Comput Complexity 1 (1991), 67–90 1016-3328/91/010067-24 \$1.50+0.20

ON THE DECIDABILITY OF SPARSE UNIVARIATE POLYNOMIAL INTERPOLATION

ALLAN BORODIN AND PRASOON TIWARI

Abstract. We consider the problem of determining whether or not there exists a sparse univariate polynomial that interpolates a given set $S = \{(x_i, y_i)\}$ of points. Several important cases are resolved, e.g., the case when the x_i's are all positive rational numbers. But the general problem remains open.

Key words. Algorithm, decidability, interpolation, sparse polynomial, complexity.

Subject classifications. 68Q40.

1. Introduction

In this paper, we study a family of decision problems arising from sparse univariate polynomial interpolation through a given set of points. We use the following notion of sparsity throughout the paper.

DEFINITION. A polynomial $a(x) = \sum_{i=0}^{d} a_i x^i$ is said to be *t-sparse* if at most t of its coefficients $a_0, a_1, \ldots, a_d$ are nonzero.

A typical interpolation problem in this family is of the following form: Given a set $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, m\} \subset \mathsf{E}^2$ of m points, is there a t-sparse polynomial $c(x) = \sum_{i=1}^{t} c_i x^{e_i} \in \mathsf{F}[x]$, such that $c(x_i) = y_i$ for $i = 1, 2, \ldots, m$. Here, the e_i's are distinct non-negative integers, and E and F are one of $\mathbf{Q}$ or $\mathbf{Z}$. (In this paper, $\mathbf{Z}$ and $\mathbf{Q}$ denote the set of integers and the set of rational numbers, respectively.) We note that there is no apparent *a priori* bound on e_t.

Perhaps the best known problem in this family is the case $\mathsf{E} = \mathsf{F} = \mathbf{Q}$, $m = t$, and we are looking for a degree $t - 1$ polynomial through the given set of t points. It appears that the general problem has not been well studied. In fact, it is not at all clear that, stated in its generality, whether or not the sparse interpolation problem is a decidable problem; and if it *is* decidable, whether it is feasible. Of course, the problem is decidable if an *a priori* bound on the degree is available, but even in this case, it is not clear that the problem is feasible.

While the problem is of intrinsic mathematical interest, we are also motivated by the question whether there exists an efficient rational method (i.e., a method using only the operations $+, -, *, /$, and $>$) for sparse polynomial interpolation. We note that the black-box interpolation method of Ben-Or and Tiwari [1] requires root-finding (or at least the truncation operation). In contrast, the methods of Grigoriev and Karpinski [6] and Tiwari [12] are rational, but are not as efficient. Moreover, the black-box allows the user to evaluate the unknown polynomial at arbitrary points; this may not be possible in many situations. Of course, before constructing an interpolating polynomial, one has to decide whether such a polynomial exists. Since we do not impose an *a priori* bound on the degree of the interpolating polynomial, it is not clear that a decision procedure exists for the general problem.

Sparse univariate interpolation problems arise naturally in the context of learnability of sparse univariate polynomials [8]. Learnability of various concept classes in several models has been extensively studied in recent years, and many instances of learnable concept classes are known. However, in many cases, methods for constructing a consistent hypothesis are not known.

The paper is organized as follows. In Section 2, we state various interpolation problems in the family, and summarize their current status. In Section 3, we present a decision procedure for the case $\mathsf{E} = \mathbb{Z}$, $\mathsf{F} = \mathbb{Z}$, and $t \geq m$. This procedure is based on a very nice property of the Lagrange interpolation polynomial (see Theorem 3.1).

In Subsection 4.1, we present a decision procedure for the case $\mathsf{E} \in \{\mathbb{Q}, \mathbb{Z}\}$, $\mathsf{F} = \mathbb{Z}$, and $\#\{x_i \mid x_i > 1, i = 1, 2, \ldots, m\} \geq t$. This decision procedure is based on the fact (proved in Theorem 4.3 and Corollary 4.4) that the degree of *any* sparse interpolating integer polynomial is bounded from above. (Furthermore, this bound depends only on S and can be computed easily.) In Subsection 4.2, we study the case $\mathsf{E} = \mathsf{F} = \mathbb{Q}$, and present the main result of this paper. Note that Lagrange interpolation always provides a solution for the case $m \leq t$. In the case $m > t$, we present a decision procedure for the case $\#\{x_i \mid x_i > 0, i = 1, 2, \ldots, m\} > t$. As in Subsection 4.1, this decision procedure is based on the fact that the degree of *any* sparse interpolating integer polynomial is bounded from above. (Furthermore, this bound depends only on S and can be computed easily.) These degree bounds are established using a key property of the generalized Vandermonde matrices.

In contrast, we show in Section 5 that in some cases the sparse interpolating polynomial must necessarily have large degree. We conclude by listing several open problems in Section 6.

We believe that the general sparse polynomial interpolation problem (without any constraints on x_i's) is decidable. However, our methods fall short of establishing this assertion. We also believe that the case of negative x_i's is more than just a technicality; in fact, mathematically, it appears to be fundamentally different.

2. Discussion of some decision problems

We say that a problem is a proper decision problem if both positive and negative instances of the problem can be constructed. Descartes's rule of signs, discussed below, is used in this section to construct negative instances for several interpolation problems.

DEFINITION. Let $a_1, a_2, \ldots, a_m$ be a sequence of real numbers. This sequence is said to have a sign variation at position i if $a_i a_{i+1} < 0$. The total number of sign variations in a sequence is determined by dropping all the zeros from the sequence, and then counting the number of positions that have a sign variation.

Descartes' Rule of Signs [7]: Let $c(x) = \sum_{i=1}^{t} c_i x^{e_i}$, where all the c_i's are real and e_i's are distinct non-negative integers. Let n be the number of positive real roots of $c(x)$, counted with multiplicity, and let s be the number of sign variations in the sequence $\{c_1, c_2, \ldots, c_t\}$. Then $s - n$ is a non-negative even integer. (Note that this implies that the number of distinct positive real roots is strictly less than the sparsity t.)

In the following discussion, we establish that the family of decision problems under consideration contains some proper ones. In all cases, it is easy to construct a positive instance for any values of m and t by picking a t-sparse polynomial in $\mathsf{E}[x]$, and sampling it at any set of m sample points. Let us begin with the case when $\mathsf{E} = \mathsf{F} = \mathbf{Z}$; i.e., given a set $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, m\} \subset \mathbf{Z}^2$ of m integer points, determine if there is a t-sparse polynomial which passes through these points. A negative instance for any value of $t \geq 2$ and any m, is the set $\{(0,0), (2,1), \ldots\}$, where the unspecified values can be chosen arbitrarily. A less transparent negative instance for this problem, given by Corollary 3.3, is $\{(0,0), (2,0), (4,4)\}$.

Next, consider the case when $\mathsf{E} = \mathbf{Z}$ and $\mathsf{F} = \mathbf{Q}$. Every instance of this problem can be solved in a straightforward manner by Lagrange interpolation, provided $t \geq m$. In the case when $t < m$, Descartes' rule of signs implies that $\{(1,0), (2,0), \ldots, (t,0), (t+1,1)\}$ is a negative instance.

Now consider the case when $\mathsf{E} = \mathbf{Q}$ and $\mathsf{F} = \mathbf{Z}$. When $t < m$, then $\{(x_1,0), (x_2,0), \ldots, (x_t,0), (x_{t+1},1)\}$, for distinct positive x_i's, is a negative instance. When $t \geq m$, the negative instance for the case $\mathsf{E} = \mathsf{Z}$ and $\mathsf{F} = \mathsf{Z}$ (from the paragraph before last) is also a negative instance for this case. Finally, we consider the case $\mathsf{E} = \mathsf{F} = \mathsf{Q}$. This problem is always solvable for $t \geq m$ by Lagrange interpolation. When $t < m$, $\{(x_1, 0), (x_2, 0), \ldots, (x_t, 0), (x_{t+1}, 1)\}$, for distinct positive x_i's, is a negative instance.

Table 1 summarizes the current status of the family of decision problems studied in this paper.

