonresidue a modulo p when F is $\mathbb{Q}_p$ or $GF(p)$ and $p \neq 2$ is a prime (and use that $p - a = 1 + 1 + \cdots + 1$). Finally, a suitable irreducible polynomial exists for any other specific finite field. ∎

Remark. The construction of the proceeding proof can be improved in the case of specific fields. For example, over the fields $\mathbb{Q}$ and $\mathbb{R}$, any number of equations can be encoded into a single equation while only doubling the formula size, when using that the multivariate polynomial $x_1^2 + x_2^2 + \cdots + x_r^2$ takes the value zero only when $x_1 = x_2 = \cdots = x_r = 0$. A similar property is satisfied by the arithmetic formula $1 - (1 - x_1^{q-1})(1 - x_2^{q-1}) \cdots (1 - x_r^{q-1})$ with respect to a fixed finite field $GF(q)$.

LEMMA 14. *Let F be a fixed field that is not algebraically closed. Then there exists a finite set of constants $E \subseteq F$ such that given an existential sentence ϕ of length m, then we can in time $m^{O(1)}$ construct an $n \times n$ matrix M with entries from $E \cup \{x_1, ..., x_t\}$ such that $F \models \phi$ iff $\exists(\mathbf{a}) \in F^t . \det M(\mathbf{a}) = 0$. The set of constants $E = \{0, 1\}$ suffices for any of the fields $\mathbb{Q}, \mathbb{R}, \mathbb{Q}_p, GF(q)$.*

Proof. Let an existential sentence be given. First we remove all negations, disjunctions and conjunctions using the constructions of Lemmas 9, 10, and 13 to obtain the single equation $p(x_1, ..., x_t) = 0$. By Proposition 6, we may find a matrix M such that $\det M = p(x_1, ..., x_t)$. ∎

TABLE 7

Construction for Elimination of $\wedge$

Rewrite rule
$p_1(\mathbf{x}) = 0 \wedge p_2(\mathbf{x}) = 0 \wedge p_3(\mathbf{x}) = 0 \wedge \cdots \wedge p_{2k-1}(\mathbf{x}) = 0 \wedge p_{2k}(\mathbf{x}) = 0$
$\rightarrow g(p_1(\mathbf{x}), p_2(\mathbf{x})) = 0 \wedge g(p_3(\mathbf{x}), p_4(\mathbf{x})) = 0 \wedge \cdots \wedge g(p_{2k-1}(\mathbf{x}), p_{2k}(\mathbf{x})) = 0$

COROLLARY 15. *Let F be one of the fields $\mathbb{Q}, \mathbb{R}, \mathbb{Q}_p$, or $GF(q)$. The decision problem* SING *for $S = F$ and $E = \{0, 1\}$ is NP-hard. If F is any field that is not algebraically closed, then there is a finite set $E \subseteq F$ such that the decision problem* SING *for $S = F$ and E is NP-hard.*

Proof. Immediate from Lemma 14 and Proposition 8. ∎

9. DECISION PROBLEMS OVER FINITE FIELDS

For finite fields, we have a stronger result.

THEOREM 16. *Let F be a fixed finite field $GF(q)$. For $S = F$ and $\{0, 1\} \subseteq E \subseteq GF(q)$, the decision problems* MAXRANK, NONSING, MINRANK, *and* SING *are all NP-complete.*

Proof. Clearly, these problems are in *NP*, since we may nondeterministically guess an assignment to the variables and compute the rank of the resulting constant matrix in polynomial time.

The *NP*-hardness of SING (and MINRANK) follows from Corollary 15.

To prove the *NP*-hardness of NONSING (and MAXRANK) we observe from Lemmas 9, 10, 13 and Proposition 8 that it is *NP*-hard to decide whether $GF(q) \models \exists \mathbf{x}.p(\mathbf{x}) = 0$ when given an arithmetic formula $p(\mathbf{x})$.

For the finite field $GF(q)$, it is well known that the function $x \mapsto x^{q-1}$ maps 0 to 0 and maps any nonzero number to 1. Therefore, it is also *NP*-hard to decide whether $GF(q) \models \exists \mathbf{x}.1 - p(\mathbf{x})^{q-1} \neq 0$, where $p(\mathbf{x})^{q-1}$ is shorthand for the formula $p(\mathbf{x}) \cdot p(\mathbf{x}) \cdots p(\mathbf{x})$, whose length is only a constant factor larger than the length of $p(\mathbf{x})$ when q is fixed.

