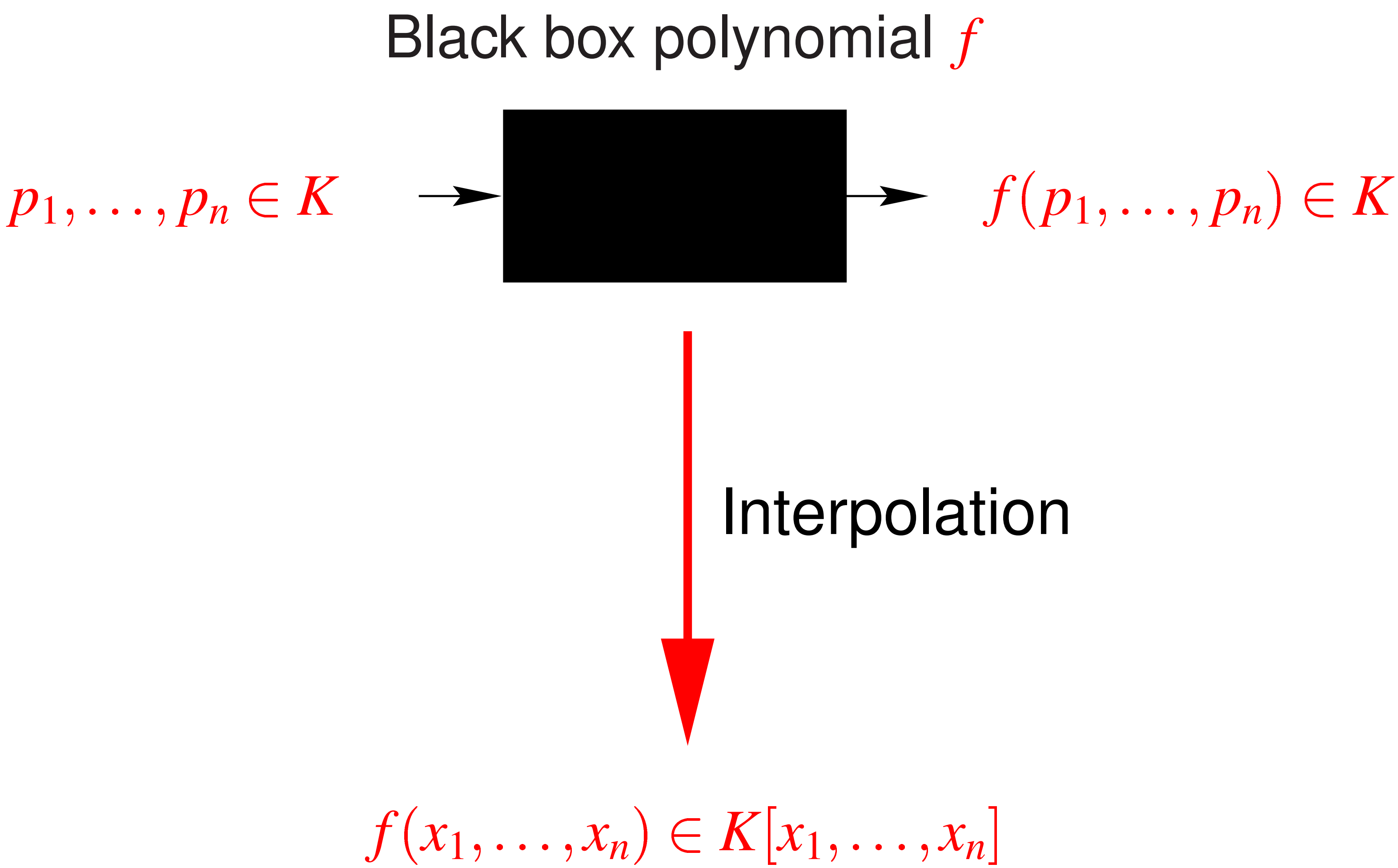
# Early Termination Strategies in Sparse Interpolation Algorithms

Wen-shin Lee
North Carolina State University
www.wen-shin.com

July 10, 2001

# Black box polynomial interpolation




What if $f(x_1, \ldots, x_n)$ is sparse?

