J. Symbolic Computation, vol. 134, May–June 2026; accepted for publication, Sep. 1, 2025.
Special Issue on ISSAC 2024, Wen-shin Lee and Jonathan Hauenstein, editors.
https://doi.org/10.1016/j.jsc.2025.102507

Early Termination for Sparse Interpolation of Polynomials in Chebyshev Bases

Erich L. Kaltofen
Department of Mathematics, North Carolina State University,
Raleigh, North Carolina 27695-8205, USA
Department of Computer Science, Duke University,
Durham, North Carolina 27708-0129, USA
kaltofen@ncsu.edu; https://kaltofen.math.ncsu.edu/
erich.kaltofen@duke.edu; https://users.cs.duke.edu/~elk27

Zhi-Hong Yang
School of Mathematics and Statistics, Central South University
Changsha, Hunan 410083, China
yangzhihong@csu.edu.cn

Abstract

We show that the early termination algorithm in [Kaltofen and Lee, JSC, vol. 36, nr. 3–4, 2003] for interpolating a polynomial that is a linear combination of t Chebyshev polynomials of the first kind can be modified to use $2t+1$ randomized evaluation points; Kaltofen and Lee required $2t+2$ randomized evaluation points. Our variants work for scalar fields of any characteristic. The number $2t+1$ of evaluations matches that of the early termination version of the Prony sparse interpolation algorithm for the standard basis of powers of the variable [Kaltofen, Lee and Lobo, Proc. ISSAC 2000].

Our interpolation algorithm can compute the term locator polynomial in $O(t^2)$ field arithmetic operations while storing $O(t)$ intermediate field elements by Heinig's Toeplitz solver with singular sections [Heinig and Rost, "Algebraic Methods for Toeplitz-like Matrices and Operators," Birkhäuser, 1984]. We describe a slight modification for the Levinson-Durbin-Heinig algorithm that mirrors the Berlekamp-Massey algorithm for Hankel matrices.

1. Introduction

Early termination in sparse polynomial interpolation with the de Prony's/BCH-decoding technique uses randomization to determine the sparsity and to minimize the number of polynomial values [Kaltofen and Lee 2003]. For a polynomial f of sparsity t one evaluates $2t+1$ points and recovers the interpolant with high probability, where the probability of failure is $O(\frac{t^3 \deg(f)}{|S|})$, and where S is a finite set with $|S|$ scalar elements from which random values are sampled. If the probability of failure is to be guaranteed, an upper bound $D \geq \deg(f)$ is required on input, $t \leq D+1$, and S can be chosen sufficiently large. The big-Oh constant can be lowered for a coefficient field K of characteristic $\neq 2$ [Hao, Kaltofen, and Zhi 2016, Section 4.1]. The sparsity is with respect to the standard basis of powers of the polynomial variable.

Sparsity with respect to Chebyshev polynomials is also studied, because a linear combination of t Chebyshev-1 (1-st kind) polynomials

$$f(x) = \sum_{j=1}^{t} \widehat{c}_j \, T_{\delta_j}(x),\ 0 \leq \delta_1 < \delta_2 < \cdots < \delta_t,\ \widehat{c}_j \neq 0 \text{ for all } 1 \leq j \leq t, \tag{1}$$

can represent a periodic harmonic function by the Chebyshev-1 identity $T_n(\cos(\theta)) = \cos(n\theta)$ (see (19) for the scalars $\widehat{c}_j$ vs. c_j). Early termination by randomization for the sparse interpolation algorithm in [Lakshman Y. N. and Saunders 1995] is analyzed for scalar fields of characteristic 0 in [Kaltofen and Lee 2003, Theorem 14]. Chebyshev-1 polynomials are defined for all scalar fields of characteristic $\neq 2$, and an analysis is given in [Arnold and Kaltofen 2015, Theorem 4.3] by expanding the Toeplitz-plus-Hankel systems in [Lakshman Y. N. and Saunders 1995; Kaltofen and Lee 2003] to symmetric Toeplitz systems of twice the dimension. Both the [Kaltofen and Lee 2003] and the [Arnold and Kaltofen 2015] analyses give early termination algorithms that in the worst case require $2t+2$ interpolation points. In Section 3 we show that with techniques from [Hao, Kaltofen, and Zhi 2016], the analysis in [Arnold and Kaltofen 2015] can be tightened to require $2t+1$ interpolation points and match the Prony early interpolation algorithm.

Sparse polynomial interpolation in Chebyshev basis with the Prony algorithm is based on the cosine identity: $T_n(\frac{1}{2}(y+\frac{1}{y})) = \frac{1}{2}(y^n + \frac{1}{y^n})$, where the variable y represents the transcendental complex function $e^{i\theta}$. The resulting sparse Laurent polynomials are in the subspace spanned by the inverse-invariant basis $y^n + \frac{1}{y^n}$ (2,3,19) [Arnold and Kaltofen 2015, Section 4]. If one chooses the basis polynomial 1 for $n = 0$ (3), sparse polynomials in inverse-invariant basis are also defined for a coefficient field of characteristic 2. By using the evaluation points from [Kaltofen and Yang 2024], whose algorithms can correct errors in the values, we prove an early termination theorem for the inverse-invariant basis which uses $2t+1$ evaluation points for all scalar coefficient fields K (Theorem 2.1). As in [Kaltofen and Lee 2003] vs. [Hao, Kaltofen, and Zhi 2016], the big-Oh in the upper bound of the failure probability is larger: see the Remarks 2.1 and 3.1.

In practice, one sets a threshold value $\zeta \geq 1$ and requires the termination condition to be valid for ζ new evaluation points (48). For dense interpolation, the lowering of the failure probability is analyzed [Kaltofen and Lee 2003, Theorem 1], but for sparse interpolation the case $\zeta \geq 2$ is not analyzed. Nonetheless, look-ahead can avoid sporadic false early termination conditions; see (49), Example 4.1 and [Imamoglu, Kaltofen, and Yang, Z. 2018]. False early termination may also be discovered when the term locator polynomial does not yield term values [Kaltofen and Yang, Z.-H. 2021, Algorithm 2 *Try Prony's algorithm*]. The algorithmic problem arises how to efficiently compute the next candidate for the sparsity t. In the Prony algorithm for standard powers basis the jump-ahead can be handled by the 1969 Berlekamp-Massey algorithm. Our early termination algorithm with thresholds for sparse Chebyshev interpolants can use Heinig's 1983 extension of the Levinson-Durbin algorithm for Toeplitz matrices with singular sections [Heinig 1983], [Heinig and Rost 1984, Part I, Chapter 6]. Both the Berlekamp-Massey and the Heinig algorithms are quadratic-time and linear-space. We describe the Heinig algorithm in Section 4 and note that the asymptotically fast Toeplitz/Hankel solvers seem not applicable in our setting if one restricts to $2t + \zeta$ evaluation points (see Remark 4.3).

2. An early termination theorem for Laurent polynomials in inverse-invariant basis

Let

$$g(y) = \sum_{j=1}^{t} c_j\, M_{\delta_j}(y) \in \mathsf{K}[y, \frac{1}{y}],\ 0 \le \delta_1 < \delta_2 < \cdots < \delta_t,\ c_j \ne 0 \text{ for all } 1 \le j \le t, \tag{2}$$

where K is a field and

$$M_\delta(y) \stackrel{\text{def}}{=} \begin{Bmatrix} y^\delta + \dfrac{1}{y^\delta} & \text{if } \delta > 0, \\ 1 & \text{if } \delta = 0, \end{Bmatrix} \quad \operatorname{iideg}(g) \stackrel{\text{def}}{=} \delta_t, \quad t_g \stackrel{\text{def}}{=} \begin{Bmatrix} 2t-1 & \text{if } \delta_1 = 0, \\ 2t & \text{if } \delta_1 > 0. \end{Bmatrix} \tag{3}$$

The Laurent polynomial g is inverse-invariant: for all $\xi \in \mathsf{K}$, $\xi \ne 0$: $g(\xi) = g(1/\xi)$. The reverse is true for infinite fields K, for instance, the algebraic closure $\bar{\mathsf{K}}$ of K: if $h(y) - h(1/y) \ne 0$ for an $h \in \mathsf{K}[y, \frac{1}{y}]$ then there exists a $\xi \in \bar{\mathsf{K}}, \xi \ne 0$ such that $h(\xi) \ne h(1/\xi)$. Therefore the inverse-invariant Laurent polynomials form a subring of the ring of Laurent polynomials $\mathsf{K}[y, \frac{1}{y}]$ with a vector space basis $(M_n)_{n \ge 0}$. In (2) t is the sparsity in basis $(M_n)_{n\ge 0}$ and t_g is the sparsity in power basis $(y^n)_{n \in \mathbb{Z}}$.

Let $\alpha_i = g(y^i)$; note that $\alpha_i = \alpha_{-i}$ for all $i \in \mathbb{Z}$. We will show that the following Hankel matrix $\mathcal{H}_\tau^{\text{odd}}$ in (4) is non-singular for all τ with $t_g = 1, 2, \ldots, t_g$:

$$\mathcal{H}_\tau^{\text{odd}} = \begin{bmatrix} \alpha_{-(2\tau-3)} & \alpha_{-(2\tau-5)} & \cdots & \alpha_{-1} & \alpha_1 \\ \alpha_{-(2\tau-5)} & \alpha_{-(2\tau-7)} & \cdots & \alpha_1 & \alpha_3 \\ \vdots & \iddots & & \vdots & \vdots \\ \alpha_1 & \cdots & & & \alpha_{2\tau-1} \end{bmatrix} \in \mathsf{K}[y, \frac{1}{y}]^{\tau\times\tau}. \tag{4}$$

We denote the power basis terms of $g(y)$ (see (2)) in degree order:

$$\beta_1 = y^{-\delta_t}, \beta_2 = y^{-\delta_{t-1}}, \ldots, \beta_{t_g-1} = y^{\delta_{t-1}}, \beta_{t_g} = y^{\delta_t}, \quad \deg(\beta_j) = -\deg(\beta_{t_g-(j-1)}), \tag{5}$$

and let

$$\eta_1 = -\delta_t, \eta_2 = -\delta_{t-1}, \ldots, \eta_{t_g-1} = \delta_{t-1}, \eta_{t_g} = \delta_t. \tag{6}$$

Note that a term β_j with $j \le t_g/2$ has a negative degree if $\delta_t > 0$. The matrix $\mathcal{H}_\tau^{\text{odd}}$ has a factorization

$$\mathcal{H}_\tau^{\text{odd}} = \mathcal{B}_\tau C \bar{\mathcal{B}}_\tau,$$

where $\mathcal{B}_\tau$, C, and $\bar{\mathcal{B}}_\tau$ are matrices as shown in Figure 1.

Figure 1: Factorization of the Hankel matrix $\mathcal{H}_\tau^{\mathrm{odd}}$

$$\mathcal{B}_\tau = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_{t_g} \\ \beta_1^3 & \beta_2^3 & \dots & \beta_{t_g}^3 \\ \vdots & \vdots & \vdots & \vdots \\ \beta_1^{2\tau-1} & \beta_2^{2\tau-1} & \dots & \beta_{t_g}^{2\tau-1} \end{bmatrix} \in \mathsf{K}[y, \frac{1}{y}]^{\tau\times t_g},$$

$$C = \begin{bmatrix} c_t & 0 & & \dots & 0 \\ 0 & c_{t-1} & & & 0 \\ \vdots & & \ddots & & \vdots \\ & & & c_{t-1} & \\ 0 & 0 & & \dots & c_t \end{bmatrix} \in \mathsf{K}^{t_g\times t_g},\ C\,[\tfrac{t_g+1}{2}, \tfrac{t_g+1}{2}] = c_1 \text{ if } \delta_1 = 0,$$

$$\bar{\mathcal{B}}_\tau = \begin{bmatrix} \beta_1^{2(1-\tau)} & \dots & \beta_1^{-2} & 1 \\ \beta_2^{2(1-\tau)} & \dots & \beta_2^{-2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{t_g}^{2(1-\tau)} & \dots & \beta_{t_g}^{-2} & 1 \end{bmatrix} \in \mathsf{K}[y, \frac{1}{y}]^{t_g\times\tau}. \quad (7)$$

Theorem 2.1. *The determinant of $\mathcal{H}_\tau^{\mathrm{odd}}$ in (4) is non-zero for $\tau = 1, 2, \dots, t_g$.*

Proof. Let J, L be subsets of $\{1, 2, \dots, t_g\}$ with cardinality τ and $J = \{j_1, j_2, \dots, j_\tau\}$. We use the following notation: for a matrix A we denote by $A_{J,L}$ the submatrix of A consisting rows in J and columns in L. Let

$$I = \{1, 2, \dots, \tau\} \quad \text{and} \quad P_J = \prod_{k=1}^{\tau} C\,[j_k, j_k], \tag{8}$$

where C is the diagonal matrix in Figure 1 and $C\,[j_k, j_k]$ is the value in the j_k-th row and j_k-th column of C. By the Cauchy-Binet formula, we have

$$\begin{aligned} \det(\mathcal{H}_\tau^{\mathrm{odd}}) &= \sum_J \sum_L (\mathcal{B}_\tau)_{I,J} (C)_{J,L} (\bar{\mathcal{B}}_\tau)_{L,I} = \sum_J (\mathcal{B}_\tau)_{I,J} (C)_{J,J} (\bar{\mathcal{B}}_\tau)_{J,I} \\ &= \sum_J P_J\, (\beta_{j_1}\beta_{j_2}\dots\beta_{j_\tau})^{2(1-\tau)+1} \det\Big(\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \beta_{j_1}^2 & \beta_{j_2}^2 & \cdots & \beta_{j_\tau}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{j_1}^{2(\tau-1)} & \beta_{j_2}^{2(\tau-1)} & \dots & \beta_{j_\tau}^{2(\tau-1)} \end{bmatrix} \Big)^2 \\ &= \sum_J P_J\, (\beta_{j_1}\beta_{j_2}\dots\beta_{j_\tau})^{2(1-\tau)+1} \prod_{1\le v<u\le\tau} (\beta_{j_u}^2 - \beta_{j_v}^2)^2. \end{aligned} \tag{9}$$

