Computing the Characteristic Polynomial of Generic Toeplitz-like and Hankel-like Matrices

Pierre Karpman
Univ. Grenoble Alpes, CNRS
LJK, UMR 5224
Grenoble, France

Clément Pernet
Univ. Grenoble Alpes, CNRS
LJK, UMR 5224
Grenoble, France

Hippolyte Signargout
Univ. Lyon, ENS de Lyon, CNRS, Inria, UCBL
LIP UMR 5668 and LJK UMR 5224
Lyon, France

Gilles Villard
Univ. Lyon, CNRS, ENS de Lyon, Inria, UCBL
LIP UMR 5668
Lyon, France

ABSTRACT

New algorithms are presented for computing annihilating polynomials of Toeplitz, Hankel, and more generally Toeplitz+Hankel-like matrices over a field. Our approach follows works on Coppersmith's block Wiedemann method with structured projections, which have been recently successfully applied for computing the bivariate resultant. A first baby steps/giant steps approach—directly derived using known techniques on structured matrices—gives a randomized Monte Carlo algorithm for the minimal polynomial of an $n \times n$ Toeplitz or Hankel-like matrix of displacement rank α using $\tilde{O}\left(n^{\omega-c(\omega)}\alpha^{c(\omega)}\right)$ arithmetic operations, where ω is the exponent of matrix multiplication and $c(2.373) \approx 0.523$ for the best known value of ω. For generic Toeplitz+Hankel-like matrices a second algorithm computes the characteristic polynomial; in particular, when the displacement rank is considered constant, its cost is $\tilde{O}\left(n^{2-1/\omega}\right)$. Previous algorithms required $O\left(n^2\right)$ operations while the exponents presented here are respectively less than 1.86 and 1.58 with the best known estimate for ω.

CCS CONCEPTS

• **Computing methodologies → Linear algebra algorithms**.

KEYWORDS

Characteristic polynomial; minimal polynomial; Toeplitz matrix; Hankel matrix; Toeplitz+Hankel-like matrix.

ACM Reference Format:
Pierre Karpman, Clément Pernet, Hippolyte Signargout, and Gilles Villard. 2021. Computing the Characteristic Polynomial of Generic Toeplitz-like and Hankel-like Matrices. In *Proceedings of the 2021 International Symposium on Symbolic and Algebraic Computation (ISSAC '21), July 18–23, 2021, Virtual Event, Russian Federation.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3452143.3465542

ISSAC '21, July 18–23, 2021, Virtual Event, Russian Federation

ACM ISBN 978-1-4503-8382-0/21/07…$15.00
https://doi.org/10.1145/3452143.3465542

1 INTRODUCTION

We consider the problem of computing the minimal or the characteristic polynomial of Toeplitz-like and Hankel-like matrices, which include Toeplitz and Hankel ones. The necessary definitions about those structures are given in Section 2.

Throughout the paper $T \in \mathbb{K}^{n\times n}$ is non-singular and either Toeplitz-like or Hankel-like, where $\mathbb{K}$ is a commutative field. The structure is parameterized by the displacement rank $1 \le \alpha \le n$ of T [12, 21]. In particular a Toeplitz or a Hankel matrix has displacement rank $\alpha = 2$.

The determinant of T can be computed in $\tilde{O}\left(\alpha^{\omega-1}n\right)$ operations in $\mathbb{K}$, where $\omega \le 3$ is a feasible exponent for square $n \times n$ matrix multiplication. For the best known value of ω one can take $\omega \approx 2.373$ [1, 20]. When T has generic rank profile (the leading principal submatrices are non singular) a complexity bound $\tilde{O}\left(\alpha^2 n\right)$ for the determinant is derived from [21, Cor. 5.3.3]. In the general case, for ensuring the rank profile one uses rank-regularization techniques initially developed in [13, 16] that lead to randomized Las Vegas algorithms assuming that the cardinality of $\mathbb{K}$ is large enough; see [21, Sec. 5] and [3] for detailed studies in our context. Taking advantage of fast matrix multiplication is possible using the results in [3, 4], where fundamental matrix operations, including the determinant, are performed in time $\tilde{O}\left(\alpha^{\omega-1}n\right)$ for a wide spectrum of displacement structures. In this approach the determinant is revealed by the recursive factorization of the inverse.

