Fast Interpolation Algorithms for Sparse Polynomials with Respect to the Size of Coefficients

Alexander L. Chistov* **Marek Karpinski†**

Abstract

In this paper we consider the interpolation of sparse polynomials in two different oracle models taking into account the size of coefficients only.

*St. Petersburg Institute for Informatics and Automation of the Academy of Sciences of Russia, and Department of Computer Science, University of Bonn, 53117 Bonn. Research supported by the Volkswagen–Stiftung, Program on Computational Complexity.

†Department of Computer Science, University of Bonn, 53117 Bonn, and International Computer Science Institute, Berkeley, California, E Mail: marek@cs.uni-bonn.de. Research supported in part by the DFG Grant KA 673/4–1, by the ESPRIT BR Grants 7097 and ECU030, and by the Volkswagen–Stiftung.

Introduction

The models considered so far require exact computations, see [3],[4],[5], but in practice exact computations of values of sparse polynomials are very difficult. Indeed, we cannot even compute values of sparse polynomials in small integer points such as 2,3,.... Since the lengths of the values will be exponential in the size of input in the general case.

We suggest two models which afford to avoid these difficulties. The first is the interpolation with the modular oracle, see section 1, and the second the interpolation with the oracle which gives the real (or complex) evaluations of values of the considered polynomial and these evaluations have polynomial in the size of input lengths, see section 2. The simple proof of theorem 2 of section 2 was found after the discussion with D.Yu. Grigoriev by the authors and independently by S.A. Evdokimov.

In this paper for an integer a we define the bitwise length

$$l(a) = \min\{s \in \mathbb{Z} : |a| \leq 2^{s-1}\},$$

and if $q \in \mathbb{Q}$ then $l(q) = l(q_1) + l(q_2)$ where $q = q_1/q_2$; $q_1, q_2 \in \mathbb{Z}$, $GCD(q_1, q_2) = 1$.

1 Fast modular interpolation of sparse polynomials

Let $f \in \mathbb{Q}[X_1, \ldots, X_n]$ be a polynomial, $f = \sum_{(i_1,\ldots,i_n)\in I} f_{i_1,\ldots,i_n} X_1^{i_1} \ldots X_n^{i_n}$ where $0 \neq f_{i_1,\ldots,i_n} \in \mathbb{Q}$, $I \subset \mathbb{Z}_+^n$, $\#I = t$, $l \geq \max_{(i_1,\ldots,i_n)\in I} l(f_{i_1,\ldots,i_n})$, $\deg_{X_i}(f) <$ for every i. Therefore, the size of f is less than $tnl(\log(d) + 1)$.

We consider the following oracle:

INPUT: $(\overline{a}, p)$ where p is a prime number, $\overline{a} \in \mathbb{F}_p^n = (\mathbb{Z}/p\mathbb{Z})^n$.

OUTPUT: $f(\overline{a}) = f(a) mod\ p \in \mathbb{F}_p \cup \{*\}$, where $a \in \mathbb{Z}^n, a \bmod p = \overline{a}$ and $f(\overline{a}) = *$ iff there exists $(i_1, \ldots, i_n) \in I$ such that $f_{i_1,\ldots,i_n} = f'_{i_1,\ldots,i_n}/p^\alpha$, $1 \leq \alpha \in \mathbb{Z}$, p does not divide the numerator and the denominator of $f'_{i_1,\ldots,i_n}$.

We suppose that the working time of this oracle for input $(\overline{a}, p)$ is polynomial in $\log p, t, n, l, \log d$.

REMARK. One can consider also a slightly different oracle for which $f(\overline{a}) \in \mathbb{F}_p \cup \{\infty\}$ and $f(\overline{a}) = \infty$ iff $f(a) = q_1/(q_2 p^\alpha)$ with $1 \leq \alpha \in \mathbb{Z}$, $q_1, q_2 \in \mathbb{Z}$, $GCD(q_1, q_2, p) = 1$. For this oracle one can also prove the formulated below theorem. The proof is almost without changes.

THEOREM 1. *Using the oracle described one can reconstruct f in time polynomial in $t, l, n, \log d$.*

PROOF Consider at first the case when $n = 1, f \in \mathbb{Q}[\mathbb{X}]$. We need the following auxiliary algorithm.