The equivalence between P3 and P7 can be established as follows. Let $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, m\} \subset \mathsf{Q}^2$ be an instance of P7. Let λ be the least common multiple of the denominators of x_i's and y_i's. Define $\hat{x}_i = \lambda x_i$ and $\hat{y}_i = \lambda y_i$, and consider the instance of P3 given by $\hat{S} = \{(\hat{x}_i, \hat{y}_i) \mid i = 1, 2, \ldots, m\}$. Let $\hat{c}(x) = \sum_{i=1}^{t} c_i x^{e_i}$ be a solution for this instance of P3. Then, $c(x) = \sum_{i=1}^{t} c_i \lambda^{e_i - 1} x^{e_i}$ is a solution for the above mentioned instance of P7. Since every instance of P3 is also an instance of P7, P3 and P7 are equivalent problems.

In a similar manner, one can argue that P4 and P8 are equivalent problems.

3. A decision procedure for integer polynomials when $t \geq m$

In this section, we present an efficient decision procedure for the case when $\mathsf{E} = \mathsf{F} = \mathsf{Z}$ and $t \geq m$ (Problem P1). The corresponding problem P5 for $\mathsf{E} = \mathsf{Q}$ remains open.

3.1. The case $t \geq m$ and $\mathsf{E} = \mathsf{F} = \mathsf{Z}$. First consider the case $t = m$. Corollary 3.2 below is the key to an efficient decision procedure for this case.

THEOREM 3.1. *Given the set $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, t\} \subset \mathsf{Z}^2$, if there exists a polynomial with* integer *coefficients that interpolates all the points of S, then the degree $t - 1$ polynomial obtained by Lagrange interpolation has* integer *coefficients.*

PROOF. Let $f(x) = \sum_{i=0}^{e} f_i x^i \in \mathsf{Z}[x]$ be such that $f(x_i) = y_i$ for $i = 1, 2, \ldots, t$, and let $a(x)$ be the degree $t-1$ polynomial obtained by Lagrange interpolation. Define $p(x) = \prod_{i=1}^{t}(x - x_i)$. Observe that all coefficients of $p(x)$ are integers, and that its leading coefficient is one. Divide $f(x)$ by $p(x)$ to obtain $q(x)$ and $r(x)$ such that $f(x) = q(x)p(x) + r(x)$ where $r(x)$ is a polynomial of degree $t-1$ or less. Since the leading coefficient of $p(x)$ is one, all coefficients of $q(x)$ and

Problem	Type of sample points	Type of coefficients	t vs. m	Status
P1	Integer	Integer	$t \geq m$	A proper decision problem. Efficiently decidable. (See Subsection 3.1.)
P2	Integer	Integer	$t < m$	A proper decision problem. Decidable if t of the x_i's are at least two. (See Subsection 4.1.)
P3	Integer	Rational	$t \geq m$	Equivalent to P7 below.
P4	Integer	Rational	$t < m$	Equivalent to P8 below.
P5	Rational	Integer	$t \geq m$	A proper decision problem. Decidable if $t = m$ and $x_i > 1$ for all i. (See Subsections 3.2 and 4.1.)
P6	Rational	Integer	$t < m$	A proper decision problem. Decidable if t of the x_i's are greater than 1. (See Subsection 4.1.)
P7	Rational	Rational	$t \geq m$	Always solvable by Lagrange interpolation.
P8	Rational	Rational	$t < m$	A proper decision problem. Decidable if $t + 1$ of the x_i's are positive. (See Subsection 4.2.)

Table 1: Status of problems considered in this paper.

$r(x)$ are integers. Furthermore, $a(x)$ and $r(x)$ are both polynomials of degree at most $t-1$, and they agree on t points. Therefore, $a(x) = r(x)$. □

COROLLARY 3.2. *If the degree $m-1$ polynomial obtained on Lagrange interpolation through a set $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, m\} \subset \mathbf{Z}^2$ of m integer points does not have integer coefficients, then no polynomial with integer coefficients interpolates these points.*

This corollary yields a simple decision procedure for the case $t \geq m$: Perform Lagrange interpolation on the points in the set S. If the resulting polynomial has integer coefficients, then it is a t-sparse solution to our interpolation problem. Otherwise, this instance of the interpolation problem has no solution. Corollary 3.2 also yields a negative instance to the interpolation problem under consideration:

COROLLARY 3.3. *There is no polynomial with integer coefficients that passes through the points $(0,0)$, $(2,0)$, and $(4,4)$.*

PROOF. Lagrange interpolation through these points yields $\frac{1}{2}x(x-2)$. □

3.2. Further comments on interpolating an integer polynomial when $t \geq m$. In the following we show that there is a set S of integer points with the following properties: S has m points, there is a 1-sparse rational polynomial (namely, $\frac{1}{2}x^m$) that interpolates S, and there is an m-sparse polynomial with integer coefficients that interpolates S, but there is no t-sparse polynomial with integer coefficients that interpolates S, for $t < m$. Thus the sparsity guaranteed by Theorem 3.1 is optimal for this set S.

THEOREM 3.4. *For all m, the Lagrange interpolation on the point set $S = \{(2i, 2^{m-1}i^m) \mid i = 1, 2, \ldots, m\}$ yields a degree $m-1$ polynomial all of whose coefficients are nonzero integers.*

PROOF. Fix an arbitrary m for this proof. Let $p(x) = \sum_{i=0}^{m} 2^{m-i}\alpha_i x^i = \prod_{i=1}^{m}(x-2i)$. Note that all the α_i's are *nonzero integers*. In the rest of this proof, we show that the Lagrange interpolation through the point set S yields the degree $m-1$ polynomial $q(x) = -\sum_{i=0}^{m-1} 2^{m-i-1}\alpha_i x^i$. This is clearly sufficient in order to establish the theorem.

Define $r(x) = \frac{1}{2}x^m \bmod p(x)$. By comparing coefficients of the leading terms, it is easy to check that, in fact, $r(x) = \frac{1}{2}x^m - \frac{1}{2}p(x)$. Since all coefficients of $p(x)$, except the leading one, are nonzero even integers, all coefficients of

$r(x)$ are nonzero integers. Finally, since both $r(x)$ and $q(x)$ are degree $m-1$ polynomials, and they agree on m points, they are in fact the same polynomial. □

The following theorem asserts that the sparsity given by Theorem 3.4 is the best possible.

THEOREM 3.5. *For all m, and $t < m$, there is no t-sparse polynomial with integer coefficients, that passes through all points in the set $S = \{(2i, 2^{m-1}i^m) \mid i = 1, 2, \ldots, m\}$.*

PROOF. Suppose that $c(x) = \sum_{j=1}^{t} c_j x^{e_j}$ with $c_j \in \mathbf{Z}$ is a t-sparse polynomial such that $c(2i) = 2^{m-1}i^m$ for all i. Then

$$\begin{pmatrix} 2^{e_1} & 2^{e_2} & \ldots & 2^{e_t} \\ 4^{e_1} & 4^{e_2} & \ldots & 4^{e_t} \\ \vdots & \vdots & \vdots & \vdots \\ (2m)^{e_1} & (2m)^{e_2} & \ldots & (2m)^{e_t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{pmatrix} = 2^{m-1} \begin{pmatrix} 1 \\ 2^m \\ 3^m \\ \vdots \\ m^m \end{pmatrix}.$$

Let $\mathbf{A}$ denote the matrix on the left hand side of this equation, and let $\mathbf{c}$ and $\mathbf{b}$ denote the vectors on the left and the right hand side of this equation, respectively.

Now, two cases may arise: (1) none of the e_i's equal m; or (2) there exists i such that $e_i = m$. In the first case, by Lemma 4.1 (in Subsection 4.1), rank $\mathbf{A} = t < t+1 =$ rank $[\mathbf{A} \mid \mathbf{b}]$; and therefore the system $\mathbf{Ax} = \mathbf{b}$ has no solution. In the second case, assume that $e_i = m$, and let $\hat{\mathbf{A}}$ be the $m \times (t-1)$ matrix obtained by deleting the i^{th} column of $\mathbf{A}$. Similarly, let $\hat{\mathbf{c}}$ be the vector obtained by deleting the i^{th} component of $\mathbf{c}$. The vector $\hat{\mathbf{c}}$ is a solution to the linear system $\hat{\mathbf{A}}\mathbf{x} = \alpha\mathbf{b}$, where $\alpha = 2^m(\frac{1}{2} - c_i)$.

