Algorithms for Computing the Sparsest Shifts of Polynomials via the Berlekamp/Massey Algorithm

Wen-shin Lee
University of Waterloo
www.wen-shin.com

University of Waterloo

Joint work with

Mark Giesbrecht
University of Waterloo
www.uwaterloo.ca/˜mwg

Erich Kaltofen
North Carolina State University
www.kaltofen.net

Consider the polynomial:

$$f(x_1,\ldots,x_n) = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$$

in shifted basis $y_j = x_j + s_j$: $\displaystyle\sum_{i=1}^{t(s)} c_i \underbrace{(x_1+s_1)}_{y_1}{}^{e_{1,i}} \cdots \underbrace{(x_n+s_n)}_{y_n}{}^{e_{n,i}}$

$t(s)$ depends on $s = (s_1,\ldots,s_n)$

Questions:

- Find a sparsest shift of f within set S:

  $s = (s_1,\ldots,s_n) \in S$ and $t(s)$ is minimized.

- T-sparse shifts of f within set S:

  $s = (s_1,\ldots,s_n) \in S$ and $t(s) \le T$.

Example:

$$
\begin{aligned}
f(x_1,x_2) = {} & x_1^5x_2^2 - 6x_1^5x_2 + 9x_1^5 + 10x_1^4x_2^2 - 60x_1^4x_2 + 90x_1^4 + 40x_1^3x_2^2 \\
& - 240x_1^3x_2 + 360x_1^3 + 80x_1^2x_2^2 - 480x_1^2x_2 + 720x_1^2 \\
& + 80x_1x_2^2 - 480x_1x_2 + 720x_1 + 32x_2^2 - 192x_2 + 288
\end{aligned}
$$

$$
= (\underbrace{x_1 + 2}_{y_1})^5 (\underbrace{x_2 - 3}_{y_2})^2
$$

$(2,-3)$ is a sparsest shift of $f(x_1,x_2)$

Previous research

Grigoriev, Karpinski (1993): Wronskians vs. shifted T-sparsities

Based on Wronskians that *require evaluating derivatives:*

Lakshman, Saunders (1994, 1996):

Univariate T-sparse shifts

Grigoriev, Lakshman (1995, 2000):

Multivariate T-sparse shifts

Early termination Ben-Or/Tiwari sparse interpolation

Kaltofen, Lee, Lobo (2000)

Interpolate: $f = \sum_{i=1}^{t} c_i x_1^{e_{1,i}} \cdots x_n^{e_{n,i}}$

- Compute minimal generator Λ of linearly recurring sequence

$$f(p_1, \dots, p_n),\ f(p_1^2, \dots, p_n^2),\ \dots,\ f(p_1^i, \dots, p_n^i),\ \dots$$

by Berlekamp/Massey algorithm with distinct random p_j.

> **The Berlekamp/Massey algorithm:**
> Process elements from field; compute "discrepancies" Δ_i.
> When $\Delta_i = 0$ at $i > 2L$, $i = 2t+1$ and Λ is determined with high probability.

- Recover terms in f by finding roots of Λ.
- Locate coefficients for terms in f.

Early termination Ben-Or/Tiwari on a shifted basis

Given $(s_1, \ldots, s_n)$ and $y_j = x_j + s_j$, interpolate $f = \sum_{i=1}^{t(s)} c_i y_1^{e_{1,i}} \cdots y_n^{e_{n,i}}$

- Compute minimal generator Λ of

$$f(p_1 - s_1, \ldots, p_n - s_n), \ \ldots, \ f(p_1^i - s_1, \ \ldots, \ p_n^i - s_n), \ \ldots$$

by Berlekamp/Massey algorithm.

The Berlekamp/Massey algorithm:
Process elements from field; compute “discrepancies” Δ_i.
When $\Delta_i = 0$ at $i > 2L$, $i = 2t(s) + 1$ and Λ is determined with high probability.

Leave p_j, s_j as symbols: $p_j \longrightarrow y_j, \quad s_j \longrightarrow z_j$

Interpolating on a symbolically shifted basis

With z_j indeterminates and $y_j = x_j + z_j$:

$$f(y_1 - z_1, \ldots, y_n - z_n), \ldots, f(y_1^i - z_1, \ldots, y_n^i - z_n), \ldots,$$

perform fraction-free Berlekamp/Massey algorithm.