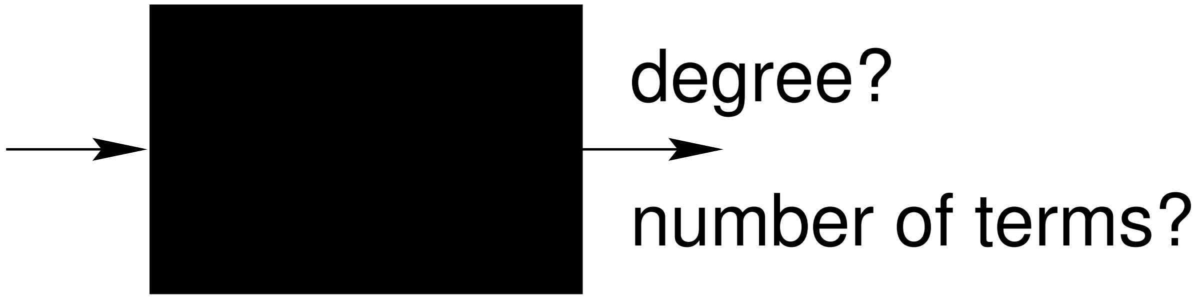
## Zippel's probabilistic interpolation.

- 😐 Interpolate variable by variable.
- ☺ Sensitive to the sparsity in the multivariate case.
- ☹ Still interpolates each variable densely.
- ☹ Needs an upper bound for the degree in each variable.
- ☺ Does not need a large modulus. $O(\max_i \deg(f(x_i)))$

## Ben-Or's/Tiwari's deterministic algorithm.

- 😐 Interpolate all variables at once.
- ☺ Sensitive to the sparsity of the the target polynomial.
- ☹ Needs an upper bound for the number of terms.
- ☹ Might need a very large modulus. $O(\max_{\mathbf{e}} p_1^{e_1} \cdots p_n^{e_n})$, $p_i$ the $i$-th prime, $e_i = \deg(f(x_i))$.

# Early termination strategy




What if ***an upper bound of degree*** and ***an upper bound of the number of terms*** of the target polynomial are ***not known?***

- Guess and check.

  And double the guess if fails.

- ***Early termination strategy.***

  Interpolate the polynomial at a random point, when the polynomial stops changing, it is done with high probability.

## Why early termination?

- Save time and space.
- A useful tool for controlling intermediate expression swell in computer algebra.
- Sensitive to the sparsity of the target polynomial without knowing any bounds on degree or number of terms.

# Early termination in Newton interpolation (12–16)

- Interpolate $f(x)$ on a sequence of random values:

$$p_0,\ p_1,\ p_2,\ \ldots,\ p_i,\ \ldots$$

- Threshold $\eta$ can improve the probability of correctness.

$f(x)$ is interpolated with respect to the mixed power base:

$$1,\ (x-p_0),\ (x-p_0)(x-p_1),\ \ldots$$

## Early termination in Ben-Or/Tiwari algorithm (24–33)

- Perform the Berlekamp/Massey algorithm on
$$f(p),\ f(p^2),\ f(p^3),\ \ldots,\ f(p^i),\ \ldots$$ with $p$ random.
- Recover the terms and coefficients.
- Threshold $\zeta$ can improve the probability, but its analysis is complicated.

$f(x)$ is interpolated with respect to the power base:
$$1,\ x,\ x^2,\ x^3,\ \ldots$$ in the univariate case.

# Early termination in sparse Pochhammer interpolation (37–43)

$$x^{\overline{n}} = x(x+1)\cdots(x+n-1)$$
$$f(x) = \sum_{j=1}^{t} c_j x^{\overline{e}_j}$$
$$f^{(i)}(x) = \sum_{j=1}^{t} e_j^i c_j x^{\overline{e}_j}$$

- Perform the Berlekamp/Massey algorithm on

  $f^{(0)}(p),\ f^{(1)}(p),\ f^{(2)}(p),\ \ldots,\ f^{(i)}(p),\ \ldots$ with $p$ random,

  $f^{(i)}(p)$ can be obtained from $f(p),\ f(p+1),\ \ldots,\ f(p+i)$.
- Recover $x^{\overline{e}_j}$ and $c_j$.
- Threshold $\zeta$ can improve the probability, but its analysis is complicated.