We fix the order of the indices in J as $j_1 > j_2 > \cdots > j_\tau$ and denote

$$S_J = P_J(\beta_{j_1}\beta_{j_2}\dots\beta_{j_\tau})^{2(1-\tau)+1} \prod_{1\le v<u\le\tau} (\beta_{j_u}^2 - \beta_{j_v}^2)^2. \tag{10}$$

It follows from (5) that $\deg(\beta_{j_i}) > \deg(\beta_{j_{i+1}})$, and therefore the highest degree term of S_J is

$$S_J^{\text{high}} = (\beta_{j_1}\beta_{j_2}\cdots\beta_{j_\tau})^{3-2\tau}\beta_{j_1}^{4(\tau-1)}\beta_{j_2}^{4(\tau-2)}\cdots\beta_{j_{\tau-1}}^{4}. \tag{11}$$

Let $D_J = \deg(S_J^{\text{high}})$, then by (5) and (6)

$$D_J = \sum_{i=1}^{\tau}(3-2\tau)\eta_{j_i} + \sum_{i=1}^{\tau-1}4(\tau-i)\eta_{j_i} = \sum_{i=1}^{\tau}(2\tau+3-4i)\eta_{j_i}. \tag{12}$$

We maximize D_J by choosing the largest positive term degrees η_{j_i} for the largest positive integers $(2\tau+3-4i)$ and the smallest negative term degrees η_{j_i} for the smallest negative integers $(2\tau+3-4i)$, that is, let

$$J_{\max} = \begin{cases} \{t_g, (t_g-1), \ldots, (t_g-\frac{\tau}{2}+1), \frac{\tau}{2}, (\frac{\tau}{2}-1), \ldots, 1\}, & \text{if } \tau \text{ is even,} \\ \{t_g, (t_g-1), \ldots, (t_g-\frac{\tau-1}{2}), (\frac{\tau-1}{2}), (\frac{\tau-1}{2}-1), \ldots, 1\}, & \text{if } \tau \text{ is odd.} \end{cases} \tag{13}$$

We claim that $D_{J_{\max}} > D_J$ for all $J \neq J_{\max}$. Therefore, $S_{J_{\max}}^{\text{high}}$ is the leading term of $\det(\mathcal{H}_\tau^{\text{odd}})$, which concludes the theorem.

Proof of the claim: Since $t_g \geq j_1 > j_2 > \cdots > j_\tau \geq 1$ and $j_1, j_2, \ldots, j_\tau$ are all integers, we have

$$\tau - i + 1 \leq j_i \leq t_g - i + 1, \tag{14}$$

and

$$\eta_{\tau-i+1} \leq \eta_{j_i} \leq \eta_{t_g-i+1}. \tag{15}$$

If τ is odd,

$$D_{J_{\max}} - D_J = \sum_{i=1}^{(\tau+1)/2}(2\tau+3-4i)(\eta_{t_g-(i-1)} - \eta_{j_i}) + \sum_{i=(\tau+3)/2}^{\tau}(2\tau+3-4i)(\eta_{\tau-i+1} - \eta_{j_i}) \tag{16}$$

which is non-negative by (15). If $J \neq J_{\max}$, then there is at least one positive summand in (16) because $2\tau+3-4i \neq 0$. If τ is even.

$$D_{J_{\max}} - D_J = \sum_{i=1}^{\tau/2}(2\tau+3-4i)(\eta_{t_g-(i-1)} - \eta_{j_i}) + \sum_{i=\tau/2+1}^{\tau}(2\tau+3-4i)(\eta_{\tau-i+1} - \eta_{j_i}), \tag{17}$$

which is non-negative by (15) and the equality is attained only if $J = J_{\max}$ because $2\tau+3-4i \neq 0$. $\square$

Remark 2.1. We have the upper bound for the inverse-invariant degree (as defined in (3))

$$\operatorname{iideg}(G) \leq t_g(t_g+1)\left(\frac{1}{3}t_g\,\delta_t + \frac{1}{6}\delta_t - \frac{1}{24}t_g^2 + \frac{1}{8}t_g - \frac{1}{12}\right) \quad \text{with} \quad G = \prod_{\tau=1}^{t_g}\det(\mathcal{H}_\tau^{\text{odd}}),$$

for the interpolant $g(y)$ in (2) and t_g in (3). We note that for a finite set S with $|S|$ scalar elements and $0 \notin S$, the probability $\operatorname{Prob}(G(\omega) = 0 \mid \omega \in S) \leq \operatorname{iideg}(G)/|S|$ provided that $\forall \omega \in S \Longrightarrow 1/\omega \notin S$. $\square$

3. An early termination theorem for Chebyshev-1 basis over fields of characteristic $\neq 2$

If the coefficient field K does not contain $\mathbb{Z}_2$, which are the integers modulo 2, as is the case for the recurrence that defines Chebyshev-1 polynomials, one can use the standard Prony evaluations $a_i = g(\omega^i)$ for some $\omega \in \mathsf{K}$. The entries of the matrix $\mathcal{H}_\tau^{\mathrm{odd}}$ in (18) are $\alpha_{2i+1} = g(y^{2i+1})$ for $i = -\tau+1, -\tau+2, \ldots, \tau-1$, where $g(y)$ is the Laurent polynomial in (2). In this section, we consider the symmetric Hankel matrix with entries $\alpha_i = g(y^i)$ for $i = -\tau+1, -\tau+2, \ldots, \tau-1$. Let

$$\mathcal{H}_\tau = \begin{bmatrix} \alpha_{1-\tau} & \alpha_{2-\tau} & \cdots & \alpha_{-1} & \alpha_0 \\ \alpha_{2-\tau} & \alpha_{3-\tau} & \iddots & \alpha_0 & \alpha_1 \\ \vdots & \iddots & \iddots & \vdots & \vdots \\ \alpha_{-1} & \iddots & & & \alpha_{\tau-2} \\ \alpha_0 & \cdots & & \alpha_{\tau-2} & \alpha_{\tau-1} \end{bmatrix} \in \mathsf{K}[y, \frac{1}{y}]^{\tau\times\tau}, \quad \tau \geq 1. \tag{18}$$

We have for f in (1) and g in (2):

$$f\Big(\frac{1}{2}\big(y^{\delta_j} + \frac{1}{y^{\delta_j}}\big)\Big) = g(y), \quad \text{with } \widehat{c}_j = 2c_j \text{ except } \widehat{c}_1 = c_1 \text{ if } \delta_1 = 0. \tag{19}$$

Theorem 4.3 in [Arnold and Kaltofen 2015] states the following:

(a) If $2 \leq \tau \leq t_g$ and τ is even, then $\det(\mathcal{H}_\tau) \neq 0$.

(b) If $1 \leq \tau \leq t_g$, τ is odd, and $\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i \neq 0$, then $\det(\mathcal{H}_\tau) \neq 0$.

By the relation between $\widehat{c}_i$ and c_i in (19), we write the sum $\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i$ as

$$\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i = \left\{ \begin{array}{ll} \displaystyle\sum_{i=1}^{t-(\tau-1)/2} 2c_i, & \text{if } \delta_1 > 0, \\ c_1 + \displaystyle\sum_{i=2}^{t-(\tau-1)/2} 2c_i, & \text{if } \delta_1 = 0. \end{array} \right\} \tag{20}$$

With the technique in the proof of [Hao, Kaltofen, and Zhi 2016, Theorem 4.1], we show that in the Case (b), where $1 \leq \tau \leq t_g$ and τ is odd, the condition $\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i \neq 0$ can be removed for all $3 \leq \tau \leq t_g$, except in one case when the determinant of $\mathcal{H}_\tau$ is always $= 0$ (Theorem 3.1, Item (iii), when τ is odd).

Theorem 3.1. *The matrices $\mathcal{H}_\tau$ in (18) satisfy the following singularity properties.*

(i) *Let K be a field of characteristic $\neq 2$ and let τ be an integer with $2 \leq \tau \leq t_g$. Then* $\det(\mathcal{H}_\tau) \neq 0$.

(ii) *Let K be a field of characteristic $= 2$ and let τ be an integer with $1 \leq \tau \leq t_g$ and let $\delta_1 = 0$. Then* $\det(\mathcal{H}_\tau) \neq 0$.

(iii) *Let* K *be a field of characteristic* $= 2$ *and let* τ *be an integer with* $1 \le \tau \le t_g$ *and let* $\delta_1 > 0$. *Then* $\det(\mathcal{H}_\tau) \neq 0$ *if* τ *is even and* $\det(\mathcal{H}_\tau) = 0$ *if* τ *is odd.*

(iv) *Let* K *be an arbitrary field and let* τ *be an integer with* $\tau > t_g$. *Then* $\det(\mathcal{H}_\tau) = 0$.

Proof. Case (i). According to [Arnold and Kaltofen 2015, Theorem 4.3], it remains to prove that for an odd integer τ with $3 \le \tau \le t_g$, $\det(\mathcal{H}_\tau) \neq 0$ even if $\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i = 0$ (20). By [Arnold and Kaltofen 2015, Eq. (24)],

$$\det(\mathcal{H}_\tau) = \sum_{J=\{j_1,\dots,j_\tau\}} P_J\,(\beta_{j_1}\cdots\beta_{j_\tau})^{1-\tau} \prod_{1\le v<u\le\tau} (\beta_{j_u}-\beta_{j_v})^2, \tag{21}$$

where the β's are defined in (5) and $P_J = \prod_{\sigma=1}^{\tau} C\,[j_\sigma, j_\sigma]$ is defined as in (8):

$$C = \text{diag-matrix}(c_t, c_{t-1}, \dots, c_{t-1}, c_t) \in \mathsf{K}^{t_g\times t_g} \text{ with } C\,[\tfrac{t_g+1}{2}, \tfrac{t_g+1}{2}] = c_1 \text{ if } \delta_1 = 0.$$

Note that the diagonal of C is the coefficient vector of g in (19). We order of the elements in J as $j_1 > j_2 > \cdots > j_\tau$. Let

$$S_J = P_J\,(\beta_{j_1}\cdots\beta_{j_\tau})^{1-\tau} \prod_{1\le v<u\le\tau} (\beta_{j_u}-\beta_{j_v})^2. \tag{22}$$

For odd τ, the highest degree term of S_J is

$$\begin{aligned} S_J^{\text{high}} &= (\beta_{j_1}\cdots\beta_{j_\tau})^{1-\tau} \beta_{j_1}^{2(\tau-1)}\beta_{j_2}^{2(\tau-2)}\cdots\beta_{j_{(\tau+1)/2}}^{(\tau-1)}\cdots\beta_{j_{\tau-2}}^{4}\beta_{j_{\tau-1}}^{2} \\ &= \beta_{j_1}^{\tau-1}\beta_{j_2}^{\tau-3}\cdots\beta_{j_{(\tau-1)/2}}^{2}\beta_{j_{(\tau+1)/2}}^{0}\beta_{j_{(\tau+3)/2}}^{-2}\cdots\beta_{j_{\tau-1}}^{3-\tau}\beta_{j_\tau}^{1-\tau}. \end{aligned} \tag{23}$$

Then $\deg(S_J^{\text{high}}) = \sum_{i=1}^{\tau}(\tau+1-2i)\eta_{j_i}$ where $\eta_{j_i} = \deg(\beta_{j_i})$. Let

$$S_J^{(2)} = \frac{S_J^{\text{high}}\beta_{j_{(\tau+1)/2}}}{\beta_{j_{(\tau-1)/2}}}, \quad S_J^{(3)} = \frac{S_J^{\text{high}}\beta_{j_{(\tau+3)/2}}}{\beta_{j_{(\tau+1)/2}}}, \tag{24}$$

and

$$\begin{aligned} J' &= \big\{t_g, t_g-1, \dots, (t_g - \frac{\tau-1}{2}+1), (t_g-\frac{\tau-1}{2}), (\frac{\tau-1}{2}), (\frac{\tau-1}{2}-1), \dots, 1\big\}, \\ J'' &= \big\{t_g, t_g-1, \dots, (t_g - \frac{\tau-1}{2}+1), (\frac{\tau+1}{2}), (\frac{\tau-1}{2}), (\frac{\tau-1}{2}-1), \dots, 1\big\}. \end{aligned} \tag{25}$$

By the degrees of β_{j_i} in y (5), one can verify that

$$\deg(S_{J'}^{(2)}) = \deg(S_{J''}^{(3)}) = 3\delta_{t-(\tau-1)/2+1} + \delta_{t-(\tau-1)/2} + \sum_{i=1}^{(\tau-1)/2-1} 2(\tau+1-2i)\delta_{t-i+1}. \tag{26}$$

Moreover, since $C\,[i,i] = C\,[t_g+1-i, t_g+1-i]$ for all $1 \le i \le t_g$, the coefficient of $S_{J'}^{(2)}$ equals to that of $S_{J''}^{(3)}$:

$$P_{J'} = C\,[\tfrac{\tau+1}{2}, \tfrac{\tau+1}{2}] \prod_{i=1}^{(\tau-1)/2} (C\,[i,i])^2 = P_{J''}. \tag{27}$$

Let $D_2 = \deg(S^{(2)}_{J'})$. We prove that if $\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i = 0$ then y^{D_2} is the leading term of $\det(\mathcal{H}_\tau)$, and the coefficient of y^{D_2} is $-(2P_{J'} + 2P_{J''}) = -4P_{J'}$, which is non-zero since the characteristic of K is $\neq 2$.