> The fraction-free Berlekamp/Massey algorithm:
> Process elements from an integral domain.
> $\Delta_i(z_1, \ldots, z_n, y_1, \ldots, y_n)$ are polynomials in $z_1, \ldots, z_n, y_1, \ldots, y_n$.

A sparsest shift $s = (s_1, \ldots, s_n)$ minimize i such that

$$\Delta_i(s_1, \ldots, s_n, y_1, \ldots, y_n) = 0 \text{ at } i > 2L,$$

that is when $i = 2t(s) + 1$.

Fraction-free Berlekamp/Massey algorithm

Input: $a_0, a_1, \ldots \in$ D, an integral domain.

(1) $\bar{\Lambda}_0^{(rev)} \leftarrow 1; \quad B_0 \leftarrow 0; \quad L_0 \leftarrow 0; \quad \Delta \leftarrow 1; \quad g \leftarrow 1; \quad h \leftarrow 1;$

For $i = 1, 2, \ldots$ **Do**

(2) $\Delta_i \leftarrow \bar{\lambda}_s a_{i-1} + \bar{\lambda}_{s-1} a_{i-2} + \cdots + \bar{\lambda}_0 a_{i-s-1};$

If $\Delta_i = 0$ then

$$\bar{\Lambda}_i^{(rev)} \leftarrow \bar{\Lambda}_{i-1}^{(rev)}; \quad B_i \leftarrow z \cdot B_{i-1}; \quad L_i \leftarrow L_{i-1};$$

(3) If $\Delta_i \neq 0$ and $2L_{i-1} < i$ then

$$\delta \leftarrow i - L_{i-1}; \quad \bar{\Lambda}_i^{(rev)} \leftarrow -\Delta \cdot \Delta_i^{\delta} \cdot \bar{\Lambda}_{i-1}^{(rev)} + \Delta_i^{\delta+1} \cdot z \cdot B_{i-1};$$

$$B_i \leftarrow \bar{\Lambda}_{i-1}^{(rev)}; \quad L_i \leftarrow i - L_{i-1}; \quad \Delta \leftarrow \Delta_i;$$

(4) If $\Delta_i \neq 0$ and $2L_{i-1} \geq i$ then

$$\bar{\Lambda}_i^{(rev)} \leftarrow \bar{\Lambda}_{i-1}^{(rev)} - (\Delta_i/\Delta) \cdot z \cdot B_{i-1}; \quad B_i \leftarrow z \cdot B_{i-1}; \quad L_i \leftarrow L_{i-1};$$

(5) If $2L_{i-1} = i$ then

$$\bar{\Lambda}_i^{(rev)} \leftarrow (-1)^{\delta+1}/(g \cdot h^{\delta}) \cdot \bar{\Lambda}_i^{(rev)}; \quad g \leftarrow \Delta; \quad h \leftarrow h^{1-\delta} \cdot g^{\delta};$$

End For;

Sparsest shifts: All $s = (s_1, \ldots, s_n)$ minimize i such that

$$\Delta_i(s_1, \ldots, s_n, y_1, \ldots, y_n) = 0.$$

Furthermore, $\Delta_j(s_1, \ldots, s_n, y_1, \ldots, y_n) = 0$ **for** $j \geq i$.

T**-sparse shifts:** All $s = (s_1, \ldots, s_n)$ such that

$$\Delta_{2T+1}(s_1, \ldots, s_n, y_1, \ldots, y_n) = 0.$$

Special “tricks” in the univariate case:

Each discrepency can be factored:

$$\Delta_i(z,y) = \underbrace{g_i(z)}_{\text{content}} \cdot \underbrace{\varphi_i(z,y)}_{\text{primitive part}}$$

Find first Δ_i with $g_i(z)$ non-trivial in z and solve $g_i(z) = 0$.

