

Error-Correcting Sparse Interpolation in the Chebyshev Basis

Andrew Arnold*

University of Waterloo
a4arnold@uwaterloo.ca
AndrewArnold.ca

Erich Kaltofen

North Carolina State University
kaltofen@math.ncsu.edu
kaltofen.us

ISSAC '15, Bath, UK
2015-07-08

Introduction

Let $T_n(x)$ be the n th Chebyshev polynomial of the first kind:

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

E.g. $T_2 = 2x^2 - 1$, $T_3 = 4x^3 - 3x$, $T_4 = 8x^4 - 8x^2 + 1$, ...

Useful Properties:

- ▶ $T_n(T_m(x)) = T_{mn}(x)$
- ▶ $T_m(x)T_n(x) = \frac{1}{2}(T_{m+n}(x) + T_{|m-n|}(x))$
- ▶ $T_n\left(\frac{x+x^{-1}}{2}\right) = \frac{1}{2}(x^n + x^{-n})$
- ▶ Over \mathbb{R} , for $|\xi| > 1$, $T_n(\xi) \neq 0$

We are interested in the interpolation and error-correcting properties of sparse Chebyshev polynomials (i.e. polynomials that are sparse in the Chebyshev basis).

Contributions

Covered in this talk

An error-correcting black-box interpolation algorithm for sparse Chebyshev polynomials.

Not covered in this talk

An alternative sparse Chebyshev interpolation algorithm for $f \in K[x]$, $\text{char}(K) \neq 2$, that reduces the problem to sparse interpolation in the power basis, i.e. $1, x, x^2, x^3, \dots$

- ▶ Allows for **early termination** (Kaltofen, Lee; 2003), such that we can (probabilistically) interpolate f with t terms, with cost sensitive to t even when bounds for t are not supplied as input.

The Problem

- Suppose $f \in \mathbb{R}[x]$ is of the form

$$f = \sum_{i=1}^t c_i \underbrace{T_{\delta_i}(x)}_{c_i \neq 0 \text{ and with } \delta_1 > \delta_2 > \dots > \delta_t}$$

- We are given a **black box** \blacksquare for f . For $j = 0, 1, \dots, L-1$, we query \blacksquare with x_j and get back a_j , where either $a_j = f(x_j)$ or a_j is an erroneous evaluation.

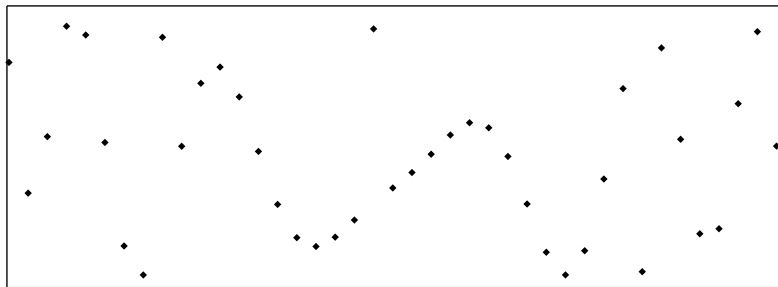


- Problem:** reconstruct f and identify errors, given \blacksquare and bounds

$$B \geq t, \quad D \geq \deg(f), \quad E \geq \#\{\text{errors}\}$$

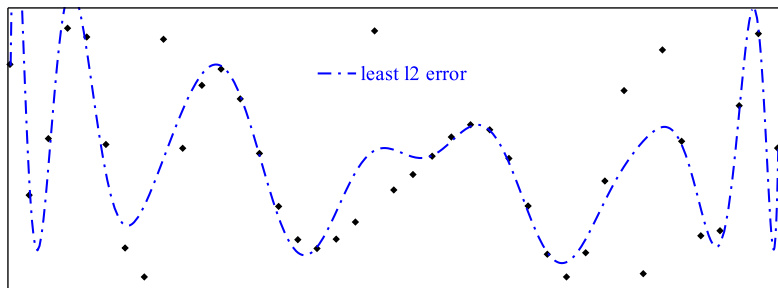
while minimizing, L the number of queries to \blacksquare .