# Early termination in sparse Chebyshev interpolation (46–59)

$$T_0(x) = 1, \quad T_1(x) = x$$
$$i \geq 2: \quad T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x)$$
$$f(x) = \textstyle\sum_{j=1}^{t} c_j T_{\delta_j}(x)$$

- With $p$ random and $a_i = f(T_i(p))$, solve the first

$$\begin{bmatrix} 2a_0 & 2a_1 & \ldots & 2a_t \\ 2a_1 & a_2 + a_0 & \ldots & a_{t+1} + a_{t-1} \\ \vdots & \vdots & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 2a_t & a_{t+1} + a_{t-1} & \ldots & a_{2t} + a_0 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{t-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

- Recover $T_{\delta_j}(x)$ and $c_j$ from $\Lambda(z) = z^t + \lambda_{t-1} z^{t-1} + \cdots + \lambda_0$.

## Solve a symmetric Toeplitz-plus-Hankel matrix in quadratic time

Any deterministic quadratic time algorithm requires:

all principal leading submatrices nonsingular, which means for $f(x) = \sum_{j=1}^{t} c_j T_{\delta_j}(x)$ with $1 \le i \le t$, it is necessary

$$c_1 + \cdots + c_i \neq 0$$

## Fix with randomization (50–51)

Pick a random $p_c$ and consider $f(x) + p_c = \sum_{j=1}^{t} \tilde{c}_j T_{\tilde{\delta}_j}(x)$:

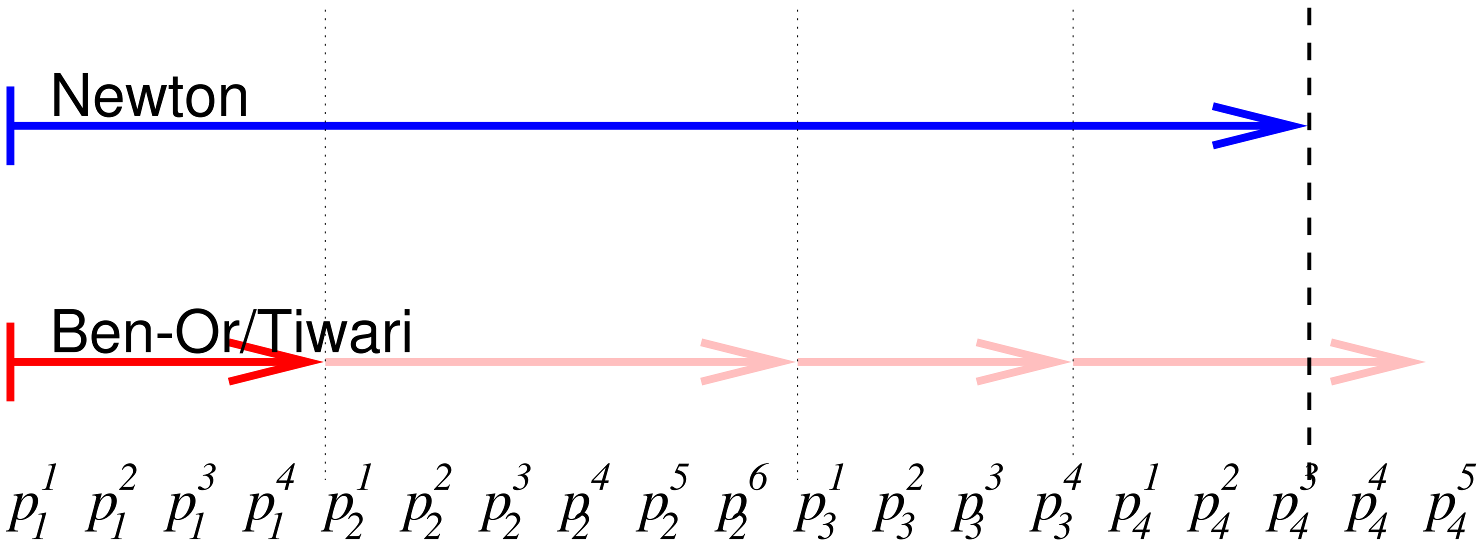
for $i$ with $1 \le i \le t$, it is correct with high probability

$$\tilde{c}_1 + \cdots + \tilde{c}_i \neq 0$$

Interpolate $f(x) + p_c$ and remove $p_c$ at the end.

# Early termination in racing algorithms (60–75)

Newton v.s. Ben-Or/Tiwari:




Newton v.s. sparse Pochhammer:

$$f(p),\ f(p+1),\ \ldots,\ f(p+i),\ \ldots$$

Newton v.s. sparse Chebyshev:

$$f(T_0(p)),\ f(T_1(p)),\ \ldots,\ f(T_i(p)),\ \ldots$$

# Embed the racing algorithms into Zippel's algorithm (83–85)

☺ Univariate interpolations within Zippel can also be sparse.