To complete the *NP*-hardness proof, use Proposition 6. ∎

10. UPPER BOUNDS FOR MINRANK OVER A FIELD

In this section, we prove that the minrank problem over a field is no harder than deciding the corresponding existential first-order theory. Combined with our earlier results, this implies that the decision problem MINRANK is in fact equivalent (under polynomial-time transformations) to deciding the corresponding existential first-order theory. In addition, we inherit the upper bounds of Table 4.

We start by giving the reduction for matrices that use only constants 0 and 1, and afterwards we extend the result to more general constants.

LEMMA 17. *Let F be a field. Given an $n \times n$ matrix M with entries from $\{0, 1\} \cup \{x_1, x_2, ..., x_t\}$ and some $k \leqslant n$, we may in time $n^{O(1)}$ construct an existential sentence φ such that* $\operatorname{minrank}_F(M) \leqslant k$ *if and only if* $F \models \varphi$.

Proof. Given $(n \times n)$ matrix M with variables $x_1, x_2, ..., x_t$ and constants from $\{0, 1\}$, we express (in a first-order existential sentence) the assertion that some k columns of M span all columns of M when appropriate values are substituted for $x_1, x_2, ..., x_t$.

We are going to use that a matrix $\mathbf{C}$ with entries from F has rank at most k precisely when there is a nonsingular $(n \times n)$ matrix $\mathbf{A}$ such that the first $(n-k)$ columns of the matrix product $\mathbf{CA}$ are all identically zero. The nonsingularity of $\mathbf{A}$ is ensured by demanding that $\mathbf{A}$ have an inverse, i.e., that there exists some $(n \times n)$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}$. Using matrix notation our first-order sentence can be expressed

$$\begin{aligned} &\operatorname{minrank}_F(M) \leqslant k \\ &\Updownarrow \\ &\exists x_1, \ldots, x_t \in F \, \exists \mathbf{A}, \mathbf{B} \in F^{n^2}. \\ &\quad \text{(i)} \quad \mathbf{AB} = \mathbf{I}, \text{ and} \\ &\quad \text{(ii)} \quad \text{the first } (n-k) \text{ columns of } M\mathbf{A} \text{ are all zero.} \end{aligned}$$

Since matrix multiplication can be expressed by a formula of size $O(n^3)$, the above sentence using matrix notation leads to a proper existential sentence of size $O(n^3)$ that is equivalent to the minrank problem as stated in the theorem. ▮

The proof of Lemma 17 we have just presented uses an improvement, suggested by von zur Gathen, of our original proof. It is based on an idea of Borodin, von zur Gathen, and Hopcroft [2].

We restricted the constants in our existential sentences to 0 and 1 in order to apply the upper bounds of Table 4. However, an analogue of Lemma 17 does actually hold for the minrank problem over matrices containing algebraic constants, because algebraic constants can be defined by short first-order sentences:

- Over any field, the constant 2 is defined by

$$\varphi(x) \equiv x = 1 + 1.$$

- Over a field with characteristic different from 2, the constant $-3/2$ is defined by

$$\varphi(x) \equiv x \cdot (1+1) + 1 + 1 + 1 = 0.$$

- Over $\mathbb{R}$, the constant $\sqrt{2}$ is defined by

$$\varphi(x) \equiv \exists y . x \cdot x = 1 + 1 \wedge y \cdot y = x.$$

(The last part ensures that we get the positive of the two square roots.)

- Over any field, the constant 15 is defined by

$$\varphi(x) \equiv \exists y \, \exists z \, \exists w . x = 1 + y + z + w \wedge y = 1 + 1 \wedge z = y + y \wedge w = z + z.$$

(We use a repeated doubling strategy to make the defining formula have length proportional to the usual binary representation of the integer 15.)

- Over $\mathbb{C}$, the constants i and $-i$ are defined by

$$\varphi(x, y) \equiv x \cdot x + 1 = 0 \wedge y \cdot y + 1 = 0 \wedge x + y = 0.$$

(Note that i and $-i$ cannot be defined separately, since i alone can only be defined up to conjugation; the only nontrivial isomorphism on $\mathbb{C}$.)

If F is a field, define its *prime field* to be the intersection of all subfields of F [10, Section V.5]. Clearly, the prime field underlying $\mathbb{C}$ and $\mathbb{R}$ is $\mathbb{Q}$, and $GF(q)$ is a finite-dimensional algebraic extension of its underlying prime field (which is $GF(p)$ for some prime p). For a field F let A_F be the set of all numbers that are algebraic over the prime field underlying F.