Claim 1: *Let $\widehat{J'} = J' \setminus \{t_g - \frac{\tau-1}{2}\}$ (25) and $J = \{j_1, \ldots, j_\tau\}$ with $j_1 > j_2 > \cdots > j_\tau$. If $J \setminus \{j_{(\tau+1)/2}\} \neq \widehat{J'}$, then $\deg(S_J) < \deg(S^{(2)}_{J'})$.*

Proof of Claim 1. Since $J \setminus \{j_{(\tau+1)/2}\} \neq \widehat{J'}$, there exists an index σ such that either $1 \leq \sigma \leq (\tau-1)/2$ with $j_\sigma \neq t_g - \sigma + 1$, or $(\tau+3)/2 \leq \sigma \leq \tau$ with $j_\sigma \neq \tau - \sigma + 1$. Assume that the first case occurs. Let

$$\sigma_{\min} = \min\{\sigma \mid 1 \leq \sigma \leq (\tau-1)/2,\ j_\sigma \neq t_g - \sigma + 1\}.$$

Since $j_\sigma \leq t_g - \sigma + 1$ (14) and $j_{\sigma+1} < j_\sigma$, by the definition of $\sigma_{\min}$ we have $j_{\sigma_{\min}} < t_g - \sigma_{\min} + 1$. Therefore

$$j_\kappa \leq t_g - \kappa \text{ for all } \sigma_{\min} \leq \kappa \leq (\tau-1)/2,$$

which implies that for the η's in (6) that

$$\eta_{j_\kappa} \leq \eta_{t_g-\kappa} \text{ for all } \sigma_{\min} \leq \kappa \leq (\tau-1)/2.$$

In particular, taking $\kappa = (\tau-1)/2$, we have $\eta_{j_{(\tau-1)/2}} \leq \eta_{t_g-(\tau-1)/2}$. Therefore, it follows from (15) that

$$\begin{aligned}
&\deg(S^{\text{high}}_J) - \deg(S^{\text{high}}_{J'}) \\
&\quad= \sum_{\sigma=1}^{(\tau-1)/2} (\tau+1-2\sigma)(\eta_{j_\sigma} - \eta_{t_g-\sigma+1}) + \sum_{\sigma=(\tau+3)/2}^{\tau} (\tau+1-2\sigma)(\eta_{j_\sigma} - \eta_{\tau-\sigma+1}) \\
&\quad\leq\ 0 + 2\left(\eta_{j_{(\tau-1)/2}} - \eta_{t_g-(\tau-1)/2+1}\right) + 0 \leq 2\left(\eta_{t_g-(\tau-1)/2} - \eta_{t_g-(\tau-1)/2+1}\right),
\end{aligned}$$

Then

$$\begin{aligned}
&\deg(S_J) - \deg(S^{(2)}_{J'}) \\
&\quad= \deg(S^{\text{high}}_J) - \left(\deg(S^{\text{high}}_{J'}) + \deg(\beta_{t_g-(\tau-1)/2}) - \deg(\beta_{t_g-(\tau-1)/2+1})\right) \\
&\quad\leq\ 2\left(\eta_{t_g-(\tau-1)/2} - \eta_{t_g-(\tau-1)/2+1}\right) - \left(\eta_{t_g-(\tau-1)/2} - \eta_{t_g-(\tau-1)/2+1}\right) < 0.
\end{aligned}$$

For the second case, where there exists an index σ such that $(\tau+3)/2 \leq \sigma \leq \tau$ with $j_\sigma \neq \tau - \sigma + 1$, we let

$$\sigma_{\max} = \max\{\sigma \mid (\tau+3)/2 \leq \sigma \leq \tau, j_\sigma \neq \tau - \sigma + 1\}.$$

Since $j_\sigma \geq \tau - \sigma + 1$ (14) and $j_\sigma > j_{\sigma+1}$, we have $j_\sigma > \tau - \sigma + 1$ for all $(\tau+3)/2 \leq \sigma \leq \sigma_{\max}$, which leads to

$$\eta_{j_\sigma} \geq \eta_{\tau-\sigma+2} \text{ for all } (\tau+3)/2 \leq \sigma \leq \sigma_{\max}.$$

In particular, $\eta_{j_{(\tau+3)/2}} \geq \eta_{\tau-(\tau+3)/2+2}$. Therefore,

$$\deg(S^{\text{high}}_J) - \deg(S^{\text{high}}_{J''}) \leq -2\left(\eta_{j_{(\tau+3)/2}} - \eta_{\tau-(\tau+3)/2+1}\right) \leq -2\left(\eta_{\tau-(\tau+3)/2+2} - \eta_{(\tau-1)/2}\right),$$

where the first inequality is due to (15). Then

$$\begin{aligned}
&\deg(S_J) - \deg(S_{J''}^{(3)}) \\
&\quad = \deg(S_J^{\text{high}}) - \left(\deg(S_{J''}^{\text{high}}) + \deg(\beta_{(\tau-1)/2}) - \deg(\beta_{(\tau+1)/2})\right) \\
&\quad \le -2\,(\eta_{\tau-(\tau+3)/2+2} - \eta_{(\tau-1)/2}) - (\eta_{(\tau-1)/2} - \eta_{(\tau+1)/2}) < 0.
\end{aligned}$$

The proof of Claim 1 is concluded. It remains to consider the summand S_J in (22) where $J \setminus \{j_{(\tau+1)/2}\} = \widehat{J}'$.

Claim 2: *Let $\widehat{J}' = J' \setminus \{t_g - \frac{\tau-1}{2}\}$ (25), and let $J = \{j_1, \ldots, j_\tau\}$ with $j_1 > j_2 > \cdots > j_\tau$ and $J \setminus \{j_{(\tau+1)/2}\} = \widehat{J}'$. Let $W_J = \beta_{j_1}^{w_1} \cdots \beta_{j_\tau}^{w_\tau}$ be a term of S_J with W_J not equal $S_{J'}^{(2)}$ and W_J not equal $S_{J''}^{(3)}$ (24), as a monomial in the variables $\beta_1, \ldots, \beta_{t_g}$. Then $\deg_y(W_J) \ge \deg_y(S_{J'}^{(2)}) = \deg_y(S_{J''}^{(3)})$ (26) $\Longrightarrow w_{(\tau+1)/2} = 0$.*

Proof of Claim 2. We prove the claim by showing that

(1) if $w_{(\tau+1)/2} > 0$, then $\deg_y(S_{J'}^{(2)}) \ge \deg_y(W_J)$, and $\deg_y(S_{J'}^{(2)}) = \deg_y(W_J) \Longleftrightarrow W_J$ is equal to $S_{J'}^{(2)}$ as a monomial in the variables $\beta_1, \ldots, \beta_{t_g}$;

(2) if $w_{(\tau+1)/2} < 0$, then $\deg_y(S_{J''}^{(3)}) \ge \deg_y(W_J)$, and $\deg(S_{J''}^{(3)}) = \deg(W_J) \Longleftrightarrow W_J$ is equal to $S_{J''}^{(3)}$ as a monomial in the variables $\beta_1, \ldots, \beta_{t_g}$.

Note that $(\beta_{j_1} \cdots \beta_{j_\tau})^{\tau-1} W_J$ is a term of the product $\prod_{1 \le v < u \le \tau} (\beta_{j_u} - \beta_{j_v})^2$. As shown in the proof of [Hao, Kaltofen, and Zhi 2016, Eq. (45)], the following two inequalities hold for all $1 \le \mu \le \tau$:

$$\sum_{\sigma=1}^{\mu} (w_\sigma + \tau - 1) \le \sum_{\sigma=1}^{\mu} 2\,(\tau - \sigma),$$

$$\sum_{\sigma=1}^{\mu} (w_{\tau-\sigma+1} + \tau - 1) \ge \sum_{\sigma=1}^{\mu} 2\,(\sigma - 1),$$

which are respectively equivalent to

$$\sum_{\sigma=1}^{\mu} w_\sigma \le \sum_{\sigma=1}^{\mu} (\tau + 1 - 2\sigma), \tag{28}$$

$$\sum_{\sigma=1}^{\mu} w_{\tau-\sigma+1} \ge \sum_{\sigma=1}^{\mu} (-\tau - 1 + 2\sigma). \tag{29}$$

Taking $\mu = (\tau+1)/2$ in (28) and $\mu = (\tau-1)/2$ in (29), we obtain

$$w_{(\tau+1)/2} + \sum_{\sigma=1}^{(\tau-1)/2} (w_\sigma - w_{\tau-\sigma+1}) \le \sum_{\sigma=1}^{(\tau-1)/2} 2(\tau + 1 - 2\sigma). \tag{30}$$

Since $j_1 > j_2 > \cdots > j_\tau$ and $J \setminus \{j_{(\tau+1)/2}\} = J' \setminus \{t_g - (\tau-1)/2\}$ (25),

$$\deg_y(S_{J'}^{\text{high}}) - \deg_y(W_J) \overset{(5)}{=} \Big(\sum_{\sigma=1}^{(\tau-1)/2} \big(2(\tau+1-2\sigma) - (w_\sigma - w_{\tau-\sigma+1})\big)\delta_{t-\sigma+1} \Big) - w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}}. \tag{31}$$

Then using the right-hand side of (31) and combining (44) in Lemma 3.2 and the inequalities (30) and (15) we have

$$\begin{aligned}
&\deg_y(S_{J'}^{\text{high}}) - \deg_y(W_J) \\
&\overset{(44)}{\geq} \Big(\sum_{\sigma=1}^{(\tau-1)/2} \big(2(\tau+1-2\sigma) - (w_\sigma - w_{\tau-\sigma+1})\big)\delta_{t-(\tau-1)/2+1} \Big) - w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}} \\
&\overset{(30)}{\geq} w_{(\tau+1)/2}\delta_{t-(\tau-1)/2+1} - w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}} \\
&\overset{(15)}{\geq} w_{(\tau+1)/2}\delta_{t-(\tau-1)/2+1} - w_{(\tau+1)/2}\eta_{t_g-(\tau+1)/2+1} \overset{(6)}{=} w_{(\tau+1)/2}(\delta_{t-(\tau-1)/2+1} - \delta_{t-(\tau-1)/2}).
\end{aligned} \tag{32}$$

As $\delta_t > \delta_{t-1} > \cdots > \delta_{t-(\tau-1)/2+1}$, by Lemma 3.2, the first equality in (32) holds if and only if the coefficient of $\delta_{t-\sigma+1}$ in (31) is zero for $\sigma = 1, 2, \ldots, (\tau-1)/2 - 1$, that is,

$$\forall \sigma, 1 \leq \sigma \leq (\tau-1)/2 - 1\colon w_\sigma - w_{\tau-\sigma+1} = 2(\tau+1-2\sigma). \tag{33}$$

The second equality in (32) holds if and only if

$$w_{(\tau+1)/2} + \sum_{\sigma=1}^{(\tau-1)/2} (w_\sigma - w_{\tau-\sigma+1}) = \sum_{\sigma=1}^{(\tau-1)/2} 2(\tau+1-2\sigma). \tag{34}$$

The third equality in (32) holds if and only if

$$j_{(\tau+1)/2} = t_g - (\tau+1)/2 + 1. \tag{35}$$

Assume that $w_{(\tau+1)/2} > 0$, then $w_{(\tau+1)/2} \geq 1$. By (32), we have

$$\begin{aligned}
&\deg_y(S_{J'}^{(2)}) - \deg_y(W_J) \\
&\quad = \deg_y(S_{J'}^{\text{high}}) - \deg_y(W_J) + \eta_{t_g-(\tau-1)/2} - \eta_{t_g-(\tau-1)/2+1} \\
&\quad \geq w_{(\tau+1)/2}(\delta_{t-(\tau-1)/2+1} - \delta_{t-(\tau-1)/2}) + \delta_{t-(\tau-1)/2} - \delta_{t-(\tau-1)/2+1} \geq 0.
\end{aligned} \tag{36}$$

If $\deg_y(S_{J'}^{(2)}) - \deg_y(W_J) = 0$, then $w_{(\tau+1)/2} = 1$ and all equalities in (32) must hold. Therefore, the equations (33), (34), and (35) hold. Together with (28) and (29), we have $w_{(\tau-1)/2} - w_{(\tau+3)/2} = 3$ and $w_\sigma = -w_{\tau-\sigma+1} = \tau+1-2\sigma$ for all $1 \leq \sigma \leq (\tau-1)/2 - 1$. On the other hand, $\sum_{\sigma=1}^{\tau} w_\sigma = 0$, which leads to $w_{(\tau-1)/2} + w_{(\tau+3)/2} = -w_{(\tau+1)/2} = -1$. Therefore, $w_{(\tau-1)/2} = 1$ and $w_{(\tau+3)/2} = -2$. In that case, W_J is the same as $S_{J'}^{(2)}$ as a monomial in $\beta_1, \ldots, \beta_{t_g}$, which concludes Claim 2, Part (1).

