On Computing Sparse Shifts for Univariate Polynomials

Lakshman Y. N.*
Department of Mathematics and Computer Science
Drexel University
Philadelphia, PA 19104
e-mail: `lakshman@mcs.drexel.edu`

B. David Saunders†
Department of Computer and Information Sciences
University of Delaware
Newark, DE 19716
e-mail: `saunders@cis.udel.edu`

(Extended Abstract)

Introduction

In this paper, we consider the problem of computing t-sparse shifts for univariate polynomials. Given a polynomial $f(x) \in \mathcal{F}[x]$ of degree d (where $\mathcal{F}$ is a field of characteristic 0), consider the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$ for some $\alpha \in \mathcal{K}$, an extension of $\mathcal{F}$, i.e.,

$$f(x) = \sum_{i=0}^{d} \mathrm{F}_i (x-\alpha)^i.$$

Let t be a positive integer $\leq d$. We say that α is a t-sparse shift for $f(x)$ (or, $f(x)$ is t-sparse in the shifted basis $1, x-\alpha, (x-\alpha)^2, \ldots$) if *at most* t of the coefficients F_i are non-zero.

The main problem that we address is: given an $f(x)$ and t as above, can we efficiently compute a t-sparse shift for $f(x)$ if one exists? We construct an efficient algorithm for solving this problem and answer several related questions of interest, such as: When is a shift unique? When is a shift rational (meaning when does $\alpha \in \mathcal{F}$?)? How many evaluations does one need to distinguish 2 polynomials that are shifted t-sparse (with respect to possibly different shifts)? As an application of our algorithm, we construct a polynomial time algorithm for computing sparse decompositions of univariate polynomials. The problem is the following: given a polynomial $f(x) \in \mathcal{F}[x]$, and a positive integer t, find polynomials $g_1, g_2, \ldots, g_r \in \mathcal{F}[x]$ (if they exist) each of which is t-sparse (has no more than t terms in the standard representation), such that $f(x) = g_1(g_2(\ldots(g_r(x))))$.

Recently, there has been much interest in the design of efficient algorithms for computing sparse representations for various classes of functions such as polynomials, rational functions, and algebraic functions (Grigoriev-Karpinski 1987, Clausen et al 1988, Ben-Or&Tiwari 1988, Kaltofen-Lakshman 1988, Borodin-Tiwari 1990, Grigoriev-Karpinski-Singer 1990,1991a,1991b, Mansour 1992). 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 1992a). Sparse univariate polynomial interpolation in the Chebyshev basis is treated in (Lakshman-Saunders 1992) and a different generalized sparse interpolation algorithm that works for a class of bases can be found in (Grigoriev-Karpinski-Singer 1991b).

Computing t-sparse shifts and the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$ was raised as an open problem in (Borodin-Tiwari 1990). Subsequently, the problem was taken up in a recent paper by Grigoriev and Karpinski (Grigoriev-Karpinski 1993) where they provide an elegant algebraic criterion to be satisfied by any t-sparse shift for $f(x)$. We use that paper as our starting point for further refinement. The main departure is in our assumption that $f(x)$ is given *explicitly* (what we really need are the values of the first $2t$ derivatives of $f(x)$ at *two* distinct points, different from α). This is indeed a much stronger requirement than the requirement in (Grigoriev-Karpinski 1993) that only a black box for $f(x)$ be given. However, the algorithm of Grigoriev and Karpinski, in many cases, requires one to do more work than if one were to first interpolate $f(x)$ (dense) from its black box and then use the algorithm developed in this paper to compute a sparse shift. Furthermore, we are able to write down explicit conditions for the existence of a t-sparse shift for $f(x)$. This might provide some hints on attacking more general problems of this type. If $f(x)$ is given by a straight-line program, then a straight-line program for the first $2t$ derivatives can be obtained by a polynomial sized blow-up in the length of the straight-line program for computing $f(x)$ (Kaltofen 1987). Then the algorithm developed in this paper can be used to compute a sparse shift for $f(x)$, provided one exists. The main contributions of this

*Work supported by NSF grant CCR–9203062
†Work supported by NSF grant CCR-9123666.

paper are:

- a condition on the uniqueness and the rationality of the t-sparse shift;
- an efficient, randomized algorithm for computing a t-sparse shift α and the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$; Our algorithm performs $O(t^5 + t^2 d)$ field operations when the shift is unique. The algorithm can be made deterministic at the cost of an additional multiplicative factor of $O(td)$ to the cost.
- a criterion for distinguishing 2 polynomials, each sparse with respect to a shift, given a bound for the shifts.
- a polynomial time algorithm for computing sparse decompositions of univariate polynomials over the rationals (in general, over fields of characteristic 0).

In section 1 we develop the condition on the uniqueness and the rationality of the t-sparse shift; in section 2, an efficient algorithm for computing a sparse shift α is given, and in section 3 we provide a criterion for distinguishing 2 polynomials which are sparse with respect to bounded shifts. Section 4 contains a description of the sparse decomposition algorithm. Proofs of several lemmas are omitted from this abstract. Complete proofs of all the lemmas and theorems can be found in the full paper (Lakshman-Saunders 1993).