PROPOSITION 18. *Let P be a prime field. Let $\{e_1, ..., e_t\} \subseteq A_p$. Let F be the smallest extension field containing all the constants $\{e_1, ..., e_t\}$. Let a standard representation of F as a k-dimensional vector space over P (with vector arithmetic defined using an irreducible polynomial) be given. Let the representation of the constants $\{e_1, ..., e_t\}$ as vectors of binary numbers be given. It is possible to construct an existential first-order formula $\varphi(x_1, ..., x_t)$ defining $\{e_1, ..., e_t\}$ in time polynomial in the combined bit length of all the constant representations.*

Proof. Left to the reader. ▮

The generalization of Lemma 17 is the following.

LEMMA 19. *Let F be a field. Let F' be a finite-dimensional algebraic extension of the prime field underlying F. Let $E \subseteq F' (\subseteq A_F)$. Given an $n \times n$ matrix M with entries from $E \cup \{x_1, x_2, ..., x_t\}$, and some $k \leqslant n$, we may in time $(ns)^{O(1)}$ construct an existential sentence φ such that*

$$\text{minrank}_F(M) \leqslant k \qquad \textit{iff} \quad F \models \varphi,$$

where s denotes the maximum bit length of the representation of an entry in M (using binary numbers/quotients for prime field elements and vectors of these for algebraic numbers).

Proof. Use the construction from the proof of Lemma 17 combined with the construction of Proposition 18. ▮

COROLLARY 20. *Let F be a field. Let F' be a finite dimensional algebraic extension of the prime field underlying F. Let $S = F$ and let $\{0, 1\} \subseteq E \subseteq F'$. The decision problem* MINRANK *is equivalent (under polynomial-time transformations) to deciding* ETh(F).

If F is one of the fields $\mathbb{Q}$, $\mathbb{R}$, or a p-adic field $\mathbb{Q}_p$, then the decision problems SING *and* MINRANK *are equivalent by polynomial-time transformation.*

If F is a fixed p-adic field $\mathbb{Q}_p$, then the decision problem MINRANK *is solvable in EXPEXPSPACE.*

If F is one of the fields $\mathbb{R}$ and $\mathbb{C}$ then the decision problem MINRANK *is in PSPACE.*

Proof. Immediate from Lemmas 19, 11, 14, and the bounds cited in Table 4. ▮

11. TIGHT APPROXIMATION OF MINRANK IS *NP*-HARD

In this section, we consider the approximation problem (parametrized with $\varepsilon > 0$) associated with the minrank problem:

$(1+\varepsilon)$-APXMINRANK

Let R be a commutative ring. Let $E, S \subseteq R$.

Input: a matrix $M = M(x_1, ..., x_t)$ with entries in $E \cup \{x_1, ..., x_r\}$.

Output: some $a_1, ..., a_t \in S$ such that

$$\text{rank } M(a_1, ..., a_t) \leqslant (1+\varepsilon) \cdot \text{minrank}_S(M).$$

We prove that $(1+\varepsilon)$-APXMINRANK is *NP*-hard for ε sufficiently small, when R is $\mathbb{Z}$ or a field. The tool will be a reduction from the approximation version of EXACT3SAT. Consider the problem:

$(1-\varepsilon)$-MAXEXACT3SAT

Input: a conjunction of clauses $C = C_1 \wedge \cdots \wedge C_k$, where each clause contains exactly three distinct literals $C_i = (l_{i1} \vee l_{i2} \vee l_{i3})$, and each literal is one of the Boolean variables $\{y_1, ..., y_r\}$ or its negation.

For $(b_1, ..., b_r) \in \{0, 1\}^r$, let numb $(C, b_1, ..., b_r)$ be the number of clauses in C that are satisfied under the assignment $y_i \mapsto b_i$, and let

$$\text{maxnumb}(C) = \max_{(b_1, ..., b_r) \in \{0, 1\}^r} \text{numb}(C, b_1, ..., b_r).$$

Output: some truth assignment $b_1, b_2, ..., b_r \in \{0, 1\}$ such that

$$\text{numb}(C, b_1, ..., b_r) \geqslant (1-\varepsilon) \cdot \text{maxnumb}(C).$$

PROPOSITION 21. *For $\varepsilon < \frac{1}{8}$ there is no polynomial-time algorithm for* $(1-\varepsilon)$-MAXEXACT3SAT *unless* $P = NP$.