The characteristic polynomial $\det(xI_n - T)$ of T is a polynomial of degree n. Using an evaluation-interpolation scheme it follows that it can be computed in $\tilde{O}\left(\alpha^{\omega-1}n^2\right)$ operations in $\mathbb{K}$. We also refer to [21, Ch. 7] for a Newton-Structured iteration scheme in time $\tilde{O}\left(\alpha^2 n^2\right)$.

For a Toeplitz or Hankel matrix these complexity bounds for computing the characteristic polynomial were quadratic; our contribution establishes an improved bound $\tilde{O}\left(n^{2-1/\omega}\right)$ for generic matrices (given in compressed form), which is sub-quadratic including when using $\omega = 3$. We build on the results of [25] where especially the case of a Sylvester matrix was treated, and show that the approach can be generalized to larger displacement rank families. In particular, the Hankel-(like) case requires the use of sophisticated techniques in order to handle the Toeplitz+Hankel structure [7, 9] and its generalizations [21].

The algorithms we propose fit into the broad family of Coppersmith's block Wiedemann algorithms ; we refer to [17] for the necessary material and detailed considerations on the approach. Another interpretation in terms of structured lifting and matrix fraction reconstruction is given in [25].

From $T \in \mathbb{K}^{n\times n}$, the problem is to compute the determinant (or a divisor) of the characteristic matrix $M(x) = xI_n - T$. For $1 \le m \le n$ and well chosen projection matrices V and W in $\mathbb{K}^{n\times m}$, the principle is to reconstruct an irreducible fraction description $P(x)Q^{-1}(x)$ of $V^{\mathsf{T}}M(x)^{-1}W \in \mathbb{K}(x)^{m\times m}$, where $P, Q \in \mathbb{K}[x]^{m\times m}$, from a truncated series expansion of the fraction. The denominator matrix Q carries information on the Smith normal form of $M(x)$ [17, Thm. 2.12]. Using random V and W allows to recover the minimal polynomial of T from the largest invariant factor of $M(x)$, and for a generic matrix T the characteristic polynomial is obtained [17, 25].

The matrix Q is computed from a truncation $S^{(m)} \in \mathbb{K}[x]^{m\times m}$ of the series expansion of $V^{\mathsf{T}}M(x)^{-1}W$,

$$S^{(m)}(x) = -\sum_{k\ge 0}^{2\lceil n/m\rceil} V^{\mathsf{T}}(T^{-k-1})Wx^k, \tag{1}$$

using for example matrix fraction reconstruction [2, 6]. We will not detail these latter aspects in this paper since they can be found elsewhere in the literature: see [17, 25] for the general techniques involved; [24, Cor. 6.4] for the power series truncation; and [18] for alternative fraction reconstruction possibilities. The results we need on matrix polynomials are recalled in Section 3.

We focus on the computation of the power series terms $H_k = V^{\mathsf{T}}(T^{-1})^kW$ of Eq. (1). The idea for improving the complexity bounds is to use structured projections V and W in order to speed up the computation of the expansion, as has been done in [5, 25]. A typical choice is such that the matrix product by V and W is reduced. The central difficulty is to show that the algorithm remains correct; special choices for V and W could prevent a fraction reconstruction with appropriate cost, or give a denominator matrix Q with too little information on the invariant structure of T.

For a generic input matrix and our best exponent, in Section 5 we follow the choice of [25] and work with $V = W = X$ where $X = \begin{pmatrix} I_m & 0\end{pmatrix}^{\mathsf{T}} \in \mathbb{K}^{n\times m}$. An $n \times n$ Toeplitz or a Hankel matrix is defined by $2n-1$ elements of $\mathbb{K}$, and our algorithm is correct except on a certain hypersurface of $\mathbb{K}^{2n-1}$. The same way, a Toeplitz-like or Hankel-like matrix of displacement rank α is defined by the $2n\alpha$ coefficients of its generators, and our algorithm is correct for all values of $\mathbb{K}^{2n\alpha}$ except for a hypersurface. If T is Hankel, the matrix $M(x) = xI_n - T$ is Toeplitz+Hankel and the algorithm involves a compressed form that generalizes the use of generators associated to displacement operators [9, 21]. The algorithm computes a compressed representation of $M(x)^{-1}$ modulo $x^{2\lceil n/m\rceil+1}$, and exploits its structure to truncate it into a compressed representation of $S^{(m)}(x) = X^{\mathsf{T}}M(x)^{-1}X \mod x^{2\lceil n/m\rceil+1}$ at no cost. The parameter m can be optimised to get an algorithm using $\tilde{O}\left(n^{2-1/\omega}\right)$ operations when the displacement rank is considered constant.

