Sparse Polynomial Interpolation in Non-standard Bases

Lakshman Y. N.[1]

Department of Mathematics and Computer Science
Drexel University
Philadelphia, PA 19104

B. David Saunders[2]

Department of Computer and Information Sciences
University of Delaware
Newark, DE 19716
e-mail: `lakshman@mcs.drexel.edu, saunders@udel.edu`

Abstract

In this paper, we consider the problem of interpolating univariate polynomials over a field of characteristic zero that are sparse in (a) the Pochhammer basis or, (b) the Chebyshev basis. The polynomials are assumed to be given by black boxes, i.e., one can obtain the value of a polynomial at any point by querying its black box. We describe efficient new algorithms for these problems. Our algorithms may be regarded as generalizations of Ben-Or and Tiwari's (1988) algorithm (based on the BCH decoding algorithm) for interpolating polynomials that are sparse in the standard basis. The arithmetic complexity of the algorithms is $O(t^2 + t \log d)$ which is also the complexity of the univariate version of the Ben-Or and Tiwari algorithm. That algorithm and those presented here also share the requirement of $2t$ evaluation points.

Key words: polynomial interpolation, sparsity, Chebyshev polynomial, Pochhamer basis, BCH codes.

AMS(MOS) subject classifications: 68Q40, 12Y05, 41A05.

Abbreviated title: Sparse interpolation, non-standard bases.

[1]Work supported in part by NSF grant CCR–9203062
[2]Work supported in part by NSF grant CCR-9123666.

Introduction

In this paper, we consider the problem of efficiently interpolating polynomials given by black boxes that have sparse representations in various bases. Usually, one considers a polynomial $f(x) = \sum_{i=0}^{n} a_i x^i$ as t-sparse if at most t of the coefficients a_i are non-zero. The problem of interpolating a sparse polynomial from a list of values at specific points, or from values given by a black box that evaluates the polynomial is of great importance in computational algebra. Sparse interpolation has received significant attention in several recent papers (Ben-or and Tiwari (1988), Clausen et al (1988), Kaltofen and Lakshman (1988), Grigoriev and Karpinski (1987), Grigoriev, Karpinski, and Singer (1990, 1991a, 1991b), Borodin and Tiwari (1990), Zippel (1990)). In particular, the problem of interpolating a t-sparse univariate polynomial given a black box for evaluating the polynomial can be efficiently solved using Ben-Or and Tiwari's adaptation of the BCH decoding algorithm.

Sparse interpolation is well recognized as a very useful tool for controlling intermediate expression swell in computer algebra (Zippel 1990, Kaltofen-Trager 1990). Sparse polynomials and rational functions can be evaluated quickly and that makes them attractive to several applications and an interesting line of research is to try to infer properties of sparse polynomials (such as divisibility, existence of non-trivial gcd, the existence of real roots etc.) from their values at a small number points (see Grigoriev-Karpinski-Odlyzko 1992). Traditionally, "sparse polynomial (or rational function)" is taken to mean a polynomial (or rational function) with a "few terms" where the "terms" are power products of the variables involved. One can reasonably ask for sparse representations for polynomials in other bases such as the Chebyshev polynomials or the shifted power basis $1, x-\alpha, (x-\alpha)^2, \ldots$. In this paper, we consider two classes of polynomials – those which are sparse in the Chebyshev basis and those which are sparse in the Pochhammer basis. A polynomial $f(x) = \sum_{i=0}^{n} a_i T_i(x)$ is t-sparse in the Chebyshev basis if at most t of the coefficients a_i are non-zero, where $T_i(x)$ is the i-th Chebyshev polynomial. Sparse Pochhammer polynomials are defined similarly. We provide algorithms for interpolating such polynomials efficiently given a black box that evaluates the polynomial. The algorithms require as input an upper bound t on the number of non-zero terms in the interpolating polynomial. The arithmetic complexity of the algorithms is $O(t^2 + t \log d)$ which is also the complexity of the univariate version of the Ben-Or and Tiwari algorithm. That algorithm and those presented here also share the requirement of $2t$ evaluation points. However, when one considers the bit complexity (when the ground field is $\mathbf{Q}$) the Ben-Or and Tiwari algorithm and our sparse Chebyshev algorithm involve intermediate values of bit size $O(td)$, whereas in the sparse Pochhammer algorithm intermediate values grow to a bit length of $O(d \log(t))$.

Several recent results concerning sparse interpolation are similar in spirit to the BCH-decoding algorithm. In (Dress and Grabmeier 1991), it is shown that many of the sparse interpolation algorithms can be formulated and proven correct in the context of t-sparse sums of abelian monoids. A somewhat different framework in Grigoriev, Karpinski, and Singer (1991a) in terms of t-sparse sums of eigenfunctions of operators is shown to provide several generalizations of the sparse interpolation algorithm of Ben-Or and Tiwari. In fact, an efficient algorithm for interpolating polynomials sparse in the Pochhammer basis follows from an observation in (Grigoriev, Karpinski, and Singer (1991a)) pointing out that t-sparse sums of eigenfunctions of of the operator $x\Delta(f) = x(f(x) - f(x-1))$ correspond to polynomials that are t-sparse in the Pochhammer basis. We use the Pochhammer case as a motivating example for our discussion of the Chebyshev case.