For the polynomial

$$f(x) = \frac{5}{7}x^5 + \frac{75}{14}x^4 + \frac{225}{14}x^3 + \frac{675}{28}x^2 + \frac{2025}{112}x + \frac{1215}{224}$$

$$= \frac{5}{7}(x + \frac{3}{2})^5$$

$$
\begin{aligned}
\Delta_3 = {}&(-3+2z)\frac{25}{25088}y(y-1)^2(32805-1280y^{13}z^3+43740y+298112y^5z^4-616464y^7z\\
&+2560y^{14}z^2+753984y^7z^2-469800y^3z+453600z^4-1001376y^4z+408240yz^2\\
&-1441536y^4z^3-1080000y^2z^3+116640y^2+408240z^2-204120yz+1010880y^2z^2\\
&-453600yz^3+1645056y^4z^2-991008y^6z+302400yz^4+840240y^3z^2-704160y^8z\\
&+1327968y^6z^2+710656y^4z^4+111780y^3-544320z^3-174960z-524880y^2z\\
&+254016y^4+187272y^5-241920z^5+298728y^6+261360y^8+691200y^2z^4\\
&+145728y^{1}0+135864y^9+56016y^{12}-331344y^9z-828000y^3z^3+484800y^3z^4\\
&+304704y^9z^2-82176y^{1}2z-907008y^6z^3+720000y^8z^2-336640y^8z^3+35456y^{12}z^2\\
&-661392y^5z+957312y^5z^2-2560y^{15}z+237568y^{10}z^2-723072y^5z^3-73728y^{10}z^3\\
&+3840y^{15}+1280y^{16}-311808y^{10}z-120960yz^5+200664y^7+70272y^{11}\\
&-168320y^3z^5+138624y^7z^4-458624y^7z^3+86912y^{11}z^2-128384y^9z^3+32000y^3z^6\\
&-134592y^{11}z-62720y^5z^5-26560y^{13}z-186880z^5y^4-5120z^7y^2+56320z^6y^2\\
&+20480z^6y^4-53760z^5y^6-2560z^5y^8+2560z^6y^6-264960z^5y^2+323968z^4y^6\\
&+65280z^4y^8+7168z^4y^{1}0-12800y^{1}4z+26880yz^6-2560yz^7-4096y^{12}z^3+1280z^8\\
&+80640z^6-15360z^7+5120y^5z^6+23424y^9z^4-16640y^7z^5-2560y^3z^7-19712y^{11}z^3\\
&-1280y^9k^5+512y^{11}z^4+10880y^{13}z^2+19680y^{13}+13440y^{14}+256z^4y^{12})
\end{aligned}
$$

Symbolic algorithm: $\gcd(g_i, g_{i+1}) = g_i(z)$, content of $\gcd(\Delta_i, \Delta_{i+1})$.

$$
\begin{aligned}
\gcd(\Delta_i, \Delta_{i+1}) = {} & (2z-3)\frac{1}{2560}y(y-1)^2(32805 + 256z^4y^{12} + 43740y + 298112y^5z^4 \\
& - 1280y^{13}z^3 + 2560y^{14}z^2 - 469800y^3z - 616464y^7z + 753984y^7z^2 + 453600z^4 \\
& - 1001376y^4z + 408240yz^2 - 1441536y^4z^3 - 1080000y^2z^3 + 116640y^2 + 408240z^2 \\
& - 204120yz + 1010880y^2z^2 - 453600yz^3 + 1645056y^4z^2 - 991008y^6z + 302400yz^4 \\
& + 840240y^3z^2 - 704160y^8z + 1327968y^6z^2 + 710656y^4z^4 + 111780y^3 - 544320z^3 \\
& - 174960z - 524880y^2z + 254016y^4 + 187272y^5 - 241920z^5 + 298728y^6 + 261360y^8 \\
& + 691200y^2z^4 + 145728y^{10} + 135864y^9 + 56016y^{12} - 331344y^9z - 828000y^3z^3 \\
& + 484800y^3z^4 + 304704y^9z^2 - 82176y^{12}z - 907008y^6z^3 + 720000y^8z^2 - 336640y^8z^3 \\
& + 35456y^{12}z^2 - 661392y^5z + 957312y^5z^2 - 2560y^{15}z + 237568y^{10}z^2 - 723072y^5z^3 \\
& - 73728y^{10}z^3 + 3840y^{15} + 1280y^{16} - 311808y^{10}z - 120960yz^5 + 138624y^7z^4 \\
& - 168320y^3z^5 + 200664y^7 + 70272y^{11} + 1280z^8 + 80640z^6 - 15360z^7 - 458624y^7z^3 \\
& + 86912y^{11}z^2 - 128384y^9z^3 + 32000y^3z^6 - 134592y^{11}z - 62720y^5z^5 - 26560y^{13}z \\
& - 186880z^5y^4 - 5120z^7y^2 + 56320z^6y^2 + 20480z^6y^4 - 53760z^5y^6 - 2560z^5y^8 \\
& + 2560z^6y^6 - 264960z^5y^2 + 323968z^4y^6 + 65280z^4y^8 + 7168z^4y^{10} - 12800y^{14}z \\
& + 26880yz^6 - 2560yz^7 - 4096y^{12}z^3 + 5120y^5z^6 + 23424y^9z^4 - 16640y^7z^5 - 2560y^3z^7 \\
& - 19712y^{11}z^3 - 1280y^9z^5 + 512y^{11}z^4 + 10880y^{13}z^2 + 19680y^{13} + 13440y^{14})
\end{aligned}
$$