Before considering the fast algorithm for the generic case, in Section 4 we consider the baby steps/giant steps algorithm of [17]. Indeed, thanks to the incorporation of fast matrix multiplication in basis structured matrix operations [3, 4], the overall approach with dense projections V and W already allows a slight exponent improvement. Taking into account that the input matrix T is structured, a direct cost analysis of the algorithm of [17] improves on the quadratic cost for Toeplitz and Hankel matrices as soon as one takes $\omega < 3$. However it is unclear to us how to compute the characteristic polynomial in this case (see the related Open Problem 3 in [14]). The algorithm we propose is randomized Monte Carlo and we compute the minimal polynomial in $\tilde{O}\left(n^{\omega-c(\omega)}\right)$ operations with $c(\omega) = \frac{\omega-1}{5-\omega}$. For Toeplitz-like and Hankel-like matrices with displacement rank α, the cost is multiplied by $\tilde{O}\left(\alpha^{c(\omega)}\right)$.

Notation. Indices of matrix and vectors start from zero. The vectors of the n-dimensional canonical basis are denoted by $e_0^n, \ldots, e_{n-1}^n$. For a matrix M, $M_{i,j}$ denotes the coefficient (i, j) of this matrix, $M_{i,*}$, its row of index i and $M_{*,j}$ its column of index j.

2 RANK DISPLACEMENT STRUCTURES

A wide range of structured matrices are efficiently described by the action of a displacement operator [12]. There are two types of such operators: the Sylvester operators of the form

$$\nabla_{M,N} : A \mapsto MA - AN,$$

and the Stein operators of the form

$$\Delta_{M,N} : A \mapsto A - MAN,$$

where M and N are fixed matrices. A Toeplitz matrix T is defined by $2n-1$ coefficients $t_{-n+1}, \ldots, t_{n-1} \in \mathbb{K}$ such that $T = (t_{i-j})_{i,j}$. Its image through $\Delta_{Z_n,Z_n^{\mathsf{T}}}$, where $Z_n = (\delta_{i,j+1})_{0\le i,j\le n-1}$ has rank at most 2. Similarly, a Hankel matrix H is defined by $2n-1$ coefficients $h_0, \ldots, h_{2n-2}$ such that $H = (h_{i+j})_{i,j}$ and its image through $\nabla_{Z_n,Z_{n,1}^{\mathsf{T}}}$, where $Z_{n,1} = Z_n + e_0^n e_{n-1}^{n\mathsf{T}}$ has rank at most 2.

As a generalization, the class of Toeplitz-like (resp. Hankel-like) matrices is defined as those matrices whose image through $\Delta_{Z_n,Z_n^{\mathsf{T}}}$ (resp. $\nabla_{Z_n,Z_{n,1}^{\mathsf{T}}}$) has a bounded rank α [8, 21], called the *displacement rank*, and can be represented by a product GH^{T}, where $G, H \in \mathbb{K}^{n\times\alpha}$ are called *generators*. These operators are non-singular and a matrix can be uniquely recovered from its generators.

Lastly, any sum of a Toeplitz and a Hankel matrix, (forming the class of Toeplitz+Hankel matrices) has an image of rank at most 4 through the displacement operator ∇_{U_n,U_n} where $U_n = Z_n + Z_n^{\mathsf{T}}$ [7]. This operator is singular and the low rank image does not suffice to uniquely reconstruct the initial matrix: additional data (usually a first or a last column) is required for a unique reconstruction [9, 21]. The class of Toeplitz+Hankel-like matrices is formed by those matrices whose image through ∇_{U_n,U_n} has a bounded rank.