The problem of efficiently interpolating polynomials that are sparse in the Chebyshev basis was stated as an open problem in Borodin and Tiwari (1990). Our algorithm provides a solution and it is another generalization of the algorithm of Ben-Or and Tiwari. It uses properties shared by the standard power basis and the Chebyshev polynomials and appears to be different from the generalizations presented in (Dress and Grabmeier (1991), Grigoriev, Karpinski, and Singer (1991a)). While the general structure of our algorithm appears similar to the algorithms falling into the Dress-Grabmeier framework, the details are quite different. In particular, the centerpiece of the algorithms in the Dress-Grabmeier framework is a certain Toeplitz matrix where as in our algorithm, a similar role is played by a matrix that is *the sum of a Toeplitz matrix and a Hankel matrix.* It would be interesting to reconcile our algorithm with the Dress and Grabmeier or the Grigoriev, Karpinski, and Singer framework.

The rest of the paper is organized as follows. In section 1, we motivate our discussion by briefly stating the algorithm for interpolating polynomials sparse in the Pochhammer basis. In section 2, we describe our algorithm for interpolating polynomials sparse in the Chebyshev basis. Section 3 provides an analysis of the complexity of the algorithm which is followed by some additional remarks in the last section.

1. Sparse Interpolation in the Pochhammer Basis

The Pochhammer symbol $x^{\overline{n}}$ denotes the rising factorial power

$$x(x+1)\dots(x+n-1)$$

for any integer $n \geq 0$. It is easily shown that $x^{\overline{i}},\ i = 0, 1, 2, \dots$ is a $\mathbf{Q}$-basis for the polynomial ring $\mathbf{Q}[x]$. A polynomial $f(x)$ is t-sparse in the Pochhammer basis (of rising powers) if and only if $\exists f_i \in \mathbf{Q},\ e_i \in \mathcal{Z}_+,\ i = 1, \dots, t$, such that

$$f(x) = \sum_{k=1}^{t} f_k x^{\overline{e_k}}.$$

We assume that $e_1 > e_2 > \dots > e_t$. Suppose we are given a black box that returns the value of $f(x)$ at any $x \in \mathbf{Q}$ and a bound t on the number of non-zero terms in the representation of $f(x)$ in the Pochhammer basis. Then, we can interpolate $f(x)$ from its values at $x = 1, 2, \ldots, 2t$. The algorithm follows from an observation in (Grigoriev, Karpinski, Singer 1991a) which places this problem in the framework of Dress and Grabmeier. It is an elementary fact that the finite difference operator behaves like the derivative on the Pochhammer symbol. More precisely, let

$$\Delta(f(x)) = f(x+1) - f(x).$$

We have

$$\begin{aligned} \Delta(x^{\overline{n}}) &= (x+1)^{\overline{n}} - x^{\overline{n}} \\ &= n(x+1)^{\overline{n-1}}. \end{aligned}$$

Let

$$f^{(i)}(x) = \sum_{k=1}^{t} e_k^i f_k x^{\overline{e_k}}, \text{ for } i = 0, 1, \ldots$$

Note that

$$x\Delta(f^{(i)}(x)) = f^{(i+1)}(x).$$

We can compute $f^{(i)}(a)$ for $i = 0, 1, \ldots, 2t-1$ by repeated applications of the above recurrence from the values $f(a+i),\ i = 0, 1, \ldots, 2t-1$.

Consider the polynomial $\Psi(z)$ of degree t whose roots are the $e_i,\ i = 1, \ldots, t$, i.e.,

$$\Psi(z) = \prod_{i=1}^{t}(z - e_i).$$

Let

$$\Psi(z) = z^t + \psi_{t-1}z^{t-1} + \ldots + \psi_1 z + \psi_0.$$

The $f^{(i)}$ and ψ_i satisfy the following linear relations. Assume that $\psi_t = 1$.

Lemma 1 *For* $j = 0, \ldots, t-1, \quad \sum_{i=0}^{t} \psi_i f^{(i+j)}(a) = 0.$

Proof: This is a special case of Theorem 1 in Dress and Grabmeier (1991) (see page 62 in that paper). □

Furthermore, Theorem 1 in Dress and Grabmeier (1991) shows that if t is the number of non-zero terms in the representation of $f(x)$ in the Pochhammer basis, then, the $t \times t$ matrix $\mathcal{F}$ with $\mathcal{F}_{i,j} = f^{(i+j-2)}(a)$ is non-singular for any $a > 0$. Using this, one can compute the polynomial $\Psi(z)$ from $f(a), f(a+1), \ldots, f(a+2t-1)$. The roots of $\Psi(z)$ are the "Pochhammer exponents" of $f(x)$ knowing which one can easily compute the coefficients f_k. We omit details to avoid repetition but state the algorithm for the sake of completeness.

Algorithm Sparse-Pochhammer-Interpolation

Input: A black box for evaluating a polynomial $f(x)$. $f(x)$ is known to be t-sparse in the Pochhammer basis.
Output: A list of pairs $(f_i, e_i), i = 1, \ldots, t$ such that $f(x) = \sum_{i=1}^{t} f_i x^{\overline{e_i}}$.

1. *Query the black box to obtain the values of $f(x)$ at $x = a, a+1, \ldots, a+2t-1$ for some $a \neq 0$. Compute $f^{(i)}(a)$ for $i = 0, 1, \ldots, 2t-1$ from $f(a+i)$, $i = 0, 1, \ldots, 2t-1$.*

2. *Solve the system of linear equations $\mathcal{F}\vec{\psi} = -\vec{\mathsf{f}}$. If*

$$\vec{\psi} = \begin{pmatrix} \psi_0 & \psi_1 & \ldots & \psi_{t-1} \end{pmatrix}^T,$$

then, the auxiliary polynomial $\Psi(z)$ is given by

$$\Psi(z) = z^t + \psi_{t-1} z^{t-1} + \ldots + \psi_1 z + \phi_0.$$

3. *Find the integer roots of $\Psi(z)$. The roots are the exponents of $f(x)$, i.e., $e_1, e_2, \ldots, e_t$.*