1. Uniqueness of Sparse Shifts

Consider the Taylor expansion of $f(x)$ about α, namely,

$$\begin{aligned} f(x) &= f(\alpha) + f^{(1)}(\alpha)(x-\alpha) + f^{(2)}(\alpha)(x-\alpha)^2/2! \\ &\quad + \cdots + f^{(d)}(\alpha)(x-\alpha)^d/d! \end{aligned}$$

where $f^{(i)}(\alpha)$ denotes $\frac{d^i}{dx^i} f(x)\Big|_{x=\alpha}$. If α is a t-sparse shift for $f(x)$, i.e., there exist $\delta_1 = d > \delta_2 > \ldots > \delta_t \in \mathcal{Z}_+$ and $\mathrm{F}_1, \mathrm{F}_2, \ldots, \mathrm{F}_t \in \mathcal{K}$ such that $f(x) = \mathrm{F}_1(x-\alpha)^{\delta_1} + \mathrm{F}_2(x-\alpha)^{\delta_2} + \ldots + \mathrm{F}_t(x-\alpha)^{\delta_t}$, then, comparing it with the above Taylor expansion, we notice that at least $d+1-t$ of the derivatives of $f(x)$ at $x = \alpha$ must be zero. The vanishing of a derivative at a point is a linear constraint on the coefficients of $f(x)$ and there can be at most $d+1$ linearly independent constraints on the coefficients of $f(x)$. In what follows, we show that if $t \leq (d+1)/2$, then asking $f(x)$ to be t-sparse in 2 distinct shifted bases means imposing more constraints than can be satisfied on the coefficients of $f(x)$ and therefore $f(x)$ can be t-sparse in at most one shifted basis. We prove this intuitive observation in theorem 1. As a corollary, we conclude that such a t-sparse shift α must be rational. First, two requisite lemmas.

Lemma 1 *Let $t \leq m < d$ be positive integers such that $t + m \leq d$. Let L be a non-increasing function mapping the interval $(0, \ldots, d)$ to $(1, \ldots, t)$. Finally, let $i_1 < i_2 < \ldots < i_m$ be a subsequence of $(0, \ldots, d)$. Then there is an s such that $\mathrm{L}(i_s) = m + 1 - s$.*

Proof: Consider the last k such that $\mathrm{L}(i_k) \leq m+1-k$. Such a k exists, since of necessity $\mathrm{L}(i_1) \leq t \leq m = (m+1) - 1$ and $\mathrm{L}(i_m) \geq 1 = (m+1) - m$. For $k = m$ the result follows immediately. If $k < m$ then we have $\mathrm{L}(i_k) \leq (m+1) - k$ and $\mathrm{L}(i_{k+1}) > (m+1) - (k+1) = m - k$. But L is non-increasing, so we must have $\mathrm{L}(i_k) = (m+1) - k = \mathrm{L}(i_{k+1})$. ◻

We also need the non-singularity of a particular matrix. Let $z^{\underline{n}}$ denote the polynomial $z(z-1)(z-2)\ldots(z-n+1)$ for $n \geq 1$ and $z^{\underline{0}}$ denote 1. It is well known that the set of polynomials $\{z^{\underline{0}}, z^{\underline{1}}, z^{\underline{2}}, z^{\underline{3}}, \ldots\}$ is an $\mathcal{F}$-basis for the polynomial ring $\mathcal{F}[z]$ and this basis is known as the *falling Pochhammer basis*. For a positive integer i, $j^{\underline{i}}$ denotes $j(j-1)(j-2)\ldots(j-i+1)$.

Lemma 2 *Let integer sequences (p_l) and (q_l) be given such that $p_1 > p_2 > \ldots > p_r > 0$, $q_1 > q_2 > \ldots > q_r = 0$, and $p_l > q_l$ for $l = 1, \ldots, r$. Then $\det(A)$ is non-zero, where A is the $r \times r$ matrix*

$$A = (a_{i,j}) = (p_i^{\underline{q_j}}), \quad i = 1, \ldots, r, \quad j = 1, \ldots, r.$$

Proof Idea: If A is singular, then there is some

$$\widehat{f} = (\widehat{f}_1, \widehat{f}_2, \ldots, \widehat{f}_r)^{Tr},$$

in Kernel(A) and an $s \leq r$ such that $\widehat{f}_s \neq 0, \widehat{f}_{s+1} = \ldots = \widehat{f}_r = 0$ and $A_s(\widehat{f}_1, \ldots, \widehat{f}_s)^{Tr} = 0$ where A_s is the $s \times s$ submatrix made from the first s rows and columns of A. Looking at it differently, this means that the polynomial $\zeta(z) = \widehat{f}_1 z^{\underline{q_1}} + \widehat{f}_2 z^{\underline{q_2}} + \ldots + \widehat{f}_s z^{\underline{q_s}}$ has $p_1, p_2, \ldots, p_s$ as its roots. However, as the polynomial $\zeta(z)$ can have at most $s-1$ sign variations in the above representation, it cannot have s roots satisfying conditions above. This follows from an analogue of Descartes' rule of signs for polynomials expressed in the falling Pochhammer basis. The details of this fairly extensive argument may be found in the full paper (Lakshman-Saunders 1993). ◻

Theorem 1 *Let $f(x) \in \mathcal{F}[x]$ be of degree d and let $t \leq (d+1)/2$. If there exists an α in some algebraic extension $\mathcal{K}$ of $\mathcal{F}$ such that $f(x)$ is t-sparse in the shifted power basis $1, (x-\alpha), (x-\alpha)^2, \ldots$, then the shift α is unique.*

Proof: Suppose, for the sake of contradiction, that $f(x)$ is t-sparse with respect to two distinct shifts α, β in an extension $\mathcal{K}$. This implies that at least some $d+1-t$ of the derivatives of $f(x)$ vanish at $x = \alpha$ and some $d+1-t$ of the derivatives of $f(x)$ vanish at $x = \beta$.