Interpolating with Errors - Example



E.g. $\deg(f) \leq D = 19$, $\leq B = 3$ terms, $\leq E = 5$ errors

Interpolating with Errors - Example

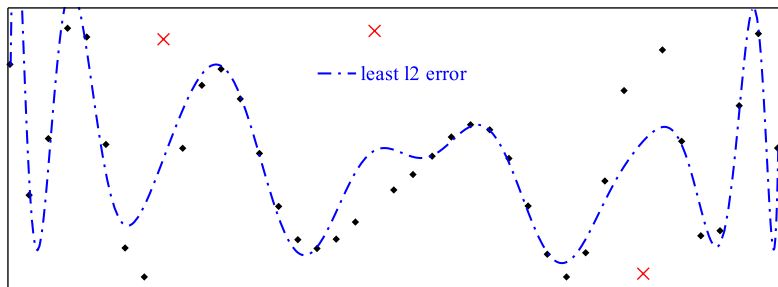


E.g. $\deg(f) \leq D = 19$, $\leq B = 3$ terms, $\leq E = 5$ errors

- ▶ Minimizing ℓ_2 -error gives a **dense approximation** for the model,

$$\begin{aligned} & 0.786462T_{19} - 0.253808T_{19} - 0.270838T_{18} + 0.101009T_{16} + 0.206344T_{15} - 0.135857T_{15} - 0.076361T_{14} + \\ & 0.051550T_{12} - 0.699793T_{12} + 0.003612T_{10} - 0.473865T_{10} + 0.352537T_8 - 0.307681T_8 - 1.054240T_7 + \\ & 0.753950T_5 - 0.112232T_5 - 1.388821T_4 + 1.025795T_2 + 1.364547T_1 + 3.325460T_0 \end{aligned}$$

Interpolating with Errors - Example



E.g. $\deg(f) \leq D = 19$, $\leq B = 3$ terms, $\leq E = 5$ errors

- ▶ Minimizing ℓ_2 -error gives a **dense approximation** for the model,

$$\begin{aligned} & 0.786462T_{19} - 0.253808T_{19} - 0.270838T_{18} + 0.101009T_{16} + 0.206344T_{15} - 0.135857T_{15} - 0.076361T_{14} + \\ & 0.051550T_{12} - 0.699793T_{12} + 0.003612T_{10} - 0.473865T_{10} + 0.352537T_8 - 0.307681T_8 - 1.054240T_7 + \\ & 0.753950T_5 - 0.112232T_5 - 1.388821T_4 + 1.025795T_2 + 1.364547T_1 + 3.325460T_0 \end{aligned}$$

- ▶ But if we identify 3 **errors**...

Interpolating with Errors - Example



E.g. $\deg(f) \leq D = 19$, $\leq B = 3$ terms, $\leq E = 5$ errors

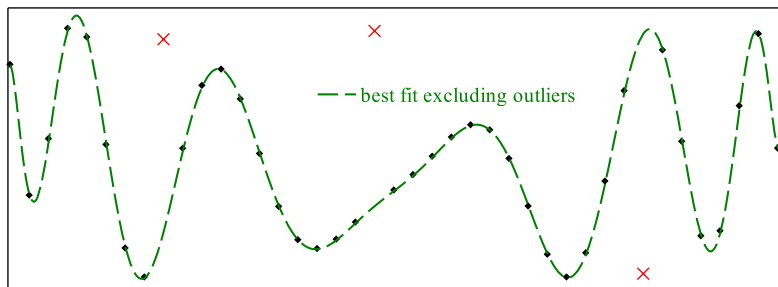
- ▶ Minimizing ℓ_2 -error gives a **dense approximation** for the model,

$$0.786462T_{19} - 0.253808T_{19} - 0.270838T_{18} + 0.101009T_{16} + 0.206344T_{15} - 0.135857T_{15} - 0.076361T_{14} + \\ 0.051550T_{12} - 0.699793T_{12} + 0.003612T_{10} - 0.473865T_{10} + 0.352537T_8 - 0.307681T_8 - 1.054240T_7 + \\ 0.753950T_5 - 0.112232T_5 - 1.388821T_4 + 1.025795T_2 + 1.364547T_1 + 3.325460T_0$$