4. *Solve the linear system of equations $\mathcal{V}\vec{\mathsf{f}} = \vec{\mathsf{F}}$ to obtain $f_1 a^{\overline{e_1}}, \ldots, f_t a^{\overline{e_t}}$. Since we know the e_i and a, the coefficients f_i are easily determined. Output the list of pairs $[(f_1, e_1), \ldots, (f_t, e_t)]$.*

Analysis of the Sparse-Pochhammer-Interpolation algorithm: While the sparse Pochhammer interpolation algorithm and the algorithm of Ben-Or and Tiwari for interpolating polynomials sparse in the standard power basis both illustrate the Dress-Grabmeier framework, the details are quite different. The sparse Pochhammer algorithm differs in a fundamental way from the algorithm of Ben-Or and Tiwari. The counterpart of the auxiliary polynomial $\Psi(z)$ in Ben-Or and Tiwari's algorithm has as roots a^{e_i}, whereas here, the auxiliary polynomial $\Psi(z)$ has as roots the *degrees* e_i of the non-zero terms in the interpolating polynomial. This has implications on how we actually compute the roots of the auxiliary polynomial. Also, this algorithm uses a completely different set of evaluation points (importantly, the evaluation points are smaller).

The analysis will be brief, pointing out the effect of the differences just mentioned. We assume unit cost for a black box query. The $f^{(i)}(a)$ for $i = 0, 1. \ldots, 2t-1$ can be computed from $f(a+i)$, $i = 0, 1, \ldots, 2t-1$ using $O(t^2)$ field operations. The Toeplitz system of equations $\mathcal{F}\vec{\psi} = -\vec{\mathsf{f}}$ can be solved using $O(t^2)$ field operations using the Berlekamp-Massey algorithm (Kaltofen and Lakshman (1988)) to find the coefficients of the auxiliary polynomial $\Psi(z)$. We can use the Cantor-Zassenhaus algorithm to factor $\Psi(z)$ modulo a suitable prime p and use Hensel lifting to recover the integer roots. The cost is $O(t^2 + t \log d)$, where d is the degree of the interpolating polynomial (see, for instance Loos (1983)). The final step, of solving a transposed Vandermonde system of equations can be done using the algorithm of Zippel (Zippel (1990)) in $O(t^2)$ field operations. We can now add up the costs of the 4 steps and we have:

Theorem 1 *The sparse Pochhammer interpolation algorithm performs $O(t^2+t\log d)$ field operations to interpolate a t-Pochhammer-sparse polynomial given by a black box. Here, d is the degree of the polynomial being interpolated (not part of the input). The algorithm makes $2t$ queries to the black box evaluating the polynomial. The numerator and denominator of the rational numbers that arise during the algorithm have $O(d\log(t))$ bits.* □

2. Sparse Interpolation in the Chebyshev Basis

In this section, we describe an efficient algorithm for interpolating polynomials in $\mathbf{Q}[x]$ that are sparse in the Chebyshev basis, where $\mathbf{Q}$ is the field of rational numbers. The algorithm is entirely rational, so it applies over any field of characteristic zero. The analysis in section 3 in terms of arithmetic operations would apply in the general characteristic zero case as well, provided the evaluation point a is rational and $a \geq 2$. The bit complexity analysis, of course, is specific to the base field $\mathbf{Q}$. Let $T_n(x)$ denote the n-th Chebyshev polynomial of the first kind:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \text{ for } n \geq 2.$$

It is well known that $T_i(x),\ i = 0, 1, 2, \ldots$ is a $\mathbf{Q}$-basis for the polynomial ring $\mathbf{Q}[x]$.

Let $f(x) = \sum_{i=0}^{m} f_i T_i(x)$ be a polynomial in the ring $\mathbf{Q}[x]$. As usual, we say that $f(x)$ is t-sparse (in the Chebyshev basis) if and only if at most t of the f_i's are non-zero.

Suppose we are given a black box that returns the value of $f(a)$ for any $a \in \mathbf{Q}$. We present an algorithm that interpolates $f(x)$ from its values at $2t$ specially chosen points. Our algorithm may be regarded as a generalization of the Ben-Or and Tiwari algorithm (Ben-Or and Tiwari (1988)). We identify two crucial properties that are required for the generalization to work. In the context of Chebyshev polynomials, they are:

- for $m, n \geq 0$, $T_n(T_m(x)) = T_{mn}(x) = T_m(T_n(x))$, i.e., Chebyshev polynomials commute with respect to composition.
- for $m, n \geq 0$, $T_m(x)T_n(x) = 1/2(T_{m+n}(x) + T_{|m-n|}(x))$, i.e., the product of two Chebyshev polynomials is a *fixed* linear form in two others.

These properties are easy consequences of well known properties of the Chebyshev polynomials. Before describing our interpolation algorithm, we need a statement of uniqueness of a sparse interpolating polynomial.