2.1 Product of Structured Matrices

Proposition 2.1 ([3, Theorem 1.2]). *Let $A \in \mathbb{K}^{n\times n}$ be a Toeplitz-like or Hankel-like matrix with displacement rank α given by its generators and $B \in \mathbb{K}^{n\times m}$ be a dense matrix. The multiplication of A by B can be computed in $\tilde{O}\left(n\max(\alpha, m)\min(\alpha, m)^{\omega-2}\right)$ operations in $\mathbb{K}$.*

Proposition 2.2. *Let $A, B \in \mathbb{K}^{n\times n}$ be two Toeplitz-like matrices of displacement rank α and β respectively, then their product AB is a*

Toeplitz-like matrix of displacement rank at most $\alpha+\beta+1$. Furthermore, given generators for A and B w.r.t. $\Delta_{Z_n,Z_n^\mathsf{T}}$, one can compute generators for AB w.r.t. the same operator in $\tilde{O}\left(n(\alpha+\beta)^{\omega-1}\right)$ field operations.

Proof. Let G_A, H_A and G_B, H_B be the generators of A and B respectively. They satisfy A-$Z_nAZ_n^\mathsf{T} = G_AH_A^\mathsf{T}$ and $B - Z_nBZ_n^\mathsf{T} = G_BH_B^\mathsf{T}$. Consequently

$$\begin{aligned} AB = & (Z_nAZ_n^\mathsf{T} + G_AH_A^\mathsf{T})(Z_nBZ_n^\mathsf{T} + G_BH_B^\mathsf{T}) \\ = & Z_nABZ_n^\mathsf{T} - Z_nA_{*,n-1}B_{n-1,*}Z_n^\mathsf{T} + (Z_nAZ_n^\mathsf{T}G_B)H_B^\mathsf{T} \\ & + G_A(H_A^\mathsf{T}Z_nBZ_n^\mathsf{T} + H_A^\mathsf{T}G_BH_B^\mathsf{T}), \end{aligned}$$

and therefore $AB - Z_nABZ_n^\mathsf{T} = G_{AB}H_{AB}^\mathsf{T}$ for

$$G_{AB} = \left(\begin{array}{c|c|c} G_A & Z_nAZ_n^\mathsf{T}G_B & -Z_nA_{*,n-1} \end{array} \right)$$

$$H_{AB} = \left(\begin{array}{c|c|c} Z_nB^\mathsf{T}Z_n^\mathsf{T}H_A + H_BG_B^\mathsf{T}H_A & H_B & Z_nB_{n-1,*}^\mathsf{T} \end{array} \right),$$

thus showing that AB has displacement rank at most $\alpha+\beta+1$.

Computing these generators involves applying A on a dense $n\times\beta$ matrix and B on a dense $\alpha\times n$ matrix, and computing the product of an $\alpha\times n$ by an $n\times\beta$ matrix and the product of an $\alpha\times\beta$ by a $\beta\times n$ matrix. Using [3, Thm 1.2], these cost $\tilde{O}\left(n(\alpha+\beta)^{\omega-1}\right)$ field operations. □

Proposition 2.3. *Let $A, B \in \mathbb{K}^{n\times n}$ be two Hankel-like matrices of displacement rank α and β respectively, then their product AB is a Toeplitz-like matrix of displacement rank at most $\alpha+\beta+2$. Furthermore, given generators for A and B w.r.t. $\nabla_{Z_n,Z_{n,1}^\mathsf{T}}$, generators for AB w.r.t. $\Delta_{Z_n,Z_n^\mathsf{T}}$ can be computed in $\tilde{O}\left(n(\alpha+\beta)^{\omega-1}\right)$.*

Proof. Let G_A, H_A and G_B, H_B be the generators of A and B respectively, satisfying $Z_nA - AZ_{n,1}^\mathsf{T} = G_AH_A^\mathsf{T}$ and $Z_nB - BZ_{n,1}^\mathsf{T} = G_BH_B^\mathsf{T}$. Using a similar reasoning as for Proposition 2.2 we can deduce that

$AB - Z_nABZ_n^\mathsf{T} = G_{AB}H_{AB}^\mathsf{T}$ for

$$G_{AB} = \left(\begin{array}{c|c|c|c} G_A & AZ_{n,1}^\mathsf{T}G_B & A_{*,n-1} & AZ_{n,1}^\mathsf{T}B_{*,n-1} \end{array} \right)$$

$$H_{AB} = \left(\begin{array}{c|c|c|c} \left(H_BG_B^\mathsf{T} - B^\mathsf{T}Z_n^\mathsf{T} + e_0^nB_{*,n-1}^\mathsf{T}\right)H_A & H_B & B_{n-1,*}^\mathsf{T} & e_0^n \end{array} \right),$$

thus showing that AB has displacement rank at most $\alpha+\beta+2$. Computing these generators again costs $\tilde{O}\left(n(\alpha+\beta)^{\omega-1}\right)$ field operations. □