- ▶ But if we identify 3 **errors**...we get a sparse fit

$$T_{15} - 2T_{11} + T_2$$

Interpolating with Errors - Example



E.g. $\deg(f) \leq D = 19$, $\leq B = 3$ terms, $\leq E = 5$ errors

- ▶ Minimizing ℓ_2 -error gives a **dense approximation** for the model,

$$0.786462T_{19} - 0.253808T_{19} - 0.270838T_{18} + 0.101009T_{16} + 0.206344T_{15} - 0.135857T_{15} - 0.076361T_{14} + \\ 0.051550T_{12} - 0.699793T_{12} + 0.003612T_{10} - 0.473865T_{10} + 0.352537T_8 - 0.307681T_8 - 1.054240T_7 + \\ 0.753950T_5 - 0.112232T_5 - 1.388821T_4 + 1.025795T_2 + 1.364547T_1 + 3.325460T_0$$

- ▶ But if we identify 3 **errors**...we get a sparse fit

$$T_{15} - 2T_{11} + T_2$$

Related Work - Sparse Interpolation in Power Basis

Theorem (Prony; 1795)

Let K be a field, $f(x) = \sum_{i=1}^t c_i x^{e_i} \in K[x]$, $\omega \in K$, and

$$\Phi(y) = \prod_{i=1}^t (y - \omega^{e_i}) = y^t + \sum_{i=0}^{t-1} \phi_i y^i.$$

Then for $a_j = f(\omega^j)$,

$$\underbrace{\begin{bmatrix} a_0 & a_1 & \cdots & a_{t-1} \\ a_1 & a_2 & \cdots & a_t \\ \vdots & \vdots & \ddots & \vdots \\ a_{t-1} & a_t & \cdots & a_{2t-2} \end{bmatrix}}_{=\mathcal{H}_t} \begin{bmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_{t-1} \end{bmatrix} = \begin{bmatrix} a_t \\ a_{t+1} \\ \vdots \\ a_{2t-1} \end{bmatrix}.$$

- ▶ i.e., exponents of f are encoded in a solution to a [Hankel](#) system.
- ▶ **Corollary:** $\mathcal{H}_{t'}$ is singular for $t' > t$.

Prony's Method for Interpolation in the Power Basis¹

Inputs: Error-free \blacksquare for t -sparse $f \in K[x]$, bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\omega \in K$ of order $> D$. Let $x_j = \omega^j$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, 1, \dots, 2B - 1$.

¹See also Ben-Or–Tiwari; 1988

Prony's Method for Interpolation in the Power Basis¹

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{K}[x]$, bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\omega \in \mathbb{K}$ of order $> D$. Let $x_j = \omega^j$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, 1, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that \mathcal{H}_t is nonsingular.

¹See also Ben-Or–Tiwari; 1988

Prony's Method for Interpolation in the Power Basis¹

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{K}[x]$, bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\omega \in \mathbb{K}$ of order $> D$. Let $x_j = \omega^j$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, 1, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that \mathcal{H}_t is nonsingular.
3. Solve size- t Hankel system to get $\Phi(y)$ with roots $\omega^{e_i}, 1 \leq i \leq t$.

¹See also Ben-Or–Tiwari; 1988

Prony's Method for Interpolation in the Power Basis¹

Inputs: Error-free \blacksquare for t -sparse $f \in K[x]$, bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\omega \in K$ of order $> D$. Let $x_j = \omega^j$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, 1, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that \mathcal{H}_t is nonsingular.
3. Solve size- t Hankel system to get $\Phi(y)$ with roots ω^{e_i} , $1 \leq i \leq t$.
4. Factor Φ to get ω^{e_i} . Determine e_i from ω^{e_i} for $1 \leq i \leq t$.

¹See also Ben-Or–Tiwari; 1988

Prony's Method for Interpolation in the Power Basis¹

Inputs: Error-free \blacksquare for t -sparse $f \in K[x]$, bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\omega \in K$ of order $> D$. Let $x_j = \omega^j$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, 1, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that \mathcal{H}_t is nonsingular.
3. Solve size- t Hankel system to get $\Phi(y)$ with roots ω^{e_i} , $1 \leq i \leq t$.
4. Factor Φ to get ω^{e_i} . Determine e_i from ω^{e_i} for $1 \leq i \leq t$.
5. Obtain the coefficients c_i as the solution to the **Vandermonde system**

$$\begin{bmatrix} x_0^{e_1} & x_0^{e_2} & \cdots & x_0^{e_t} \\ x_1^{e_1} & x_1^{e_2} & \cdots & x_1^{e_t} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t-1}^{e_1} & x_{t-1}^{e_2} & \cdots & x_{t-1}^{e_t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{t-1} \end{bmatrix},$$

and output $f = \sum_{i=1}^t c_i x^{e_i}$.