Lemma 2 *Let $f(x)$ be a t-sparse polynomial (in the Chebyshev basis). If for some $a > 1$, $f(T_i(a)) = 0$ for $i = 0, 1, 2, \ldots, t-1$, then $f(x)$ is identically zero.*

Proof: Since $f(x)$ is t-sparse, there exist $f_i \in \mathbf{Q}$ and $\delta_i \in \mathcal{Z}_+$, for $i = 1, \ldots, t$ such that

$$f(x) = f_1 T_{\delta_1}(x) + f_2 T_{\delta_2}(x) + \ldots + f_t T_{\delta_t}(x).$$

Now, we can rewrite the fact that $f(T_i(a)) = 0$ for $i = 0, 1, 2, \ldots, t-1$ as $\mathcal{V}\vec{\mathsf{f}} = \vec{\mathbf{0}}$ where

$$\mathcal{V} = \begin{pmatrix} T_{\delta_1}(T_0(a)) & T_{\delta_2}(T_0(a)) & \ldots & T_{\delta_t}(T_0(a)) \\ T_{\delta_1}(T_1(a)) & T_{\delta_2}(T_1(a)) & \ldots & T_{\delta_t}(T_1(a)) \\ \vdots & \vdots & \ddots & \vdots \\ T_{\delta_1}(T_{t-1}(a)) & T_{\delta_2}(T_{t-1}(a)) & \ldots & T_{\delta_t}(T_{t-1}(a)) \end{pmatrix}, \qquad \vec{\mathsf{f}} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_t \end{pmatrix}.$$

If f is not identically zero, then $\mathcal{V}$ is singular. Let $\vec{c} = (c_0, c_1, \ldots, c_{t-1})^T$ be such that $\vec{c}^T\mathcal{V} = 0$. Consider the polynomial

$$C(x) = c_0 T_0(x) + c_1 T_1(x) + \ldots + c_{t-1} T_{t-1}(x)$$

of degree at most $t-1$. Using the fact that $T_{\delta_j}(T_i(a)) = T_i(T_{\delta_j}(a))$, re-write $\mathcal{V}$ as

$$\mathcal{V} = \begin{pmatrix} T_0(T_{\delta_1}(a)) & T_0(T_{\delta_2}(a)) & \ldots & T_0(T_{\delta_t}(a)) \\ T_1(T_{\delta_1}(a)) & T_1(T_{\delta_2}(a)) & \ldots & T_1(T_{\delta_t}(a)) \\ \vdots & \vdots & \ddots & \vdots \\ T_{t-1}(T_{\delta_1}(a)) & T_{t-1}(T_{\delta_2}(a)) & \ldots & T_{t-1}(T_{\delta_t}(a)) \end{pmatrix}$$

Now, $\vec{c}^T\mathcal{V} = 0$ implies that $T_{\delta_1}(a), T_{\delta_2}(a), \ldots, T_{\delta_t}(a)$ are all roots of $C(x)$. These t values are distinct (in fact, the values $T_i(a)$ are strictly monotonic increasing if $a > 1$, which is straightforward to check from the defining recurrence). But $C(x)$ is of degree $t-1$ and cannot have t roots. Hence, $\mathcal{V}$ cannot be singular, and f must be identically zero. □

As a consequence, we have

Corollary 1 *If $f(x)$ and $g(x)$ are two distinct t-sparse polynomials, then, for each $a > 1$, there is an i, $0 \le i \le 2t-1$ such that $f(T_i(a)) \ne g(T_i(a))$.*

Proof: Consider $h(x) = f(x) - g(x)$. The polynomial $h(x)$ is $2t$-sparse and if the corollary is not true, then, $h(T_i(a)) = 0$ for $i = 0, \ldots, 2t-1$. Apply previous lemma. □

We can now describe the sparse interpolation algorithm. As before, $f(x) = \sum_{i=1}^{t} f_i T_{\delta_i}(x)$. The algorithm determines the δ_i first. Once the δ_i are known, it is a simple task to compute the coefficients f_i. Let $a_i = f(T_i(a)), i = 0, 1, \ldots, 2t-1$ for some $a > 1$. Consider the polynomial of degree t whose roots are $T_{\delta_i}(a)$ represented in the Chebyshev basis, normalized to have a leading coefficient of 1, i.e.,

$$\Phi(z) = T_t(z) + \phi_{t-1}T_{t-1}(z) + \ldots + \phi_0 T_0(z)$$

and

$$\Phi(T_{\delta_i}(a)) = 0 \quad \text{for} \quad i = 1, \ldots, t.$$

Our strategy is to compute $\Phi(z)$ and then find its integer roots. First, we show that the coefficients of $\Phi(z)$ satisfy the following linear relations:

Lemma 3

$$\sum_{j=0}^{t-1} \phi_j \times (a_{i+j} + a_{|j-i|}) = -(a_{i+t} + a_{|t-i|})$$

for $i = 0, 1, 2, \ldots$

Proof: Assume $\phi_t = 1$.

$$\begin{aligned}
(\sum_{j=0}^{t-1} \phi_j a_{j+i}) + a_{t+i} &= \sum_{j=0}^{t} \phi_j (\sum_{l=1}^{t} f_l T_{\delta_l}(T_{j+i}(a))) \\
&= \sum_{l=1}^{t} f_l (\sum_{j=0}^{t} \phi_j T_{\delta_l}(T_{j+i}(a))) = \sum_{l=1}^{t} f_l (\sum_{j=0}^{t} \phi_j T_{j+i}(T_{\delta_l}(a))) \\
&= \sum_{l=1}^{t} f_l (\sum_{j=0}^{t} \phi_j (2T_j(T_{\delta_l}(a))T_i(T_{\delta_l}(a)) - T_{|j-i|}(T_{\delta_l}(a)))) \\
&= \sum_{l=1}^{t} f_l [\Phi(T_{\delta_l}(a)) \times 2T_i(T_{\delta_l}(a)) - (\sum_{j=0}^{t} \phi_j T_{|j-i|}(T_{\delta_l}(a)))] \\
&= -\sum_{l=1}^{t} f_l (\sum_{j=0}^{t} \phi_j T_{\delta_l}(T_{|j-i|}(a))) \\
= -\phi_j (\sum_{l=1}^{t} f_l T_{\delta_l}(T_{|j-i|}(a))) &= -\sum_{j=0}^{t} \phi_j a_{|j-i|} = -(\sum_{j=0}^{t-1} \phi_j a_{|j-i|}) - a_{|t-i|}.
\end{aligned}$$

