Interpolation of Sparse Rational Functions Without Knowing Bounds on Exponents

Dima Yu. Grigoriev
Steklov Institute of Mathematics
Soviet Academy of Sciences
Leningrad 191011

Marek Karpinski *
Dept. of Computer Science
University of Bonn
5300 Bonn 1
and
International Computer Science Institute
Berkeley, California

Michael F. Singer †
Dept. of Mathematics
N. C. State University
Raleigh, NC 27695

Abstract. We present the first algorithm for the (*black box*) interpolation of t-sparse, n-variate rational functions without knowing bounds on exponents of their sparse representation with the number of queries independent on exponents. In fact, the algorithm uses $O(nt^t)$ queries to the *black box*, and can be implemented for a fixed t in a polynomially bounded storage (or polynomial parallel time).

Introduction.

A t-sparse rational function is a function that can be written as a quotient of two polynomials, each containing at most t terms. We show in this paper that, if we are given a *black box* to evaluate a t-sparse rational function f with integer coefficients, then one can bound the exponents appearing in a t-sparse representation of f by making $2(t+1)^t - 1$ black box evaluations in the univariate case and $O(nt^t)$ black box evaluations in the n-variable case. Using this, we also give the first algorithm, depending in exponent only on t (!), for interpolation of t-sparse rational functions without knowing bounds on exponents and show that for fixed t this problem is in polynomial parallel time (sequential storage).

The authors were motivated by the question whether the recent parallel deterministic sparse interpolation algorithms ([GK 87], [BT 88], [GKS 88], [KL 88]) could be generalized to the rational function case without knowing an *a priori* bound on the exponents of their defining polynomials; and also by its natural connection to the seminal problem of Strassen ([S 73]) of computing the numerator and denumerator of rational functions. It was not known before whether the number of queries needed for the problem was *recursive* in n and t. Approximative unbounded degree interpolation arises also naturally in the issues of computational learnability of sparse rational functions (cf. [KW 89]).

For the corresponding versions of bounded degree rational interpolation (where the bound on the degree is part of the input) see [S 73], [K 86], [BT 88], [KT 88]. Another version of unbounded degree univariate polynomial interpolation is studied in [BT 89].

To bound the exponents appearing in some t-sparse representation of a t-sparse rational function $f(X)$ of one variable, we will proceed as follows. We consider representations of $f(X)$ of the form $(\sum_{i=1}^{t} a_i X^{\alpha_i})/(\sum_{i=1}^{t} b_i X^{\beta_i})$, where the a_i and b_i are real numbers and the α_i and β_i are non-negative real numbers. Such a function is called a *quasirational function*. We show that for t-sparse $f(X)$ the α_i and β_i must satisfy a system S of polynomial equalities and inequalities whose coefficients depend on the value of $f(X)$ at $2(t+1)^t - 1$ points. By evaluating the black box for $f(X)$ at these points, we can determine this system. Using the results of [GV 88], we can bound a *real* solution of this system. Using the fact that $f(X)$ is a t-sparse *rational* function, we are then able to bound an integer solution of S and this gives our desired bound. The detailed complexity analysis of the algorithm will be given in the final version of the paper.

The rest of this paper is organized as follows.

In Section 1 we give a formal definition of quasirational functions and prove some basic facts about

*Supported in part by the Leibniz Center for Research in Computer Science, by the DFG Grant KA 673/2-1 and by the SERC Grant GR-E 68297
†Supported in part by NSF Grant DMS-8803109

these functions. In Section 2, we describe some elementary properties of right euclidean rings. An example of such a ring is $F[\mathcal{D}]$, where F is the field of quasirational functions of one variable and $\mathcal{D}$ is the operator defined by $\mathcal{D}(f(X)) = f(pX)$ for some fixed prime p. For this ring, we are able to derive an analogue of the Sylvester matrix and the resultant. In Section 3 we use this to obtain the system S and the bound for the exponents appearing in a t-sparse representation of a t-sparse rational function. In Section 4 we show how the results of section 3 can be used to obtain a bound on the exponents of a t-sparse rational function of several variables. In Section 5 we describe an algorithm to interpolate t-sparse rational functions and give complexity bounds.

1. Quasirational Functions.

A finite sum of the form

$$\sum_I c_I \mathbf{X}^I$$

where $I = (\alpha_1, \ldots, \alpha_n)$, $0 \leq \alpha_i \in \mathbb{R}$, $\mathbf{X}^I = X_1^{\alpha_1} \cdot \ldots \cdot X_n^{\alpha_n}$, $c_I \in \mathbb{R}$ is called a *quasipolynomial* of n variables. Denote by $\mathbb{R}< X_1, \ldots X_n >$ the ring of quasipolynomials of n variables.