Let us express the derivatives of $f(x)$ at $x = \alpha$ in terms of the derivatives of $f(x)$ at $x = \beta$ using Taylor expansions:

$$\begin{aligned} f^{(i)}(\alpha) &= f^{(i)}(\beta) + f^{(i+1)}(\beta)(\alpha-\beta) + \cdots + \\ &\quad f^{(d)}(\beta)(\alpha-\beta)^{d-i}/(d-i)!, \; i = 0, \ldots, d. \quad (1) \end{aligned}$$

It will be convenient to convert this triangular system to an equivalent Vandermonde-like system via diagonal scalings. Multiply each row i by $(d-i)!(\beta-\alpha)^i$; then factor $(\alpha-\beta)^j$ from column j, obtaining

$$(d-i)!(\alpha-\beta)^i f^{(i)}(\alpha) = \sum_{l=0}^{d} ((d-i)^{\underline{(d-l)}}) \times (\alpha-\beta)^l f^{(l)}(\beta).$$

Note that $(d-i)! = (d-i)^{\underline{(d-l)}}(d-l)!$. Next we restrict to only the vanishing derivatives of $f(x)$ at $x = \alpha$. With the aid of lemma 1, we can extract a square system of equations in which the rows correspond to some *vanishing* derivatives at α and the columns correspond to all the *non-vanishing* derivatives at β on which they depend. Let $j_1, \ldots, j_t$ be the sequence of indices such that $f^{(j_k)}(\beta) \neq 0$, and let $\mathrm{LENGTH}(f^{(i)}(\alpha))$ be the number of corresponding non-zero terms in the Taylor expansion of $f^{(i)}(\alpha)$ in

(1). Apply lemma 1 to the non-increasing function $\text{L}(k) = \text{LENGTH}(f^{(i_k)}(\alpha))$ and the sequence $i_1 < \ldots < i_m$ of indices such that $f^{(i_k)}(\alpha) = 0$. Choose the largest s such that $\text{LENGTH}(f^{(i_s)}(\alpha)) = m + 1 - s$. Let $r = m + 1 - s$. The r equations corresponding to $f^{(i_s)}(\alpha)$ through $f^{(i_m)}(\alpha)$ then form the following square system

$$\begin{pmatrix} p_1^{\underline{q_1}} & p_1^{\underline{q_2}} & \cdots & p_1^{\underline{q_r}} \\ p_2^{\underline{q_1}} & p_2^{\underline{q_2}} & \cdots & p_2^{\underline{q_r}} \\ \vdots & \vdots & \ddots & \vdots \\ p_r^{\underline{q_1}} & p_r^{\underline{q_2}} & \cdots & p_r^{\underline{q_r}} \end{pmatrix} \begin{pmatrix} (\alpha-\beta)^{d-q_1} f^{(d-q_1)}(\beta) \\ (\alpha-\beta)^{d-q_2} f^{(d-q_2)}(\beta) \\ \vdots \\ (\alpha-\beta)^{d-q_r} f^{(d-q_r)}(\beta) \end{pmatrix} = 0,$$

where $p_k = i_{m+1-k}$ and $d - q_k = j_{(t-r+k)}$, for $k = 1, \ldots, r$. By lemma 2, this $r \times r$ matrix is non-singular so that the vector must be zero. The vector is zero only if $\alpha = \beta$, in other words, only if the shifts are identical. □

The above argument is made for characteristic zero. It is also valid for sufficiently large positive characteristic (greater than d and non-divisor of the determinant). It is an interesting question to what extent the results of this paper go through for positive characteristic.

Corollary 1 *Let $f(x) \in \mathcal{F}[x]$ be of degree d and let $t \leq (d+1)/2$. If α (in any extension $\mathcal{K}$ of $\mathcal{F}$) is a t-sparse shift (hence the unique t-sparse shift) for $f(x)$, then $\alpha \in \mathcal{F}$.*

Proof Idea: Were α algebraic over $\mathcal{F}$, its conjugates would also necessarily be t-sparse shifts, contradicting uniqueness. □

2. An Algorithm to Determine Sparse Shifts

Let $f(x) = \text{F}_1(x-\alpha)^{\delta_1} + \text{F}_2(x-\alpha)^{\delta_2} + \ldots + \text{F}_t(x-\alpha)^{\delta_t}$ where $\delta_1 > \delta_2 > \ldots > \delta_t$, and let $\Phi(z) = \phi_t z^{\underline{t}} + \phi_{t-1} z^{\underline{t-1}} + \ldots + \phi_1 z^{\underline{1}} + \phi_0 z^{\underline{0}}$ be the monic polynomial of degree t whose roots are precisely $\delta_1, \delta_2, \ldots, \delta_t$. We have expressed $\Phi(z)$ as a monic polynomial in the falling Pochhammer basis for reasons that will be explained shortly. Consider the following set of equalities:

$$\begin{aligned} &\phi_i f^{(i)}(x)(x-\alpha) \\ &\quad = \phi_i \delta_1^{\underline{i}} \text{F}_1(x-\alpha)^{\delta_1} + \phi_i \delta_2^{\underline{i}} \text{F}_2 (x-\alpha)^{\delta_2} + \cdots \\ &\qquad + \phi_i \delta_t^{\underline{i}} \text{F}_t (x-\alpha)^{\delta_t}, \quad i = 0, \ldots, t \end{aligned} \tag{2}$$