¹See also Ben-Or-Tiwari; 1988

Related Work - Sparse Chebyshev Interpolation

Theorem (Lakshman, Saunders; 1995)

Let $f(x) = \sum_{i=1}^t c_i T_{\delta_i}(x) \in \mathbb{R}[x]$, $\xi > 1$, and

$$\Phi(y) = \prod_{i=1}^t (y - T_{\delta_i}(\xi)) = y^t + \sum_{i=0}^{t-1} \phi_i y^i.$$

Then for $a_j = f(T_j(\xi))$,

$$\left(\underbrace{[a_{i+j}]_{i,j=0}^{t-1}}_{\mathcal{H}_t} + \underbrace{[a_{|i-j|}]_{i,j=0}^{t-1}}_{\mathcal{T}_t} \right) [\phi_i]_{i=0}^t = [a_{t+i} + a_{t-i}]_{i=0}^{t-1}.$$

- ▶ i.e., indices δ_i are encoded in solution to a *Hankel+Toeplitz* system.
- ▶ One can show $\mathcal{H}_{t'} + \mathcal{T}_{t'}$ is nonsingular for $t' > t$.

Lakshman–Saunders Sparse Chebyshev Interpolation

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{R}[x]$; bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\xi > 1$. Let $x_j = T_j(\xi)$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, \dots, 2B - 1$.

Lakshman–Saunders Sparse Chebyshev Interpolation

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{R}[x]$; bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\xi > 1$. Let $x_j = T_j(\xi)$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that $\mathcal{H}_t + \mathcal{T}_t$ is nonsingular.

Lakshman–Saunders Sparse Chebyshev Interpolation

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{R}[x]$; bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\xi > 1$. Let $x_j = T_j(\xi)$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that $\mathcal{H}_t + \mathcal{T}_t$ is nonsingular.
3. Solve Hankel+Toeplitz system to obtain $\Phi(y)$ with roots $T_{\delta_j}(\xi)$.

Lakshman–Saunders Sparse Chebyshev Interpolation

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{R}[x]$; bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\xi > 1$. Let $x_j = T_j(\xi)$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that $\mathcal{H}_t + \mathcal{T}_t$ is nonsingular.
3. Solve Hankel+Toeplitz system to obtain $\Phi(y)$ with roots $T_{\delta_j}(\xi)$.
4. Factor $\Phi(y)$ to get $T_{\delta_i}(\xi)$. Determine δ_i from T_{δ_i} for $1 \leq i \leq t$.

Lakshman–Saunders Sparse Chebyshev Interpolation

Inputs: Error-free \blacksquare for t -sparse $f \in \mathbb{R}[x]$; bounds $B \geq t$, $D \geq \deg(f)$

1. Choose $\xi > 1$. Let $x_j = T_j(\xi)$.
Evaluate $a_j = \blacksquare(x_j)$ for $j = 0, \dots, 2B - 1$.
2. Determine $t \leq B$ as the largest value such that $\mathcal{H}_t + \mathcal{T}_t$ is nonsingular.
3. Solve Hankel+Toeplitz system to obtain $\Phi(y)$ with roots $T_{\delta_i}(\xi)$.
4. Factor $\Phi(y)$ to get $T_{\delta_i}(\xi)$. Determine δ_i from T_{δ_i} for $1 \leq i \leq t$.
5. Obtain coefficients c_i as a solution to the linear system

$$\left[T_{\delta_i}(x_{j-1}) \right]_{i,j=1}^t \left[c_i \right]_{i=1}^t = \left[a_{i-1} \right]_{i=1}^t,$$

$$\text{output } f = \sum_{i=1}^t c_i T_{\delta_i}.$$

Block-Decoding Error Correction

We can use Prony (or **Lakshman–Saunders**) as in order to employ error-correcting interpolation in the power (or **Chebyshev**) basis.

Block Majority-Vote Decoding

- ▶ Run Prony or Lakshman–Saunders on $(2E + 1)$ blocks of $2B$ evaluations.
- ▶ A majority of blocks will produce the true interpolant f .