Hence, $\sum_{j=0}^{t-1} \phi_j (a_{i+j} + a_{|j-i|}) = -(a_{i+t} + a_{|t-i|})$. □

The above lemma says that the coefficients of $\Phi(z)$ satisfy the system of equations

$$\mathcal{A}\vec{\phi} = -\vec{\alpha}$$

where

$$\mathcal{A} = \begin{pmatrix} 2a_0 & 2a_1 & \dots & 2a_{t-1} \\ 2a_1 & a_2 + a_0 & \dots & a_t + a_{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ 2a_{t-1} & a_t + a_{t-2} & \dots & a_{2t-2} + a_0 \end{pmatrix}, \quad \vec{\phi} = \begin{pmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_{t-1} \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} 2a_t \\ a_{t+1} + a_{t-1} \\ \vdots \\ a_{2t-1} + a_1 \end{pmatrix}.$$

Notice that $\mathcal{A}_{i,j} = a_{i+j} + a_{|i-j|}$ is the sum of a Toeplitz matrix and a symmetric Hankel matrix. Next, we show that $\mathcal{A}$ is non-singular.

Lemma 4 *$\mathcal{A}$ is non-singular.*

Proof: Consider the (transposed) *Vandermonde-like* matrix $\mathcal{V}$ and diagonal matrix $\mathcal{B}$.

$$\mathcal{V} = \begin{pmatrix} T_{\delta_1}(T_0(a)) & T_{\delta_2}(T_0(a)) & \dots & T_{\delta_t}(T_0(a)) \\ T_{\delta_1}(T_1(a)) & T_{\delta_2}(T_1(a)) & \dots & T_{\delta_t}(T_1(a)) \\ \vdots & \vdots & \ddots & \vdots \\ T_{\delta_1}(T_{t-1}(a)) & T_{\delta_2}(T_{t-1}(a)) & \dots & T_{\delta_t}(T_{t-1}(a)) \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} 2f_1 & 0 & \dots & 0 \\ 0 & 2f_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 2f_t \end{pmatrix}.$$

By arguments similar to those in lemma 1, $\mathcal{V}$ is non-singular. Since f has exactly t terms, no f_i is zero, hence $\mathcal{B}$ is non-singular. We claim that $\mathcal{A} = \mathcal{V}\mathcal{B}\mathcal{V}^T$, thereby proving that $\mathcal{A}$ is non-singular. The (i,j)-th element of the product $\mathcal{V}\mathcal{B}\mathcal{V}^{tr}$ is

$$\begin{aligned} & \sum_{l=1}^{t} 2f_l T_{\delta_l}(T_i(a)) T_{\delta_l}(T_j(a)) = \sum_{l=1}^{t} 2f_l T_i(T_{\delta_l}(a)) T_j(T_{\delta_l}(a)) \\ = \ & \sum_{l=1}^{t} f_l(T_{i+j}(T_{\delta_l}(a)) + T_{|i-j|}(T_{\delta_l}(a))) = \sum_{l=1}^{t} f_l(T_{\delta_l}(T_{i+j}(a)) + T_{\delta_l}(T_{|i-j|}(a))) \\ = \ & a_{i+j} + a_{|i-j|} = \mathcal{A}_{i,j}. \end{aligned}$$

□

From the above lemmas we see that we can compute the coefficients of $\Phi(z)$ by solving the linear system of equations given by $\mathcal{A}\vec{\phi} = -\vec{\alpha}$. The roots of $\Phi(z)$ are integers with special properties and they can be determined without using numerical root finding algorithms. The roots of $\Phi(z)$ are $T_{\delta_i(a)}$ and we can determine the "Chebyshev exponents" δ_i from $T_{\delta_i}(a)$. The coefficients f_i of f can be determined by solving the linear system of equations given by

$$\mathcal{V}\vec{\mathsf{f}} = \vec{\mathsf{a}}$$

where $\mathcal{V}$ is the matrix in the previous lemma and

$$\vec{\mathsf{f}} = (\ f_1 \quad f_2 \quad \dots \quad f_t\ )^T, \qquad \vec{\mathsf{a}} = (\ a_0 \quad a_1 \quad \dots \quad a_{t-1}\ )^T.$$

To summarize, we now collect all the steps of our interpolation algorithm.

Algorithm Sparse-Chebyshev-Interpolation

Input: A black box for evaluating a polynomial $f(x)$ known to be t-sparse in the Chebyshev basis.
Output: A list of pairs $(f_i, \delta_i), i = 1, \ldots, t$ such that $f(x) = \sum_{i=1}^{t} f_i T_{\delta_i}(x)$.

1. *Query the black box to obtain the values of $f(x)$ at the $2t$ points $x = T_i(a),\ i = 0, 1, \ldots, 2t-1$ for some $a > 1$. Let $a_i = f(T_i(a))$.*

2. *Solve the system of linear equations $\mathcal{A}\vec{\phi} = -\vec{\alpha}$. If*

$$\vec{\phi} = (\ \phi_0 \quad \phi_1 \quad \ldots \quad \phi_{t-1}\ )^T,$$

then, the auxiliary polynomial $\Phi(z)$ is given by

$$\Phi(z) = z^t + \phi_{t-1} z^{t-1} + \ldots + \phi_1 z + \phi_0.$$

3. *Find the integer roots of $\Phi(z)$. The roots are $T_{\delta_1}(a), T_{\delta_2}(a), \ldots, T_{\delta_t}(a)$. Find the "Chebyshev powers" $\delta_1, \delta_2, \ldots, \delta_t$ from $T_{\delta_1}(a), T_{\delta_2}(a), \ldots, T_{\delta_t}(a)$.*

4. *Solve the linear system of equations $\mathcal{V}\vec{\mathsf{f}} = \vec{\mathsf{a}}$ to obtain the coefficients $f_1, \ldots, f_t$ of f. Output the list of pairs $[(f_1, \delta_1), \ldots, (f_t, \delta_t)]$.*

3. Analysis of the Chebyshev Interpolation Algorithm

In our analysis, we first count the number of field operations ($+, -, \times, /$ involving rational numbers) performed by the algorithm in the worst case. We then bound the bit-sizes of the rational numbers that might arise during the computation, thus giving upper bounds on the number of bit operations performed by the algorithm. We also assume that querying the black box for the value of $f(x)$ is a unit cost operation.