We now prove Claim 2, Part (2), which assumes that $w_{(\tau+1)/2} < 0$, namely, $w_{(\tau+1)/2} \leq -1$. Taking $\mu = (\tau-1)/2$ in (28) and $\mu = (\tau+1)/2$ in (29), we get

$$-w_{(\tau+1)/2} + \sum_{\sigma=1}^{(\tau-1)/2} (w_\sigma - w_{\tau-\sigma+1}) \leq \sum_{\sigma=1}^{(\tau-1)/2} 2(\tau+1-2\sigma). \tag{37}$$

Since $-w_{(\tau+1)/2} > 0$ and, by (15), $\eta_{j_{(\tau+1)/2}} \geq \eta_{\tau-(\tau+1)/2+1} = -\delta_{t-(\tau-1)/2}$ (see (6)), we have

$$-w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}} \geq w_{(\tau+1)/2}\delta_{t-(\tau-1)/2}. \tag{38}$$

Again, because $j_1 > j_2 > \cdots > j_\tau$ and $J \setminus \{j_{(\tau+1)/2}\} = J' \setminus \{t_g - (\tau-1)/2\} = J'' \setminus \{(\tau+1)/2\}$ (see (25)), we have

$$\deg_y(S^{\text{high}}_{J''}) - \deg_y(W_J) \overset{(5)}{=} \Big(\sum_{\sigma=1}^{(\tau-1)/2} (2(\tau+1-2\sigma) - (w_\sigma - w_{\tau-\sigma+1}))\delta_{t-\sigma+1} \Big) - w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}}.$$

Then it follows from (44) in Lemma 3.2 and the inequalities (37) and (38) that

$$\begin{aligned}
&\deg_y(S^{\text{high}}_{J''}) - \deg_y(W_J) \\
&\overset{(44)}{\geq} \Big(\sum_{\sigma=1}^{(\tau-1)/2} (2(\tau+1-2\sigma)(w_\sigma - w_{\tau-\sigma+1}))\delta_{t-(\tau-1)/2+1} \Big) - w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}} \\
&\overset{(37)}{\geq} -w_{(\tau+1)/2}\delta_{t-(\tau-1)/2+1} - w_{(\tau+1)/2}\eta_{j_{(\tau+1)/2}} \\
&\overset{(38)}{\geq} w_{(\tau+1)/2}(\delta_{t-(\tau-1)/2} - \delta_{t-(\tau-1)/2+1}).
\end{aligned} \tag{39}$$

By the same argument below (32), the first equality of (39) holds if and only if (33) holds. The second equality of (39) holds if and only if

$$-w_{(\tau+1)/2} + \sum_{\sigma=1}^{(\tau-1)/2} (w_\sigma - w_{\tau-\sigma+1}) = \sum_{\sigma=1}^{(\tau-1)/2} 2(\tau+1-2\sigma). \tag{40}$$

By (15), The third equality of (39) holds if and only if

$$j_{(\tau+1)/2} = (\tau+1)/2. \tag{41}$$

Then using (39), we have

$$\begin{aligned}
&\deg_y(S^{(3)}_{J''}) - \deg_y(W_J) \\
&\quad = \deg_y(S^{\text{high}}_{J''}) - \deg_y(W_J) + \eta_{(\tau-1)/2} - \eta_{(\tau+1)/2} \\
&\quad \geq w_{(\tau+1)/2}(\delta_{t-(\tau-1)/2} - \delta_{t-(\tau-1)/2+1}) - \delta_{t-(\tau-1)/2+1} + \delta_{t-(\tau+1)/2+1} \geq 0.
\end{aligned} \tag{42}$$

If $\deg_y(S^{(3)}_{J''}) = \deg_y(W_J)$, then $w_{(\tau+1)/2} = -1$ and all equalities in (39) must hold. Hence (33), (40), and (41) hold, which leads to $w_{(\tau-1)/2} - w_{(\tau+3)/2} = 3$. Moreover, $w_{(\tau-1)/2} + w_{(\tau+3)/2} = -w_{(\tau+1)/2} = 1$. Then $w_{(\tau-1)/2} = 2$ and $w_{(\tau+3)/2} = -1$, and therefore W_J is equal to $S^{(3)}_{J''}$ as a monomial in $\beta_1, \ldots, \beta_{t_g}$, which concludes the proof of Claim 2, Part (2) and finishes the proof of Claim 2.

Let $J = \{j_1, \ldots, j_\tau\}$ and $W_J = \beta_{j_1}^{w_1}\beta_{j_2}^{w_2}\cdots\beta_{j_\tau}^{w_\tau}$ be a term in the expansion of $\det(\mathcal{H}_\tau)$ (21). Assume that $\deg_y(W_J) \geq \deg_y(S_{J'}^{(2)})$ and W_J is not equal to $S_{J'}^{(2)}$ or $S_{J''}^{(3)}$ as a monomial in the variables $\beta_1, \ldots, \beta_{t_g}$. By Claim 1 and Claim 2, we have $J \setminus \{j_{(\tau+1)/2}\} = \widehat{J}'$ and $w_{(\tau+1)/2} = 0$. Let $y^{D'} = W_J$, namely, $D' = \sum_{\sigma=1}^{\tau} w_\sigma \eta_{j_\sigma}$, by (14) we have

$$(\tau+1)/2 \leq j_{(\tau+1)/2} \leq t_g - (\tau-1)/2.$$

Then the coefficient of $y^{D'}$ is

$$Q_{D'} = \sum_{J\setminus j_{(\tau+1)/2}=\widehat{J}'} P_J = \Big(\prod_{\sigma=1,\sigma\neq(\tau+1)/2}^{\tau} C\,[j_\sigma, j_\sigma]\Big) \sum_{\sigma=(\tau+1)/2}^{t_g-(\tau-1)/2} C\,[j_\sigma, j_\sigma].$$

We write $Q_{D'}$ in $c_1, \ldots, c_t$:

$$Q_{D'} = \left\{\begin{array}{ll} c_t^2 c_{t-1}^2 \cdots c_{t-(\tau-1)/2+1}^2 \displaystyle\sum_{i=1}^{t-(\tau-1)/2} 2c_i, & \text{if } \delta_1 > 0, \\ c_t^2 c_{t-1}^2 \cdots c_{t-(\tau-1)/2+1}^2 \Big(c_1 + \displaystyle\sum_{i=2}^{t-(\tau-1)/2} 2c_i\Big), & \text{if } \delta_1 = 0. \end{array}\right\} \tag{43}$$

Therefore, by the assumption that $\sum_{i=1}^{t-(\tau-1)/2} \widehat{c}_i = 0$ and (20), the coefficient $Q_{D'} = 0$, which implies that the degree of $\det(\mathcal{H}_\tau)$ in y is $\leq \deg_y(S_{J'}^{(2)}) = D_2$ (see (26)). Since the coefficient of y^{D_2} is $-4P_{J'} \neq 0$ (27), y^{D_2} is the leading term of $\det(\mathcal{H}_\tau)$.

Case (ii). The proof in [Arnold and Kaltofen 2015, Theorem 4.3] for the case where τ is an even integer with $2 \leq \tau \leq t_g$, showing that $\det(\mathcal{H}_\tau) \neq 0$, does not depend on the characteristic of K, and applies when the characteristic of K is $= 2$. Now let τ be an odd integer with $1 \leq \tau \leq t_g$. The determinant $\det(\mathcal{H}_\tau)$ has a term $y^{D'} = S_{J'}^{\text{high}}$ (see (23) and (25)) whose coefficient is $Q_{D'}$ (43). Since $\delta_1 = 0$, $(Q_{D'} \mod 2) = c_1 \prod_{i=t-(\tau-1)/2+1}^{t} c_i^2 \neq 0$, we conclude that $\det(\mathcal{H}_\tau) \neq 0$.

Case (iii). If τ is an even integer with $2 \leq \tau \leq t_g$, as in Case (ii), $\det(\mathcal{H}_\tau) \neq 0$ because the proof in [Arnold and Kaltofen 2015, Theorem 4.3] for even τ remains valid when the characteristic of K is $= 2$. Now let τ be an odd integer with $1 \leq \tau \leq t_g$.

Let $J = \{j_1, j_2, \ldots, j_\tau\}$ with $j_1 > j_2 > \cdots > j_\tau$ and let $W_J = \beta_{j_1}^{w_1}\beta_{j_2}^{w_2}\cdots\beta_{j_\tau}^{w_\tau}$ be a term in the expansion of $\det(\mathcal{H}_\tau)$. We set $L = \{\ell_\tau, \ell_{\tau-1}, \ldots, \ell_2, \ell_1\}$ with $\ell_\sigma = t_g + 1 - j_\sigma$, then $\ell_\tau > \ell_{\tau-1} > \cdots > \ell_2 > \ell_1$. Let $D_J = \deg_y(W_J)$ and $W_L = \beta_{\ell_1}^{-w_1}\beta_{\ell_2}^{-w_2}\cdots\beta_{\ell_\tau}^{-w_\tau}$, we have

$$D_J = \sum_{\sigma=1}^{\tau} w_\sigma \eta_{j_\sigma} = \sum_{\sigma=1}^{\tau} (-w_\sigma)(-\eta_{j_\sigma}) = \sum_{\sigma=1}^{\tau} (-w_\sigma)\eta_{t_g+1-j_\sigma} = \sum_{\sigma=1}^{\tau} (-w_\sigma)\eta_{\ell_\sigma},$$

which implies that W_J and W_L have the same degree in y. Moreover, W_L is also a term in the expansion of $\det(\mathcal{H}_\tau)$, and W_L is uniquely determined by W_J from our setting. The coefficients of W_J and W_L are also equal:

$$P_J = \prod_{\sigma=1}^{\tau} C\,[j_\sigma, j_\sigma] = \prod_{\sigma=1}^{\tau} C\,[t_g+1-j_\sigma, t_g+1-j_\sigma] = \prod_{\sigma=1}^{\tau} C\,[\ell_\sigma, \ell_\sigma] = P_L.$$

Finally, $\delta_1 > 0$ implies $t_g = 2t$, and thus $\ell_{(\tau+1)/2} = t_g + 1 - j_{(\tau+1)/2} \neq j_{(\tau+1)/2}$, which shows that L and J are two distinct sets. Therefore, W_J and W_L are two distinct terms in the variables $\beta_1, \beta_2, \ldots, \beta_{t_g}$. Consequently, $P_J W_J + P_L W_L = 2P_J y^{D_J}$, which concludes that $\det(\mathcal{H}_\tau) = 0$ when K is a field of characteristic $= 2$.

Case (iv). $\det(\mathcal{H}_\tau) = 0$ for all $\tau > t_g$ because $(\alpha_i)_{i\in\mathbb{Z}}$ is linearly generated by a polynomial of degree t_g. $\square$

Lemma 3.2. *Let $k \geq 2$, $d_1 > d_2 > \cdots > d_k > 0$ and $\sum_{i=1}^{\mu} w_i \geq \sum_{i=1}^{\mu} w_i'$ for all $2 \leq \mu \leq k$, then*

$$\sum_{i=1}^{\mu}(w_i - w_i')d_i \geq \sum_{i=1}^{\mu}(w_i - w_i')d_\mu, \ \forall \mu : 1 \leq \mu \leq k, \tag{44}$$

and the equality holds if and only if $w_i = w_i'$ for all $1 \leq i \leq \mu - 1$. In particular, $\sum_{i=1}^{\mu} w_i d_i = \sum_{i=1}^{\mu} w_i' d_i$ if and only if $w_i = w_i'$ for all $1 \leq i \leq \mu$.

Proof. We prove the lemma by induction on μ. For $\mu = 2$ we have

$$(w_1 - w_1')d_1 + (w_2 - w_2')d_2 \geq (w_1 - w_1' + w_2 - w_2')d_2$$

and the equality holds if and only if $w_1 = w_1'$, because $d_1 > d_2 > 0$. Assume that (44) holds for $\mu - 1$, then

$$\begin{aligned}
\sum_{i=1}^{\mu}(w_i - w_i')d_i &= w_\mu d_\mu - w_\mu' d_\mu + \sum_{i=1}^{\mu-1}(w_i - w_i')d_i \\
&\geq w_\mu d_\mu - w_\mu' d_\mu + \sum_{i=1}^{\mu-1}(w_i - w_i')d_{\mu-1} \\
&\geq \sum_{i=1}^{\mu}(w_i - w_i')d_\mu,
\end{aligned}$$

where the equality in the second line holds if and only if $w_i = w_i'$ for $1 \leq i \leq \mu - 2$ by the induction hypothesis. Assuming that $w_i = w_i'$ for $1 \leq i \leq \mu - 2$, the equality in the third line holds if and only if $w_{\mu-1} = w_{\mu-1}'$ since $d_{\mu-1} > d_\mu > 0$. $\square$

Remark 3.1. If $\det(\mathcal{H}_\tau) \neq 0$ for all $1 \leq \tau \leq t_g$, we have the upper bound for the inverse-invariant degree

$$\operatorname{iideg}\Big(\prod_{\tau=1}^{t_g}\det(\mathcal{H}_\tau)\Big) \leq t_g(t_g+2)\Big(\frac{1}{6}t_g\,\delta_t - \frac{1}{12}\delta_t - \frac{1}{48}t_g^2 + \frac{1}{12}t_g - \frac{1}{12}\Big)$$

for the interpolant $g(y)$ in (2) and t_g in (3). The bound is asymptotically $1/2$ of that in Remark 2.1. $\square$

4. The Levinson–Durbin–Heinig Algorithm

For a Hankel matrix H_τ,

$$H_\tau = \begin{bmatrix} a_{1-\tau} & a_{2-\tau} & \dots & a_0 \\ a_{2-\tau} & a_{3-\tau} & \iddots & a_1 \\ \vdots & \vdots & \iddots & \vdots \\ a_0 & a_1 & \dots & a_{\tau-1} \end{bmatrix} \in \mathsf{K}^{\tau\times\tau}, \quad H_\tau[i,j] = a_{i+j-(\tau+1)}, \tag{45}$$

the matrix $A_\tau = H_\tau J_\tau$ with $J_\tau = \begin{bmatrix} 0 & \dots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \dots & 0 \end{bmatrix} \in \mathsf{K}^{\tau\times\tau}$ is a Toeplitz matrix:

$$A_\tau = H_\tau J_\tau = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \dots & a_{-(\tau-1)} \\ a_1 & a_0 & a_{-1} & \dots & a_{-(\tau-2)} \\ a_2 & a_1 & a_0 & \dots & a_{-(\tau-3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{\tau-1} & a_{\tau-2} & a_{\tau-3} & \dots & a_0 \end{bmatrix} \in \mathsf{K}^{\tau\times\tau}, A_\tau[i,j] = a_{i-j}.^{1} \tag{46}$$

Notation: By $A_\tau[i_1 \dots i_2, j_1 \dots j_2]$ we denote the submatrix of A_τ (46) formed by rows $i_1 \le i \le i_2$ and columns $j_1 \le j \le j_2$; by $A_\tau[i,*]$ we denote the i-th row and by $A_\tau[*,j]$ we denote the j-th column; $\langle\cdot,\cdot\rangle$ denotes the scalar product of two vectors. In (45,46) a_i denotes the Hankel/Toeplitz entry, which for our sparse interpolation problem is $f(\omega^i)$ or $f(\omega^{2i+1})$.