If $\alpha \neq 0$, then by an argument similar to that in the first case, this system has no solution. On the other hand, if $\alpha = 0$, then the i^{th} component of $\mathbf{c}$ is not an integer. □

It is worth pointing out that Theorem 3.1 does not hold if we replace the condition "$S \subset \mathbf{Z}^2$" by "$S \subset \mathbf{Q}^2$". Specifically, consider

$$f(x) = \frac{1}{2}x^{t-1} + \prod_{i=1}^{t}(x - x_i),$$

where $x_1 = 1/2$ and $x_i = 2i$ for $2 \leq i \leq t$. Then $f(x) \in \mathbb{Z}[x]$, and the points $(x_i, f(x_i))$ for $1 \leq i \leq t$ are interpolated by $\frac{1}{2}x^{t-1} \in \mathbb{Q}[x]$. Therefore, the methods of this section do not imply a decision procedure for Problem P5. However, the method of Corollary 4.4 (Subsection 4.1), will provide a (not so efficient) procedure for the case $\mathsf{E} = \mathbb{Q}$, $\mathsf{F} = \mathbb{Z}$ when $t \leq m$ and at least t of the sample points $x_1, x_2, \ldots, x_m$ are greater than one (or less than -1). Thus, in particular, Problem P5 is decidable when $t = m$ and all the sample points satisfy the condition just stated.

4. Decision procedures for integer and rational polynomials when $t < m$

In this section, we consider problems P2, P4, P6, and P8. We are able to provide decision procedures if sufficiently many of the sample points are positive (or negative). The situation when this condition is not met appears to be more complex as we discuss in Subsection 4.3.

4.1. A decision procedure for integer polynomials. In this subsection, we present a decision procedure for the case when $\mathsf{E} = \mathsf{F} = \mathbb{Z}$ and $x_i \geq 1$, i.e., $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, m\} \subset \mathbb{Z}_{\geq 1} \times \mathbb{Z}$, where $\mathbb{Z}_{\geq 1} = \{a \in \mathbb{Z} \mid a \geq 1\}$. We should also note that while the decidability result of Subsection 4.2 subsumes the result of this subsection, the results here are still useful for two reasons. First, our degree bound in the integer polynomial case is substantially better than the corresponding result for rational polynomials. Second, the degree bound in the integer polynomial case also applies when $t = m$ (thus providing a partial result for Problem P5).

Our decision procedure is based on the following two facts:

(i) Given an increasing sequence of non-negative integers $0 \leq e_1 < e_2 < \ldots < e_t$, there is an efficient procedure to check if there exists a polynomial $c(x) = \sum_{i=1}^{t} c_i x^{e_i}$, $c_i \in \mathbb{Z}$, such that $c(x_j) = y_j$ for $i = 1, 2, \ldots, m$.

(ii) Given the set S, one can compute an upper bound D on the degree of any polynomial having at most t nonzero integer coefficients which passes through the points of S.

Assuming facts (i) and (ii) above, the decision procedure can be stated easily:

Step 1: Compute the upper bound D;

Step 2: Try all possible degree sequences $0 \leq e_1 < e_2 < \ldots < e_t \leq D$ to determine if there is a polynomial $c(x) = \sum_{i=1}^{t} c_i x^{e_i}$, $c_i \in \mathbf{Z}$, passing through all the points of S.

In order to perform the test in (i) above, note that if such a polynomial exists, then the coefficients c_i's are a solution to the following system:

$$\begin{pmatrix} x_1^{e_1} & x_1^{e_2} & \ldots & x_1^{e_t} \\ x_2^{e_1} & x_2^{e_2} & \ldots & x_2^{e_t} \\ \vdots & \vdots & \vdots & \vdots \\ x_m^{e_1} & x_m^{e_2} & \ldots & x_m^{e_t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}. \tag{4.1}$$

Let us denote the matrix on the left hand side of this system of equations by $\mathbf{V}$. Matrix $\mathbf{V}$ is called a *generalized Vandermonde* matrix [4]. It is known (see Lemma 4.1 below) that any $t \times t$ minor of $\mathbf{V}$ is nonsingular provided $x_i > 0$ and $x_i \neq x_j$ for $i \neq j$. Therefore, the above system has at most one solution. If there is a solution, it can be easily determined and checked for integrality.

In order to compute the upper bound D of (ii) above, we will restrict our attention to the first t equalities in (4.1). Let $\mathbf{X}$ be the matrix consisting of the first t rows of $\mathbf{V}$. We will prove that for large enough value of e_t, the system

$$\mathbf{X} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{pmatrix}$$

has no *integer* solution. Observe that

$$c_t = \frac{\det \mathbf{Z}}{\det \mathbf{X}}, \tag{4.2}$$

where

$$\mathbf{Z} = \begin{pmatrix} x_1^{e_1} & x_1^{e_2} & \ldots & x_1^{e_{t-1}} & y_1 \\ x_2^{e_1} & x_2^{e_2} & \ldots & x_2^{e_{t-1}} & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_t^{e_1} & x_t^{e_2} & \ldots & x_t^{e_{t-1}} & y_t \end{pmatrix}. \tag{4.3}$$

Since $c_t \in \mathbf{Z}$ and $c_t \neq 0$, we have $|c_t| \geq 1$. We shall determine a bound D such that $e_t > D$ implies $|c_t| < 1$, thus proving that the degree of $c(x)$ is bounded by D.

Next, we use Lemma 4.1 stated below, in order to prove a lower bound on the determinant of generalized Vandermonde matrices.

LEMMA 4.1. *[4] There is a nonzero polynomial $P(\mathbf{x})$ in $x_1, x_2, \ldots, x_t$ with positive integer coefficients such that*

$$\frac{\det(\mathbf{X})}{\prod_{i>j}(x_i - x_j)} = P(\mathbf{x}).$$

COROLLARY 4.2. *If $0 \leq x_1 \leq x_2 \leq \cdots \leq x_t$, then*

$$\frac{\det \mathbf{X}}{\prod_{i>j}(x_i - x_j)} \geq x_t^{e_t-t+1} x_{t-1}^{e_{t-1}-t+2} \ldots x_2^{e_2-1} x_1^{e_1}.$$

We will use the following lexicographic ordering of the monomials in the proof of Corollary 4.2.

DEFINITION. Let $\{x_t^{i_t} x_{t-1}^{i_{t-1}} \ldots x_1^{i_1} \mid i_j \geq 0\}$ be the set of all the monomials in $x_t, x_{t-1}, \ldots, x_1$. A monomial $x_t^{i_t} x_{t-1}^{i_{t-1}} \ldots x_1^{i_1}$ is said to be lexicographically greater than another monomial $x_t^{j_t} x_{t-1}^{j_{t-1}} \ldots x_1^{j_1}$ if and only if there exists some l such that $i_t = j_t, \ldots, i_{l+1} = j_{l+1}$, and $i_l > j_l$.

PROOF OF COROLLARY 4.2. In order to prove this corollary, it is sufficient to prove that $x_t^{e_t-t+1} x_{t-1}^{e_{t-1}-t+2} \ldots x_2^{e_2-1} x_1^{e_1}$ is a term in the polynomial $P(\mathbf{x})$ with coefficient at least one.

The lexicographically largest monomial in $\det \mathbf{X}$ is the product of the lexicographically largest monomials in $\prod_{i>j}(x_i - x_j)$ and $P(\mathbf{x})$, respectively. The lexicographically largest monomials in $\det \mathbf{X}$ and $\prod_{i>j}(x_i - x_j)$ are $x_t^{e_t} x_{t-1}^{e_{t-1}} \ldots x_1^{e_1}$ and $x_t^{t-1} x_{t-1}^{t-2} \ldots x_2^1$ respectively. Therefore, $x_t^{e_t-t+1} x_{t-1}^{e_{t-1}-t+2} \ldots x_2^{e_2-1} x_1^{e_1}$ is the lexicographically largest monomial in $P(\mathbf{x})$. Moreover, by comparing coefficients, it is clear that this monomial appears in $P(\mathbf{x})$ with coefficient 1. □

Returning to the interpolation problem at hand, using Corollary 4.2, now we prove the existence of an upper bound D on the degree of any interpolating polynomial. Moreover, D can be computed easily.

THEOREM 4.3. *If $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, t\} \subset \mathbb{Z}_{\geq 2} \times \mathbb{Z}$ and $c(x)$ is a t-sparse polynomial in $\mathbb{Z}[x]$ that passes through all the points of S, then*

$$\deg c(x) \leq D = \log_2 \beta + t^2 \log_2 \alpha + 2,$$

where $\alpha = \max_i x_i$, and $\beta = \max_i |y_i|$.