A ratio of two quasipolynomials is called a *quasirational* function. If the number of terms in the sum is at most t, we say that the quasipolynomial is t-sparse. If a quasirational function can be represented as a ratio of two t-sparse quasipolynomials, we say that it is also t-sparse. We use the expressions "polynomial" or "rational function" in this usual sense, that is for quasipolynomials or quasirational functions with non-negative integer exponents in their terms.

We assume that we are given a "black box" representing an n-variable rational function f with integer coefficients into which we can put points with rational coefficients. The output of the black box is either the value of the function at this point or some special sign, e.g. "∞", if the denominator of the irreducible representation of the function vanishes at this point (a representation $f = g/h$, $g, h \in \mathbb{R}[X_1, \ldots, X_n]$, is irreducible if g and h are relatively prime). In what follows, we will sometimes obtain in intermediate steps a representation of a rational function in the form of a quasirational function. Nevertheless, our aim is to obtain a representation of a rational function in the usual form, provided that it is t-sparse.

We will need a zero test for t-sparse rational functions. This is similar to well known zero tests for t-sparse polynomials (cf. [GK 87], [GKS 88], [BT 88]). Recall that if $M_1, \ldots, M_t$ are distinct positive numbers, then any $t \times t$ subdeterminant of the $(2t-1) \times t$ matrix $(M_i^j)_{1 \leq i \leq t,\, 1 \leq j \leq 2t-1}$ is non-singular (c.f. [EI 76]).

To test if a t-sparse rational function f is identically zero, use its black box to evaluate f at the $2t-1$ points $P^j = (p_1^j, \ldots, p_n^j)$, $1 \leq j \leq 2t-1$, where the $p_1, \ldots, p_n$ are distinct primes. Since the black box gives output based on an irreducible representation of f, we see that any zero of the denominator of such a representation is a zero of the denominator of a t-sparse representation of f. Using the remark about the matrix (M_i^j) above we see that the denominator can vanish at, at most, $t-1$ of these points. The same concerns the numerator. Therefore, the t-sparse function f is not identically zero if and only if the black box outputs a number different from 0 and ∞ at one of the points P^j.

The next result concerns different t-sparse representations of a quasirational function f. This result can be thought of as saying that, under suitable hypotheses, two such representations can only differ in certain redundant terms that can be eliminated. If g is a quasipolynomial, we denote by $\mathrm{ord}_{X_i}(g)$ the least power of X_i occurring in g. We call a representation $g_1/h_1 = f$ *normalized* if for each i, $1 \leq i \leq n$, $\min(\mathrm{ord}_{X_i}(g_1), \mathrm{ord}_{X_i}(h_1)) = 0$. For an arbitrary g_1/h_1, there is a unique monomial M such that $(g_1/M)/(h_1/M)$ is normalized. We call the latter representation the *normalization* of g_1/h_1.

Lemma 1. Assume that g_1/h_1 is a t-sparse representation of a quasirational function and $g_2/h_2 = g_1/h_1$ is another t-sparse normalized representation. Let $d = \max_i \{\deg_{X_i}(g_1), \deg_{X_i}(h_1)\}$. We can delete some terms from g_2 and h_2 obtaining $\bar{g}_2, \bar{h}_2$ so that

$$\bar{g}_2/\bar{h}_2 = g_1/h_1$$

and

$$\max_i \{\deg_{X_i}(\bar{\bar{g}}_2), \deg_{X_i}(\bar{\bar{h}}_2)\} \leq 2td,$$

where $\bar{\bar{g}}_2/\bar{\bar{h}}_2$ is the normalization of $\bar{g}_2/\bar{h}_2$.

PROOF.

Write

$$g_2 = \sum_i g_2^{(i)} X_1^{\beta_i}, \quad h_2 = \sum_i h_2^{(j)} X_1^{\gamma_j}$$

where $\beta_1 < \beta_2 < \ldots, \gamma_1 < \gamma_2 < \ldots$ and $0 \neq g_2^{(i)} \in \mathbb{R} < X_2, \ldots, X_n >$, $0 \neq h_2^{(j)} \in \mathbb{R} < X_2, \ldots, X_n >$. We can assume that $\beta_1 = 0$ (the argument is similar if $\gamma_1 = 0$). In this case we see from $g_2 h_1 = h_2 g_1$, that $\gamma_1 \leq d$. If $\beta_{i+1} - \beta_i \leq d$ for each i, then $\deg_{X_1} g_2 \leq (t-1)d$. This would imply $\deg_{X_1} h_2 \leq \deg_{X_1} g_2 h_1 \leq td$ and we would be done.

If $\gamma_{i+1} - \gamma_i \leq d$ for each i, then $\deg_{X_1} h_2 \leq td$. This would imply that $\deg_{X_1} g_2 \leq \deg_{X_1} h_2 g_1 \leq td + d \leq 2td$ and we would be done. Therefore we can assume that there is a minimal number s such that $\beta_{i_0} \leq s - d < s < \beta_{i_0+1}$ and $\gamma_{j_0} \leq s - d < s < \gamma_{j_0+1}$ for suitable i_0, j_0. Since $\beta = 0$ and $\gamma_1 \leq d$, we have that $s \leq (t+1)d \leq 2td$.