AUXILIARY ALGORITHM:
INPUT: $s \in \mathbb{N}$
DESCRIPTION: Find by enumerating a minimal prime $p_s = p \equiv 1 \; mod \, s$. Find by enumerating $\zeta \in \mathbb{F}_p$ such that $\zeta^p = 1, \zeta \neq 1$. Compute using the oracle $f(\zeta^t)$, $0 \leq t < s$. If $f(\zeta^t) \neq *$ for all t then solve the linear system

$$\sum_{0 \leq j < s} \lambda_j \zeta^{tj} = f(\zeta^t) \qquad 0 \leq t < s$$

and find $\lambda_0, \ldots, \lambda_{s-1} \in \mathbb{F}_p$.
OUTPUT: (i) The element $\lambda^{(s)} = \lambda = \sum_{j \in J_s} \lambda_j \sigma^j \in \mathbb{F}_p[\sigma]$ where $\mathbb{F}_p[\sigma] = \mathbb{F}_p[\mathbb{X}]/(\mathbb{X}^s - 1)$ is the group algebra of the cyclic group of the order s, $\sigma = X \; mod \, (X^s - 1)$, and the set $J_s = \{j : \lambda_j \neq 0, \; 0 \leq j < s\}$.

(ii) The symbol $*$ if $f(\zeta^t) = *$ for some $0 \leq t < s$.

We shall identify J_s with the subset of $\mathbb{Z}/s\mathbb{Z}$.

Note that
(1) by Linnik's theorem, see [6] $p \leq s^c$ where c is constant,
(2)

$$\lambda_j = \sum_{\{i:\, i \; mod \, s = j \; \& \; i \in I\}} f_i \; mod \, p$$

for every $j \in J_s$,
(3) $f_i = 0$ if $i \; mod \, s \notin J_s$.

Denote $I_{s,j} = \{i : \; i \; mod \, s = j \; \& \; i \in I\}$ for every $0 \leq j < s$.

MAIN ALGORITHM (for $n = 1$)
Find the finite set S consisting of successive primes $2, 3, \ldots$ such that

$$\prod_{s \in S} s > \max\{2^{l+1}, d\} d^{t(t-1)/2} \, 2^{2lt}$$

For every $s \in S$ apply the auxiliary algorithm to the input s. Let S_1 be the subset of $s \in S$ such that the auxiliary algorithm with the input s has output (i). Set

$$\alpha = \max_{s \in S_1} \# J_s$$

$$S_2 = \{s \in S_1 : \# J_s = \alpha\}$$

LEMMA (i) $\alpha = t$
(ii) $\prod_{s \in S_2} s > \max\{2^{l+1}, d\} \, 2^{tl}$
(iii) $\# I_{s,j} = 1$ for every $s \in S_2$ and $j \in J_s$.

PROOF $\alpha \leq t$, and $\alpha = t$ implies (iii).
Note that $\prod_{s \in S \setminus S_1} s \leq LCM_{s \in S \setminus S_1}\{p_s\} \leq LCM_{0 \leq i \leq deg(f)}\{denominator(f_i)\} \leq$

2^{lt}, since s and p_s are primes and $p_s \equiv 1 \; mod \; s$. So $\prod_{s \in S_1} s \geq (\prod_{s \in S} s)/2^{lt} > \max\{2^{l+1}, d\} d^{t(t-1)/2} 2^{lt}$. Let $N = \prod_{i_1 > i_2;\, i_1, i_2 \in I} (i_1 - i_2)$. Then $N < d^{t(t-1)/2}$. The conditions $s \in S_1$ and s does not divide N imply $\alpha = t$, since $\lambda_j = \sum_{i \in I_{s,j}} f_i \; mod \, p$. We have $\prod_{s \in S_1, s|N} s < N$, since $s \in S_1$ are different primes. Therefore, $\prod_{s \in S_2} s \geq (\prod_{s \in S_1} s)/(\prod_{s \in S_1, s|N} s) > (\prod_{s \in S_1} s)/d^{t(t-1)/2} > \max\{2^{l+1}, d\} 2^{lt}$. Lemma is proved.

Fix $s_0 \in S_2$. For every $s \in S_2, s \neq s_0$ apply the auxiliary algorithm to the input ss_0. Denote by S_3 the subset of $s \in S_2$ such that the auxiliary algorithm with input ss_0 has output (i), i.e. for every $s \in S_3$ we get in output of the auxiliary algorithm $\lambda^{(ss_0)}$ and J_{ss_0}.

Note that
$\prod_{s \in S_2 \setminus S_3} s \leq LCM_{s \in S_2 \setminus S_3} \{p_{ss_0}\} \leq LCM_{0 \leq i \leq deg(f)} \{denominator(f_i)\} < 2^{lt}$, since s and p_{ss_0} are primes and $p_{ss_0} \equiv 1 \; mod \; (ss_0)$.
Therefore, $\prod_{s \in S_3} s \geq (\prod_{s \in S_2} s)/2^{lt} > \max\{2^{l+1}, d\}$.