Adding the left and right sides of the above set of equalities, we get

$$\begin{aligned} &\sum_{i=0}^{t} \phi_i f^{(i)}(x)(x-\alpha)^i \\ &\quad = \Phi(\delta_1)\text{F}_1(x-\alpha)^{\delta_1} + \Phi(\delta_2)\text{F}_2(x-\alpha)^{\delta_2} + \cdots \\ &\qquad + \Phi(\delta_t)\text{F}_t(x-\alpha)^{\delta_t} = 0. \end{aligned} \tag{3}$$

Since the $t+1$ polynomials $(x-\alpha)^i f^{(i)}(x), i = 0, \ldots, t$, satisfy a linear relation, their Wronskian vanishes (this is classical; see, for example, (Muir 1960, page 664)). The converse is also true as we show in the next theorem. We denote Wronskians of this type by $\mathcal{W}_k(x,z)$, i.e., $\mathcal{W}_k(x,z)$ is the determinant of the $k \times k$ matrix $W_k(x,z)$ where

$$\begin{aligned} W_k(x,z) &= (w_{i,j}) = ((d^i/dx^i)(g_j(x,z))), \\ &\text{for } i = 0, \ldots, k-1, \quad j = 0, \ldots, k-1, \end{aligned}$$

and $g_i(x,z) = (x-z)^i f^{(i)}(x)$. In particular, $\mathcal{W}_{t+1}(x,\alpha)$ denotes the classical Wronskian of $(x-\alpha)^i f^{(i)}(x), i = 0, \ldots, t$.

Theorem 2 *For any $\alpha \in \mathcal{K}$, $f(x)$ is t-sparse in the α-shifted power basis for $\mathcal{K}[x]$ if* and only if *the Wronskian $\mathcal{W}_{t+1}(x,\alpha) = 0$.*

Proof: $(\Longrightarrow)$: Well-known (see (Muir 1960, page 664)).

$(\Longleftarrow)$: Since $g_0(x,\alpha), g_1(x,\alpha), \ldots, g_t(x,\alpha) \in \mathcal{K}[x]$, $\mathcal{W}_{t+1}(x,\alpha) = 0$ implies that $g_0(x,\alpha), g_1(x,\alpha), \ldots, g_t(x,\alpha)$ are $\mathcal{K}$-linearly dependent (Kaplanski, 1957). In other words, $\exists c_0, c_1, \ldots, c_t \in \mathcal{K}$ s.t. $\sum_{i=0}^{t} c_i g_i(x,\alpha) = 0$. Let us look at the Taylor expansions of $g_0(x,\alpha), g_1(x,\alpha), \ldots, g_t(x,\alpha)$ about $x = \alpha$:

$$\begin{aligned} g_i(x,\alpha) &= \frac{f^{(d)}(\alpha)}{(d-i)!}(x-\alpha)^d + \frac{f^{(d-1)}(\alpha)}{(d-i-1)!}(x-\alpha)^{d-1} \\ &\quad + \cdots + f^{(i)}(\alpha)(x-\alpha)^i, \quad i = 0, \ldots, t. \end{aligned} \tag{4}$$

If $f(x)$ is not t-sparse in the α-shifted power basis, at least $t+1$ derivatives of $f(x)$ are non-zero at α. Consider the $(t+1) \times (t+1)$ matrix

$$\begin{aligned} V_{t+1}(\alpha) &= (v_{i,j}) = (f^{(\delta_j)}(\alpha)/(\delta_j - i)!), \\ &\text{for } i = 0, \ldots, t, \quad j = 1, \ldots, t+1, \end{aligned}$$

where the columns correspond to some $t+1$ non-vanishing derivatives of $f(x)$ at $x = \alpha$. The linear relation

$$\sum_{i=0}^{t} c_i g_i(x,\alpha) = 0$$

implies that $V_{t+1}(\alpha)$ is singular. But it is easy to see that after factoring out $f^{(\delta_j)}(\alpha)/\delta_j!$ from the j-th column, we are left with the matrix $(\delta_j^{\underline{i-1}})$, which can be row-reduced to a Vandermonde matrix. Since $\delta_j \neq \delta_k$, $V_{t+1}(\alpha)$ cannot be singular. Therefore, it cannot happen that at least $t+1$ derivatives of $f(x)$ at α are non-zero, i.e., $f(x)$ must be t-sparse in the α-shifted power basis. □

Corollary 2 *If α is the* unique *t-sparse shift for $f(x)$, and the representation of $f(x)$ in the α-shifted power basis has exactly t terms, then, the system of linear equations*

$$W_t(b,\alpha) \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} g_t(b,\alpha) \\ (dg_t/dx)(b,\alpha) \\ \vdots \\ (d^{t-1} g_t/dx^{t-1})(b,\alpha) \end{pmatrix} \tag{5}$$

has $(\phi_0, \phi_1, \ldots, \phi_{t-1})^{Tr}$ (the vector of coefficients of $\Phi(z)$) as its unique *solution for any $b \neq \alpha$.* □

Corollary 3 *For generic choices of $(b_1, b_2) \in \bar{\mathcal{F}}^2$, if $t \leq (d+1)/2$, then* $\gcd(\mathcal{W}_{t+1}(b_1,z), \mathcal{W}_{t+1}(b_2,z)) = c(z-\alpha)^{\sigma}$, *for some $\sigma \geq 0$.*

Here "generic" means that (b_1, b_2) belongs to a Zariski open set in $\bar{\mathcal{F}}^2$.