Let

$$\tilde{g}_2 = \sum_{i \leq i_0} g_2^{(i)} X_1^{\beta_i}, \quad \tilde{h}_2 = \sum_{j \leq j_0} h_2^{(j)} X_1^{\gamma_j}.$$

If one compares the coefficients of X_1^ρ, $\rho \leq s$, in $g_2 h_1 = h_2 g_1$, one can see that $\tilde{g}_2 h_1 = \tilde{h}_2 g_1$ so $\tilde{g}_2/\tilde{h}_2 = g_1/h_1$.

We now take the normalization $\tilde{\tilde{g}}_2/\tilde{\tilde{h}}_2$ of $\tilde{g}_2/\tilde{h}_2$ and apply considerations similar to those above to $\tilde{\tilde{g}}_2/\tilde{\tilde{h}}_2$ with X_2 playing the role of X_1. At the end of this process we obtain the normalized representation $\bar{\bar{g}}_2/\bar{\bar{h}}_2$. It corresponds to a pre-normalized $\bar{g}_2/\bar{h}_2$ that satisfies the conclusion of the lemma. □

Corollary. If, in the above Lemma, we assume $g_2, h_2 \in \mathbb{R}[X_1, \ldots, X_n]$ are polynomial, then we can conclude that $\bar{g}_2, \bar{h}_2 \in \mathbb{R}[X_1, \ldots, X_n]$ as well.

We note that in Lemma 1 and its corollary, $\bar{g}_2$ and $\bar{h}_2$ are obtained by eliminating terms of sufficiently high degree and keeping lower order terms in g_2 and h_2.

2. Right Euclidean Rings (a digest).

Let F be a field and let $\mathcal{D} : F^+ \to F^+$ be a homomorphism with respect to the additive structure of F. Let $F[\mathcal{D}]$ be the subring of HOM (F^+, F^+) generated by F (acting on F^+ by multiplication) and $\mathcal{D}$. We assume that each element $a \neq 0$ from $F[\mathcal{D}]$ can be uniquely represented in the form $a = \sum_{0 \leq i \leq m} \alpha_i \mathcal{D}^i$ where $\alpha_i \in F$ and $\alpha_m \neq 0$. We denote the integer m by $\deg(a)$ and adopt the convention that $\deg(0) = -\infty$.

We furthermore assume that for $a, b \in F[\mathcal{D}]$, $\deg(ab) = \deg(a) + \deg(b)$. This assumption is equivalent to the statement that for each α in F there are unique α_1, α_2 in F, with $\alpha_1 \neq 0$, such that $\mathcal{D} \cdot \alpha = \alpha_1 \mathcal{D} + \alpha_2$. We can conclude that there exists right Euclidean division in $F[\mathcal{D}]$, that is, for any $a, b \in F[\mathcal{D}]$ $b \neq 0$, there exist unique b_1, b_2 with $\deg(b_2) < \deg(b)$ such that $a = b_1 b + b_2$. This leads to a right Euclidean algorithm and a notion of greatest common right divisor ($\mathrm{gcrd}(a, b)$) of two elements a and b, which can be represented in the form $\mathrm{gcrd}(a, b) = a_1 a + b_1 b$ for some $a_1, b_1 \in F[\mathcal{D}]$. Furthermore $a = a_0 \mathrm{gcrd}(a, b)$ and $b = b_0 \mathrm{gcrd}(a, b)$ for some $a_0, b_0 \in F[\mathcal{D}]$ (c.f. [O 33]).

Let $\deg(a) = m$ and $\deg(b) = k$ and consider

$$\mathcal{D}^i(a) = \sum_{0 \leq j \leq m+i} a_j^{(i)} \mathcal{D}^j, \quad \mathcal{D}^l(b) = \sum_{0 \leq j \leq k+l} b_j^{(l)} \mathcal{D}^j$$

for $0 \leq i \leq k-1, 0 \leq l \leq m-1$. Let S be the $(m+k) \times (m+k)$ matrix whose columns correspond to the operators $\mathcal{D}^{k+m-1}, \ldots, \mathcal{D}^2, \mathcal{D}, 1$ and whose rows contain the coefficients of the operators in $\mathcal{D}^i(a)$, $0 \leq i \leq k-1$ and $\mathcal{D}^l(b)$, $0 \leq l \leq m-1$ (S resembles the Sylvester matrix [VDW 66]; for differential operators a similar object is described in [G 89]). As in [G 89] Lemma 12, one can show:

Lemma 2. $\deg(\mathrm{gdrc}(a, b)) = n - \mathrm{rank}(S)$.

In what follows we restrict ourselves to the case where F is the field of quasirational functions in one variable and $\mathcal{D}$ is the operator defined by $\mathcal{D}(X^\alpha) = (pX)^\alpha$, where p is some fixed prime number. Note that $\mathcal{D} \cdot f = \mathcal{D}(f) \cdot \mathcal{D}$.