Our early termination algorithm samples until a Toeplitz matrix A_{r+1} becomes singular, in which case the Laurent sparsity is $t_g = r$ with high probability. At this point all A_τ for $\tau = 1, \dots, r$ are non-singular and one has vectors

$$A_r\,\varphi^{(r)} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_1^{(r)} \in \mathsf{K}^r, \quad A_r\,\chi^{(r)} = \begin{bmatrix} A_{r+1}[1,r+1] \\ A_{r+1}[2,r+1] \\ \vdots \\ A_{r+1}[r,r+1] \end{bmatrix} = \begin{bmatrix} a_{-r} \\ a_{-(r-1)} \\ \vdots \\ a_{-1} \end{bmatrix} \in \mathsf{K}^r. \tag{47}$$

In (47) $\varphi^{(r)}$ is the first column of A_r^{-1} and is called the forward vector and the vector $\chi^{(r)}$ generates the next column, which are called the Yule-Walker equations. In order to increase the probability that one has determined the sparsity t_g, and to skip over a sporadic false singular submatrix, we shall sample further until

$$\langle A_{r+\zeta}[r+1+k, 1 \dots r],\, \chi^{(r)}\rangle = A_{r+\zeta}[r+1+k, r+1] \text{ for all } k = 0, \dots, \zeta-1, \tag{48}$$

where $\zeta \ge 1$ is an input look-ahead count. If in the range of (48) one discovers for $\zeta \ge 2$ a discrepancy

$$\exists \ell, 1 \le \ell \le \zeta-1\colon \langle A_{r+\zeta}[r+1+\ell, 1 \dots r],\, \chi^{(r)}\rangle \ne A_{r+\zeta}[r+1+\ell, r+1], \tag{49}$$

the matrix dimension r cannot be the sparsity t_g, and the interpolation algorithm computes the next non-singular A_{r+k}.

[1] We use the letter A for the Toeplitz matrices because T_n is the n-degree Chebyshev-1 polynomial.

Figure 2: The matrix $A_{n+1} \in \mathsf{K}^{(n+1)\times(n+1)}$ with $n+1 = r+\ell+m$

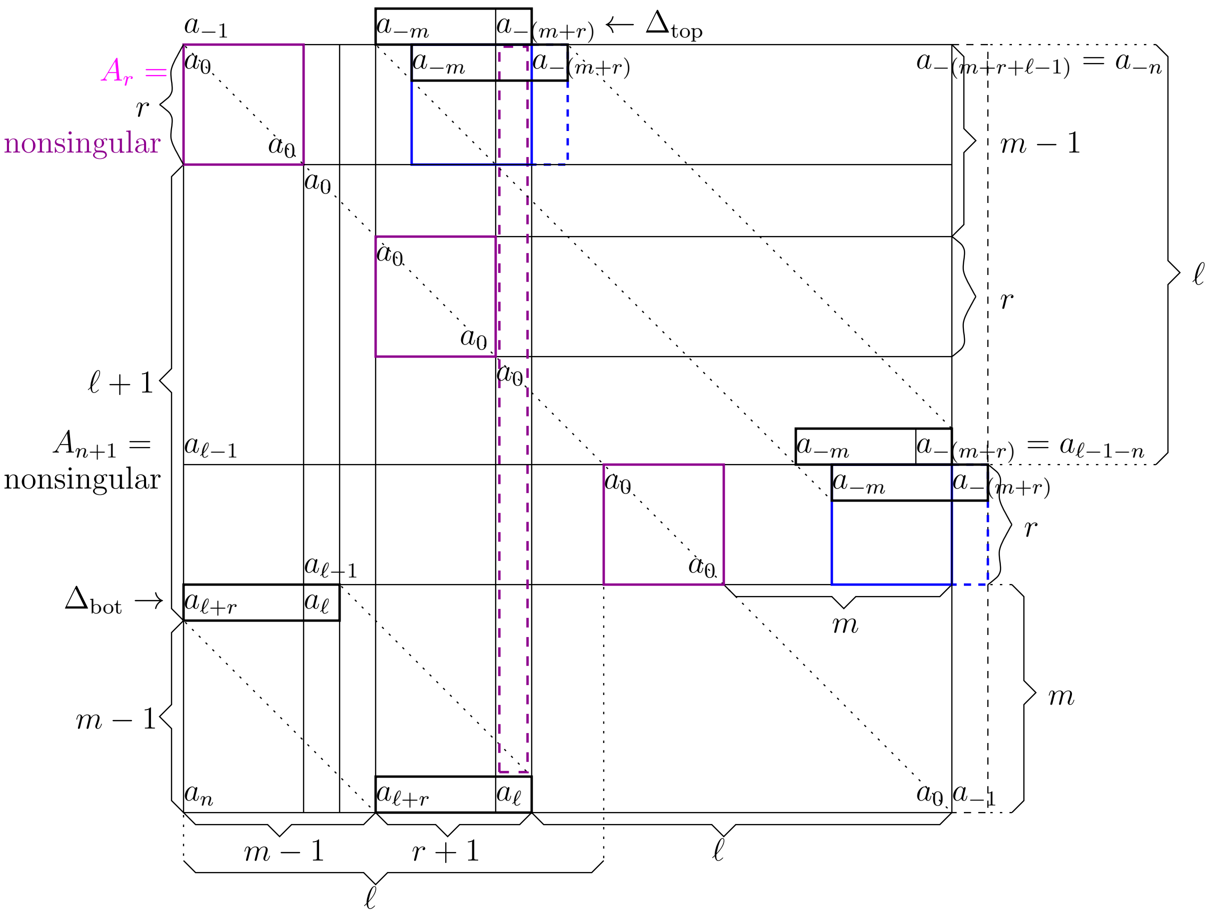
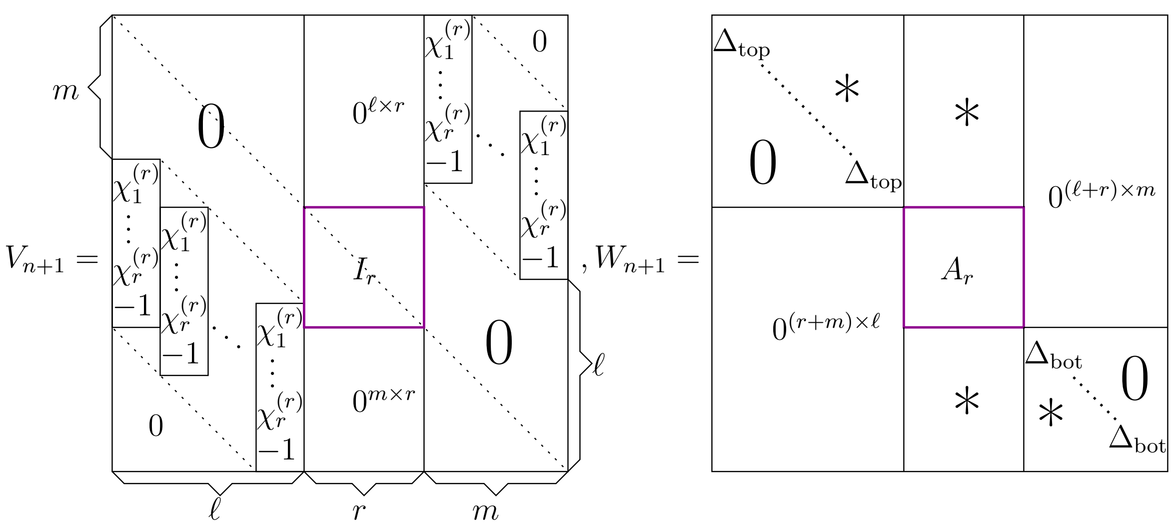


Figure 3: The matrices $V_{n+1}, W_{n+1} \in \mathsf{K}^{(n+1)\times(n+1)}$ with $A_{n+1}V_{n+1} = W_{n+1}$



A false column relation $\chi^{(r)}$ may be discovered in subsequent steps in the Prony algorithm even if there are ζ discrepancies $= 0$ (48). In [Kaltofen and Yang, Z.-H. 2021] we explicitly describe in Algorithm 1 *Try Prony's algorithm* how one can diagnose that a sequence of values does not yield an interpolant. Again, the algorithm jumps forward to the next non-singular A_{r+k}.

The classical Levinson-Durbin algorithm computes the first singular A_τ and Georg Heinig's 1983 algorithm computes the subsequent non-singular $A_{\tau+k}$'s [Heinig 1983], [Heinig and Rost 1984, Part I, Chapter 6], both in quadratic time and linear space. The algorithm for symmetric Toeplitz matrices is also in [Delsarte, Genin, and Kamp 1985, see Note in I. Introduction for the relation to Heinig's algorithm]. We define the Iohvidov indices and the corresponding discrepancies as

$$\ell = \min\Big\{i \mid i \geq 0, a_i \neq \sum_{k=1}^{r} a_{r+1+i-k}\chi_k^{(r)}\Big\}, \quad \Delta_{\text{bot}} = \Big(\sum_{k=1}^{r} a_{r+1+\ell-k}\chi_k^{(r)}\Big) - a_\ell \neq 0, \qquad (50)$$

where ℓ is as in (49), and

$$m = \min\Big\{i \mid i \geq 1, a_{-(i+r)} \neq \sum_{k=1}^{r} a_{-(i+k-1)}\chi_k^{(r)}\Big\},$$
$$\Delta_{\text{top}} = \Big(\sum_{k=1}^{r} a_{-(m+k-1)}\chi_k^{(r)}\Big) - a_{-(m+r)} \neq 0. \qquad (51)$$

In Figure 2 (compare with [Heinig and Rost 1984, p. 105, Picture 6.1]) we display the matrix properties of ℓ and m: The second purple submatrix from the left is A_r which has slid down the diagonal to start at row m and column m. The vector $\chi^{(r)}$ generates column $m+r$ of A_{n+1} for all except the last entry: $A_{n+1}[1\dots n,\, m\dots(m+r-1)]\,\chi^{(r)} = A_{n+1}[1\dots n,\, m+r]$, which is indicated by the purple-dashed column. There is a discrepancy in row $n+1$: $\langle A_{n+1}[n+1,\, m\dots(m+r-1)],\, \chi^{(r)}\rangle = A_{n+1}[n+1,\, m+r] + \Delta_{\text{bot}}$ (50), and a discrepancy in the row above the first row, which diagonally shifts down to row 1: $\langle A_{n+1}[1,\, (m+1)\dots(m+r)],\, \chi^{(r)}\rangle = A_{n+1}[1, m+r+1] + \Delta_{\text{top}}$ (51). Those segments in the rows are highlighted by a thicker black border. That purple-dashed linearly dependent column, which can be diagonally shifted up, proves for $\ell \geq 1$ that all $A_{r+1}, \dots, A_n$ are singular. If $\ell = 0$ the matrix A_{r+1} is non-singular because its last column is linearly independent of its first r columns and the index m is not computed.