Single projection: project y to a value.

Two (or more) sequences: $\gcd(\Delta_i(p), \Delta_i(q)) = g_i(z)$ in

$$\Delta_1(z,p), \Delta_2(z,p), \ldots, \Delta_i(z,p), \ldots$$
$$\Delta_1(z,q), \Delta_2(z,q), \ldots, \Delta_i(z,q), \ldots$$

$2 \longrightarrow y$:

$$\Delta_3(z,2) = \frac{25}{12544}(2z-3)(1280z^8 - 61440z^7 + 1271040z^6 - 14768640z^5 + 105231456z^4 - 470393856z^3 + 1288338576z^2 - 1978582176z + 1306530941)$$

$3 \longrightarrow y$:

$$\Delta_3(z,3) = \frac{75}{6272}(2z-3)(1280z^8 - 138240z^7 + 6301440z^6 - 155485440z^5 + 2229386976z^4 - 18789350976z^3 + 91595514096z^2 - 239699154096z + 261797684661)$$

$$\gcd(\Delta_3(z,2), \Delta_3(z,3)) = g_3(z) = (z - \frac{3}{2})$$

One sequence: $\gcd(\Delta_i(z,p),\ldots,\Delta_{2\delta}(z,p)) = g_i(z)$, $\delta \geq \deg(f)$.

$2 \longrightarrow y:$

$$\gcd(\Delta_3(z,2),\Delta_4(z,2),\ldots,\Delta_{10}(z,2)) = g_3(z) = z - \frac{3}{2}$$

$$\gcd(\Delta_3(z,2),\Delta_4(z,2),\Delta_5(z,2)) = g_3(z) = z - \frac{3}{2}$$

Question: $\gcd(\Delta_i(z,p),\ldots,\Delta_\zeta(z,p)) = g_i(z)$ for $\zeta < 2\deg(f)$?

Two projections

$f \in \mathbb{Q}[x]$ $\qquad f = \frac{5}{7}(x+\frac{3}{2})^5$

2 black box probes

`DenominatorAndContent`

$\tilde{f}$: μ-primitive part of f $\qquad \tilde{f} = (2x+3)^5$

`Berlekamp/Massey` $\qquad \tilde{f}(3^i - 7), \quad \tilde{f}(5^i - 7)$

$$\Delta_i(z,y) = (az-b)^e \cdot \varphi(z,y)$$

$$\Delta_3(7,3) = \frac{459683859624}{3125}$$

$$\Delta_3(7,5) = 8141185896456800$$

$\gcd(a,b) = 1$

`FindLinearFactor`

$$\gcd(\mathrm{numer}(\Delta_3(7,3)), \mathrm{numer}(\Delta_3(7,5))) = 2^3 \cdot 11$$

$az - b$ $\qquad z = b/a = 3/2$, since $3/2 \equiv 7 \bmod 11$

Remarks

Two projections vs. “real” sparse interpolation: sparse interpolation w.r.t. sparsest shifted basis.

Sparsest shifts for a set of polynomials $f_1, \ldots, f_m \in D[x_1, \ldots, x_n]$: consider

$$G(x_1, \ldots, x_n, z_0) = f_1 + z_0 f_2 + \cdots + z_0^{m-1} f_{m-1} + z_0^{m-1} f_m.$$

Any given power basis

Chebyshev, Pochhammer bases: cf. Kaltofen, Lee (2002)

Further extensions