Proof. See Håstad [8]. ∎

To prove the nonapproximability of minrank, we need a special type of reduction first defined by Papadimitriou and Yannakakis [16]. Since we only use the reduction in a single case, we specialize the definition to the concrete application.

Given $E, S \subseteq R$, MAXEXACT3SAT is said to *L-reduce* to APXMINRANK with parameters α, β, if there exist two polynomial time computable functions f and g such that, for a given instance C of MAXEXACT3SAT,

1. Algorithm f produces matrix M with entries in $E \cup \{x_1, ..., x_t\}$ such that

$$\text{minrank}_S(M) \leqslant \alpha \cdot \text{maxnumb}(C);$$

2. Given any substitution $(a_1, a_2, ..., a_t) \in S^t$ for the variables in M, g produces a truth assignment $(b_1, b_2, ..., b_r) \in \{0, 1\}^r$ such that

$$|\text{maxnumb}(C) - \text{numb}(C, b_1, b_2, ..., b_r)| \leqslant \beta \cdot |\text{minrank}_S(M) - \text{rank } M(a_1, a_2, ..., , a_t)|.$$

L-reduction preserves approximability.

PROPOSITION 22. *Let $E, S \subseteq R$ be given. If* MAXEXACT3SAT *L-reduces to* APXMINRANK *with parameters $\alpha, \beta \geqslant 0$ and $(1+\varepsilon)$-*APXMINRANK *has a polynomial time solution then $(1-\alpha\beta\varepsilon)$-*MAXEXACT3SAT *has a polynomial time solution.*

Proof. The polynomial time solution for $(1-\alpha\beta\varepsilon)$-MAXEXACT3SAT works as follows: Given an instance $C(y_1, ..., y_r)$ of MAXEXACT3SAT, compute an instance $M(x_1, ..., x_t)$ of APXMINRANK using the function f. Find a substitution $(a_1, ..., a_t)$ for $(x_1, ..., x_t)$ using the polynomial time solution for $(1+\varepsilon)$-APXMINRANK and transform this substitution into a truth assignment $(b_1, ..., b_r)$ for $(y_1, ..., y_r)$ using the function g. We verify the $(1-\alpha\beta\varepsilon)$ bound by a computation.

$$\begin{aligned} |\text{maxnumb}(C) - \text{numb}(C, b_1, ..., b_r)| &\leqslant \beta \cdot |\text{minrank}_S(M) - \text{rank } M(a_1, ..., a_t)| \\ &\leqslant \beta\varepsilon \cdot \text{minrank}_S M \\ &\leqslant \alpha\beta\varepsilon \cdot \text{maxnumb}(C). \quad \blacksquare \end{aligned}$$

LEMMA 23. *Let R be a commutative ring without zero divisors, and let $\{0, 1\} \subseteq S \subseteq R$ and $E = \{0, 1\}$.* MAXEXACT3SAT *L-reduces to* APXMINRANK *with parameters $\alpha = \frac{65}{7}$ and $\beta = 1$.*

Proof. First, we describe the function f. Assume we have an instance of MAXEXACT3SAT, viz. a conjunction of clauses $C = C_1 \wedge \cdots \wedge C_k$, where each clause contains three distinct literals $C_i = (l_{i1} \vee l_{i2} \vee l_{i3})$ and each literal is one of the Boolean variables $\{y_1, ..., y_r\}$ or its negation.

For each clause C_i, there will be a 12×12 matrix M_i, containing four smaller 3×3 matrices down the diagonal and zeros elsewhere. The four smaller matrices are one for each of the three variables occurring in the clause and one for the clause itself.

Each Boolean variable y_j is represented by two arithmetic variables x_{j1} and x_{j2}. The variable x_{j1} being zero represents y_j being true, and x_{j2} being zero represents y_j being false. We can ensure that not both of x_{j1} and x_{j2} are zero by requiring

$$x_{j1} + x_{j2} = 1. \tag{4}$$

We allow the case that neither x_{j1} nor x_{j2} is zero.

For each of the three variables occurring in a clause, there will be a matrix ensuring (4); i.e., for $s=1, 2, 3$, if $l_{is}=y_j$ or $l_{is}=\overline{y_j}$ then

$$A_{is}=\begin{bmatrix} 1 & x_{j1} & x_{j2} \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

The matrix A_{is} always has rank at least 2 and has rank exactly 2 when (4) is satisfied, since $\det A_{is}=1-x_{j1}-x_{j2}$.