Lemma 3. If $f \in F$ and $\mathcal{D}(f) = f$, then $f \in \mathbb{R}$.

PROOF. If $\mathcal{D}(f) = f$, then $f(X) = f(pX) = f(p^2X) = \ldots$. The zero test of section 1 implies that $f(X) = f(YX)$ for a new variable Y. If $f = g/h$ let

$$g = \sum a_i X^{\alpha_i}, \; h = \sum b_i X^{\beta_i}$$

$0 \leq \alpha_1 < \alpha_2 < \ldots$ $0 \leq \beta_1 < \beta_2 < \ldots$, and $a_i, b_i \in \mathbb{R}$. Since

$$g(YX)h(X) = g(X)h(YX),$$

we can conclude, by comparing coefficients of the corresponding monomials in X and Y, that $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots$ and $a_i b_j = a_j b_i$ for all i, j. Therefore $f \in \mathbb{R}$. □

Lemma 4. If $y_1, \ldots, y_n \in F$, then $y_1, \ldots, y_n$ are linearly dependent over $\mathbb{R}$ if and only if

$$W(y_1, \ldots, y_n) = \det \begin{bmatrix} y_1(x) & \ldots & y_n(x) \\ y_1(px) & \ldots & y_n(px) \\ \vdots & & \vdots \\ y_1(p^{n-1}x) & \ldots & y_n(p^{n-1}x) \end{bmatrix} = 0$$

PROOF. If $y_1, \ldots, y_n$ are linearly dependent over $\mathbb{R}$ then, clearly, $W(y_1, \ldots, y_n) = 0$. Now assume that $W(y_1, \ldots, y_n) = 0$. In this case there exist $f_1, \ldots, f_n \in F$, not all zero, such that

$$f_1 y_1 + \ldots + f_n y_n = f_1 \mathcal{D} y_1 + \ldots + f_n \mathcal{D} y_n = \ldots$$

$$= f_1 \mathcal{D}^{n-1} y_1 + \ldots + f_n \mathcal{D}^{n-1} y_n.$$

We may assume $f_1 = 1$. Applying $\mathcal{D}$ to each of these equations, we have

$$\mathcal{D}^i y_1 + \mathcal{D} f_2 \mathcal{D}^i y_1 + \ldots + \mathcal{D} f_n \mathcal{D}^i y_n = 0$$

for $i = 1, \ldots, n$. This implies that

$$(f_2 - \mathcal{D} f_2)\mathcal{D}^i y_2 + \ldots + (f_n - \mathcal{D} f_n)\mathcal{D}^i y_n = 0$$

for $i = 1, \ldots, n-1$. Either $f_i - \mathcal{D} f_i = 0$ for $i = 2, \ldots, n$, in which case we are done by Lemma 3, or by induction there exist $\alpha_2, \ldots, \alpha_n \in \mathbb{R}$, not all zero, such that $\alpha_2 \mathcal{D} y_2 + \ldots + \alpha_n \mathcal{D} y_n = 0$. Therefore $\mathcal{D}(\alpha_2 y_2 + \ldots + \alpha y_n) = 0$ so $\alpha_2 y_2 + \ldots + \alpha_n y_n = 0$. □

Corollary. Let $L = \sum_{i=0}^{t} a_i \mathcal{D}^i$ with $a_i \in F$, not all zero. The dimension of the $\mathbb{R}$-vectorspace of solutions in F of $Ly = 0$ is at most t.

PROOF. Let $y_1, \ldots, y_{t+1}$ be solutions of $Ly = 0$. We then have

$$(a_0, \ldots, a_t) \begin{bmatrix} y_1 & \ldots & y_{t+1} \\ \mathcal{D} y_1 & \ldots & \mathcal{D} y_{t+1} \\ \vdots & & \vdots \\ \mathcal{D}^t y_1 & \ldots & \mathcal{D}^t y_{t+1} \end{bmatrix} = 0$$

Lemma 4 implies that $y_1, \ldots, y_{t+1}$ are linearly dependent. □

Lemma 5. Let $L = \sum_{j=0}^{t} a_j \mathcal{D}^j$ with $a_i \in \mathbb{R}$ and assume that $P_L(z) = a_t z^t + \ldots + a_0 \in \mathbb{R}[z]$ has t distinct roots ≥ 1, say $p^{\alpha_1}, \ldots, p^{\alpha_t}$. Then $\{X^{\alpha_1}, \ldots, X^{\alpha_t}\}$ is a base for the space of solutions of $Ly = 0$. □

PROOF. One easily sees that $L(X^{\alpha_i}) = 0$ for $i = 1, \ldots, t$. The functions $X^{\alpha_1}, \ldots, X^{\alpha_t}$ are linearly independent over $\mathbb{R}$, so by the corollary to Lemma 4 they must be a basis of the space of solutions. □

Lemma 6. Let L be as in Lemma 5 and assume that $L = L_1 \cdot L_2$ where $L_1 = \sum_{j=0}^{t-s} b_j^{(1)} \mathcal{D}^j$ and $L_2 = \sum_{j=0}^{s} b_j^{(2)} \mathcal{D}^j$ with $b_j^{(i)} \in F$. Then the space of solutions in F of $L_2(y) = 0$ has dimension s.