PROOF. Let $c(x) = \sum_{i=1}^{t} c_i x^{e_i}$. Consider (4.2) and assume, without loss of generality, that $x_t > x_{t-1} > \ldots > x_1 > 1$. By the definition of $\mathbf{Z}$ in (4.3),

$$|\det \mathbf{Z}| \leq (t!)\beta x_t^{e_{t-1}} x_{t-1}^{e_{t-2}} \ldots x_2^{e_1}.$$

Therefore, by (4.2) and Corollary 4.2,

$$\begin{aligned} 1 \le |c_t| &\le \frac{(t!)\beta x_t^{e_{t-1}} x_{t-1}^{e_{t-2}} \dots x_2^{e_1}}{\prod_{i>j}(x_i - x_j) x_t^{e_t - t+1} x_{t-1}^{e_{t-1}-t+2} \dots x_2^{e_2-1} x_1^{e_1}} \\ &\le \frac{3\beta}{x_t^{e_t - e_{t-1} - t+1} x_{t-1}^{e_{t-1} - e_{t-2} - t+2} \dots x_2^{e_2 - e_1 - 1} x_1^{e_1}} \end{aligned} \tag{4.4}$$

Here, we have used the fact that, for any t and any set of t increasing integers x_i, $(t!) \le 3\prod_{i>j}(x_i - x_j)$. Since $x_i \ge 2$ for $i = 1, 2, \dots, t$, we have

$$\begin{aligned} 1 &\le |c_t| \le \frac{3\beta\alpha^{t^2}}{2^{e_t}}, \\ e_t &\le \log_2 \beta + t^2 \log_2 \alpha + 2. \quad \square \end{aligned}$$

For $b = 0$ or $b = 1$, we write $\mathbb{Q}_{>b} = \{a \in \mathbb{Q} \mid a > b\}$.

COROLLARY 4.4. *Given a set $S = \{(x_i, y_i) \mid i = 1, 2, \dots, t\} \subset \mathbb{Q}_{>1} \times \mathbb{Q}$, there exists a bound D such that if $c(x)$ is a t-sparse polynomial in $\mathbb{Z}[x]$ that passes through all the points of S, then*

$$\deg c(x) \le D = \log_\delta \beta + t^2 \log_\delta \alpha + \log_\delta(t!) - \log_\delta \left\{ \prod_{i>j} \mid x_i - x_j \mid \right\};$$

where $\alpha = \max_i x_i$, $\beta = \max_i \mid y_i \mid$, and $\delta = \min_i x_i$.

PROOF. Let $c(x) = \sum_{i=1}^{t} c_i x^{e_i}$. Instead of (4.4) in the proof of Theorem 4.3, we have

$$1 \le |c_t| \le \frac{(t!)\beta\alpha^{t^2}}{\delta^{e_t} \prod_{i>j} \mid x_i - x_j \mid}. \quad \square$$

4.2. A decision procedure for rational polynomials. In this section, we study the case $\mathsf{E} = \mathsf{F} = \mathbb{Q}$. Since the case $t \ge m$ is easily solved using Lagrange interpolation, we restrict our attention to the case $t < m$.

We present a decision procedure for the case $x_i > 0$ for $i = 1, 2, \dots m$ (and $t < m$). As in Subsection 4.1, this decision procedure is based on an upper bound on the degree of sparse interpolating polynomials.

THEOREM 4.5. *Given a set $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, t+1\} \subset \mathbb{Q}_{>0} \times \mathbb{Q}$, there exists a bound D such that if $c(x)$ is a t-sparse polynomial in $\mathbb{Q}[x]$ that interpolates the set S, then*

$$\deg c(x) \leq D = \left(t + \log_\gamma \left[(t!)2^{t+1}\lambda\beta(\alpha\lambda)^{t^2}\right]\right)\left(1 + \log_\gamma[\alpha\lambda]\right)^{t-1},$$

where $\alpha = \max_i x_i$, $\beta = \max_i |y_i|$, $\gamma = \min\{x_j/x_i : x_j > x_i\}$, and λ is the least common multiple of the denominator of x_i's and y_i's (in the reduced form).

PROOF. Let $c(x) = \sum_{i=1}^{t} c_i x^{e_i} \in \mathbb{Q}[x]$ where $0 \leq e_1 < e_2 < \cdots < e_t$, and $c_t \neq 0$. In this proof, we compute bounds D_i, depending only on S and t, such that $e_i \leq D_i$ for $1 \leq i \leq t$. Suppose that we have already computed bounds $D_1, D_2, \ldots, D_{k-1}$ where $1 \leq k \leq t$. ($k = 1$ corresponds to the case when we have not computed any of the bounds.) We now describe a method for determining D_k such that $e_k \leq D_k$.

Define

$$\mathbf{Z} = \begin{pmatrix} y_1 & x_1^{e_1} & x_1^{e_2} & \cdots & x_1^{e_t} \\ y_2 & x_2^{e_1} & x_2^{e_2} & \cdots & x_2^{e_t} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{t+1} & x_{t+1}^{e_1} & x_{t+1}^{e_2} & \cdots & x_{t+1}^{e_t} \end{pmatrix}. \tag{4.5}$$

Notice the $\mathbf{Z}$ is a $(t+1) \times (t+1)$ matrix and $\det \mathbf{Z} = 0$, since $c(x_i) = y_i$ for $i = 1, 2, \ldots, t+1$. Our aim is to find D_k such that $e_k > D_k$ implies $\det \mathbf{Z} \neq 0$. For this purpose, we need the following definition.

DEFINITION. Given a sequence $0 < i_1 < i_2 < \cdots < i_l \leq t+1$, let $\overline{i_1, i_2, \ldots, i_l}$ denote the sequence obtained upon deleting $i_1, i_2, \ldots, i_l$ from the sequence $1, 2, \ldots, t+1$. Given two sequences of integers $0 < i_1 < i_2 < \cdots < i_l \leq t+1$ and $0 < j_1 < j_2 < \cdots < j_q \leq t+1$, and a matrix $\mathbf{A}$, let $\mathbf{A}_{j_1, j_2, \ldots, j_q}^{i_1, i_2, \ldots, i_l}$ denote the submatrix of $\mathbf{A}$ consisting of rows $i_1, i_2, \ldots, i_l$ and columns $j_1, j_2, \ldots, j_q$. Note that $\mathbf{A}_{j_1, j_2, \ldots, j_q}^{\overline{i_1, i_2, \ldots, i_l}}$ denotes the submatrix of $\mathbf{A}$ consisting of rows *other* than $i_1, i_2, \ldots, i_l$ and columns $j_1, j_2, \ldots, j_q$.

We also make use of the Laplace expansion of the determinant [10]. It states that:

$$\det \mathbf{A} = (-1)^{\frac{k(k-1)}{2}} \sum_{0 < i_1 < i_2 < \cdots < i_k \leq t+1} (-1)^{\sum_{j=1}^{k} i_j} \det \mathbf{A}_{1,2,\ldots,k}^{i_1, i_2, \ldots, i_k} \det \mathbf{A}_{k+1, k+2, \ldots, t+1}^{\overline{i_1, i_2, \ldots, i_k}},$$

for any $(t+1) \times (t+1)$ matrix $\mathbf{A}$. Applying this identity to $\mathbf{Z}$ we obtain:

$$\det \mathbf{Z} = \sum_{0 < i_1 < i_2 < \cdots < i_k \leq t+1} \pm \det \mathbf{Z}_{1,2,\ldots,k}^{i_1, i_2, \ldots, i_k} \det \mathbf{Z}_{k+1, k+2, \ldots, t+1}^{\overline{i_1, i_2, \ldots, i_k}}. \tag{4.6}$$

Returning to the proof, let λ denote the least common multiple of the denominators of the x_i's and the y_i's. Then, for any sequence $0 < i_1 < i_2 \cdots < i_k < t+1$, either $\det \mathbf{Z}_{1,2,\ldots,k}^{i_1,i_2,\ldots,i_k} = 0$ or $\mid \det \mathbf{Z}_{1,2,\ldots,k}^{i_1,i_2,\ldots,i_k} \mid \geq \left(\frac{1}{\lambda}\right)^{1+\sum_{j=1}^{k-1} D_j}$.