If $C_i=(y_{j1} \vee y_{j2} \vee y_{j3})$, the fourth matrix will be

$$B_i=\begin{bmatrix} x_{j_1 1} & 1 & 0 \\ 0 & x_{j_2 1} & 1 \\ 0 & 0 & x_{j_3 1} \end{bmatrix}.$$

(If $\overline{y_j}$ occurs in C_i instead of y_j, then replace x_{j1} with x_{j2} in matrix B_i.)

The matrix B_i always has rank at least 2 and has rank exactly 2 when $x_{j_1 1}=0$ or $x_{j_2 1}=0$ or $x_{j_3 1}=0$.

If we let $\operatorname{diag}(M_1, ..., M_k)$ denote the large matrix having the smaller matrices $M_1, ..., M_k$ consecutively down the main diagonal and zeros elsewhere, then the function f returns the matrix

$$M=\operatorname{diag}(M_1, ..., M_k), \qquad \text{where} \quad M_i=\operatorname{diag}(A_{i1}, A_{i2}, A_{i3}, B_i).$$

Clearly, f can be computed in polynomial time.

Clearly, $\operatorname{minrank}_S(M) \leqslant k \cdot (4 \cdot 2)+(k-\operatorname{maxnumb}(C))=9k-\operatorname{maxnumb}(C)$. We know that $\operatorname{maxnumb}(C) \geqslant 7k/8$, since the expected fraction of true clauses using a random truth assignment is at least $\frac{7}{8}$. Combining, we get that

$$\begin{aligned} \operatorname{minrank}_S(M) &\leqslant 9k-\operatorname{maxnumb}(C) \\ &\leqslant 9 \cdot \tfrac{8}{7} \operatorname{maxnumb}(C)-\operatorname{maxnumb}(C) \\ &\leqslant \tfrac{65}{7} \operatorname{maxnumb}(C), \end{aligned}$$

which proves the assertion about α.

We still need to describe the function g. Let a substitution $a_{11}, a_{12}, ..., a_{r1}, a_{r2} \in S^{2r}$ for the arithmetic variables in M be given. Construct a truth assignment $b_1, ..., b_r$ for the Boolean variables in C as follows. If $a_{j1}=0$ then let $b_j=1$; otherwise if $a_{j2}=0$ then let $b_j=0$. But if both $a_{j1} \neq 0$ and $a_{j2} \neq 0$ then let b_j take an arbitrary value. Clearly, g can be computed in polynomial time.

If clause C_i is not satisfied under the truth assignment $b_1, ..., b_r$, then matrix M_i will have rank at least 9 under the substitution $a_{11}, a_{12}, ..., a_{r1}, a_{r2}$, because either $a_{j1}=a_{j2}=0$ for some variable y_j occurring in C_i, and then one of A_{is} will have rank 3, or matrix B_i will have rank 3.

Therefore, $k - \text{numb}(C, b_1, ..., b_r) \leqslant \text{rank } M(a_{11}, a_{12}, ..., a_{r1}, a_{r2}) - 8k$ which, combined with our earlier inequality, $\text{minrank}_S(M) \leqslant 9k - \text{maxnumb}(C)$, implies

$$\begin{aligned}
&\text{maxnumb}(C) - \text{numb}(C, b_1, ..., b_r) \\
&\quad \leqslant 9k - \text{minrank}_S(M) + \text{rank } M(a_{11}, a_{12}, ..., a_{r1}, a_{r2}) - k - 8k \\
&\quad = \text{rank } M(a_{11}, a_{12}, ..., a_{r1}, a_{r2}) - \text{minrank}_S(M),
\end{aligned}$$

which proves the assertion about β. ▮

THEOREM 24. *Let R be a commutative ring without zero divisors, and let $\{0, 1\} \subseteq S \subseteq R$ and $E = \{0, 1\}$. For $\varepsilon < \frac{7}{520} \approx 0.0134615$ there is no polynomial time solution for $(1 + \varepsilon)$-*APXMINRANK *unless $P = NP$.*

Proof. Combine Propositions 21 and 22 with Lemma 23. ▮

12. THE CASE WHEN EACH VARIABLE OCCURS EXACTLY ONCE

In previous sections we have been considering matrices $M = M(x_1, x_2, ..., x_t)$ with entries in $E \cup \{x_1, x_2, ..., x_t\}$, and each variable can occur arbitrarily often in M. In this section and the next, we restrict our attention to matrices where each variable occurs exactly once, and we call such matrices *eveo*.