PROOF. Let V be the solution space of $Ly = 0$. By Lemma 5, this has dimension t. L_2 maps V into the solution space of $L_1 y = 0$, which has dimension at most $t - s$ by Lemma 4. Therefore the dimension of the solution space of $L_2 y = 0$ is at least s and so by Lemma 4, it must equal s. □

3. Bounding the Exponents of a Sparse Univariate Rational Function

Lemma 5 in the previous section allows us to characterize t-sparse quasipolynomials g as those quasipolynomials for which there exits an operator of degree t, $L = \sum_{j=0}^{t} a_j \mathcal{D}^j$, with $P_L(z) = a_t z^t + \ldots + a_0 \in \mathbb{R}[z]$ having distinct real roots ≥ 1, such that $Lg = 0$. Therefore a t-sparse quasirational function f is a quasirational function for which there exists a quasipolynomial h and operators of degree t, L_1 and L_2 as above such that $L_1(h) = 0$ and $L_2(hf) = 0$. $L_1(y)$ and $L_2(yf)$ will therefore have a common solution. The results of section 2 allow us to eliminate y using the determinant of the Sylvester matrix. This determinant is a quasirational function and, by evaluating at sufficiently many points, we obtain (together with the conditions that the α_i, β_j are distinct and ≥ 1) a system of polynomial inequalities that must be satisfied by the exponents appearing in f. We will then bound a *real* solution of this system using [GV 88] and, assuming that f is a univariate rational function, we can use Lemma 1 to bound the exponents of f.

We now proceed more formally. Let $f = \frac{g}{h}$ be a t-sparse quasirational function of one variable where $g = \sum_{i=1}^{t} a_i X^{\alpha_i}$ and $h = \sum_{i=1}^{t} b_i X^{\beta_i}$ are t-sparse quasipolynomials. Let $G(z) = c_0 + c_1 z + \ldots + z^t$ be the unique monic polynomial whose roots are $p^{\alpha_1}, \ldots, p^{\alpha_t}$ and let $H(z) = d_0 + d_1 z + \ldots + z^t$ be the unique monic polynomial whose roots are $p^{\beta_1}, \ldots, p^{\beta_t}$. Consider the operators $L_G = \sum_{i=0}^{t} c_i \mathcal{D}^i$ and $L_H = \sum_{i=0}^{t} d_i \mathcal{D}^i$ (where $d_t = c_t = 1$). We then have $L_H(h) = 0$ and $L_G(fh) = 0$. Therefore $L_H(y) = 0$ and $L_G(fy) \equiv \tilde{L}_G(y) = 0$ have a non-zero common solution $y = h$ in F (note that the coefficients of $\tilde{L}_G$ are $\mathbb{R}$-linear combinations of $f, \mathcal{D}f, \ldots, \mathcal{D}^t f$). Consider the Sylvester matrix $S = S(c_0, c_1, \ldots, c_{t-1}, d_0, \ldots, d_{t-1}, f)$ of L_H

and $\tilde{L}_G$. By Lemma 2, $\det(S) = 0$ (note that $\det S$ is a quasirational function).

Conversely if $\det(S) = 0$, then Lemma 2 implies that $\deg(\gcrd(L_H, \tilde{L}_G)) \geq 1$. Since the coefficients of L_H satisfy the hypotheses of Lemma 5 and $\gcrd(L_H, \tilde{L}_G)$ divides L_H, L_H and $\tilde{L}_G$ will have a common non-zero solution h_0 in F (by Lemma 6). Lemma 5 then implies that f is a t-sparse quasirational function because h_0 and $h_0 f$ are both t-sparse quasipolynomials, again by Lemma 5. We have therefore proved the following lemma.

Lemma 7. A quasirational function f is t-sparse if and only if there exist real numbers $\bar{c}_0, \ldots, \bar{c}_{t-1}, \bar{d}_0, \ldots, \bar{d}_{t-1}$ such that

(i) $\det(S(\bar{c}_0, \ldots, \bar{c}_{t-1}, \bar{d}_0, \ldots, \bar{d}_{t-1}, f)) = 0$, and