Block List Decoding

- ▶ Run Prony or Lakshman–Saunders on $(E + 1)$ blocks of $2B$ evaluations.
- ▶ One block must produce the true interpolant.

Can we uniquely decode or list decode f with fewer evaluations than block decoding?

Subsequence List Decoding (Kaltofen, Pernet; ISSAC '14)

- ▶ Prony's algorithm generalizes for evaluations $\blacksquare(\rho\omega^j), 0 \leq j < 2B$.

Subsequence List Decoding (Kaltofen, Pernet; ISSAC '14)

- ▶ Prony's algorithm generalizes for evaluations $\blacksquare(\rho\omega^j), 0 \leq j < 2B$.
- ▶ We can query over $a_j = \blacksquare(\omega^j), j = 0, 1, 2, \dots, L - 1$, and run Prony's algorithm over every geometric subsequence of length $2B$, i.e., $a_r, a_{r+s}, a_{r+2s}, \dots, a_{r+2Bs}$ for any $r \geq 0, s \geq 1$ with $r + 2Bs < L$.

Subsequence List Decoding (Kaltofen, Pernet; ISSAC '14)

- ▶ Prony's algorithm generalizes for evaluations $\mathbf{a}(\rho\omega^j), 0 \leq j < 2B$.
- ▶ We can query over $a_j = \mathbf{a}(\omega^j), j = 0, 1, 2, \dots, L - 1$, and run Prony's algorithm over every geometric subsequence of length $2B$, i.e., $a_r, a_{r+s}, a_{r+2s}, \dots, a_{r+2Bs}$ for any $r \geq 0, s \geq 1$ with $r + 2Bs < L$.
- ▶ Kaltofen & Pernet showed that this outperforms block list decoding. i.e., we can always choose $L \leq (E + 1)2B$.

Subsequence List Decoding (Kaltofen, Pernet; ISSAC '14)

- ▶ Prony's algorithm generalizes for evaluations $\blacksquare(\rho\omega^j), 0 \leq j < 2B$.
- ▶ We can query over $a_j = \blacksquare(\omega^j), j = 0, 1, 2, \dots, L - 1$, and run Prony's algorithm over every geometric subsequence of length $2B$, i.e., $a_r, a_{r+s}, a_{r+2s}, \dots, a_{r+2Bs}$ for any $r \geq 0, s \geq 1$ with $r + 2Bs < L$.
- ▶ Kaltofen & Pernet showed that this outperforms block list decoding. i.e., we can always choose $L \leq (E + 1)2B$.
- ▶ For $f \in \mathbb{R}[x], \omega > 0$, and $L \geq 2B + 2E$, if a t -sparse g satisfies $g(\omega^j) = \blacksquare(\omega^j)$ for $0 \leq j < L$ with $\leq E$ exceptions, then by Descartes' rule of signs, g must be f .
 \Rightarrow list decoder uniquely determines f if we use at least $2B + 2E$ evaluations.

Generalization of Lakshman–Saunders Algorithm

In order to extend subsequence list decoding to the Chebyshev basis, we need to be able to interpolate f over choices of subsequences.

Lemma (A, Kaltofen; 2015)

Let $f(x) = \sum_{i=1}^t c_i T_{\delta_i}(x) \in \mathbb{R}[x]$ and fix $r, s \in \mathbb{Z}$, and let

$$\Phi(y) = \prod_{i=1}^t (y - T_{s\delta_i}(\xi)) = y^t + \sum_{i=0}^{t-1} \phi_i y^i.$$

Then the sequence $a_j = f(T_j(\xi)), j = 0, 1, \dots$, where $\xi \in \mathbb{R}$ satisfies

$$\underbrace{[a_{|r+(i+j)s|} + a_{|r+(i-j)s|}]_{i,j=0}^{t-1}}_{A(r,s)} [\phi_i]_{i=0}^{t-1} = [a_{|r+(i+t)s|} + a_{|r+(i-t)s|}]_{i=0}^{t-1}$$

This requires evaluations $a_{|rj+s|}$ for $-B \leq j < 2B$.

Generalization of Lakshman–Saunders Algorithm

In order to interpolate f from evaluation points $a_{|rj+s|}$, $-B \leq j < 2B$, we require that $A^{(r,s)} = [a_{|r+(i+j)s|} + a_{|r+(i-j)s|}]_{i,j=0}^{t-1}$ is nonsingular.