Without loss of generality, assume that $0 < x_1 < x_2 < \cdots < x_{t+1}$. We now argue that there exists D_k such that if $e_k > D_k$, then there is a term on the right hand side of (4.6) that dominates the sum of the absolute values of all the other terms. In fact, this dominant term can be identified as follows: let $r_0 = 0$, and for $i > 0$ let r_i be the rank of $\mathbf{Z}_{1,2,\ldots,k}^{1,2,\ldots,i}$. If $r_{t+1} < k$, then the first k columns of $\mathbf{Z}$ are linearly dependent, and the last $k-1$ are linearly dependent by Lemma 4.1; therefore, there is a $(k-1)$-sparse polynomial, say $p(x)$, with exponents $e_1, e_2, \ldots, e_{k-1}$, that interpolates the set S. But then $c_t = 0$ because $c(x) - p(x)$ is the zero polynomial. (It is t-sparse, and it has $(t+1)$ positive zeros.) This contradicts the fact that $c_t \neq 0$, and implies that $r_{t+1} = k$. Next, define u_j such that $r_{u_j} = j$ but $r_{u_j - 1} = j - 1$. Let $\mathbf{W}$ denote the $k \times k$ submatrix $\mathbf{Z}_{1,2,\ldots,k}^{u_1,u_2,\ldots,u_k}$. By the choice of u_j's, $\det \mathbf{W} \neq 0$. In the rest of this proof, we establish that the term corresponding to the sequence $u_1, u_2, \ldots, u_k$ is the dominant term on the right hand side of (4.6).

Define $v_1, v_2, \ldots, v_{t+1-k}$ to be the sequence $\overline{u_1, u_2, \ldots, u_k}$. The term in (4.6), corresponding to the sequence $u_1, u_2, \ldots, u_k$, equals
$\det \mathbf{W} \det \mathbf{Z}_{k+1,k+2,\ldots,t+1}^{v_1,v_2,\ldots,v_{t+1-k}}$. By Lemma 4.1,

$$\det \mathbf{Z}_{k+1,k+2,\ldots,t+1}^{v_1,v_2,\ldots,v_{t+1-k}} = P(x_{v_1}, x_{v_2}, \ldots, x_{v_{t+1-k}}) \prod_{i>j} (x_{v_i} - x_{v_j}),$$

where $P(z_1, z_2, \ldots, z_{t+1-k})$ is a nonzero polynomial with *positive* integer coefficients. Also note that $P(z_1, z_2, \ldots, z_{t+1-k})$ is a homogeneous polynomial of degree $(\sum_{i=k}^{t} e_i) - \binom{t+1-k}{2}$, and that all exponents appearing in $P(z_1, \ldots, z_{t+1-k})$ are at least as large as $e_k - (t-k)$.

Suppose that $\det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k} \det \mathbf{Z}_{k+1,k+2,\ldots,t+1}^{\overline{a_1,a_2,\ldots,a_k}}$ is *another nonzero* term on the right hand side of (4.6); and let $b_1, b_2, \ldots, b_{t+1-k}$ be the sequence $\overline{a_1, a_2, \ldots, a_k}$. Observe that if $u_i > a_i$ for any i, then $\det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k} = 0$. Since $\det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k} \neq 0$, $u_i \leq a_i$ and therefore, $v_j \geq b_j$. Since $(u_1, u_2, \ldots, u_k) \neq (a_1, a_2, \ldots, a_k)$, there exists q such that $u_q < a_q$. Therefore, there exists a $\hat{q}$ such that $v_{\hat{q}} > b_{\hat{q}}$.

In order to complete the proof, we now establish that, for sufficiently large value of e_k, the following inequality holds:

$$\mid \det W \det \mathbf{Z}_{k+1,k+2,\ldots,t+1}^{v_1,v_2,\ldots,v_{t+1-k}} \mid \geq \binom{t+1}{k} \mid \det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k} \det \mathbf{Z}_{k+1,k+2,\ldots,t+1}^{b_1,b_2,\ldots,b_{t+1-k}} \mid .$$

But we can rewrite the last inequality as follows:

$$\frac{P(x_{v_1}, x_{v_2}, \ldots, x_{v_{t+1-k}})}{P(x_{b_1}, x_{b_2}, \ldots, x_{b_{t+1-k}})} \geq \binom{t+1}{k} \frac{\prod_{i>j}(x_{b_i} - x_{b_j})}{\prod_{i>j}(x_{v_i} - x_{v_j})} \frac{|\det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k}|}{|\det W|}$$

Consider a monomial $c \cdot z_1^{s_1} z_2^{s_2} \cdots z_{t+1-k}^{s_{t+1-k}}$ in $P(z_1, z_2, \ldots, z_{t+1-k})$. We establish the last inequality by showing that, for sufficiently large e_k, the following holds:

$$\frac{x_{v_1}^{s_1} x_{v_2}^{s_2} \cdots x_{v_{t+1-k}}^{s_{t+1-k}}}{x_{b_1}^{s_1} x_{b_2}^{s_2} \cdots x_{b_{t+1-k}}^{b_{t+1-k}}} \geq \binom{t+1}{k} \lambda^{(1+\sum_{i=1}^{k-1} D_i)} \frac{\prod_{i>j}(x_{b_i} - x_{b_j})}{\prod_{i>j}(x_{v_i} - x_{v_j})} \; |\det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k}|$$

Now $\frac{x_j}{x_i} \geq \gamma$ whenever $j > i$, there exists q such that $v_q > b_q$, and $s_i > e_k - t$. Hence

$$\left(\frac{x_{v_1}}{x_{b_1}}\right)^{s_1} \left(\frac{x_{v_2}}{x_{b_2}}\right)^{s_2} \cdots \left(\frac{x_{v_{t+1-k}}}{x_{b_{t+1-k}}}\right)^{s_{t+1-k}} \geq \gamma^{(e_k - t)}.$$

Since $|\det \mathbf{Z}_{1,2,\ldots,k}^{a_1,a_2,\ldots,a_k}| \leq (t!)(\alpha^{(\sum_{i=1}^{k-1} D_i)} \beta)$, it is sufficient to establish

$$\gamma^{(e_k - t)} \geq (t!) \binom{t+1}{k} \lambda^{(1+\sum_{i=1}^{k-1} D_i)} \frac{\prod_{i>j}(x_{b_i} - x_{b_j})}{\prod_{i>j}(x_{v_i} - x_{v_j})} (\alpha^{(\sum_{i=1}^{k-1} D_i)} \beta).$$

Therefore, it is sufficient to choose

$$D_k = t + \log_\gamma \left\{ (t!) 2^{t+1} \lambda \beta \, (\alpha\lambda)^{t^2 + \sum_{i=1}^{k-1} D_i} \right\}.$$

This is a recurrence of the form

$$A_k = B(\sum_{i=1}^{k-1} A_i) + C.$$

It is easy to check that $A_k = (1+B)^{k-1} C$ solves this recurrence. Therefore,

$$D_k = \left(t + \log_\gamma \left[(t!) 2^{t+1} \lambda \beta (\alpha\lambda)^{t^2}\right]\right) \left(1 + \log_\gamma [\alpha\lambda]\right)^{k-1}$$

is the desired bound. □

Observe that we are unable to establish whether this rational polynomial interpolation problem is in NP, whereas, by the results of Subsection 4.1, the corresponding problem for integer polynomials is in NP.

Finally, we observe that although Theorems 4.3 and 4.5 guarantee that there are only finitely many interpolating polynomials for the cases considered (see also Subsection 4.3), it is certainly not the case that there is at most one t-sparse interpolating polynomial as we shall now observe.