DEFINITION. A polynomial $p(x_1, x_2, ..., x_t)$ is said to be *multi-affine* over a field F if every variable occurs with degree 0 or 1 in every term.

For example, $2xyz + 3z + 4x + 5$ is multi-affine over $\mathbb{Q}$. Note that the determinant of an eveo matrix is multi-affine. The following lemmas will prove useful.

LEMMA 25. *Let p be a multi-affine polynomial over a field F. Then p is identically zero over F iff p is the zero polynomial.*

Proof. This is implied by Lemma 1(i), when using $V = \{0, 1\}$ in the statement of that Lemma. ▮

Note that Theorem 25 is *not* necessarily true for polynomials in which variables occur with higher degree; for example, the polynomial $x^2 - x$ is not the zero polynomial, but it is identically zero over $GF(2)$.

COROLLARY 26. *A multi-affine polynomial is identically zero over a field F iff it is identically zero over some extension field $F' \supseteq F$.*

LEMMA 27. *A multi-affine function over a field is either constant or takes all values in the field.*

Proof. This may be proved by induction in the number of variables. We omit the details. ▮

THEOREM 28. *For all fields F, and all eveo matrices M, we can compute $\text{maxrank}_F(M)$ in random polynomial time.*

Proof. We mimic the proof of Theorem 2. Let M be an $n \times n$ eveo matrix. If the field F has at least $2n$ elements, then the proof goes through essentially unchanged, with V any subset of F of cardinality $2n$. Otherwise, choose an appropriate field extension F' with at least $2n$ elements. By Corollary 26 a minor is not identically zero over F' iff it is not identically zero over F, so we may compute maxrank over F' instead of over F. ▮

Now recall the singularity problem.

THEOREM 29. *If F is a field and M is an eveo matrix, then the decision problem* SING *is in the complexity class RP.*

Proof. By Lemmas 25 and 27, it is enough to ensure that the determinant $\det M$ is not a nonzero constant polynomial. Mimic the proof of Theorem 5, using Corollary 26, if necessary, to extend the base field. ▮

13. THE MINRANK PROBLEM FOR ROW-PARTITIONABLE MATRICES

In this section we show that the minrank problem is solvable in deterministic polynomial time if the matrix has a certain special form, in which each variable appears only once and there is a division between the variable and nonvariable entries.

More formally, let M be an $m \times n$ matrix with entries chosen from $E \cup \{x_1, x_2, ..., x_1\}$. We say that M is *row-partitionable* if

(a) each variable x_i occurs exactly once in M; and

(b) for each row i there exists an index k_i such that $a_{ij} \in E$ if $1 \leqslant j \leqslant k_i$, and $a_{ij} \notin E$ if $k_i \leqslant j \leqslant n$.

As an example, the matrix is row-partitionable:

$$M = \begin{bmatrix} 3 & 7 & -2 & x_1 & x_2 \\ 2 & 4 & x_3 & x_4 & x_5 \\ -3 & 5 & 6 & 2 & x_6 \\ 7 & 2 & 9 & 1 & 4 \end{bmatrix}.$$

The main motivation for this subproblem comes from the theory of *rational series*; for an introduction to this area see [1]. Let f be a formal power series in noncommuting variables over a field F. Then f is said to be *rational* if it can be expressed using the operations sum, product, and quasi-inverse (the map sending $x \mapsto 1/(1-x)$). The series f is said to be *recognizable* if the coefficient of the term corresponding to w (which is written as (f, w)) can be computed as follows: there is a matrix-valued homomorphism μ, a row matrix λ, and a column matrix γ such that $(f, w) = \lambda\mu(w)\gamma$. A well-known theorem due to Schützenberger (e.g., [1, Theorem 6.1]) proves that a formal power series is rational iff it is recognizable. In this case the dimension of the smallest possible matrix representation (the dimension of the square matrix $\gamma\lambda$) is an invariant called the *rank* of the rational series. The following problem now arises [9, 19]: given a (not necessarily rational) formal

power series f, compute the smallest possible rank $R_f(n)$ of any rational series agreeing with f on all terms of total degree at most n.