Construct the mappings

$$\beta_0 : J_{ss_0} \longrightarrow J_{s_0} \quad and$$
$$\beta_s : J_{ss_0} \longrightarrow J_s,$$

which are reductions $mod \, s_0$ and $mod \, s$ respectively for every $s \in S_3, s \neq s_0$. The mappings β_0 and β_s are bijective, since $\# J_{ss_0} = \# J_s = t$ and $\beta_0(J_{ss_0}) = J_{s_0}$, $\beta_s(J_{ss_0}) = J_s$ by (2), see above.

Using chinese reminders theorem find minimal $u_j \in \mathbb{Z}$, $\nvdash \leq \approxeq \daleth <$, $\daleth \in \mathbb{J}_{\sim\mu}$ such that

$$u_j \; mod \; s \; = \beta_s \beta_0^{-1}(j),$$
$$u_j \; mod \; s_0 = j$$

for all $j \in S_3$. It is possible, since $\prod_{s \in S_3} s > d$.

We have $I = \{u_j : j \in J_{s_0}\}$ by (2). Again applying chinese reminders theorem find f_i, $i \in I$ from the conditions

$$f_i \; mod \; p_s = \lambda_j \quad , \quad |f_i| < 2^l,$$

where $j = i \; mod \; s \in J_s$ for every $s \in S_3$. It is possible since $LCM_{s \in S_3}\{p_s\} > \prod_{s \in S_3} s > 2^{l+1}$.

Thus, we can reconstruct f in the required time in the case $n = 1$. The case of n variables is reduced to $n = 1$ by the substitution $X_i = X^{d^{i-1}}, 1 \leq i \leq n$. Denote $\bar{f} = f(X, X^d, \ldots, X^{d^{n-1}})$. The oracle for f gives the oracle for $\bar{f}$. So we can reconstruct $\bar{f}$ in time polynomial in $t, l, \log(nd^n) + 1$, i.e. polynomial in $t, n, l, \log(d) + 1$. Then knowing $\bar{f}$ one can easy find f. The theorem is proved.

2 Fast interpolation of sparse polynomials with real and complex coefficient

Let f be the same as in section 1. Consider the following oracle
INPUT $(a_1, \ldots, a_n) \in \mathbb{Q}^{\ltimes}$ and polynomials P_1, P_2 in 4 variables with integer coefficients.

OUTPUT (i) $u \in \mathbb{Q}$ such that $|f(a) - u| < 2^{-P_1(t,l,n,\log d)}$,
(ii) the symbol $*$ if u does not exist.
The working time of this oracle is polynomial in $\sum_{1 \le i \le n} l(a_i)$, t, l, n.

THEOREM 2. *Using the oracle described one can reconstruct f in time polynomial in* $l, n, t \log d$.

PROOF Consider the case $n = 1$, $f = \sum_{1 \le i \le t} f_i X^{b_i}$. Let $\epsilon > 0$. Consider the expansion

$$f(1+\epsilon) = \sum_i f_i + \epsilon \sum_i f_i b_i + \epsilon^2 \sum_i f_i \binom{b_i}{2} + \cdots$$

Choose ϵ and the oracle such that we can find from this expansion $2t$ terms $\sum_i f_i \binom{b_i}{j}$, $0 \le j \le 2t$ with the exactness 2^{-2lt-1}. It is possible, since $|\sum_i f_i \binom{b_i}{j}| < t2^{tl} d^j$. For example, one can take $\epsilon = (2^{2lt+l+2} d^{2t} t^2)^{-1}$. So we can find $q_j \in \mathbb{Q}$ such that $|\sum_i f_i \binom{b_i}{j} - q_j| < 1/2^{2lt+1}$ and $l(q_j) < P(t,l,n,\log d)$ for some polynomial P. But the denominator of each $\sum_i f_i \binom{b_i}{j} = u_j$ is less than 2^{lt}. So u_j is the uniquely determined appropriate traction in the expansion of q_j in the chain fraction. It can be found in time polynomial in t, l, n, $\log d$.

Thus, we can find all u_j, $0 \le j \le 2t$, and, therefore, all v_j, $1 \le j \le 2t$, where

$$\sum_{1 \le i \le t} f_i b_i^j = v_j \quad , \quad 1 \le j \le 2t.$$

Now we can find from this system, as it is well known in the theory of interpolation of sparse polynomials, all f_i and b_i, $1 \le i \le t$.