Given positive $x_1, x_2, \ldots, x_m$, $e_1, e_2, \ldots, e_t$, and $f_1, f_2, \ldots, f_t$, $e_i \neq f_j$, we describe a method for constructing $y_1, y_2, \ldots, y_m$ such that there exist two polynomials $c(x) = \sum_{i=1}^t c_i x^{e_i}$, and $\hat{c}(x) = \sum_{i=1}^t \hat{c}_i x^{f_i}$ satisfying $c(x_i) = \hat{c}(x_i) = y_i$. Define

$$\mathbf{u}_i = \begin{bmatrix} x_1^{e_i} \\ x_2^{e_i} \\ \vdots \\ x_m^{e_i} \end{bmatrix}, \qquad \mathbf{v}_i = \begin{bmatrix} x_1^{f_i} \\ x_2^{f_i} \\ \vdots \\ x_m^{f_i} \end{bmatrix} \in \mathbf{Q}^m.$$

The set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_t\}$ is linearly dependent when $2t > m$. (When $m \geq 2t$, Descartes' rule of signs shows that there can only be one solution if all sample points are positive.) Therefore, there exist α_i and β_i, $i = 1, 2, \ldots, m$, not all zero, satisfying

$$\sum_{i=1}^t \alpha_i \mathbf{u}_i + \sum_{i=1}^t \beta_i \mathbf{v}_i = 0.$$

Define $\mathbf{y} = \sum_{i=1}^t \alpha_i \mathbf{u}_i = -\sum_{i=1}^t \beta_i \mathbf{v}_i$. (Note that $\mathbf{y}$ is in the intersection of two t-dimensional subspaces.) Since $\mathbf{u}_i$'s (and $\mathbf{v}_i$'s) are linearly independent, $\mathbf{y} \neq 0$. Choose y_i to be the i^{th} component of $\mathbf{y}$, $c_i = \alpha_i$, $\hat{c}_i = -\beta_i$, and check that the desired conditions are satisfied. In fact, when $m = t + k$ for any $k \leq t - 1$, one can construct a $\mathbf{y}$ such that there exist $\ell = \lfloor \frac{t-1}{k} \rfloor + 1$ interpolating polynomials by choosing $\mathbf{y}$ to be in the intersection of ℓ subspaces, each of dimension t.

4.3. Comments on sparse interpolation when the sample points have mixed signs. The decision procedure of Theorem 4.5 requires that at least $t+1$ of the sample points be positive. It is also sufficient that at least $t+1$ of the sample points be negative. This follows since Theorem 4.5 is based on the singularity of the matrix $\mathbf{Z}$ in (4.5), and if $x_1, x_2, \ldots, x_{t+1}$ are all negative, then we can replace each x_i by $-x_i$ without changing the absolute value of $\det \mathbf{Z}$. Thus, the singularity is not affected and the degree bound D still applies. A similar remark applies to the results of Subsection 4.1.

Theorem 4.5 has a purely mathematical (i.e., non-computational) consequence. Namely, when $t < m$, and at least $t + 1$ of the sample points are positive (or negative) then there are only a finite number of t-sparse interpolating polynomials. (The degree bound implies that we need to consider only finitely many exponent sequences. Since the generalized Vandermonde matrix is nonsingular, each exponent sequence can yield at most one interpolating polynomial.) In contrast, if $t = m$ and $\{x_i\}$ are positive, then every exponent sequence again yields a nonsingular generalized Vandermonde matrix and hence a rational interpolating polynomial, assuming at least one y_i is not equal to zero; that is, there are infinitely many t $(=m)$ sparse interpolating polynomials.

We now show that the situation is quite different when the positivity (or negativity) constraint on the sample points is not met. The simplest example is when $m = t + 1$, and $x_1 = 0$, $y_1 = 0$, and all the remaining x_i's are positive (or negative). This implies $e_1 \neq 0$ but otherwise, any exponent sequence can be used to interpolate the given set of points.

As another example, we construct a point set S with (say) one negative x_i, such that one can find *sparse* polynomials of arbitrarily large degree passing through the points of S. Again, let $m = t + 1$. Choose $x_i \in \mathbb{Q}$ such that $x_i \neq 0$, $x_i \neq x_j$, and $x_i \neq -x_j$ for $i \neq j$, for $i, j = 1, 2, \ldots t$. In addition, choose $y_i \in \mathbb{Q}$, for $i = 1, 2, \ldots, t$ such that at least one of them is nonzero. Set $x_{t+1} = -x_t$ and $y_{t+1} = y_t$. Let $S = \{(x_i, y_i) \mid i = 1, 2, \ldots, t + 1\}$ and let $e_1 < e_2 < \ldots < e_{t-1} < e_t$ be any sequence of positive *even* integers. Then the system in (4.1) has a solution.

Given the previous example, one might conjecture that there are only finitely many solutions, as long as at least $t + 1$ of the sample points have different absolute values. We now show that even in this case we can construct a set S such that there are infinitely many integer polynomials that interpolate S.

Define

$$\mathbf{A} = \begin{pmatrix} (-2)^n & (-2)^{n+2} & (-2)^{n+3} \\ 3^n & 3^{n+2} & 3^{n+3} \\ 6^n & 6^{n+2} & 6^{n+3} \end{pmatrix}.$$

It is easy to check that $\det \mathbf{A} = 0$. This simple example shows that a generalized Vandermonde determinant may be zero even when $|x_i| \neq |x_j|$ for $i \neq j$. Moreover, matrix $\mathbf{A}$ shows that determinants of this form are singular even for arbitrarily large values of exponents.

Let $x_1 = -2$, $x_2 = 3$, $x_3 = 6$, and $y_1 = y_2 = y_3 = 0$. In addition, let x_4 be any positive rational number, choose $y_4 \neq 0$, and define $S = \{(x_i, y_i) \mid i = 1, 2, 3, 4\}$. For any non-negative integer n, set $e_1 = n$, $e_2 = n + 2$, and $e_3 = n + 3$. Then there is a polynomial $c(x) = \sum_{i=1}^{3} c_i x^{e_i}$ that interpolates the point set S because

$$\det\begin{pmatrix} y_1 & x_1^{e_1} & x_1^{e_2} & x_1^{e_3} \\ y_2 & x_2^{e_1} & x_2^{e_2} & x_2^{e_3} \\ y_3 & x_3^{e_1} & x_3^{e_2} & x_3^{e_3} \\ y_4 & x_4^{e_1} & x_4^{e_2} & x_4^{e_3} \end{pmatrix} = 0.$$

In light of these examples, it is not clear whether the degree of *some* sparse polynomial interpolating a set S can be bounded from above by an easily computable function of the coordinates contained in the set S.

5. Sparse interpolating polynomials may necessarily have large degree

The decision procedures of Section 4 are based on an upper bound on the degree of the candidate polynomials. As a consequence of Theorem 5.1 and Corollary 5.4 below, we will establish that in some cases, the degree of the minimum degree interpolating polynomial is large enough, so that any algorithm based on degree bounds alone (e.g., the algorithms of Section 4) would require exponential time.

THEOREM 5.1. *For any t, $d \geq t+1$, and t-sparse polynomial $c(x) = \sum_{i=1}^{t} c_i x^{e_i}$ in $\mathbb{Q}[x]$ of degree d there exists a point set $S = \{(x_i, c(x_i)) \mid i = 1, 2, \ldots, t+1\}$ such that no t-sparse polynomial of degree less than d interpolates S.*

PROOF. Choose a set $S = \{(x_i, c(x_i)) \mid i = 1, 2, \ldots, t+1\}$ arbitrarily. (Later, we will need to set the x_i's appropriately.) Suppose that $a(x) = \sum_{i=0}^{d} a_i x^i$ is a t-sparse polynomial of degree less than d that interpolates the point-set S. (Note that at most t of the a_i's are nonzero and $a_d = 0$.) We will need the dense representation of $c(x)$; let $c(x) = \sum_{i=0}^{d} f_i x^i$. Note that $c_t = f_d$, and that at most t of the f_i's are nonzero.

As in the proof of Theorem 3.1, $c(x) - a(x) = (\sum_{i=0}^{r} b_i x^i)\prod_{i=1}^{t+1}(x - x_i)$, where $r = d - t - 1$. Recall that $\prod_{i=1}^{t+1}(x - x_i) = \sum_{i=0}^{t+1} \sigma_{t+1-i} x^i$ where σ_i is the degree i elementary symmetric function in the x_j's multiplied by $(-1)^i$; with $\sigma_0 = 1$. Equating the coefficients of x^k on both sides, we get

$$f_k - a_k = \sum_{l=0}^{\min\{k,t+1\}} \sigma_{t+1-l} b_{k-l}; \text{ for } k = 0, 1, \ldots, d.$$

This set of equations can be written as $\mathbf{Ab} = \mathbf{f}$ where $\mathbf{b}$ is a $(r+1)$-component column vector given by $\mathbf{b}_i = b_{r+1-i}$, for $i = 1, 2, \ldots, r+1$; $\mathbf{f}$ is a $(d+1)$-component column vector given by $\mathbf{f}_i = f_{d+1-i} - a_{d+1-i}$, for $i = 1, 2, \ldots, d+1$; and $\mathbf{A}$ is a $(d+1) \times (r+1)$ matrix given by

$$\mathbf{A}_{ij} = \begin{cases} \sigma_{i-j} & \text{if } 0 \le i-j \le t+1, \\ 0 & \text{otherwise,} \end{cases}$$

where $i = 1, 2, \ldots, d+1$ and $j = 1, 2, \ldots, r+1$.