Solving $\mathcal{A}\vec{\phi} = -\vec{\alpha}$: The matrix $\mathcal{A}$ is the sum of a Hankel and a Toeplitz matrix of size $t \times t$. Such systems can be solved by well known techniques using $O(t^2)$ field operations (see, for example, Merchant and Parks (1982)).

Finding the integer roots of $\Phi(z)$: The special nature of the roots of $\Phi(z)$ (all are integers of the type $T_{\delta_j}(a)$) allows us to compute them without using regular root finders. Since $\Phi(z)$ is expressed in the Chebyshev basis with leading coefficient 1, ϕ_{t-1}, the coefficient of $T_{t-1}(z)$, is given by

$$\phi_{t-1} = \begin{cases} -T_{\delta_1}(a) & \text{if } t = 1 \\ -2\sum_{j=1}^{t} T_{\delta_j}(a) & \text{if } t > 1 \end{cases}$$

Let $s = -\phi_{t-1}/2$. It is easily shown from the defining recurrence that $T_n(a) = 1/2(\beta^n + \bar{\beta}^n)$ where $\beta = a + \sqrt{a^2 - 1}$ and $\bar{\beta} = a - \sqrt{a^2 - 1}$ for $n = 0, 1, 2, \ldots$ Suppose, for convenience we choose $a = 2$. Then $\beta = 2 + \sqrt{3}$ and $\bar{\beta} = 2 - \sqrt{3}$. Without any loss of generality, assume that $\delta_1 > \delta_2 > \ldots > \delta_t$. Clearly, δ_1 is the unique integer k such that

$$1/2(\beta^k + \bar{\beta}^k) \leq s < 1/2(\beta^{k+1} + \bar{\beta}^{k+1}).$$

We can find such a k by an algorithm analogous to the binary powering algorithm as follows:

> Choose $D = 2\lceil \log_2(s) \rceil$; $(1/2(\beta^D + \bar{\beta}^D) > s)$
> Compute integer arrays p, q such that
> $p_i + q_i\sqrt{3} = \beta^{2^i}$ for $i = 0, 1, \ldots$ to at most D.
> $i := 0;\ p_0 := 1;\ q_0 := 0;$
> **loop**
> $increment(i);$
> $p_i := p_{i-1}^2 + 3q_{i-1}^2;\ q_i := 2p_{i-1}q_{i-1};$
> **until** $p_i > s;$
> $decrement(i);$
> $k := 2^i;\ j := i - 1;\ \widehat{p} := p_i;\ \widehat{q} := q_i;$
> **loop**
> $tmp := \widehat{p} \times p_i + 3\widehat{q} \times q_i;$
>   **if** $(tmp \leq s)$ **then**
> $\widehat{q} := p_i \times \widehat{q} + q_i \times \widehat{p};\ \widehat{p} := tmp;\ k := k + 2^j;$ **fi;**
> $decrement(j);$
> **until** $j = 0;$
> **return** $k, \widehat{p};\ \{\widehat{p} = T_k(2)\}$

Once we know k, i.e., δ_1, we can set $s := s - T_k(2)$ and repeat the process to get δ_2 and so on. The number of field operations performed by the above algorithm is $O(\log_2 D)$ and we need to run it t times to compute all the δ_i. Therefore, the Chebyshev exponents can be computed by performing $O(t \log_2 D)$ field operations. D is chosen to be $2\lceil \log_2(s) \rceil$ and,

$$s = -\phi_{t-1}/2 = \sum_{j=1}^{t} T_{\delta_j}(a) < 1/2(\beta^{\delta_1+1} + \bar{\beta}^{\delta_1+1}) < 4^{\delta_1}.$$

Therefore, the number of bits in the integers involved in the root finding step is no more than $2\delta_1$.

Solving $\mathcal{V}\vec{\mathsf{f}} = \vec{\mathsf{a}}$: $\mathcal{V}$ is the transpose of a Vandermonde-like matrix. An $n \times n$ non-singular Vandermonde matrix can be inverted using $O(n^2)$ field operations and $O(n)$ space (Zippel (1990)). More efficient algorithms ($O(n \log^2 n \log(\log n))$ time) that use fast polynomial multiplication are also known (Canny et al 1989)). In the following discussion, we adapt Zippel's algorithm for solving the system of equations $\mathcal{V}\vec{\mathsf{f}} = \vec{\mathsf{a}}$. Consider the product

$$\mathcal{V}^T \mathcal{A} = \begin{pmatrix} T_0(T_{\delta_1}(a)) & T_1(T_{\delta_1}(a)) & \dots & T_{t-1}(T_{\delta_1}(a)) \\ T_0(T_{\delta_2}(a)) & T_1(T_{\delta_2}(a)) & \dots & T_{t-1}(T_{\delta_2}(a)) \\ \vdots & \vdots & \ddots & \vdots \\ T_0(T_{\delta_t}(a)) & T_1(T_{\delta_t}(a)) & \dots & T_{t-1}(T_{\delta_t}(a)) \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1t} \\ c_{21} & c_{22} & \dots & c_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ c_{t1} & c_{t2} & \dots & c_{tt} \end{pmatrix}.$$