Proof: It follows from theorems 1 and 2 that $\mathcal{W}_{t+1}(x,z) = c(z-\alpha)^{\sigma}\mathcal{U}(x,z)$ where $\mathcal{U}(x,z)$ is primitive as an element of $(\mathcal{F}[z])[x]$. Choose $(b_1, b_2) \in \bar{\mathcal{F}}^2$ such that it is not a root of $\text{Res}_z(\mathcal{U}(x,z), \mathcal{U}(y,z))$, the Sylvester resultant of

$$\mathcal{U}(x,z), \mathcal{U}(y,z) \in \mathcal{F}[x,y][z].$$

Clearly, for such a choice of b_1, b_2, $\gcd(\mathcal{U}(b_1,z), \mathcal{U}(b_2,z)) = 1$ and the lemma follows. □

For a different condition relating the vanishing of a Wronskian to sparse representation, see (Grigoriev-Karpinski-Singer 1991, Grigoriev-Karpinski 1993).

Algorithm Shifted-Sparse-Interpolation

Input: A polynomial $f(x) \in \mathcal{F}[x]$ of degree d and an integer $t \leq (d+1)/2$.
Output: $\alpha \in \mathcal{F}$, and a list of pairs $(\mathrm{F}_i, \delta_i), i = 1, \ldots, t$ such that $f(x) = \sum_{i=1}^{t} \mathrm{F}_i (x-\alpha)^{\delta_i}$, or, the message "$f(x)$ does not have a t-sparse shift for the given t."

1. *Compute* $\gcd(\mathcal{W}_{t+1}(b_1, z), \mathcal{W}_{t+1}(b_2, z))$ *for generic* $b_1, b_2 \in \mathcal{F}$.
   - *if the* gcd $\in \mathcal{F}$, *a t-sparse shift does not exist for* $f(x)$.
   - *if the square-free part of the* gcd *is* $c(z-\alpha)$, *then* α *is the unique t-sparse shift. Go to step 2.*
   - *if the square-free part of the* gcd *has degree* > 1, $f(x)$ *has a t-sparse shift; however, the number of terms in the shifted basis is strictly less than t. Find the smallest principle minor* $W_{k+1}(x, z)$ *of* $W_{t+1}(x, z)$ *such that the square-free part of the gcd of* $\det(W_{k+1}(b_1, z))$ *and* $\det(W_{k+1}(b_2, z))$ *is linear in z, set t equal to k and goto step 2.*
2. *Solve the system of linear equations* (5) *to get the auxiliary polynomial* $\Phi(z) = z^t + \phi_{t-1} z^{t-1} + \ldots + \phi_1 z + \phi_0$.
3. *Find the integer roots of* $\Phi(z)$. *The roots are* $\delta_1, \delta_2, \ldots, \delta_t$.
4. *Solve the linear system of equations* $\mathcal{V}_t(b_1, \alpha)\vec{\mathsf{f}} = \vec{\mathsf{a}}$ *to obtain the coefficients* $\mathrm{F}_1, \ldots, \mathrm{F}_t$ *of* f. *Output the list of pairs* $[(\mathrm{F}_1, \delta_1), \ldots, (\mathrm{F}_t, \delta_t)]$.

Remark: This algorithm can be used even when $t > (d+1)/2$. In such cases, the shift might not be unique and might be algebraic over $\mathcal{F}$. The third option in step 1 (the square-free part of the gcd has degree greater than 1) must be modified to take care of multiple shifts.

Analysis of the running time of the algorithm: We count the number of rational arithmetic operations performed by the above algorithm assuming that *classical algorithms* are used for doing polynomial arithmetic (for multiplication, gcd etc). In step 1, setting up the matrix $W_{t+1}(b_1, z)$ requires evaluating the first $2t$ derivatives of $f(x)$ at $x = b_1$ and $O(t^2 d)$ additional work. The $2t$ derivatives can be evaluated in $O(td)$ operations. The polynomial $\mathcal{W}_{t+1}(b_1, z)$ is of degree at most $\binom{t+1}{2}$ in z (this is the degree of the anti-diagonal term in $W_{t+1}(b_1, z)$). To compute this by interpolation, one needs its values at $\binom{t+1}{2} + 1$ distinct values of z and each value is to be obtained by evaluating a $(t+1) \times (t+1)$ determinant. The total work involved in this is $O(t^5)$ operations (and a similar number for computing $\mathcal{W}_{t+1}(b_2, z)$). Since they are of degree $\binom{t+1}{2}$ each, their gcd can be computed using $O(\binom{t+1}{2}^2) = O(t^4)$ operations. The square-free part of the gcd can be computed within the same bounds on the number of operations. Step 2 can be performed in $O(t^3)$ operations, step 3 in $O(t^2 \log d)$ operations (Loos 1983). The linear system to be solved in Step 4 is a very special, Vandermonde-like system and can be solved in $O(t^2)$ steps. Clearly, step 1 dominates for an overall count of $O(t^2 d + t^5)$. The polynomial $f(x)$ and its first $2t$ derivatives need to be evaluated at 2 points b_1, b_2 for a total of $4t + 2$ evaluations. In contrast, the algorithm of Grigoriev-Karpinski needs $d^{O(1)}$ evaluations and its running time is dominated by $O(td^4)$.