It can be shown that this number $R_f(n)$ is equal to the minrank of an associated Hankel-like matrix $M(f, n)$. More specifically, we have $R_f(n) = \text{minrank}_F(M(f, n))$, where the rows of $M(f, n)$ are labeled with words w of length $\leqslant n$, the columns are labeled with words x of length $\leqslant n$, and the entry in the row corresponding to w and the column corresponding to x is (f, wx) if $|wx| \leqslant n$, and it is a unique indeterminate otherwise. It is easy to see that this particular $M(f, n)$ is row-partitionable.

Consider the algorithm:

MR($M = (a_{ij})_{1 \leqslant i \leqslant m,\, 1 \leqslant j \leqslant n}$).

(1) rearrange rows so that $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_m$;

(2) if there exists u, $1 \leqslant u \leqslant k_1$, such that $a_{1u} \neq 0$,

$$\text{set } r \leftarrow 1;\ T \leftarrow \{1\}$$

else

$$\text{set } r \leftarrow 0;\ T \leftarrow \varnothing$$

(3) for $s = 2$ to m do

if the vector $(a_{s1}, a_{s2}, \ldots, a_{s,k_s})$ is not linearly dependent on $(a_{ij})_{i \in T,\, 1 \leqslant j \leqslant k_s}$

$$\text{set } r \leftarrow r + 1;\ T \leftarrow T \cup \{s\}$$

(4) return (r)

THEOREM 30. *Let F be a field. Then algorithm* MR *correctly computes* $\text{minrank}_F(M)$ *and uses* $O(m^3 n)$ *field operations.*

To prove correctness, we first observe that the reordering in step (1) cannot change $\text{minrank}_F(M)$. Next, we observe that the following invariants hold before the loop step corresponding to s is performed:

(a) for all possible assignments to the variables, the rows in the set T are linearly independent;

(b) for each assignment to the variables in the rows of T, there exists an assignment to the variables in the rows $\bar{T} = \{1, 2, \ldots, s-1\} - T$ such that each of the rows in $\bar{T}$ is dependent on a row of T.

These invariants clearly hold after step (2). We now prove by induction on s that they hold throughout the algorithm.

Suppose the invariants hold up to step $s-1$. At step s, we consider row s of M. If $(a_{s1}, \ldots, a_{s,k_s})$ is not dependent on $(a_{ij})_{i \in T,\, 1 \leqslant j \leqslant k_s}$, then for any assignment of the variables row s of M is not dependent on the rows in T, so by adding s to T we preserve part (a) of the invariant and part (b) is unaffected. If, on the other hand,

$a = (a_{s1}, ..., a_{s, k_s})$ is dependent on $M' = (a_{ij})_{i \in T, 1 \leqslant j \leqslant k_s}$, then write a as a linear combination of the rows of M'. We can then assign the variables in row s of M appropriately so that the entire row s is a linear combination of the rows of T. Then part (b) of the invariant is preserved and part (a) is unaffected. This completes the proof of correctness.

To complete the proof of the theorem, it suffices to observe that we can test to see if row s is dependent on the rows of T in at most $O(m^2 n)$ field operations, and this step is performed at most m times. ▌

ACKNOWLEDGMENTS

We thank Igor Shparlinski for discussions leading to the proof of Lemma 13. We thank Joachim von zur Gathen for suggesting a simplification of the proof of Lemma 17. Finally, we thank the referee for a careful reading of this paper.