This product is the matrix

$$\begin{pmatrix} P_1(T_{\delta_1}) & P_2(T_{\delta_1}) & \dots & P_t(T_{\delta_1}) \\ P_1(T_{\delta_2}) & P_2(T_{\delta_2}) & \dots & P_t(T_{\delta_2}) \\ \vdots & \vdots & \ddots & \vdots \\ P_1(T_{\delta_t}) & P_2(T_{\delta_t}) & \dots & P_t(T_{\delta_t}) \end{pmatrix}$$

where $P_j(z)$ is the polynomial

$$P_j(z) = c_{1j}T_0(z) + c_{2j}T_1(z) + c_{3j}T_2(z) + \ldots + c_{tj}T_{t-1}(z).$$

Notice that by choosing

$$P_j(z) = \prod_{i \neq j} \frac{z - T_{\delta_i}(a)}{T_{\delta_j}(a) - T_{\delta_i}(a)},$$

the product becomes the identity matrix. In other words, the j-th column of the inverse of $\mathcal{V}^T$ is made up of the coefficients of $P_j(z)$ *expressed in the Chebyshev basis.* Let $P_{j,i}$ denote the coefficient of $T_i(z)$ in $P_j(z)$. Since the inverse of the transpose is the transpose of the inverse, the system of equations $\mathcal{V}\vec{\mathsf{f}} = \vec{\mathsf{a}}$, i.e.,

$$\begin{pmatrix} T_{\delta_1}(T_0(a)) & T_{\delta_2}(T_0(a)) & \dots & T_{\delta_t}(T_0(a)) \\ T_{\delta_1}(T_1(a)) & T_{\delta_2}(T_1(a)) & \dots & T_{\delta_t}(T_1(a)) \\ \vdots & \vdots & \ddots & \vdots \\ T_{\delta_1}(T_{t-1}(a)) & T_{\delta_2}(T_{t-1}(a)) & \dots & T_{\delta_t}(T_{t-1}(a)) \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_t \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{t-1} \end{pmatrix}$$

has the solution

$$f_j = a_0 P_{j,0} + a_1 P_{j,1} + \ldots + a_{t-1} P_{j,t-1}, \quad \text{for} \quad j = 1, \ldots, t.$$

The coefficients $P_{j,i}$ can be computed as follows. Let

$$P(z) = \prod_{k=1}^{t} (z - T_{\delta_k}(a))$$

($P(z)$ is actually $\Phi(z)/2^{t-1}$, $\Phi(z)$ being the auxiliary polynomial computed in step 2 of the interpolation algorithm). Let

$$\widehat{P}_j(z) = P(z)/(z - T_{\delta_j}(a)),$$

expressed in the Chebyshev basis. We can compute $\widehat{P}_j(z)$ from $P(z)$ by "synthetic Chebyshev division" – use the identity

$$T_k(z) = (z - T_{\delta_j}(a)) \times 2T_{k-1}(z) + 2T_{\delta_j}(a)T_{k-1}(z) + T_{k-2}(z) \quad \text{for} \quad k > 1$$

$t-1$ times on $P(z)$. This division can be performed using $O(t)$ field operations. Let $D_j = \widehat{P}_j(T_{\delta_j}(a))$. Clearly,

$$P_j(z) = \widehat{P}_j(z)/D_j.$$

D_j can be computed using $O(t)$ field operations (evaluate $\widehat{P}_j(z)$ at $T_{\delta_j}(a)$). Therefore, we can compute all the Chebyshev coefficients of $P_j(z)$ in $O(t)$ field operations, and from them, we can compute f_j using $O(t)$ field operations. Since each f_j needs $O(t)$ field operations in this scheme, we need $O(t^2)$ field operations to compute all the f_j, $j = 1, \ldots, t$. Notice that we need $O(t)$ space to hold the coefficients of each $P_j(z)$. However, we can re-use the same space for all the $P_j(z)$. Hence, we only need $O(t)$ storage units. Also, notice that the rational numbers encountered in this scheme have numerators and denominators with $O(t \log(T_{\delta_1}(a)))$ bits. Notice that $\log(T_{\delta_1}(a) = O(d))$, d being the degree of the interpolating polynomial. We can now add up the costs of the 4 steps and we have:

Theorem 2 *The sparse Chebyshev interpolation algorithm performs $O(t^2 + t \log d)$ field operations to interpolate a t-Chebyshev-sparse polynomial given by a black box. Here, d is the degree of the polynomial being interpolated (not part of the input). The algorithm makes $2t$ queries to the black box evaluating the polynomial. The numerator and denominator of the rational numbers that arise during the algorithm have $O(td)$ bits.* □

Remark 1: We may not know the exact number of terms in the polynomial but only an upper bound $\tau \geq t$. In such a case, we can still use the algorithm, except that the matrix $\mathcal{A}$ will be singular (for $\tau > t$). The solution is to use the largest non-singular leading principle minor of $\mathcal{A}$ in place of $\mathcal{A}$. Its rank gives the exact number of terms in the interpolating polynomial. The details are worked out for the Ben-Or and Tiwari algorithm in (Kaltofen and Lakshman (1988)) and are essentially the same for this case.

Remark 2: We have used classical algorithms for solving the various subproblems in the interpolation algorithm. One can use asymptotically faster algorithms to improve the overall complexity of the interpolation algorithm to $O(t(\log^2 t \log(\log t) + \log d))$ field operations. See Kaltofen and Lakshman (1988), Canny et al (1989).