3. On Distinguishing Two Shifted Sparse Polynomials

If we are given black boxes for two polynomials that are known to be t-sparse with respect to some bounded, but unknown and possibly different shifts, the test developed in this section can be used to distinguish the two polynomials by evaluating them at $3t$ sufficiently large points. The test applies to polynomials with real coefficients and it is based on the number of sign variations in the representation of such polynomials in the standard power basis. The test hinges on how often two shifted sparse polynomials (with possibly different sparse shifts) can have the same value at real points away from their sparse shifts. We summarize our main observations in the following lemmas and theorem 3 below. Complete proofs can be found in (Lakshman-Saunders 1993).

Let $H(x) = \sum_{i=1}^{t} h_i x^i$ be t-sparse in the standard power basis. Let $G(x) = H(x+a) = \sum_{i=0}^{d} g_i x^i$ for some positive real a, where d is the degree of $H(x)$.

Lemma 3 *The number of sign variations, var(G), in the coefficient sequence of* $G(x)$ *is at most* $t-1$. □

Let $f(x)$ be t-sparse in the α shifted power basis and $g(x)$ be t-sparse in the β-shifted power basis. Without loss of generality, assume that $\alpha \geq \beta$ and let $\gamma = \alpha - \beta$.

Lemma 4 *The number of real roots of* $g(x) - f(x)$ *in the interval* (α, ∞) *is at most 3t-1.* □

Theorem 3 *Let* $f(x), g(x)$ *be polynomials with real coefficients each of which is t-sparse with respect to a real shift whose magnitude is bounded by some B. Let* $\mathcal{A} = \{a_i\}$ *be a set of* $3t$ *distinct real numbers greater than B. If* $f(a_i) = g(a_i)$ *for* $a_i \in \mathcal{A}$, *then,* $f(x) = g(x)$. □

4. An Algorithm for Finding Sparse Decompositions

Let us recall some facts about univariate polynomial decomposition from (von zur Gathen-Kozen-Landau 1987). A decomposition of $f(x) \in \mathcal{F}[x]$ is a sequence of polynomials $f_1, f_2, \ldots, f_k \in \mathcal{F}[x]$ such that $f(x) = f_1 \circ f_2 \circ \ldots \circ f_k(x)$. If there is no such decomposition with *at least* two of the f_i being of degree 2 or more, then, $f(x)$ is said to be indecomposable. A complete decomposition of $f(x)$ is a decomposition in which each component $f_i(x)$ has degree at least 2 and is indecomposable. A theorem of Ritt guarantees that a complete decomposition is essentially unique for polynomials over fields of characteristic zero. The following types of ambiguities can occur in a complete decomposition:

$$\begin{aligned} f \circ g(x) &= f(x+a) \circ (g(x) - a), \\ x^n \circ x^m &= x^m \circ x^n, \\ T_n(x) \circ T_m(x) &= T_m(x) \circ T_n(x), \end{aligned} \tag{6}$$

where $T_j(x)$ is the j-th Chebyshev polynomial of the first kind. The Chebyshev polynomials are defined as $T_0(x) = 1$, $T_1(x) = x$, $T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$ for $n > 1$. A well known result of Ritt states that the above ambiguities are the only possible ambiguities in the decomposition of a polynomial (Ritt 1922, Fried-MacRae 1969). In (von zur Gathen-Kozen-Landau 1987), an efficient algorithm for computing a complete decomposition of a polynomial

$f(x) \in \mathcal{F}[x]$ is described (the authors treat coefficient fields of characteristic ≥ 0).

A decomposition $f(x) = f_1 \circ f_2 \circ \ldots \circ f_r(x)$ is t-sparse if each component f_i is t-sparse. Such a decomposition need not be a complete decomposition and some of the f_i are allowed to have degree one. While it is true that if $f(x) \in \mathcal{F}[x]$ is decomposable over some algebraic extension of $\mathcal{F}$, then it is decomposable over $\mathcal{F}$, this is not the case with sparse decomposition – for a given t, there may exist a t-sparse decomposition of $f(x)$ over an algebraic extension of $\mathcal{F}$, but not over $\mathcal{F}$.

Consider a complete decomposition of f,

$$f(x) = f_1 \circ f_2 \circ \ldots \circ f_m(x). \tag{7}$$

Suppose f also has a t-sparse decomposition $f = g_1 \circ g_2 \circ \ldots \circ g_s$. Consider the complete decomposition of each g_i. Because of the uniqueness of the complete decomposition of f, the polynomials in the complete decomposition of each g_i must correspond to a block of indecomposable polynomials $f_l, f_{l+1}, \ldots, f_j$ in (7) (up to possible reordering of Chebyshev polynomials if they occur at the ends of the block and shifts such as $(x+a) \circ (x-a)$ introduced in the middle). The algorithm checks for all possible ways to split (7) into segments and checks existence of t-sparse shifts for each possible split. Hence, it will find a t-sparse decomposition of f if there is one. It involves repeated use of the algorithm of section 2, scanning for blocks $f_i \circ f_{i+1} \circ \ldots \circ f_j(x)$ that may be shifted t-sparse. Commuting powers of x pose no additional burden, because they do not affect sparsity. However the ambiguity arising from the commutativity of composition of Chebyshev polynomials requires more attention. It can matter in which order adjacent commuting Chebyshev polynomials in the decomposition are considered. For example, consider $T_2(T_3(h(x)))$ which is the same as $T_3(T_2(h(x)))$; it may happen that $T_2(h(x))$ is t-sparse while $T_3(h(x))$ is not). Our solution amounts to an exaustive search of all possibilities.