REFERENCES

1. J. Berstel and C. Reutenauer, "Rational Series and Their Languages," EATCS Monographs on Theoretical Computer Science, Vol. 12, Springer-Verlag, New York/Berlin, 1988.
2. A. Borodin, J. von zur Garthen, and J. Hopcroft, Fast parallel matrix and GCD computations, *Inform. Control* **52** (1982), 241–256.
3. J. Canny, Some algebraic and geometric computations in PSPACE, *in* "Proc. Twentieth ACM Symp. Theor. Comput., 1988," pp. 460–467.
4. L. Egidi, The complexity of the theory of p-adic numbers, *in* "Proc. 34th Ann. Symp. Found. Comput. Sci., 1993," pp. 412–421, [*Comput. Complexity*, to appear].
5. H. B. Enderton, "A Mathematical Introduction to Logic," Academic Press, New York, 1972.
6. J. Friedman, A note on matrix rigidity, *Combinatorica* **13** (1993), 235–239.
7. F. Q. Gouvêa, "p-adic Numbers, An Introduction," Universitext, Springer-Verlag, New York/Berlin, 1993.
8. J. Håstad, Some optimal inapproximability results, *in* "Proc. Twenty-ninth Ann. ACM Symp. Theor. Comput., 1997," pp. 1–10.
9. C. Hespel, Approximation de séries formelles par des séries rationnelles, *RAIRO Inform. Théor.* **18** (1984), 241–258.
10. T. W. Hungerford, "Algebra," Graduate Text in Mathematics, Vol. 73, Springer-Verlag, New York/Berlin, 1987.
11. D. Ierardi, Quantifier elimination in the theory of an algebraically-closed field, *in* "Proc. Twenty-first Ann. ACM Symp. Theor. Comput., 1989," pp. 138–147.
12. P. Koiran, Hilbert's nullstellensatz is in the polynomial hierarchy, *J. Complexity* **12** (1996), 273–286.
13. S. Lang, "Algebra," Addison–Wesley, Reading, MA, 1971.
14. S. V. Lokam, Spectral methods for matrix rigidity with applications to size-depth tradeoffs and communication complexity, *in* "Proc. 36th Ann. Symp. Found. Comput. Sci., 1995," pp. 6–16, 1995.
15. Y. V. Matiyasevich, "Hilbert's Tenth Problem," The MIT Press, Cambridge, MA, 1993.
16. C. H. Papadimitriou and M. Yannakakis, Optimization, approximation, and complexity classes, *J. Comput. System Sci.* **43** (1991), 425–440.
17. J. Renegar, On the computational complexity and geometry of the first-order theory of the reals. part 1: Introduction. preliminaries. the geometry of semi-algebraic sets. the decision problem for the existential theory of the reals, *J. Symbolic Comput.* **13** (1992), 255–299.

18. J. T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, *J. Assoc. Comput. Mach.* **27** (1980), 701–717.
19. J. O. Shallit, On approximation by rational series in noncommuting variables, unpublished manuscript.
20. L. Valiant, Graph-theoretic arguments in low-level complexity, *in* "6th Mathematical Foundations of Computer Science," Lecture Notes in Computer Science, Vol. 197, pp. 162–176, Springer-Verlag, New York/Berlin, 1977.
21. L. G. Valiant, Completeness classes in algebra, *in* "Proc. Eleventh Ann. ACM Symp. Theor. Comput., 1979," pp. 249–261.
22. N. Cohen, C. R. Johnson, L. Rodman, and H. Woerdeman, Ranks of completions of partial matrices, The Gohberg anniversary collection, Vol. I (Calgary, AB, 1988), pp. 165–185, *Oper. Theory: Adv. Appl.*, Vol. 40, Birkhäuser, Basel, 1989.
23. P. J. Cohen, Decision procedures for real and p-adic fields, *Comm. Pure Appl. Math.* **22** (1969), 131–151.
24. J. Edmonds, Submodular functions, matroids and certain polyhedra, *in* "Combinatorial Structures and Their Applications" (R. Guy *et al.*, Eds.), pp. 69–87, Gordon & Breach, New York, 1970.
25. J. Edmonds, Systems of distinct representatives and linear algebra, *J. Res. Nat. Bur. Standards B* **71** (1967), 241–245.
26. L. Egidi, A quantifier elimination for the theory of p-adic numbers, *Comput. Complexity* **7** (1998), 205–263.
27. J. F. Geelen, Maximum rank matrix completion, *Linear Algebra Appl.* **288** (1999), 211–217.
28. D. J. Hartfiel and R. Loewy, A determinantal version of the Frobenius–König theorem, *Linear Multilinear Algebra* **16** (1984), 155–166.
29. M. Laurent, Matrix completion problems, *in* "Encyclopedia of Optimization" (C. A. Floudas and P. M. Pardalos, Eds.), Kluwer, Dordrecht, to appear.
30. L. Lovász, On determinants, matchings, and random algorithms, *in* "Fundamentals of Computation Theory, FCT '79" (L. Budach, Ed.), pp. 565–574, Akademie-Verlag, Berlin, 1979.
31. K. Murota, Mixed matrices: Irreducibility and decomposition, *in* "Combinatorial and Graph-Theoretical Problems in Linear Algebra" (R. A. Brunaldi, S. Friedland, and V. Klee, Eds.), pp. 39–71, Springer-Verlag, New York, 1993.
32. H. J. Woerdeman, "Matrix and Operator Extensions," CWI Tract 68, Centrum voor Wiskunde en Informatica, Amsterdam, 1989.