(ii) there exist t distinct real numbers ≥ 1 that are roots of

$$G(z) = \bar{c}_0 + \ldots + \bar{c}_{t-1} z^{t-1} + z^t = 0$$

and there exist t distinct real numbers ≥ 1 that are roots of

$$H(z) = \bar{d}_0 + \ldots + \bar{d}_{t-1} z^{t-1} + z^t = 0.$$

Now assume that f is a t-sparse *rational* function whose coefficients are integers. We see that each entry of S is a t-sparse rational function. ¿From the form of the matrix, we see that $\det(S)$ is a $(t+1)^t$ sparse rational function. Therefore condition (i) is equivalent (by the zero test) to the fact that $\det(S)_{X=p_1^i}$ is either ∞ of 0 for $i = 1, \ldots, 2(t+1)^t - 1$ (p_1 is any prime). For at least $(t+1)^t$ of these points $\det(S)_{X=p_1^i}$ will be zero. Using the black box, we can determine a system of $(t+1)^t$ equations in the unknowns $c_0, \ldots, c_{t-1}, d_0, \ldots, d_{t-1}$ of degree at most $2t$, that is equivalent to the vanishing of $\det(S)$ at these points. Assume the bitsize of the values (yielded by the black box) of $\mathcal{D}^j f(p_1^i)$, $i = 1, \ldots, 2(t+1)^t - 1$, $j = 0, \ldots, t$ is at most M (that is, those that are not ∞ and therefore rational numbers). We then see that the bitsize of each coefficient in this system is at most $O(t \ln t) + tM$.

Furthermore, condition (ii) is equivalent to

(ii') On $1 \leq z \leq 1 - c_{t-1} + c_{t-2} - \ldots \pm c_0$, the polynomial $G(z) = 0$ has precisely t roots and a similar statement holds for H. In addition, the discriminants of G and H are not zero.

The first sentence of (ii') can be expressed in terms of Sturm sequences. This yields a system of $2t$ polynomial inequalities and (by the Habicht subresultant theorem [LO 83]) each inequality has degree at most $2t$ and the bitsize of the coefficients is at most $O(t \ln t)$. Similar bounds hold for the discriminants.

Therefore, under the assumption that f is a t-sparse rational function with integer coefficients, we are able to construct a system of polynomial inequalities equivalent to (i) and (ii) and bound the bitsize of the coefficients of this system. The results of [GV 88] imply that this system has a solution in a ball of radius $\exp((Mt^{t^2})^{O(1)})$.

This gives a bound on some solution $\bar{c}_0, \ldots, \bar{c}_{t-1}, \bar{d}_0, \ldots, \bar{d}_{t-1}$ which gives a t-sparse quasirational representation of f. The exponents in this representation can be bounded from these $\bar{c}_0, \ldots, \bar{c}_{t-1}, \bar{d}_0, \ldots, \bar{d}_{t-1}$ also by $\exp((Mt^{t^2})^{O(1)})$ since they are roots of the polynomials $G(z)$ and $H(z)$. By Lemma 2, the exponents in a t-sparse *rational* function representation of f are bounded by a similar number. This bound therefore can be explicitly calculated by making $2(t+1)^t - 1$ black box evaluations.

4. Bounding the Exponents of a Sparse Multivariate Rational Function.

Let

$$\begin{aligned} f(X_1, \ldots, X_n) &= \frac{g(X_1, \ldots, X_n)}{h(X_1, \ldots, X_n)} \\ &= \frac{\sum_{i=1}^t g_i(X_2, \ldots, X_n) X_1^{\alpha_i}}{\sum_{i=1}^t h_i(X_2, \ldots, X_n) X_1^{\beta_i}} \end{aligned}$$

be a normalized representation of the t-sparse rational function f.

Consider the $2t^4 + 2t^2 + 1$ points $P^j = (p_2^j, \ldots, p_n^j)$ for $1 \leq j \leq 2t^4 + 2t^2 + 1$. Let

$$\begin{aligned} f^j(X_1) &= f(X_1, p_2^j, \ldots, p_n^j) \\ &= \frac{\sum_{i=1}^t g_i(p_2^j, \ldots, p_n^j) X_1^{\alpha_i}}{\sum_{i=1}^t h_i(p_2^j, \ldots, p_n^j) X_1^{\beta_i}} \end{aligned}$$

Note that there are at most $2t^2$ points P^j for which some g_i or h_i vanishes at P^j. We call these points bad points. For a point P_j that is not bad, let D_j be the bound on the degree of some normalized t-sparse representation of $f^j(X_1)$. For these points, Lemma 2

allows is to conclude that there exist t_1, t_2 (not necessarily unique) such that

$$f^j(X_1) = \frac{\sum_{i=1}^{t_1} g_i(p_2^j, \ldots, p_n^j) X_1^{\alpha_i}}{\sum_{i=1}^{t_2} h_i(p_2^j, \ldots, p_n^j) X_1^{\beta_i}} = \frac{\bar{g}^j}{\bar{h}^j}$$

and

$$\max\{\deg_{X_1} \bar{g}^j, \deg_{X_1} \bar{h}^j\} \leq 2tD_j.$$

To each j corresponds some pair t_1, t_2. Therefore, at least $2t^2 + 1$ non-bad points P^j correspond to some pair $(\tilde{t}_1, \tilde{t}_2)$. For this pair $(\tilde{t}_1, \tilde{t}_2)$ we have that