Proposition 2.4. *Let $A \in \mathbb{K}^{n\times n}$ be a Toeplitz-like (resp. Hankel-like) matrix of displacement rank α, then for an arbitrary (resp. even)r, A^r is a Toeplitz-like matrix of displacement rank at most $(\alpha+1)r$ and its generators can be computed from the generators of A in $\tilde{O}\left(n(\alpha r)^{\omega-1}\right)$ field operations.*

Proof. Using fast exponentiation one computes A^r as:

$$A^r = \prod_{k=0}^{\lfloor \log r \rfloor} \left(A^{2^k}\right)^{l_k} \text{ where the } l_k \text{ satisfy } \sum_{k=0}^{\log r} l_k 2^k = r,$$

which only requires squarings and products between matrices of the form A^{2^k}. When A is Toeplitz-like the result is a straightforward consequence of Proposition 2.2; when it is Hankel-like the product A^2 is computed using Proposition 2.3, the remaining products are between Toeplitz-like matrices, and the result again follows from Proposition 2.2. □

2.2 Reconstruction of a Toeplitz+Hankel-like Matrix from its Generators

The operator ∇_{U_n,U_n} is defined in [21, Section 4.5] as partly regular, which means that a Toeplitz+Hankel-like matrix is completely defined by its generators *and* its *irregularity set* that may be all the entries in its first column.

A formula to recover a dense representation of the matrix from its generators and its first column is given in [21].

Theorem 2.5 ([21, Thm. 4.5.1]). *Let $M \in \mathbb{K}^{n\times n}$ be a Toeplitz+Hankel-like matrix, $G, H \in \mathbb{K}^{n\times\alpha}$ its generators and $c_0 = Me_0^n$ its first column, then*

$$M = \tau_{U_n}(c_0) - \sum_{j=0}^{\alpha-1} \tau_{U_n}(G_{*,j})\tau_{Z_n}(Z_nH_{*,j})^\mathsf{T} \tag{2}$$

where for an $n\times n$ matrix A and a vector v of length n $\tau_A(v)$ denotes the matrix of the algebra generated by A which has v as its first column.

We show that one can derive a fast reconstruction algorithm for a Toeplitz+Hankel-like matrix from Eq. (2) and first detail the structure of the various $\tau_A(v)$ matrices.

Lemma 2.6. *$\tau_{Z_n}(v)^\mathsf{T}$ is the Toeplitz upper-triangular matrix with v^T as its first row.*

Lemma 2.7. *$\tau_{U_n}(v) = \sum_{i=0}^{n-1} v_iQ_i(U_n)$ where $Q_0(x) = 1$, $Q_1(x) = x$ and $Q_{i+1}(x) = xQ_i(x) - Q_{i-1}(x)$.*

Proof. The first column of $Q_i(U_n)$ is e_i^n. □

Corollary 2.8. *Column j of $\tau_{U_n}(v)$ is $Q_j(U_n)v$.*

Proof. With Lemma 2.7 and after checking the property for $j \in \{0,1\}$, it suffices to prove $Q_i(U_n)_{*,j+1} = U_nQ_i(U_n)_{*,j} - Q_i(U_n)_{*,j-1}$. This is true for $i \in \{0,1\}$ and if it is for i and $i-1$, then

$$\begin{aligned} Q_{i+1}(U_n)_{*,j+1} = & U_n^2Q_i(U_n)_{*,j} - U_nQ_i(U_n)_{*,j-1} \\ & - U_nQ_{i-1}(U_n)_{*,j} + Q_{i-1}(U_n)_{*,i-1} \end{aligned}$$

□

From these we can write the following proposition, inspired by [7, Prop. 4.2]. It enables fast recursive reconstruction of the columns of a Toeplitz+Hankel-like matrix from the first one.