First, identify those f_i which are Chebyshev polynomials (this can be done in several simple ways, for instance, using the fact that the only polynomial solutions of the differential equation $(1-x^2)(dy/dx)^2 = m^2(1-y^2)$ are $\pm T_m(x)$).

- If $f_u, f_{u+1}, \ldots, f_v$ are all Chebyshev polynomials for some $u < v$ in (7) above, and f_{u-1}, f_{v+1} are not, we call $f_u, f_{u+1}, \ldots, f_v$ a Chebyshev block and denote it by $C_{u,v}$.
- Let $\mathcal{T}_{u,v}$ denote the power set of $\{u, u+1, \ldots, v\}$. For any index set $J \in \mathcal{T}_{u,v}$, let T_J denote the Chebyshev polynomial $f_{u_1} \circ f_{u_2} \circ \ldots \circ f_{u_l}$ where $J = \{u_1, u_2, \ldots, u_l\}$. Note that the degree of T_J is $d_J = \prod_{u_j \in I} \deg(f_{u_j})$.
- Let $h_{i,j}$ we denote $f_i \circ f_{i+1} \circ \ldots \circ f_j(x)$ for $i \leq j$ if f_i, f_j are not Chebyshev polynomials; If f_j belongs to the Chebyshev block $C_{j,k}$ for some $k \geq j$, but f_i isn't Chebyshev, then, for any $J \in \mathcal{T}_{j,k}$, let $h_{i,J}$ denote $f_i \circ f_{i+1} \circ \ldots \circ f_{j-1} \circ T_J(x)$; If f_i belongs to the Chebyshev block $C_{l,i}$ for some $l \leq i$, and f_i isn't Chebyshev, then, for any $I \in \mathcal{T}_{l,i}$, let $h_{I,j}$ denote $T_I \circ f_{i+1} \circ \ldots \circ f_{j-1} \circ f_j(x)$; If f_i, f_j are both Chebyshev but are in different Chebyshev blocks $C_{l,i}, C_{j,k}$, then for any $I \in \mathcal{T}_{l,i}$ and $J \in \mathcal{T}_{j,k}$, let $h_{I,J}$ denote $T_I \circ f_{i+1} \circ \ldots \circ f_{j-1} \circ T_J(x)$.

For each $h_{i,j}$, if there is a t-sparse shift $\alpha \in \mathcal{K}$ then, denote the t-sparse representation of $h_{i,j}$ in the α shifted basis by $g_{i,j}$ (similarly, for $h_{i,J}$, $h_{I,j}$, $h_{I,J}$, compute $g_{i,J}$, $g_{I,j}$, $g_{I,J}$ respectively). If no t-sparse shift exists for $h_{i,j}$, then check if there is a $t+1$-sparse shift $\alpha \in \mathcal{K}$ for $h_{i,j}$ *with a non-zero constant term* b. If such a shift exists, then, set $g_{i,j}$ to $(x+b) \circ G_{i,j}$ where $G_{i,j}$ is the t-sparse representation of $h_{i,j} - b$ in the α shifted basis. If not, set $g_{i,j}$ to zero (similarly for $g_{i,J}$, $g_{I,j}$, $g_{I,J}$).

The $g_{i,j}, g_{i,J}, g_{I,j}, g_{I,J}$ are the possible sparse components of f. In order to find out if f indeed has a t-sparse decomposition, we construct a directed acyclic graph $G = (V, E)$. It has the following vertices: a vertex labeled i for each f_i in the complete decomposition of f that is not in any Chebyshev block; for each Chebyshev block $C_{u,v}$, vertices labeled $V_{u,v}(I), U_{u,v}(I)$ for each $I \in \mathcal{T}_{u,v}$; a vertex labeled $m+1$. The graph has the following labeled edges:

- an edge directed from i to j if $g_{i,j-1} \neq 0$, labeled $g_{i,j-1}$.
- if $g_{i,j-1} \neq 0$, and f_{j-1} is not Chebyshev but $f_j \in C_{j,k}$ for some k, an edge directed from i to $V_{j,k}(\phi)$, labeled $g_{i,j-1}$; if $f_k \in C_{j,k}$ for some j,k but f_{k+1} is not Chebyshev, an edge directed from $U_{j,k}(\phi)$ to $k+1$ with *no* label.
- if $g_{i,J} \neq 0$ for $J \in \mathcal{T}_{j,k}$ for some j,k, an edge directed from i to $V_{j,k}(J)$, labeled $g_{i,J}$; if $g_{I,j} \neq 0$ for $I \in \mathcal{T}_{k,i}$ for some k,i, an edge directed from $U_{k,i}(J)$ to j, labeled $g_{I,j}$.
- if $g_{I,J} \neq 0$ for some $I \in \mathcal{T}_{l,i}$ and $J \in \mathcal{T}_{j,k}$, an edge directed from $U_{l,i}(I)$ to $V_{j,k}(J)$ labeled $g_{I,J}$.
- for every $(I,J) \in \mathcal{T}_{i,j} \times \mathcal{T}_{i,j}$, an edge directed from $V_{i,j}(I)$ to $U_{i,j}(J)$ if $I \cap J = \phi$ and for each $k \in \mathcal{T}_{i,j} \setminus (I \cup J)$, $g_{k,k} \neq 0$. The edge is labeled $g_{k_1,k_1} \circ \ldots \circ g_{k_r,k_r}$ (in any order) where $\{k_1, \ldots, k_r\} = \mathcal{T}_{i,j} \setminus (I \cup J)$.