$$g(X_2, \ldots, X_n) \sum_{i=1}^{t} h_i(X_2, \ldots, X_n) X_1^{\alpha^i}$$

$$- h(X_1, \ldots, X_n) \sum_{i=1}^{t} g_i(X_2, \ldots, X_n) X_1^{\beta^i} \qquad (1)$$

is zero at these $2t^2 + 1$ points. If we consider (1) as a polynomial in X_1 whose coefficients are $2t^2$ sparse polynomials in $X_2, \ldots, X_n$, we see that (1) is identically zero. This implies that f has a t-sparse representation with $\deg_{X_1} f \leq B_1 = \max_j \{2tD_j\}$. We consider this representation and let X_2 play the role of the principal variable. We apply the same construction to prove the existence of a representation with $\deg_{X_1} f \leq B_1$ and $\deg_{X_2} f \leq B_2$. In this way we are able to determine B for which these exists a t-sparse representation of f with $\deg_{X_i} f \leq B$ for $i \leq i \leq n$.

Note that as in the univariate case, $B \leq \exp((Mt^{t^2})^{O(1)})$ where M is a bound on the bitsize of $f(p_1^j, \ldots, p_n^j)$ for $1 \leq j \leq 2(t+1)^t - 1$.

5. Interpolation of Sparse Multivariate Rational Functions and the Complexity Issues.

Let f be a t-sparse multivariate rational function and let $B \leq \exp((Mt^{t^2})^{O(1)})$ be the bound obtained in the previous section. Let

$$\mathcal{A} = \{A_i = (\alpha_{i_1}, \ldots, \alpha_{i_n}) | 0 \leq \alpha_{ij} \leq B\}$$

and

$$\mathcal{B} = \{B_i = (\beta_{i_1}, \ldots, \beta_{i_n}) | 0 \leq \beta_{ij} \leq B\}.$$

Select, in parallel, 2 t-tuples $I = \{A_1, \ldots, A_t\}$, $J = \{B_1, \ldots, B_t\}$ with $A_i \in \mathcal{A}$ and $B_i \in \mathcal{B}$. We calculate $f(X)$ at $(p_1^i, \ldots, p_n^i)$ for $i = 1, \ldots, 4t^2$. For each selection of I and J we obtain the following linear system

$$f(p_1^i, \ldots, p_n^i)(b_1(p_1^{\beta_{11}} \ldots p_n^{\beta_{1n}})^i + b_1(p_1^{\beta_{t1}} \ldots p_n^{\beta_{tn}})^i)$$
$$= a_1(p_1^{\alpha_{11}} \ldots p_n^{\alpha_{1n}})^i + \ldots + a_t(p_1^{\alpha_{t1}} \ldots p_n^{\alpha_{tn}})^i \qquad (2)$$

where $1 \leq i \leq 4t^2$, in the unknowns $b_1, \ldots, b_t, a_1, \ldots, a_t$. If such a system has a solution $\bar{b}_1, \ldots, \bar{b}_t, \bar{a}_1, \ldots, \bar{a}_t$, then the zero test implies that

$$f(X_1, \ldots, X_n) =$$
$$\frac{\bar{a}_1 X_1^{\alpha_{11}} \ldots X_n^{\alpha_{1n}} + \ldots + \bar{a}_t X_1^{\alpha_{t1}} \ldots X_n^{\alpha_{tn}}}{\bar{b}_1 X_1^{\beta_{11}} \ldots X_n^{\alpha_{1n}} + \ldots + \bar{b}_t X_1^{\alpha_{t1}} \ldots X_n^{\alpha_{tn}}}$$

For some I and J we will be able to solve (2) so the algorithm terminates with a correct answer. We give now an analysis of the complexity referring the reader to the basics of algebraic computation (cf., e. g., [KR 88]). Each of $\mathcal{A}$ and $\mathcal{B}$ contain B^n terms. We select t elements from each, so there are $O(B^{nt})$ systems of type (2), each of size at most $4t^2$. This implies that the sequential time complexity is $B^{O(nt)}$ and the parallel complexity is $(nt \log B)^{O(1)}$ (cf. [BGH 82], [M 86], [KR 88]). We can further bound B in terms of the size of the output. Let $\delta = \max_i \{\deg_{X_i} f\}$ and let μ be a bound on the bitsize of the coefficients of f. Let μ_j be the bitsize of any coefficient of $f^j(X_1)$ (as in section 4). We then have that the bitsize of any output is not greater than $\mu_j + t + \delta$. Furthermore, each μ_j can be bounded by $\mu + O(t^4 \delta n \log n)$ by looking at each term of the representation of f and noting that $p_n = O(n \log n)$. Therefore

$$\begin{aligned} B_1 &\leq \exp((Mt^{t^2})^{O(1)}) \\ &\leq \exp(((\mu + t^4 \delta n \log n + \delta) t^{t^2})^{O(1)}) \\ &= \exp(((\mu + \delta n \log n) t^{t^2}))^{O(1)}) \end{aligned}$$