Remind how it can be done. Consider the linear operator $A : \mathbb{R}^{\approx} \to \mathbb{R}^{\approx}$
$\mathbb{A}((\setminus_{\nVdash}, \ldots, \setminus_{\approx})^{\mathbb{T}}) = (_{\nVdash}\setminus_{\nVdash}, \ldots, _{\approx}\setminus_{\approx})^{\mathbb{T}}$ (T denotes the transponation). The eigenvalues of A are $b_1, \ldots, b_t$. Let $F = (f_1, \ldots, f_t)^T$. Then $F, AF, \ldots, A^{t-1}F$ is a basis of $\mathbb{R}^{\approx}$. Let $\sigma : \mathbb{R}^{\ltimes} \to \mathbb{R}$ be the sum of coordinate, i.e. $\sigma((r_1, \ldots, r_t)^T) = r_1 + \ldots + r_t$. Consider the following matrix

$$(v_{i+j-2})_{i,j} = \left(\begin{array}{cccc|c} \sigma F & \sigma AF & \cdots & \sigma A^{t-1}F & \sigma A^t F \\ \sigma AF & \sigma A^2 F & \cdots & \sigma A^t F & \sigma A^{t+1} F \\ \vdots & \vdots & & \vdots & \vdots \\ \sigma A^{t-1}F & \sigma A^t F & \cdots & \sigma A^{2t-2}F & \sigma A^{2t-1}F \end{array} \right)$$

The first t columns of this matrix are linearly independent. Indeed, otherwise there exist $\lambda_1, \ldots, \lambda_t \in \mathbb{R}$, $(\lambda_{\nVdash}, \ldots, \lambda_{\approx}) \neq (\nvdash, \ldots, \nvdash)$ such that

$$(\sigma A^j)(\sum_{1 \le i < t} \lambda_i A^{i-1} F) = 0 \quad , \quad 0 \le j < t$$

i.e. $0 \neq \sum_i \lambda_i A^{i-1} F \in \bigcap_{0 \le j < t} Ker(\sigma A^j) = \{0\}$ and we get the contradiction which proves our assertion.

Therefore, there exist unique $\mu_0, \ldots, \mu_{t-1} \in \mathbb{R}$ such that $(\sigma A^j)(A^t + \sum_{0 \le i < t} \mu_i A^{i-1})(F) = 0$, $0 \le j < t$. By the same argument as above we get $(A^t + \sum_{0 \le i < t} \mu_i A^{i-1})(F) = 0$. It follows from here that

$$(A^t + \sum_{0 \le i < t} \mu_i A^{i-1})(A^j F) = 0$$

for all j. It means that $Z^t + \sum_i \mu_i Z^i$ is the characteristic polynomial of A (up to the sign). We can find μ_i solving the linear system for the linear dependence of columns of the matrix $(v_{i+j-2})_{i,j}$ and then find b_i, $1 \le i \le t$, which are roots of $Z^t + \sum_i \mu_i Z^i$. After that solving linear system we find f_i, $1 \le i \le t$. Thus we reconstruct f in the case $n = 1$.

In the case of many variables we can proceed similarly to that it was in section 1 by reduction from arbitrary n to $n = 1$. The theorem is proved.

REMARK. We can change everywhere in the definitions of f, the oracle, ... and the statement of the theorem 2 the field $\mathbb{Q}$ for the field $\mathbb{Q}[\beth]$ where $i = \sqrt{-1}$. The theorem will be true also in this case. The proof is almost without changes.

References

[1] Grigoriev D.Y., Karpinski M., Singer M.F., *The Interpolation Problem for k-Sparse Sums of Eigenfunctions of Operators.* Advances in Applied Mathematics 12 (1991) pp. 76–81.

[2] Grigoriev D.Y., Karpinski M., *Algorithms for Sparse Rational Interpolations.* Proc. ISSAC, 1991, pp. 7–13.

[3] Grigoriev D.Y., Karpinski M., Singer M.F., *Fast Parallel Algorithms for Sparse Multivariate Polynomial Interpolation over Finite Fields.* SIAM Journal of Comput. 19, #6 (1990) pp. 1059–1063.

[4] Kaltofen E., Yagati, L., *Improved Sparse Multivariate Polynomial Interpolation Algorithms.* Preprint. Rensselaer Polytechnic Institute, 1988.

[5] Karpinski, M., *Boolean Circuit Complexity of Algebraic Interpolation Problems,* Proc. CSL '88, LNCS 385 (1989), Springer Verlag, pp. 138–147.

[6] Zippel R. *Interpolating Polynomials from their Values.* Journal of Symbolic Computation 9 (1990) pp. 375–403.

[7] Prachar K. *Primzahlverteilung,* Springer Verlag, Berlin Göttingen Heidelberg, 1957.