The index ℓ may not exist, which means A_{r+i} is singular for all $i \geq 1$. If the index ℓ exists, the index m may not exist and again A_{r+i} is singular for all $i \geq 1$. The crucial property is that the elements $a_{\ell+r}$ and $a_{-(m+r)}$, which cause the discrepancies, make the matrix A_{n+1} non-singular no matter what are the values of the remaining elements $a_{\ell+r+1}, \dots,\, a_{-(m+r+1)}, \dots$ [Heinig and Rost 1984, p. 103, Lemma 6.1]. As proof we give in Figure 3 the matrix product $A_{n+1}V_{n+1} = W_{n+1}$ and show that W_{n+1} is non-singular. In the matrix W_{n+1}, the middle block-column is the same as the block-column in A_{n+1} and one can eliminate the top and bottom non-zero blocks, denoted by $\bar{A}$ and $\bar{\bar{A}}$, with the non-singular matrix A_r:

$$\begin{bmatrix} I_\ell & -\bar{A}A_r^{-1} & 0 \\ 0 & I_r & 0 \\ 0 & 0 & I_m \end{bmatrix} \begin{bmatrix} I_\ell & 0 & 0 \\ 0 & I_r & 0 \\ 0 & -\bar{\bar{A}}A_r^{-1} & I_m \end{bmatrix} \underbrace{\begin{bmatrix} U & \bar{A} & 0 \\ 0 & A_r & 0 \\ 0 & \bar{\bar{A}} & L \end{bmatrix}}_{W_{n+1} = A_{n+1}V_{n+1}} = \begin{bmatrix} U & 0 & 0 \\ 0 & A_r & 0 \\ 0 & 0 & L \end{bmatrix}. \tag{52}$$

The resulting matrix is a 3×3 block triangular matrix with non-singular blocks on the diagonal and is therefore non-singular. Note that the corner blocks U and L in W_{n+1} are triangular Toeplitz matrices. We will use the columns in the block-column 1 and the block-column 3 of V_{n+1} in the algorithm.

4.1. Algorithm Levinson-Durbin-Heinig

Input:
- ▸ $r = 0$ or $r \geq 1$ with $A_r \in \mathsf{K}^{r \times r}$ non-singular;
- ▸ if $r \geq 1$ vectors $\varphi^{(r)}$, $\chi^{(r)}$ which satisfy (47);
- ▸ optionally, a dimension bound $B \geq r + 1$.

Output:
- ▸ an integer $n \geq r$ such that $A_{n+1} \in \mathsf{K}^{(n+1)\times(n+1)}$ is non-singular; if $n \geq r+1$ then $\forall i, 1 \leq i \leq n - r$: A_{r+i} is singular.
- ▸ vectors $\varphi^{(n+1)}, \chi^{(n+1)} \in \mathsf{K}^{n+1}$ which satisfy (47) for $r = n + 1$;
- ▸ optionally, a message that all A_{r+i} are singular for $1 \leq i \leq B - r$.

LDH1: *If $r = 0$ then if $a_0 \neq 0$ then return $n = 0$; $\varphi^{(1)} \leftarrow [1/a_0]$; $\chi^{(1)} \leftarrow [a_{-1}/a_0]$.*

If $r = 0$ and $a_0 = 0$ then $A_{n+1} = A_{\ell+m} = \begin{bmatrix} 0 & \cdots & 0 & a_{-m} & \cdots & a_{-(m+\ell-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & a_{-m} \\ a_\ell & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{\ell+m-1} & \cdots & a_\ell & 0 & \cdots & 0 \end{bmatrix}$ *with*

$a_\ell \neq 0$, $a_{-m} \neq 0$; *return* $n = \ell + m - 1$, $\varphi^{(n+1)} = \begin{bmatrix} 0^m \\ \frac{1}{a_{-m}} \\ 0^{\ell-1} \end{bmatrix}$; $\chi^{(n+1)} = A_{n+1}^{-1} \begin{bmatrix} a_{-(m+\ell)} \\ \vdots \\ a_{-m} \\ 0^{m-1} \end{bmatrix}$;

the vector $\chi^{(n+1)}$ can be computed by back- and forward substitution with the triangular (Toeplitz) blocks on the block anti-diagonal. See also Steps LDH6 and LDH7 below. Faster algorithms can be used based on truncated power series inverse computations.

Optionally, if the first non-zero element in the first column is too far down or the first non-zero element in the first row is too far right so that $\ell + m > B$, all matrices up to the dimension bound B are singular.

LDH2: Now $r \geq 1$ and A_r is a non-singular matrix.

For the column relation given by $\chi^{(r)}$ compute the Iohvidov index ℓ with (50). Again no index may, optionally, be found with $r + \ell \leq B$ in which case all Toeplitz matrices up to dimension B are singular.

LDH3: If $\ell = 0$ the next non-singular matrix is A_{r+1} and the classical Levinson-Durbin algorithm can be applied.

(a) *Set* $\gamma^{(r+1)} \leftarrow \frac{1}{\Delta_{\mathrm{bot}}}\begin{bmatrix}\chi^{(r)}\\ -1\end{bmatrix}$*, where* Δ_{bot} *is defined in* (50)*.* One has $A_{r+1}\gamma^{(r+1)} = \begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} = e_{r+1}^{(r+1)}$, that is, $\gamma^{(r+1)}$ is the last column of A_{r+1}^{-1}. The vector $\gamma^{(r+1)}$ is called the backward vector.

Compute the vector $\varphi^{(r+1)}$ *as*

$$A_{r+1}\Bigl(\underbrace{\begin{bmatrix}\varphi^{(r)}\\ 0\end{bmatrix} - \Bigl\langle A_{r+1}[r+1,*], \begin{bmatrix}\varphi^{(r)}\\ 0\end{bmatrix}\Bigr\rangle \gamma^{(r+1)}}_{\varphi^{(r+1)}}\Bigr) = e_1^{(r+1)}. \tag{53}$$

(b) *Compute the vector* $\chi^{(r+1)}$ *as*

$$A_{r+1}\Bigl(\underbrace{\begin{bmatrix}0\\ \chi^{(r)}\end{bmatrix} + \Bigl(A_{r+2}[1,r+2] - \Bigl\langle A_{r+1}[1,*], \begin{bmatrix}0\\ \chi^{(r)}\end{bmatrix}\Bigr\rangle\Bigr)\varphi^{(r+1)}}_{\chi^{(r+1)}}\Bigr)$$

$$= \begin{bmatrix}A_{r+2}[1,r+2]\\ \vdots\\ A_{r+2}[r+1,r+2]\end{bmatrix} = \begin{bmatrix}a_{-(r+1)}\\ \vdots\\ a_{-1}\end{bmatrix}. \tag{54}$$

Return $n = r, \varphi^{(r+1)}, \chi^{(r+1)}$*.*

LDH4: We now have $\ell \geq 1$. *For the column relation given by* $\chi^{(r)}$ *compute the Iohvidov index* m *with* (51)*. Again no index may, optionally, be found with* $r+\ell+m \leq B$ *in which case all Toeplitz matrices up to dimension* B *are singular.*

LDH5: As proved above, we now have the next non-singular matrix A_{n+1} with $n+1 = r+\ell+m$. The structure of A_{n+1} is shown in Figure 2. The columns of the solid-blue-bordered $r\times r$ submatrix in the first r rows, $A_{n+1}[1\ldots r, (m+1)\ldots(m+r)]$, generate via the column relation $\chi^{(r)}$ the next blue-dashed column except in row 1, namely, for the subcolumn $A_{n+1}[2\ldots r, m+r+1]$. In row 1 we have the discrepancy Δ_{top}. The solid-blue-bordered $r\times r$ matrix slides down to row $\ell+1$ and column $\ell+m+1$: $A_{n+1}[(\ell+1)\ldots(\ell+r), (\ell+m+1)\ldots(n+1)] = A_{n+1}[1\ldots r, (m+1)\ldots(m+r)]$. Those rows $\ell+1,\ldots,\ell+r$ of A_{n+1} also contain A_r and we have for the $r\times(n+1)$ submatrix formed by those rows:

$$A_{n+1}[(\ell+1)\ldots(\ell+r),*]\Biggl(\underbrace{\begin{bmatrix}0^{\ell}\\ 0^{m}\\ \chi^{(r)}\end{bmatrix} - \Delta_{\mathrm{top}}\begin{bmatrix}0^{\ell}\\ \varphi^{(r)}\\ 0^{m}\end{bmatrix}}_{\widehat{\chi}}\Biggr) = \begin{bmatrix}A_{n+2}[\ell+1,n+2]\\ \vdots\\ A_{n+2}[\ell+r,n+2]\end{bmatrix}. \tag{55}$$

Here $\varphi^{(r)}$ is used to correct the generator in row $\ell+1$, in correspondence to the Berlekamp-Massey algorithm.

Compute the vector $\widehat{\chi} \in \mathsf{K}^{n+1}$ *defined in* (55), which now generates the segment of the $(n+2)$-nd column of A_{n+2} in the row positions $\ell+1,\ldots,\ell+r$. The $(n+2)$-nd column of A_{n+2} is indicated by the dashed border in Figure 2.

In [Heinig and Rost 1984, 2nd step on p. 104, Lemma 6.2] the algorithm starts with

$$\widehat{\chi} = \begin{bmatrix} 0^\ell \\ A_r^{-1} A_{n+2}\texttt{[}(\ell+1)\ldots(\ell+r), n+2\texttt{]} \\ 0^m \end{bmatrix},$$

which is computed from $\chi^{(r)}$ by iteratively shifting the right-side vector down; see Step Sh1 in Algorithm 6.1 in Section 6.

LDH6: We use the columns $\ell, \ell-1, \ldots,$ back to 1 of V_{n+1} in Figure 3 to correct $\widehat{\chi}$ starting at row ℓ going up.
For $k \leftarrow \ell, \ell-1, \ell-2, \ldots, 1$ *Do*
$\texttt{\{}\ \widehat{\Delta} \leftarrow \langle A_{n+1}\texttt{[}k,*\texttt{]}, \widehat{\chi}\rangle - A_{n+2}\texttt{[}k, n+2\texttt{]};\ \widehat{\chi} \leftarrow \widehat{\chi} - \frac{\widehat{\Delta}}{\Delta_{\text{top}}} V_{n+1}\texttt{[}*,k\texttt{]};\ \texttt{\}}$
Note that because in Figure 3 the entries in column k of W_{n+1} below the entry containing Δ_{top} all are $= 0$, the update of $\widehat{\chi}$ does not change the (row of A_{n+1})-(updated $\widehat{\chi}$) scalar products below row k and the updated generator $\widehat{\chi}$ remains correct down to row $\ell + r$.

LDH7: We use the columns $\ell+r+1, \ell+r+2, \ldots, \ell+r+m$ of V_{n+1} in Figure 3 to correct $\widehat{\chi}$ starting at row $\ell+r+1$ going down.
For $k \leftarrow \ell+r+1, \ell+r+2, \ldots, \ell+r+m$ *Do*
$\texttt{\{}\ \widehat{\Delta} \leftarrow \langle A_{n+1}\texttt{[}k,*\texttt{]}, \widehat{\chi}\rangle - A_{n+2}\texttt{[}k, n+2\texttt{]};\ \widehat{\chi} \leftarrow \widehat{\chi} - \frac{\widehat{\Delta}}{\Delta_{\text{bot}}} V_{n+1}\texttt{[}*,k\texttt{]};\ \texttt{\}}$
Note that again because in Figure 3 the entries in column k of W_{n+1} above the entry containing Δ_{bot} all are $= 0$, the update of $\widehat{\chi}$ does not change the row-vector scalar products above row k and the updated generator remains correct up to row 1. (Compare the Steps LDH6 and LDH7 with [Heinig and Rost 1984, 3rd step on p. 105].)

LDH8: $\chi^{(n+1)} \leftarrow \widehat{\chi};\ \varphi^{(n+1)} \leftarrow \frac{1}{\Delta_{\text{top}}} V_{n+1}\texttt{[}*,1\texttt{]};$ (See Eq. (6.6) on p. 105 in [Heinig and Rost 1984].) Here we use that $m+r+1 \le m+r+\ell$ because $\ell \ge 1$. Therefore the entry in row 1 which contains $a_{-(m+r)}$ is within the matrix A_{n+1} (Figure 2). Note the correspondence to the Berlekamp-Massey algorithm: the new auxiliary vector, namely $\varphi^{(n+1)}$, is the shifted (and scaled) old generator vector, which is $\begin{bmatrix} \chi^{(r)} \\ -1 \end{bmatrix}$.

Return $n, \varphi^{(n+1)}, \chi^{(n+1)}$. Note that $\gamma^{(n+1)} = \frac{1}{\Delta_{\text{bot}}} V_{n+1}\texttt{[}*, n+1\texttt{]}$.

Remark 4.1. Algorithm 4.1 requires $O((n+1-r)n)$ arithmetic operations and $O(n)$ space. For a given non-singular matrix A_{n+1}, starting from $r = 0$ the algorithm reaches dimension $n+1$ in $O(n^2)$ arithmetic operations and $O(n)$ space [Heinig and Rost 1984, Remark 6.1].