Definition

We say r, s are *valid* if $|r + is|$, $0 \leq i < B$, are distinct.

Fact

$A^{(r,s)} = \mathcal{U}\mathcal{B}\mathcal{V}$, where $\mathcal{B} = \text{diag}(2c_1, 2c_2, \dots, 2c_t)$ and

$$\mathcal{U} = [T_{|r+si|}(T_{\delta_{j+1}}(\xi))]_{i,j=0}^{t-1}, \quad \mathcal{V} = [T_{sj}(T_{\delta_{i+1}}(\xi))]_{i,j=0}^{t-1},$$

Observation: If $w\mathcal{U} = 0$ for a row vector $w \neq 0$, then

$$\sum_{i=1}^t w_i T_{|r+si|}(T_{\delta_{j+1}}(\xi)) = 0 \quad \text{for } 0 \leq j < B.$$

i.e., $\sum_{i=1}^t w_i T_{|r+si|}$ is nonzero (for valid r, s) and t -sparse with t roots $T_{\delta_1}(\xi), \dots, T_{\delta_t}(\xi)$.

A Rule of Signs for Chebyshev Basis

Theorem (Obrechhoff, 1918²)

Let $f = \sum_{i=1}^t c_i T_{\delta_i}(x) \in \mathbb{R}[x]$, and let s be the number of sign changes in (c_1, \dots, c_t) . Then f has at most s real roots ≥ 1 .

Corollary

If f is B -sparse, and $f(x) = 0$ for $x = \underbrace{x_1, \dots, x_B}_{\text{distinct, } \geq 1}$, then $f = 0$.

- ▶ $\Rightarrow \mathcal{U}, \mathcal{V}$ are nonsingular $\Rightarrow A^{(r,s)} = \mathcal{U}\mathcal{B}\mathcal{V}$ is nonsingular for valid r, s
- ▶ Generalized Lakshman–Saunders uses $\in \{2B, \dots, 3B\}$ evaluations.

Corollary

A list decoder decodes $f \in \mathbb{R}[x]$ uniquely for $L \geq 2B + 2E$ evaluations at points $x_i > 1$.

²Thanks to the anonymous referee for bringing this reference to our attention

Subsequence List Decoding Sparse Chebyshev Interpolation

- ▶ We evaluate $a_j = \blacksquare(T_j(\xi))$, for $j = 0, 1, \dots, L_{B,E} - 1$, where $L_{B,E}$ is the least value such that $a_0, \dots, a_{L_{B,E}-1}$ must have valid r, s such that $a_{|r+is|}$, $-B \leq i < 2B$ is error-free.
- ▶ We run Lakshman–Saunders algorithm over evaluation points $a_{|r+is|}$, $-B \leq i < 2B$, for all valid (r, s) , and check each resulting polynomial g if it agrees with \blacksquare for all but at most E exceptions.
- ▶ We would like, for each $B \geq 1$, that subsequence list decoding requires fewer evaluations than list decoding for sufficiently large E
- ▶ **Problem:** $L_{B,E}$ is exponentially costly to compute.

Combining Block and Subsequence List Decoding

Fact (Kaltofen, Pernet; 2014)

$$L_{1,8} \leq 17, \quad L_{2,8} \leq 34$$

- ▶ i.e., We can correct 8 errors for B -sparse f and $B = 1, 2$ with $17B$ evaluations.
- ▶ If $B = 1, 2$, and we want to correct for E errors, then we can choose $\lceil E/8 \rceil$ blocks of $17B$ evaluations, and run the subsequence list decoder on each block.
 \Rightarrow We can correct E errors with $17B \lceil E/8 \rceil$ evaluations.

Corollary

Let $L_{B,E}^{\min}$ be the least number of evaluations needed to interpolate a B -sparse polynomial with E errors. Then

$$L_{B,E}^{\min} < (E + 1)2 \text{ for } E \geq 57, \quad L_{B,E}^{\min} < (E + 1)4 \text{ for } E \geq 86.$$

\Rightarrow We can do better than block list decoding for $B = 1, 2$.

Conjecture: for every $B > 0$, $\exists E_B$ such that $L_{B,E}^{\min} < (E + 1)B$ for $E > E_B$.

Thank You