Remark 3: The algorithm that we have discussed can be viewed as a generalization of the Ben-Or and Tiwari algorithm which in turn is based on the BCH decoding algorithm. The generalization uses a crucial property of the Chebyshev polynomials, namely, that they commute with respect to composition, i.e., $T_n(T_m(z)) = T_m(T_n(z))$. It turns out that this property is rather special and the only *families* of polynomials that exhibit this property are the Chebyshev polynomials and the standard powers $\{1, z, z^2, \ldots\}$. More precisely, it is known that if in a family of polynomials $f_0(z), f_1(z), f_2(z), \ldots$ with $\deg(f_i(z)) = i$, it is true that $f_i(f_j(z)) = f_j(f_i(z))$ for any $i, j \geq 0$, then, either $f_i(z) = \lambda((\lambda^{-1}(z))^i)$ for all i, or, $f_i(z) = \lambda(T_i(\lambda^{-1}(z)))$ for all i where $\lambda(z) = az + b$ and $\lambda^{-1}(z) = (z-a)/b$ for some constants a, b (see, for instance, Rivlin (1974))[3]. The algorithm can be adapted easily to interpolate polynomials that are sparse in such "generalized" power bases or Chebyshev bases.

One can consider several possible "natural" bases and polynomials that are sparse in those bases. For example, suppose we know that the polynomial given by a black box is t-sparse in the basis $1, x-\alpha, (x-\alpha)^2, (x-\alpha)^3, \ldots$ for some unknown α. Two recent papers discuss algorithms for solving this problem (Grigoriev-Karpinski 1992, Lakshman-Saunders 1993).

Sparsity in the standard basis is related to the number of real roots of the polynomial (sparsity limits the number of variations in sign), which is not the case when we consider polynomials that are sparse in the Chebyshev or Pochhammer bases. What are the corresponding observations about sparsity in these bases?

Acknowledgements

We thank our colleagues Prof. Bobby Caviness, Prof. David Wood and Prof. Jeremy Johnson for several helpful discussions. Our thanks also to Prof. Micheal Singer of North Carolina State University for detailed comments on an earlier draft.

References

Ben Or, M. and Tiwari, P. (1988), "A deterministic algorithm for sparse multivariate polynomial interpolation," Proc. 20th Symp. Theory of Computing, ACM Press, pp. 301–309.

Borodin, A. and Tiwari, P. (1990), "On the decidability of sparse univariate polynomial interpolation,"Proc. 22nd Symp. Theory of Computing, ACM Press, pp. 535–545.

[3]We are grateful to Dr. Jeremy Johnson of Drexel University for bringing this to our attention.

Canny, J. F., Kaltofen, E., Lakshman, Y.N. (1989), "Solving systems of non-linear equations faster," Proc. ISSAC 1989, Portland, Oregon, ACM Press, pp. 121–128.

Clausen, M., Dress, A., Grabmeier, J., Karpinski, M. (1988), "On zero testing and interpolation of k-sparse multivariate polynomials over finite fields," TR 88.06.006, IBM Germany, Heidelberg Scientific Center, June 1988.

Dress, A., Grabmeier, J. (1991), "The interpolation problem for k-sparse polynomials and Character sums," Advances in Appl Math, Vol 12, pp. 57–75.

Grigoriev, D. Yu. and Karpinski, M. (1987), "The matching problem for bipartite graphs with polynomially bounded permanents is in NC," Proc. 28th IEEE Symp. Foundations Comp. Sci., pp. 166–172.

Grigoriev, D., Karpinski, M., and Singer, M. (1990), "Fast parallel algorithms for sparse multivariate polynomial interpolation over finite fields," SIAM J. Comp, Vol. 19, pp. 1059–1063.

Grigoriev, D., Karpinski, M., and Singer, M. (1991a), "The interpolation problem for k-sparse sums of eigenfunctions of operators," Advances in Appl Math, Vol 12, pp. 76–81.

Grigoriev, D., Karpinski, M., and Singer, M. (1991b), "Computational complexity of sparse rational interpolation," Tech. Report, Max-Planck-Institut für Mathematik, Bonn.

Grigoriev, D., and Karpinski, M., (1992), "A zero-test and an interpolation algorithm for the shifted sparse polynomials," Proc. AAECC 1993, Puerto-Rico, Springer-Verlag LNCS **673**, pp. 162–169.

Grigoriev, D., Karpinski, M., and Odlyzko, A.M., (1992), "Existence of short proofs of non-divisibility of sparse polynomials under the extended Riemann hypothesis," Proc. ISSAC 92, ACM Press, pp. 117–122.

Kaltofen, E. and Lakshman, Y.N. (1988), "Improved sparse multivariate polynomial interpolation algorithms," Proc. ISSAC 1988, Rome, Italy, Springer-Verlag LNCS **358**, pp. 467–474.

Lakshman, Y.N., and Saunders, B.D., (1993), "On Computing sparse shifts for Univariate polynomials," Tech. Rep., Dept. Math. Comp. Sci., Drexel University, Philadelphia.

Loos, R. (1983), "Computing rational zeros of integral polynomials by p-adic expansion," SIAM J. Comp., Vol. 12, pp. 286–293.

Merchant, G.A. and Parks, T.W. (1982), "Efficient solution of a Toeplitz-plus-Hankel matrix system of equations," IEEE Trans. Acoustics, Speech, and Signal Proc., Vol. 30, pp. 40–44.

Rivlin, T. (1974), " The Chebyshev polynomials," New York, Wiley.

Zippel, R. (1990), "Interpolating polynomials from their values," J. Symb. Comp., Vol. 9, No. 3, pp. 375–403.