Algorithm sparse decomposition:

1. *Find a complete decomposition of $f(x)$ into $f(x) = f_1 \circ f_2 \circ \ldots \circ f_m(x)$ with $f_i \in \mathcal{F}[x]$ using the algorithm of von zur Gathen et al. Identify the various Chebyshev blocks $C_{u,v}$ in the complete decomposition.*
2. *For each i,j $1 \leq i \leq j \leq m$, compute $h_{i,j}$, or all possible $h_{i,J}$, or $h_{I,j}$, or $h_{I,J}$ as the case may be (purge any duplicates in the case of $h_{i,J}, h_{I,j}$, and $h_{I,J}$). For each $h_{i,j}, h_{i,J}, h_{I,j}$, and $h_{I,J}$, compute the corresponding $g_{i,j}, g_{i,J}, g_{I,j}$, and $g_{I,J}$.*
3. *Construct the directed graph $G = (V, E)$.*
4. *If f_1 is not in any Chebyshev block, find a simple path from the vertex labeled 1 to $m+1$; if f_1 is in the Chebyshev block $C_{1,j}$, find a simple path from $V_{1,j}(\phi)$ to $m+1$. If such a path exists and the successive edges are labeled as $L_1, \ldots, L_r$, then return $L_1 \circ \ldots \circ L_r$ as a t-sparse decomposition of f. If no such path exists, f has no t-sparse decomposition.*

Brief analysis of the running time of the algorithm:

The simple observation required for showing that the algorithm runs in polynomial time (in the degree d of f) is that the number of composition factors in the complete decomposition of f (m above) is $\leq \log_2 d$. In step 1, we use the algorithm of von zur Gathen et al to find the complete decomposition of f and for each component f_i, we check whether it is a Chebyshev polynomial. Using classical polynomial arithmetic, both these tasks can be achieved in

$O(d^2)$ rational arithmetic operations. Since $\prod_{i=1}^{m} \deg(f_i) = d$ and $\deg(f_i) > 1$, $m \leq \log_2 d$. Therefore, the union of all the power-sets $\mathcal{T}_{u,v}$ corresponding to the various Chebyshev blocks contains no more than d elements. Finding t-sparse representations for each $h_{i,j}$ or $h_{i,j} - b$ in step 2 costs $O(t^2 d + t^5)d$ "field" operations in all, where the "field" may be an algebraic extension of $\mathcal{F}$ of degree $O(d)$ (we are putting no constraints on t relative to d and as a result, t-sparse shifts need not be rational). Again, using classical polynomial arithmetic, the number of rational field operations ($\mathcal{F}$-operations) for steps 2 and 3 can be bounded by $O(t^2 d + t^5)d^3$. The graph G has at most $O(d)$ vertices and $O(d^2)$ edges. Constructing and traversing it costs $O(d^2)$ basic operations. The total cost is clearly dominated by step 2 for an overall operation count of $O(t^2 d^3(d + t^3))$.

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.

Fried, M.D., and MacRae, R.E., (1969), "On the invariance of chains of fields," Ill. Jour. Math., **13**, 165–171.

von zur Gathen, J., Kozen, D., and Landau, S. (1987), "Functional decomposition of polynomials," Proc. 28th IEEE Symp. Found. Comp. Sci., Nov 1987, pp. 127–131.

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.

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. (1991), "Computational complexity of sparse rational interpolation," Tech. Report No. 8566-CS, Max-Planck-Institut für Mathematik, Bonn (to appear in SIAM J.Comp.).

Grigoriev, D., Karpinski, M., and Singer, M. (1991a), "Computational complexity of sparse real algebraic function interpolation," Tech. Report No. 8576-CS, Max-Planck-Institut für Mathematik, Bonn (to appear in Proc. MEGA-92).

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

Grigoriev, D., and Karpinski, M., (1993), "A zero-test and an interpolation algorithm for the shifted sparse polynomials," in Proc. AAECC–10, LNCS Vol. 673, pp. 162–169.

Grigoriev, D., Karpinski, M., and Odlyzko, A.M.(1992a), "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. (1987), "Single-factor Hensel lifting and its application to the straight-line complexity of certain polynomials," Proc. 19th Symp. Theory of Computing, ACM Press, pp. 443–452.

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.

Kaltofen, E., and Trager, B. (1990), "Computing with polynomials given by black boxes for their evaluations: Greatest common divisors, factorization, separation of numerators and denominators," J. Symb. Comp. **9** (1990), pp. 301-320.

Kaplanski, I. (1957), "An introduction to differential algebra," Hermann, Paris.

Lakshman, Y.N., and Saunders, B.D., (1992), "Sparse polynomial interpolation in non-standard bases," to appear in SIAM J. Comp.

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

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

Mansour, Y. (1992), "Randomized interpolation and approximation of sparse polynomials ," Proc. ICALP '92.

Muir, T. (1960, enlarged by Metzler, H.), "A treatise on the theory of determinants,", Dover Publishing Inc., New York.

Ritt, J.F., (1922), "Prime and composite polynomials," Trans. Amer. Math. Soc. **23**, pp. 51–66.

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