Therefore the sequential complexity of the algorithm is $\exp((\mu + \delta n \log n) t^{t^2})^{O(1)}$ and the parallel complexity is $((\mu + \delta n \log n) t^{t^2})^{O(1)}$. Therefore, for fixed t, interpolation of t-sparse rational functions can be done in polynomial parallel time (as well as in sequential storage, [C 81]). □

6. Further Research

It remains an interesting open problem how to improve our algorithm. Is, in view of our results, also the polynomial time solution in the size of g, h and n, t possible?

Acknowledgement.

We are indebted to Volker Strassen for motivating the problem and a number of interesting discussions.

References

[BT 88] Ben-Or, M. and Tiwari, P. A., *A Deterministic Algorithm for Sparse Multivariate Polynomial Interpolation*, Proc. 20^{th} ACM STOC (1989), pp. 301–309.

[BGH 82] Borodin, A., von zur Gathen, J. and Hopcroft, J., *Fast Parallel Matrix and GCD Computation*, Information and Control (1982), **52**, pp. 478–504.

[BT 89] Borodin, A. and Tiwari, P. A., *On the Decidability of Sparse Univariate Polynomial Interpolation*, Research Report RC 14923, IBM T. J. Watson Research Center, New York, 1989, to appear in Proc. 21^{st} ACM STOC (1990).

[C 81] Cook, S. A., *Towards a Complexity Theory of Synchronous Parallel Computation*, Ensein. Math. (1981), **27**, pp. 99-124.

[C 85] Cook, S. A., *A Taxonomy of Problems with Fast Parallel Algorithms*, Information and Control (1985), **64**, pp. 2-22.

[EI 76] Evans, R.J. and Isaacs, I.M., *Generalized Vandermonde Determinants and Roots of Unity of Prime Order*, Proc. of the AMS (1976), **58**.

[G 89] Grigoriev, D. Yu., *Factoring and Calculating the GCDs of Linear Differential Operators*, to appear in Journal of Symbolic Computation (1989).

[GK 87] Grigoriev, D. Yu., and Karpinski, M., *The Matching Problem for Bipartite Graphs with Polynomially Bounded Permanents is in NC*, Proc. 28^{th} IEEE FOCS (1987), pp. 166–172.

[GKS 88] Grigoriev, D. Yu., Karpinski, M., and Singer, M., *Fast Parallel Algorithms for Sparse Multivariate Polynomial Interpolation over Finite Fields*, University of Bonn (1988), Research Report No. 8523-CS: to appear in SIAM J. Computing (1990).

[GV 88] Grigoriev, D. Yu., and Vorobjov, N. N., *Solving Systems of Polynomial Inequalities in Subexponential Time*, Journal of Symbolic Computation (1988), **5**, pp. 37–64.

[K 86] Kaltofen, E., *Uniform Closure Properties of P-computable Functions*, Proc. 18^{th} ACM STOC (1986), pp. 330–337.

[KL 88] Kaltofen, E. and Lakshman, Y., *Improved Sparse Multivariate Polynomial Interpolation Algorithms*, Proc. ISSAC '88, LCNS, Springer-Verlag Berlin (1988).

[KT 88] Kaltofen, E. and Trager, B., *Computing with Polynomials Given by Black Boxes for Their Evaluations: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators*, 29^{th} IEEE FOCS (1988), pp. 296–305.

[KR 88] Karp, R. M. and Ramachandran, V. L., *A Survey of Parallel Algorithms for Shared-Memory Machines*, Research Report No. UCB/CSD88/407, University of California, Berkeley (1988); to appear in: Handbook of Theoretical Computer Science, North Holland (1989).

[KW 89] Karpinski, M. and Werther, T., *VC Dimension and Learnability of Sparse Polynomials and Rational Functions*, Technical Report TR-89-060, International Computer Science Institute, Berkeley (1989).

[LO 83] Loos, R., *Generalized Polynomial Remainder Sequences*, in *Computer Algebra: Symbolic and Algebraic Computation, Second Edition*, Buchberger et al eds., Springer-Verlag, Wien (1983).

[M 86] Mulmuley, K., *A Fast Parallel Algorithm to Compute the Rank of a Matrix over an Arbitrary Field*, Proc. ACM STOC (1986).

[O 33] Ore, O. *Theory of Non-Commutative Polynomials*, Ann. of Math. (1933), **34**, pp. 480–508.

[S 73] Strassen, V., *Vermeidung von Divisionen*, J. Reine und Angewandte Math. (1973), **65**, pp. 182-202.

[VDW 66] van der Waerden, B. L., *Algebra*, Part I (esp. section 34), Springer Verlag Berlin (1966).