In Step LDH3a, instead of using $\chi^{(r)}$, some versions compute $\gamma^{(r+1)}$ from $\gamma^{(r)}$: one has

$$A_{r+1} \underbrace{\left(\begin{bmatrix} 0 \\ \gamma^{(r)} \end{bmatrix} - \left\langle A_{r+1}\texttt{[}1,*\texttt{]}, \begin{bmatrix} 0 \\ \gamma^{(r)} \end{bmatrix} \right\rangle \varphi^{(r+1)} \right)}_{\gamma^{(r+1)}} = e_{r+1}^{(r+1)}. \tag{56}$$

Substituting for $\varphi^{(r+1)}$ its assigned value in (53) one obtains an equation for $\gamma^{(r+1)}$ which yields

$$\gamma^{(r+1)} \leftarrow \frac{1}{1-\nu\mu} \left(\begin{bmatrix} 0 \\ \gamma^{(r)} \end{bmatrix} + \mu \begin{bmatrix} \varphi^{(r)} \\ 0 \end{bmatrix} \right) \tag{57}$$

with $\nu = -\langle A_{r+1}[r+1,*], \left[\begin{smallmatrix}\varphi^{(r)}\\ 0\end{smallmatrix}\right] \rangle$ and $\mu = -\langle A_{r+1}[1,*], \left[\begin{smallmatrix}0\\ \gamma^{(r)}\end{smallmatrix}\right] \rangle$ and A_{r-1} non-singular $\Longrightarrow \gamma^{(r)}[r] \neq 0 \Longrightarrow \nu\mu \neq 1$. $\square$

Remark 4.2. The matrix product $A_{n+1}V_{n+1} = W_{n+1}$ in Figure 3 is adapted from [Kaltofen and Yuhasz 2013a, Proof of Lemma 2], where the Berlekamp-Massey algorithm is described as a solver for linear systems with Hankel coefficient matrices. As in [Kaltofen and Yuhasz 2013a], in the Toeplitz matrices A_τ (46) the entries a_i can be square matrices: Algorithm 4.1 works for block Toeplitz matrices subject to the restriction that the discrepancies $\Delta_{\mathrm{top}}, \Delta_{\mathrm{bot}}$, which now are matrices, always are invertible matrices. A fraction-free version of Algorithm 4.1 can be formulated in correspondence to [Kaltofen and Yuhasz 2013a]. $\square$

Remark 4.3. By (45,46), for a non-singular Toeplitz matrix A_{n+1} the linear system $A_{n+1}x = b$ with $b \in \mathsf{K}^{n+1}$ has the solution $x = J_{n+1}\tilde{x}$ with $H_{n+1}\tilde{x} = b$. Therefore, the arithmetic complexity of solving non-singular square linear system with Toeplitz and Hankel coefficient matrices is the same.

The Berlekamp-Massey algorithm [Massey 1969] steps from a non-singular leading principal Hankel submatrix H_r to the next non-singular Hankel submatrix H_{r+k} like Algorithm 4.1. The Berlekamp-Massey algorithm is a streamlined version of Kronecker's 1881 anti-diagonal algorithm for computing entries in the table of Padé rational function approximants for a power series by the extended Euclidean algorithm for polynomials; see [Kaltofen and Yuhasz 2013b, Table 1, Scalar case] for literature references. With the Lehmer-Knuth-Schönhage half-GCD algorithm for polynomials [Moenck 1973], [Strassen 1983, Proof of Theorem 4.3, procedure SCH] one has a ("super-fast") Hankel solver of arithmetic complexity $O(\mathsf{M}(n)\log(n))$ [Brent, Gustavson, and Yun 1980], where $\mathsf{M}(n)$ is the arithmetic complexity of multiplying polynomials of degree n with coefficients in the field K. One has $\mathsf{M}(n) = O(n\log(n)\,\mathrm{loglog}(n))$ for any arbitrary field K [Cantor and Kaltofen 1991], and $\mathsf{M}(n) = O(n\ \log(n))$ for fields with roots of unity such as the complex numbers.

For the system $A_{n+1}x = b$ with an arbitrary right-side vector b one can use the landmark Gohberg-Semencul formulas for A_{n+1}^{-1} [Brent, Gustavson, and Yun 1980, Section 6] (compare with [Heinig and Rost 1984, p. 133, Notes and Comments, Section 1, item 5]). If a Gohberg-Semencul representation of A_r^{-1} is also given on input, or computed from $\phi^{(r)}$ and $\gamma^{(r)}$ [Brent, Gustavson, and Yun 1980, Section 6(c)], the factorization (52) and FFT-based polynomial multiplication yields a version of Algorithm 4.1 which requires $O(\mathsf{M}(n))$ arithmetic operations. However, the computation of the Iohvidov indices ℓ (50) and m (51) then use entry values beyond a_n and a_{-n}. For ℓ, for example, one computes the polynomial product $(a_0 + a_1z + a_2z^2 + \cdots)\,(-z^r + \chi_r^{(r)}z^{r-1} + \cdots + \chi_0^{(r)})$ in blocks of $r+1$ coefficients at a cost $O(\,(\frac{\ell}{r}+1)\,\mathsf{M}(r)\,)$.

For the Hankel solver by the standard Berlekamp-Massey algorithm with running time $O(n^2)$, the solution $\tilde{x}$ for a right-side vector b can be computed iteratively [Kaltofen and Yuhasz 2013a, Section 4]. The same can be done for Algorithm 4.1; see Section 6. The classical Levinson-Durbin algorithm can also be speeded to $O(\mathsf{M}(n)\log(n))$ arithmetic cost [Bitmead and Anderson 1980], [Morf 1980] and the required genericity conditions can be realized by multiplicative randomized pre-conditioning [Kaltofen 1994], [Kaltofen 1995, Appendix A]. $\square$

Remark 4.4. For the two-sided infinite sequence of entries $\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots \in \mathsf{K}^{\mathbb{Z}}$, or one-sided if the Toeplitz matrix is symmetric, we can view Algorithm 4.1 as a sequence transformer. Each element a_i is mapped to the discrepancy computed by the algorithm: see Table 1.

Table 1: The discrepancy associated with each a_i

Index i	*Discrepancy* Δ_i
$i = r + \ell$	$\Delta_{r+\ell} = \Delta_{\text{bot}}$ (50)
Step Levinson-Durbin LDH3: $\ell = 0$	
$i = -(r+1)$	$\Delta_{-(r+1)} = \langle A_{r+1}[1,*], \left[\begin{smallmatrix} 0 \\ \chi^{(r)} \end{smallmatrix}\right] \rangle - a_{-(r+1)}$ (54)
Steps Heinig LDH4–7: $\ell \geq 1, m \geq 1$	
$-(n+1) \leq i \leq -(r+m+1)$	$\Delta_i = \widehat{\Delta}$ in Step LDH6 at loop index $k = \ell, \ell-1, \ldots, 1$ and $i = -(r+m+1+\ell-k) = -(n+2-k)$
$i = -(r+m)$	$\Delta_{-(r+m)} = \Delta_{\text{top}}$ (Step LDH5)
$-(r+m-1) \leq i \leq -(r+1)$	$\Delta_i = 0$ (51); empty if $m = 1$
$-r \leq i \leq r-1$	computed for A_r and column $A[1 \ldots r, r+1]$
$r \leq i \leq r+\ell-1$	$\Delta_i = 0$ (50)
$r+\ell+1 \leq i \leq n$	$\Delta_i = \widehat{\Delta}$ in Step LDH7 at loop index $k = \ell+r+2, \ell+r+3, \ldots, n+1$ and $i = k-1$; Note: the first discrepancy for $a_{r+\ell}$ was Δ_{bot} before

With each discrepancy, there is also associated a linear generator $\chi^{(i)}$, which for $\Delta_i \neq 0$ repairs the discrepancy. We note that the map is specific to Algorithm 4.1. One can, for example, do Step LDH7 immediately after Step 5 before Step LDH6, in which case the sequence of discrepancies changes. If there are only finitely many non-singular matrices, the largest being A_r, the map stops towards $a_{-\infty}$ at a_{-r} (no $\Delta_{\text{bot}} \neq 0$) or towards a_{∞} at $a_{r+\ell}$ (no $\Delta_{\text{top}} \neq 0$). $\square$

For our sparse interpolation algorithms in inverse-invariant basis with early termination, the fast Toeplitz/Hankel algorithms referenced in Remark 4.3 are not directly applicable. Those interpolation algorithms with threshold $\zeta \geq 1$ compute the last non-singular Toeplitz matrix $A_t = A_r$ such that (48) is satisfied:

$$a_i = \sum_{k=1}^{r} a_{r+1+i-k} \chi_k^{(r)} \text{ for all } 0 \leq i \leq \zeta - 1. \tag{58}$$

The goal is to minimize the number of values a_i. We give an example.

Example 4.1. In [Imamoglu et al. 2018, Section 4] we interpolate

$$g(y) = \frac{32768}{5281339833}(\frac{1}{y^6}+y^6) - \frac{1024}{2540327}(\frac{1}{y^5}+y^5) \\ + \frac{64}{7227}(\frac{1}{y^4}+y^4) - \frac{744}{8687}(\frac{1}{y^3}+y^3) + \frac{62}{153}(\frac{1}{y^2}+y^2) + \frac{254}{189} \quad (59)$$

at the values $a_i = a_{-i} = g(2^i)$ for $i = 0, 1, 2, \ldots$ ($\omega = 2$). The ranks of the arising (symmetric) Toeplitz matrices A_τ for $\tau = 1, 2, 3, \ldots$ are

τ	1	2	3	4	5	6	7	8	9	10	11	12	13	...
rank(A_τ)	1	2	2	2	2	2	4	6	8	10	11	11	11	...

(60)

We have for $r = 2$ the Iohvidov indices $\ell = 4$ and $m = 4$. For threshold $\zeta = 4$ the algorithm interpolates $\tilde{g}(y) = y + \frac{1}{y}$ because g was constructed such that $g(2^i) = \tilde{g}(2^i)$ for $i = 0, 1, \ldots, 5$. For threshold $\zeta = 5$ the algorithm discovers a discrepancy at $\ell = 4$ and interpolates $g(y)$. $\square$

5. Conclusion

Singularity, or with numerical values ill-conditionedness, of the arising Toeplitz matrices in the Prony algorithm for interpolating a linear sum of Chebyshev polynomials is used to determine the sparsity by randomization. Singular submatrices that do not correspond to an interpolant or an interpolant with fewer terms, as in Example 4.1, can be stepped over by look-ahead until a threshold-many singular matrices have been seen. If look-ahead produces a discrepancy, the next non-singular matrix needs to be computed. The Heinig algorithm for Toeplitz matrices with singular sections can perform that step in overall quadratic time and linear space while requiring the minimum number of interpolation points.

The Heinig algorithm, in correspondence to the Berlekamp-Massey algorithm for Hankel matrices [Kaltofen and Yuhasz 2013a], can be viewed as a sequence transformer for the entries of an infinite Toeplitz matrix. The infinite sequence is mapped to discrepancies and auxiliary vectors (see Remark 4.4), which can be used to represent the inverse of each non-singular submatrix by the Gohberg-Semencul formulas.

Acknowledgement: Zhi-Hong Yang is supported by the National Natural Science Foundation of China (12201425). We thank Karla Rost for her communications about the Heinig algorithm. We appreciate the detailed comments from the reviewers.

References

Arnold, Andrew and Kaltofen, Erich L. Error-correcting sparse interpolation in the Chebyshev basis. In *ISSAC'15 Proc. 2015 ACM Internat. Symp. Symbolic Algebraic Comput.*, pages 21–28, New York, N. Y., 2015. Association for Computing Machinery. URL: http://users.cs.duke.edu/~elk27/bibliography/15/ArKa15.pdf.

Bitmead, R. R. and Anderson, B. D. O. Asymtotically fast solution of Toeplitz and related systems of linear equations. *Linear Algebra Applic.*, 34:103–116, 1980.

Brent, R. P., Gustavson, F. G., and Yun, D. Y. Y. Fast solution of Toeplitz systems of equations and computation of Padé approximants. *J. Algorithms*, 1:259–295, 1980.

Cantor, David G. and Kaltofen, Erich. Fast multiplication of polynomials over arbitrary rings. Technical Report 87-35, Rensselaer Polytechnic Instit., Dept. Comput. Sci., Troy, N. Y., December 1987. URL: http://users.cs.duke.edu/~elk27/bibliography/87/CaKa87_techrep.pdf; Cantor and Kaltofen [1991] contains an alternate algorithm.

Cantor, D. G. and Kaltofen, E. On fast multiplication of polynomials over arbitrary algebras. *Acta Inform.*, 28(7):693–701, 1991. URL: http://users.cs.duke.edu/~elk27/bibliography/91/CaKa91.pdf; Cantor and Kaltofen [1987] contains an alternate algorithm.

Delsarte, Philippe, Genin, Yves V., and Kamp, Yves G. A generalization of the Levinson algorithm for Hermitian Toeplitz matrices with any rank profile. *IEEE Trans. Acoustics, Speech and Signal Proc.*, ASSP-33(4):964–971, August 1985.

Hao, Zhiwei, Kaltofen, Erich L., and Zhi, Lihong. Numerical sparsity determination and early termination. In Rosenkranz, Markus, editor, *ISSAC'16 Proc. 2016 ACM Internat. Symp. Symbolic Algebraic Comput.*, pages 247–254, New York, N. Y., 2016. Association for Computing Machinery. URL: http://users.cs.duke.edu/~elk27/bibliography/16/HKZ16.pdf.

Heinig, Georg. Inversion of Toeplitz and Hankel matrices with singular sections. *Wissenschaftliche Zeitschrift der Technischen Hochschule Karl-Marx-Stadt*, 25(3):326–333, 1983.

Heinig, Georg and Rost, Karla. *Algebraic Methods for Toeplitz-like Matrices and Operators.* Birkhäuser Verlag, Basel-Boston-Stuttgart, 1984.

Imamoglu, Erdal, Kaltofen, Erich L., and Yang, Zhengfeng. Sparse polynomial interpolation with arbitrary orthogonal polynomial bases. In Arreche, Carlos, editor, *ISSAC '18 Proc. 2018 ACM Internat. Symp. Symbolic Algebraic Comput.*, pages 223–230, New York, N. Y., 2018. Association for Computing Machinery. In memory of Bobby F. Caviness (3/24/1940–1/11/2018). URL: http://users.cs.duke.edu/~elk27/bibliography/18/IKY18.pdf.