☺ Reduce the magnitude of the modulus needed for the recovery of all the terms. $O(\max_{\mathbf{e}} p_1^{e_1} \cdots p_n^{e_n}) \longrightarrow O(\max_{e_i} 2^{e_i})$

## My 7 original contributions

1. Early termination proved for sparse Pochhammer interpolations
2. Early termination proved for sparse Chebyshev interpolations
3. The complications for Chebyshev bases in a symmetric Hankel-plus-Toeplitz matrix solver are eliminated with randomization
4. Early termination proved for racing algorithms
5. Thresholds
6. Maple implementation, *ProtoBox*
7. Early termination in sparse shifts

# Sparse shifts for polynomials

For univariate polynomials (Lakshman and Saunders)

- Need $f(x) = \sum_{i=0}^{d} c_i x^i$, its first $2\tau$ derivatives, and an upper bound $\tau$ for the number of terms in the sparse shift.
- Need to evaluate $f(x)$ and its first $2\tau$ derivatives at $2$ points for a total of $4\tau + 2$ evaluations.

Grigoriev and Lakshman extended to multivariate polynomials.

## Early termination in sparse shifts for black box polynomials

$f(x)$ is given as a black box polynomial, without

- its derivatives
- an upper bound $\tau$ for the number of terms in the sparse shift

Let $t$ be the number of terms in a sparsest shift of $f(x)$. After $2t+\zeta$ black box evaluations,

$$f(x) = \sum_{i=1}^{t} \tilde{c}_i (x+\alpha)^{\delta_i}$$

is returned with high probability.

# Recall the early termination of Ben-Or/Tiwari algorithm

For a given black box polynomial $f$

- Pick a random $p$.
- Perform the Berlekamp/Massey algorithm on $f(p), f(p^2), f(p^3), \ldots$.
  - Whenever the early termination condition is satisfied, that is, $\Delta = 0$ when $2L < r$ and $1 < r$, the roots of $\Lambda(z)$ correspond to the terms in $f$ with high probability.
- Recover the terms from $\Lambda(z)$.
- Determine the coefficients from the recovered terms.

# The trick for early termination in the sparse shifts

***Write down what we do not know: the sparsest shift*** $\alpha$

- Then the sparsest shift is $f(x) = \sum_{i=1}^{t} \tilde{c}_i (x+\alpha)^{\delta_i}$.

- Consider $y = x + \alpha$ and $g(\alpha, y) = f(y - \alpha)$ as follows:

$$g(\alpha, y) = f(y-\alpha) = \sum_{i=1}^{t} \tilde{c}_i y^{\delta_i}, \qquad \tilde{c}_i \in \mathbb{K}[\alpha]$$

- When $g(\alpha, y)$ evaluated at $y = p$, $g(\alpha, p) \in \mathbb{K}[\alpha]$.

## Perform the Berlekamp/Massey algorithm on polynomials

| Early termination in sparse shifts | Early termination in Ben-Or/Tiwari |
|---|---|
| • At a random $p$, perform the Berlekamp/Massey algorithm on $g(\alpha, p), g(\alpha, p^2), g(\alpha, p^3), \ldots$ $\Delta(\alpha) \in \mathbb{K}[\alpha]$ <br> • ***Find the first*** $\alpha \in \mathbb{K}$ ***such that*** $\Delta(\alpha) = 0$ ***when*** $2L < r$ ***and*** $1 < r$ | • At a random $p$, perform the Berlekamp/Massey algorithm on $f(p), f(p^2), f(p^3), \ldots$ $\Delta \in \mathbb{K}$ <br> • Early termination: the first $\Delta = 0$ when $2L < r$ and $1 < r$ |

## How to find $\alpha$ efficiently?

☹ Factorize $\Delta(\alpha)$

$2t+1$ black box probes: until $\Delta_{2t+1}(\alpha)$

😐 Compute the GCD of $\Delta_i(\alpha)$ and $\Delta_{i+1}(\alpha)$

$2t+2$ black box probes: until $\mathrm{GCD}(\Delta_{2t+1}(\alpha), \Delta_{2t+2}(\alpha))$

☺ Compute the GCD of $\Delta_i(0)$ and $\Delta_{i+1}(0)$

?

☺ ???

## Under construction