Observe that at most t rows (components) of the vector $\mathbf{f}$ depend on the nonzero coefficients of $a(x)$. Let $\mathbf{H}$ and $\mathbf{c}$ be the matrix and the vector obtained by deleting these rows from $\mathbf{A}$ and $\mathbf{f}$, respectively. (If $l < t$ rows were deleted in this process, then delete an additional $t - l$ rows from the bottom.) Then $\mathbf{H}$ is an $(r+2) \times (r+1)$ matrix, $\mathbf{c}$ is a $(r+2)$-component vector, and the vector $\mathbf{b}$ satisfies $\mathbf{Hb} = \mathbf{c}$. Notice that this system is *independent* of the actual values of the coefficients of $a(x)$, but it does *depend* on the set of exponents that appear with nonzero coefficients in $a(x)$. Depending on this set, we will end up with one of $\binom{d}{t}$ linear systems; the following argument is applicable to any of these $\binom{d}{t}$ systems.

Notice that the matrix $\mathbf{H}$ has a nice structure. In particular, if the k_i's are such that $\mathbf{H}_{i+1,i} = \sigma_{k_i}$, then it is easy to see that the matrix $\mathbf{H}$ is completely specified by the sequence $0 < k_1 \le k_2 \le \cdots \le k_{r+1}$.

The next step in the proof is to show that one can choose the x_i's so that the linear system $\mathbf{Hz} = \mathbf{c}$ is inconsistent. In order to prove this last assertion, it is sufficient to prove that $\det(\mathbf{G})$ is a nonzero polynomial in the x_i's, where $\mathbf{G}$ is the $(r+2) \times (r+2)$ matrix obtained by appending the vector $\mathbf{c}$ to the matrix $\mathbf{H}$ as the last column.

Define $k = \sum_{i=1}^{r+1} k_i$. Next, we show that the cofactor of the i^{th} entry in the last column of $\mathbf{G}$ is a nonzero homogeneous polynomial in the x_i's. We need the following definition in order to proceed with the proof.

DEFINITION. Given a vector $\mathbf{k} = [u_1, u_2, \ldots, u_n]$ of non-negative integers such that $0 \le u_1 \le u_2 \le \cdots \le u_n \le t+1$, define the $n \times n$ matrix $\mathbf{S}^{\mathbf{u}}$ by

$$\mathbf{S}^{\mathbf{u}}_{ij} = \sigma_{u_i + i - j}.$$

Continuing the proof, it is now easy to check that the cofactor of the i^{th} entry in the last column of $\mathbf{G}$ is given by:

$$Cof(\mathbf{G}_{i,r+2}) = \begin{cases} (-1)^{r+3} \det(\mathbf{S}^{[k_1, k_2, \ldots, k_{r+1}]}), \\ (-1)^{r+2+i} \det(\mathbf{S}^{[0, k_1 - 1, k_2 - 1, \ldots, k_{i-2} - 1, k_i, k_{i+1}, \ldots, k_{r+1}]}), \end{cases}$$

where the top entry is valid for $i = 1$, and the bottom entry for $2 \leq i \leq r+1$. Therefore, by Lemma 5.2 below, the cofactor of the i^{th} element in the last column of $\mathbf{G}$ is a nonzero homogeneous polynomial in the x_i's of total degree k and $k-(i-2)-k_{i-1}$, for $i = 1$ and $2 \leq i \leq r+1$, respectively. Observe that the top right element of $\mathbf{G}$ is $c_t \neq 0$. Therefore, $\det(\mathbf{G})$ is a nonzero (not necessarily homogeneous) polynomial in the x_i's, and its total degree is k.

Returning to the main line of argument, each of the $\binom{d}{t}$ systems gives rise to one such nonzero polynomial. Consequently, there is a choice of the x_i's (in Q) that makes all these polynomials nonzero, and therefore all the $\binom{d}{t}$ linear systems inconsistent. □

In the following discussion, σ_i is the elementary symmetric function in $\{x_1, x_2, \ldots, x_m\}$ of degree i multiplied by $(-1)^i$; $\sigma_0 = 1$; and $\sigma_i = 0$ for $i < 0$ or $i > m$.

LEMMA 5.2. *Let* $\mathbf{k} = [k_1, k_2, \ldots, k_n]$, $0 \leq k_1 \leq k_2 \leq \cdots \leq k_n \leq m$, *and let* $\mathbf{S}^{\mathbf{k}}$ *be given by* $(\mathbf{S}^{\mathbf{k}})_{ij} = \sigma_{k_i+i-j}$ *Then* $\det(\mathbf{S}^{\mathbf{k}})$ *is a nonzero homogeneous polynomial in the* x_i*'s of total degree* $\sum_{i=1}^{n} k_i$.

PROOF. By definition

$$\det(\mathbf{S}^{\mathbf{k}}) = \sum_{\rho} \prod_{i=1}^{n} \sigma_{k_i+i-\rho_i}, \tag{5.1}$$

where ρ runs through the permutations of $1, 2, \ldots, n$. Suppose that the product corresponding to a particular permutation π is nonzero. Then, it is a homogeneous polynomial in the x_i's of total degree $\sum_{i=1}^{n}(k_i + i - \pi_i) = \sum_{i=1}^{n} k_i \equiv k$. Consequently, all the nonzero terms on the right hand side of (5.1) have the same total degree. Therefore, *if* $\det(\mathbf{S}^{\mathbf{k}})$ is a nonzero polynomial, then its total degree in the x_i's is exactly $\sum_{i=1}^{n} k_i$. It remains to prove that $\det(\mathbf{S}^{\mathbf{k}})$ is a *nonzero* polynomial.

Define $e_i = n - j$ if $k_j < i$ but $k_{j+1} \geq i$. (e_i is the number of k_l's that are greater than or equal to i.) Observe that $e_1 \geq e_2 \geq \cdots \geq e_m$. We will complete the proof by showing that the coefficient of the monomial $x_1^{e_1} x_2^{e_2} \cdots x_m^{e_m}$ in $\det(\mathbf{S}^{\mathbf{k}})$ is nonzero. In fact, a stronger assertion is true: the monomial $x_1^{e_1} x_2^{e_2} \cdots x_m^{e_m}$ appears only in the product corresponding to the identity permutation. We prove this fact by showing that all the other monomials appearing on the right hand side of (5.1) are lexicographically larger than $x_1^{e_1} x_2^{e_2} \cdots x_m^{e_m}$. (See Subsection 4.1 for the definition of the lexicographic order on monomials.)

To begin with, note that $x_1^{e_1} x_2^{e_2} \cdots x_m^{e_m}$ is the lexicographically least monomial appearing in the product corresponding to the identity permutation. Next, let π be any permutation other than the identity, such that the corresponding product on the right hand side of (5.1) is nonzero. Suppose that $\pi_n = n, \pi_{n-1} = n-1, \ldots, \pi_{j+1} = j+1$, but $\pi_j < j$. Let $l = k_j + j - \pi_j$. It can be easily verified that for any monomial $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_m^{\lambda_m}$ appearing in $(\prod_{i=1}^n \sigma_{k_i+i-\pi_i})$, $\sum_{i=\alpha}^m \lambda_i \geq \sum_{i=\alpha}^m e_i$ for $\alpha > l$, and $\sum_{i=l}^m \lambda_i > \sum_{i=l}^m e_i$. Therefore, each monomial appearing in $(\prod_{i=1}^n \sigma_{k_i+i-\pi_i})$ is lexicographically larger than $x_1^{e_1} x_2^{e_2} \cdots x_m^{e_m}$. □

An upper bound on the x_i's appearing in the statement of Theorem 5.1 can be obtained using results on the density of zeros of multivariate polynomials:

LEMMA 5.3. *([11], see also [13]) Suppose that $Q(x_1, x_2, \ldots, x_m)$ is a nonzero multivariate polynomial of degree d_i in x_i, $1 \leq i \leq m$. Let $I_1, I_2, \ldots, I_m$ be sets of elements in the domain or the field of coefficients of Q. Then Q has at most*

$$| I_1 \times I_2 \times \cdots \times I_m | \left(\frac{d_1}{| I_1 |} + \frac{d_2}{| I_2 |} + \cdots + \frac{d_m}{| I_m |} \right)$$

zeros in the set $I_1 \times I_2 \times \cdots \times I_m$.