Kaltofen, E. Asymptotically fast solution of Toeplitz-like singular linear systems. In *Proc. 1994 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'94)*, pages 297–304, New York, N. Y., 1994. ACM Press. Journal version in Kaltofen [1995]. URL: http://users.cs.duke.edu/~elk27/bibliography/94/Ka94_issac.pdf.

Kaltofen, E. Analysis of Coppersmith's block Wiedemann algorithm for the parallel solution of sparse linear systems. *Math. Comput.*, 64(210):777–806, 1995. URL: http://users.cs.duke.edu/~elk27/bibliography/95/Ka95_mathcomp.pdf.

Kaltofen, Erich and Lee, Wen-shin. Early termination in sparse interpolation algorithms. *J. Symbolic Comput.*, 36(3–4):365–400, 2003. Special issue Internat. Symp. Symbolic Algebraic Comput. (ISSAC 2002). Guest editors: M. Giusti & L. M. Pardo. URL: http://users.cs.duke.edu/~elk27/bibliography/03/KL03.pdf.

Kaltofen, Erich L. and Yang, Zhi-Hong. Sparse interpolation with errors in Chebyshev basis beyond redundant-block decoding. *IEEE Trans. Information Theory*, 67(1):232–243, January 2021. URL: http://users.cs.duke.edu/~elk27/bibliography/19/KaYa19.pdf, https://ieeexplore.ieee.org/document/9207761, https://arxiv.org/abs/1912.05719.

Kaltofen, Erich L. and Yang, Zhi-Hong. Sparse polynomial interpolation with error correction: Higher error capacity by randomization. In Hauenstein, Jonathan, Lee, Wen-shin, and Chen, Shaoshi, editors, *ISSAC '24 Proc. 2024 ACM Internat. Symp. Symbolic Algebraic Comput.*, pages 264–273, New York, N. Y., 2024. Association for Computing Machinery. URL: http://users.cs.duke.edu/~elk27/bibliography/24/KY24.pdf, https://doi.org/10.1145/3666000.3669698.

Kaltofen, Erich and Yuhasz, George. A fraction free matrix Berlekamp/Massey algorithm. *Linear Algebra and Applications*, 439(9):2515–2526, November 2013a. URL: http://users.cs.duke.edu/~elk27/bibliography/08/KaYu08.pdf.

Kaltofen, Erich and Yuhasz, George. On the matrix Berlekamp-Massey algorithm. *ACM Trans. Algorithms*, 9(4):33:1–33:24, September 2013b. URL: http://users.cs.duke.edu/~elk27/bibliography/06/KaYu06.pdf.

Lakshman Y. N. and Saunders, B. D. Sparse polynomial interpolation in non-standard bases. *SIAM J. Comput.*, 24(2):387–397, 1995.

Massey, J. L. Shift-register synthesis and BCH decoding. *IEEE Trans. Inf. Theory*, IT-15: 122–127, 1969.

Moenck, R. T. Fast computation of GCDs. In *Proc. 5th ACM Symp. Theory Comp.*, pages 142–151, 1973.

Morf, M. Doubling algorithms for Toeplitz and related equations. In *Proc. 1980 IEEE Internat. Conf. Acoust. Speech Signal Process.*, pages 954–959. IEEE, 1980.

Strassen, V. The computational complexity of continued fractions. *SIAM J. Comput.*, 12(1): 1–27, 1983.

6. Appendix: A Toeplitz Solver for Arbitrary Right-Side Vectors

The Levinson-Durbin-Heinig Algorithm 4.1 can be extended to compute the solution $A_{n+1}^{-1} \times b^{(n+1)}$ for an arbitrary right-side vector $b^{(n+1)}$. The solution can be computed by the Gohberg-Semencul formulas at the final non-singular matrix [Brent, Gustavson, and Yun 1980, Section 6]. Instead, we compute a solution for each non-singular leading principle submatrix.

We use the following auxiliary algorithm.

6.1. Algorithm Shift Right Vector

Input: ▸ $A_r \in \mathsf{K}^{r\times r}$ non-singular for $r \geq 1$; $b_1, \ldots, b_{r+1} \in \mathsf{K}$;

▸ vectors $\varphi^{(r)}$, $\chi^{(r)}$ which satisfy (47); a vector $\gamma^{(r)} \in \mathsf{K}^r$ with $A_r\gamma^{(r)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathsf{K}^r$ (see Step LDH3a, Note in Step LDH8);

▸ a vector $\psi^{(r)} \in \mathsf{K}^r$ with $A_r\psi^{(r)} = \begin{bmatrix} b_1 \\ \vdots \\ b_r \end{bmatrix}$;

▸ optionally, $b_0 \in \mathsf{K}$.

Output: ▸ a vector $\widehat{\psi}^{(r)} \in \mathsf{K}^r$ such that $A_r\widehat{\psi}^{(r)} = \begin{bmatrix} b_2 \\ \vdots \\ b_{r+1} \end{bmatrix}$.

▸ optionally, a vector $\bar{\psi}^{(r)} \in \mathsf{K}^r$ such that $A_r\bar{\psi}^{(r)} = \begin{bmatrix} b_0 \\ \vdots \\ b_{r-1} \end{bmatrix}$.

Sh1: *Optionally, compute $\bar{\psi}^{(r)}$ by the algorithm in* [Heinig and Rost 1984, 2nd step on p. 104, Lemma 6.2].

(a) $\bar{\psi} \leftarrow \begin{bmatrix} 0 \\ \psi_1^{(r)} \\ \vdots \\ \psi_{r-1}^{(r)} \end{bmatrix} + \psi_r^{(r)}\chi^{(r)}$; note: $A_r\bar{\psi} = \begin{bmatrix} * \\ b_1 \\ \vdots \\ b_{r-1} \end{bmatrix}$.

(b) $\bar{\psi}^{(r)} \leftarrow \bar{\psi} + \big(b_0 - \langle A_r[1, 1\ldots r], \bar{\psi}\rangle\big)\,\varphi^{(r)}$.

Sh2: *If $\chi_1^{(r)} \neq 0$ do the following substeps:*

(a) $\widehat{\psi} \leftarrow \begin{bmatrix} \psi_2^{(r)} \\ \vdots \\ \psi_r^{(r)} \\ 0 \end{bmatrix} + \psi_1^{(r)}\dfrac{1}{\chi_1^{(r)}}\begin{bmatrix} -\chi_2^{(r)} \\ \vdots \\ -\chi_r^{(r)} \\ 1 \end{bmatrix}$ note: $A_r\dfrac{1}{\chi_1^{(r)}}\begin{bmatrix} -\chi_2^{(r)} \\ \vdots \\ -\chi_r^{(r)} \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_{r-1} \\ * \end{bmatrix}$, $A_r\widehat{\psi} = \begin{bmatrix} b_2 \\ \vdots \\ b_r \\ * \end{bmatrix}$.

(b) $\widehat{\psi}^{(r)} \leftarrow \widehat{\psi} + \big(b_{r+1} - \langle A_r[r, 1\ldots r], \widehat{\psi}\rangle\big)\,\gamma^{(r)}$; *return $\widehat{\psi}^{(r)}$ and, optionally, $\bar{\psi}^{(r)}$.*

Sh3: Now $\chi_1^{(r)} = 0$. *Do Steps* Sh4–Sh6; *afterwards, return $\widehat{\psi}^{(r)}$ and, optionally, $\bar{\psi}^{(r)}$.*

Sh4: We add a bottom row to A_r to have a non-singular Toeplitz matrix $\widetilde{A}_{r+1} \in \mathsf{K}^{(r+1)\times(r+1)}$. Note that there already is an $(r+1)$-st column: $A_r\chi^{(r)} = \begin{bmatrix} a_{-r} \\ \vdots \\ a_{-1} \end{bmatrix}$. If $\chi_1^{(r)} = 0$ then $\chi^{(r)}$ produces a discrepancy for row $r+1$, $\langle \widetilde{A}_{r+1}[r+1, 1\ldots r], \chi^{(r)}\rangle \neq a_0$, for all $\widetilde{a}_r = \widetilde{A}_{r+1}[r+1, 1]$, because the columns $2, \ldots, r+1$ are linearly independent: $\widetilde{A}_{r+1}[2\ldots r+1, 2\ldots r+1] = A_r$ is non-singular. The discrepancy prevents generating column $r+1$ in $\widetilde{A}_{r+1}$ and therefore $\widetilde{A}_{r+1}$ is non-singular. Note that $\chi_1^{(r)} = 0$ is possible, but then A_{r-1} is non-singular [Brent, Gustavson, and Yun 1980, Section 6(c), Lemma 4].
$\widetilde{A}_{r+1}[r+1, 1] = \widetilde{a}_r \leftarrow$ *an arbitrary element in* K; *if known,* a_r.

Sh5: Now $\widetilde{A}_{r+1}$ is non-singular. *Perform Step* LDH3a *in Algorithm* 4.1 *and from* $\chi^{(r)}$ *and* $\varphi^{(r)}$ *compute vectors* $\widetilde{\varphi}^{(r+1)}$, $\widetilde{\gamma}^{(r+1)} \in \mathsf{K}^{r+1}$ *which are the first and last column of* $\widetilde{A}_{r+1}^{-1}$. We have $\widetilde{\varphi}_1^{(r+1)} = \widetilde{\gamma}_{r+1}^{(r+1)} = \det(A_r)/\det(\widetilde{A}_{r+1}) \neq 0$.

Sh6: (a) *Compute the vector* $\widetilde{\psi}^{(r+1)} \in \mathsf{K}^{r+1}$ *as*

$$\widetilde{A}_{r+1}\underbrace{\Bigg(\begin{bmatrix}\psi^{(r)}\\ 0\end{bmatrix} + \Big(b_{r+1} - \Big\langle \widetilde{A}_{r+1}[r+1,*], \begin{bmatrix}\psi^{(r)}\\ 0\end{bmatrix}\Big\rangle\Big)\,\widetilde{\gamma}^{(r+1)}\Bigg)}_{\widetilde{\psi}^{(r+1)}} = \begin{bmatrix} b_1\\ \vdots\\ b_{r+1}\end{bmatrix};$$

(b) *Compute the vector* $\widehat{\psi}^{(r)} \in \mathsf{K}^r$ *as*

$$\widetilde{A}_{r+1}\underbrace{\Bigg(\widetilde{\psi}^{(r+1)} - \frac{\widetilde{\psi}_1^{(r+1)}}{\widetilde{\varphi}_1^{(r+1)}}\,\widetilde{\varphi}^{(r+1)}\Bigg)}_{\begin{bmatrix}0\\ \widehat{\psi}^{(r)}\end{bmatrix}} = \begin{bmatrix} *\\ b_2\\ \vdots\\ b_{r+1}\end{bmatrix}.$$

Algorithm 6.1 computes the vectors $\widehat{\psi}^{(r)}$ and $\bar{\psi}^{(r)}$ in $O(r)$ operations in K using $O(r)$ space. Algorithm 4.1 can now be extended to compute in $O((n+1-r)\,n)$ arithmetic operations with $O(n)$ additional space the solution vector $\psi^{(n+1)} = A_{n+1}^{-1}b^{(n+1)}$ for a right-side vector $b^{(n+1)} \in \mathsf{K}^{n+1}$; the solution vector $\psi^{(r)} = A_r^{-1}b^{(r)}$ with $b_i^{(r)} = b_i^{(n+1)}$ for $1 \le i \le r$ is input. If $r = 0$, Step LDH1 computes $\psi^{(n+1)}$ by solving an upper triangular and lower triangular linear system by backward and forward substitution, as for $\chi^{(n+1)}$. If A_{r+1} is non-singular, the computation is as in Step Sh6a, replacing $\widetilde{A}_{r+1}$ by A_{r+1} and $\widetilde{\gamma}^{(r+1)}$ by $\gamma^{(r+1)}$.

If $\ell \ge 1$ we extend Steps LDH5–8 as follows.

LDH5+: *By iterating Algorithm* 6.1 *ℓ-times, compute*

$$\widehat{\psi} \leftarrow \begin{bmatrix}0^\ell\\ \widehat{\psi}^{(r)}\\ 0^m\end{bmatrix} \quad \text{with} \quad A_r\widehat{\psi}^{(r)} = \begin{bmatrix}b_{1+\ell}\\ \vdots\\ b_{r+\ell}\end{bmatrix}.$$

Note that the preparatory Step Sh5 for $\chi_1^{(r)} = 0$ needs to be done only once. The cost of ℓ shifts by Algorithm 6.1 is $O(\ell\, r)$ arithmetic operations. For $\ell \ge r$, one instead can use the Gohberg-Semencul formulas [Brent, Gustavson, and Yun 1980, Section 6].

LDH6+: *For* $k \leftarrow \ell, \ell-1, \ell-2, \ldots, 1$ *Do*
$\{\ \widehat{\Delta} \leftarrow \langle A_{n+1}[k,*], \widehat{\psi}\rangle - b_k;\ \widehat{\psi} \leftarrow \widehat{\psi} - \frac{\widehat{\Delta}}{\Delta_{\text{top}}} V_{n+1}[*,k];\ \}$

LDH7+: *For* $k \leftarrow \ell+r+1, \ell+r+2, \ldots, \ell+r+m$ *Do*
$\{\ \widehat{\Delta} \leftarrow \langle A_{n+1}[k,*], \widehat{\psi}\rangle - b_k;\ \widehat{\psi} \leftarrow \widehat{\psi} - \frac{\widehat{\Delta}}{\Delta_{\text{bot}}} V_{n+1}[*,k];\ \}$

LDH8+: *Return* $\psi^{(n+1)} \leftarrow \widehat{\psi}$.