COROLLARY 5.4. *Each of the x_i's in the statement of Theorem 5.1 can be chosen to be an $O(t \log d)$ bit integers.*

PROOF. The x_i's must be chosen so that $(x_1, x_2, \ldots, x_m)$ is not a zero of the product of the $\binom{d}{t}$ polynomials (obtained as the determinant of linear systems) in the proof of Theorem 5.1. Since each of these polynomials has degree at most d in each variable, the degree of the product polynomial in each variable is at most $\binom{d}{t} d$. By Lemma 5.3, each of the x_i's can be chosen in the range from 0 to $2^{O(t \log d)}$. □

6. Concluding remarks

Clearly our present results provide only a partial picture of the sparse interpolation problem. We enumerate some of the open problems.

1. Is the general interpolation problem (Problems P1-P8) decidable?

   (a) For $t < m$, is there an easily computable degree bound D, depending on $t, x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_m$, such that if there is any t-sparse interpolating polynomial, then there is one of degree at most D?

   (b) Can the problem of computing a t-sparse interpolating polynomial be reduced to the corresponding decision problem?

2. When there are only finitely many solutions to a sparse interpolation problem, is there a nice characterization for the number of solutions?

3. What is the complexity of the general interpolation problem or of any of the subproblems P1-P8? Are all these subproblems in NP? Are any of the subproblems NP-hard or provably difficult? It is instructive to contrast Theorem 5.1 and Corollary 5.4 with Theorem 4.3. Namely, the degree bound of Theorem 4.3 can be made to grow as $\Omega(\alpha^{\frac{1}{t}})$, where $\alpha = \max_i |x_i|$. The degree bound can be forced to be at least $2t$ (by choosing α to be $(2t)^t$). Then, the decision procedure (of Subsection 4.1) tries at least $\binom{2t}{t}$ exponent sequences. Since the input description is $O(mt \log t)$ bits, the decision procedure is of exponential cost even for $m = t + 1$. We ask whether or not the decision problem of Subsection 4.1 is NP-complete. (Theorem 4.3 shows that the problem is in NP.) We do not know whether the decision problem of Subsection 4.2 is in NP.

4. Our sparse interpolation problems are posed relative to the standard basis $1, x, x^2, \ldots$. The same questions could be asked relative to any basis; for example, the Chebyshev polynomials. Using a simple shift of the x_i's, sparsity problems with respect to the basis $1, (x - \alpha), (x - \alpha)^2, \ldots$ can be shown to be equivalent to the to sparsity problems with respect to the standard basis. It is interesting to ask whether one can compute an α such that sparsity with respect to the basis $1, (x - \alpha), (x - \alpha)^2, \ldots$ is minimized. And if $S = \{(x_i, y_i)\}$ is rational, does it follow that such an optimal α is rational?

5. Is sparse multivariate polynomial interpolation decidable? If so, what is its complexity? Is sparse rational function interpolation decidable? If so, what is its complexity?

6. While we have phrased our results in terms of traditional computability theory, it is also interesting to phrase these problems in the context of the real (or complex) field computability model of [2]. All of our decision procedures can clearly be placed within this framework with the possible exception of Theorem 4.5 which uses the least common multiple λ of the denominators. In this last case, there is still a computable degree bound. λ is only used to place a lower bound on the absolute value of $\det \mathbf{W}$. Instead, such a lower bound (say, for the k^{th} exponent) can be obtained by computing $\det \mathbf{W}$ for each $\mathbf{W}$ that arises when one fixes the first $k - 1$

exponents. However, we do not know whether there is a degree bound that can be expressed in a closed form.

7. Two natural problems arise in the context of learnability theory (see Karpinski and Werther [8]).

   (a) Given $\{(x_i, y_i) \mid 1 \leq i \leq m\}$, and an $\epsilon > 0$, does there exist a t-sparse polynomial $p(x)$, such that $|p(x_i) - y_i| \leq \epsilon$ for all i. The degree bound D, given by Corollary 4.4, requires only an upper bound on the y_i's. Therefore, the method of Corollary 4.4 can be used to compute a degree bound in the case when $x_i, y_i \in \mathbf{Q}$, $x_i > 1$, and $p(x) \in \mathbf{Z}[x]$. However, the method of Theorem 4.5 does not readily extend to compute a degree bound for the case when we are looking for a polynomial $p(x) \in \mathbf{Q}[x]$. The reason is that we are unable to assert a nonzero lower bound on $\det \mathbf{W}$ in the proof of Theorem 4.5. Notice that, given a set of t monomials, the problem of determining if suitable coefficients exist (Step 2 in Section 4.1) can be formulated as the problem of determining a feasible solution to a linear program.

   (b) Given $\{(x_i, y_i, s_i) \mid s_i \in \{-1, 0, 1\}$ and $1 \leq i \leq m\}$, does there exist a t-sparse polynomial $p(x)$, such that $\mathrm{sign}(p(x_i) - y_i) = s_i$.

Acknowledgements

We are deeply indebted to Don Coppersmith, not only for significant improvements to this paper, but also for frequently aiming us in the right direction. Joachim von zur Gathen, Erich Kaltofen, John-Michael Wiley, and Lakshman Yagati simplified our original proofs of Theorems 3.1 and 3.4. We would also like to thank Volker Strassen for pointing out to us that the polynomial $P(x)$ of Lemma 4.1 is in fact a Schur function. (See, for example, [9].)

Part of this work was done while the first author was visiting IBM T. J. Watson Research Center, Yorktown Heights, and The International Computer Science Institute, Berkeley. Continuing support from the Natural Sciences and Engineering Research Council of Canada is also gratefully acknowledged.

A preliminary version of this paper appeared in the proceedings of the twenty-second Annual ACM Symposium on Theory of Computing [3].

References

[1] M. Ben-Or and P. Tiwari, A deterministic algorithm for sparse multivariate polynomial interpolation, *Proc. 20th Ann. ACM Symp. Theory of Computing*, 301-309, 1988.

[2] L. Blum, M. Shub and S. Smale, On a theory of computation over the real number; NP completeness, recursive functions and universal machines, *Proc. 29th IEEE Symp. Foundations of Comp. Sci.*, 387-397, 1988.

[3] A. Borodin and P. Tiwari, On the decidability of sparse univariate polynomial interpolation (preliminary version), *Proc. 22nd Ann. ACM Symp. Theory of Computing*, 535-545, 1990.

[4] R. J. Evans and I. M. Isaacs, Generalized Vandermonde determinants and roots of unity of prime order, *Proc. Amer. Math. Soc.* **58** (1976), 51-54.

[5] F. R. Gantmacher, *The Theory of Matrices*, K. A. Hirsch, New York, 1959.

[6] D. Yu. Grigoriev and M. Karpinski, The matching problem for bipartite graphs with polynomially bounded permanents is in NC, *Proc. 28th IEEE Symp. Foundations of Comp. Sci.*, 166-172, 1987.

[7] N. Jacobson, *Basic Algebra* I, W. H. Freeman and Company, San Francisco, 1974.

[8] M. Karpinski and T. Werther, *Learnability and VC-dimension of sparse polynomials and rational functions*, Technical Report TR-89-060, International Computer Science Institute, Berkeley, 1989.

[9] D. E. Littlewood, *The Theory of Group Characters and Matrix Representations of Groups*, Oxford, 1950.

[10] M. Marcus and H. Minc, *Basic Algebra, A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, Mass., 1964.

[11] J. T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, *J. Assoc. Comput. Mach.* **27** (1980), 701-707.

[12] P. Tiwari, *Parallel algorithms for instances of linear matroid parity with a small number of solutions*, IBM Research Report 12766, IBM T. J. Watson Research Center, New York, 1987.

[13] R. ZIPPEL, Probabilistic algorithms for sparse polynomials. *Lecture Notes in Computer Science* **72**, *Symbolic and Algebraic Computation*, 216-226, Springer-Verlag, 1979.

Manuscript received June 26, 1990

ALLAN BORODIN
Department of Computer Science
University of Toronto
Toronto, Ontario
Canada M5S 1A4
bor@theory.toronto.edu

PRASOON TIWARI
IBM T. J. Watson Research Center
P. O. Box 218
Yorktown Heights, NY 10598
USA

Current address of PRASOON TIWARI:
Department of Computer Science
1210 W. Dayton Street
University of Wisconsin-Madison
Madison, WI 53706
USA
tiwari@cs.wisc.edu