Proposition 2.9. *Let $M \in \mathbb{K}^{n\times n}$ be a Toeplitz+Hankel-like matrix, $G, H \in \mathbb{K}^{n\times\alpha}$ its generators for ∇_{U_n,U_n} and $c_0 = Me_0^n$ its first column. With the notation $c_{-1} = 0$, the columns $(c_k)_{0\le k\le n-1}$ of M follow the recursion:*

$$c_{k+1} = U_nc_k - c_{k-1} - \sum_{j=0}^{\alpha-1} H_{k,j}G_{*,j}. \tag{3}$$

Proof. Let C be the matrix defined by the recursion formula and initial conditions of Proposition 2.9, we will prove $C = M$.

By definition c_0 is the first column of M; assume now that for $i \le k$, c_i is column i of M. Using Lemma 2.6 and Corollary 2.8 on Eq. (2) that is

$$c_i = Q_i(U_n)c_0 - \sum_{j=0}^{\alpha-1}\sum_{l=0}^{i-1} H_{i-1-l,j}Q_l(U_n)G_{*,j} \tag{4}$$

and Eq. (3) can be detailed as

$$\begin{aligned} c_{k+1} &= U_n\left(Q_k(U_n)c_0 - \sum_{j=0}^{\alpha-1}\sum_{i=0}^{k-1} H_{k-1-i,j}Q_i(U_n)G_{*,j}\right) \\ &\quad - \left(Q_{k-1}(U_n)c_0 - \sum_{j=0}^{\alpha-1}\sum_{i=0}^{k-2} H_{k-2-i,j}Q_i(U_n)G_{*,j}\right) - \sum_{j=0}^{\alpha-1} H_{k,j}G_{*,j} \\ &= Q_{k+1}(U_n)c_0 - \sum_{j=0}^{\alpha-1}\sum_{i=0}^{k} H_{k-i,j}Q_i(U_n)G_{*,j} \end{aligned}$$

□

3 MATRIX POLYNOMIALS

We rely on the material from [17, 25]. For matrix polymonials and fractions the reader may refer to [11]. The rational matrix $H(x) = V^{\mathsf{T}}M(x)^{-1}W$ over $\mathbb{K}(x)$ can be written as a fraction of two polynomial matrices. A right fraction description is given by square polynomial matrices $P(x)$ and $Q(x)$ such that $H(x) = P(x)Q(x)^{-1} \in \mathbb{K}(x)^{m\times m}$, and a left description by $P_l(x)$ and $Q_l(x)$ such that $H(x) = Q_l(x)^{-1}P_l(x) \in \mathbb{K}(x)^{m\times m}$. Degrees of denominator matrices are minimized using column-reduced forms. A non-singular polynomial matrix is said to be column-reduced if its leading column coefficient matrix is non-singular [11, Sec. 6.3]. We also have the notion of irreducible and minimal fraction descriptions. If P and Q (resp. P_l and Q_l) have unimodular right (resp. left) matrix gcd's [11, Sec. 6.3] then the description is called irreducible. If Q (resp. Q_l) is column-reduced then the description is called minimal.

For a given m, define $1 \le \nu \le n$ to be the sum of the degrees of the first m largest invariant factors of $M(x) = xI_n - T$ (equivalently, the first m diagonal elements of its Smith normal form). The following will ensure that the minimal polynomial of T, which is the largest invariant factor of $M(x)$, can be computed from the Smith normal form of an appropriate denominator $Q(x)$.

Theorem 3.1. ([17, Thm. 2.12] and [24].) *Let V and W be block vectors over a sufficiently large field $\mathbb{K}$ whose entries are sampled uniformly and independently from a finite subset $S \subseteq \mathbb{K}$. Then with probability at least $1 - 2n/|S|$, $H(x) = V^{\mathsf{T}}M(x)^{-1}W$ has left and right irreducible descriptions with denominators of degree $\lceil \nu/m \rceil$, of determinantal degree ν, and whose i^{th} invariant factor (starting from the largest degree) is the i^{th} invariant factor of $M(x)$.*

The next result we need is concerned with the computation of an appropriate denominator Q as soon as the truncated power series in Eq. (1) is known. We notice that $H(x) = V^{\mathsf{T}}M(x)^{-1}W$ is strictly proper in that it tends to zero when x tends to infinity. For fraction reconstruction we use the computation of minimal approximant bases (or σ-bases) [2, 23], and the algorithm with complexity bound $\tilde{O}\left(m^{\omega-1}n\right)$ in [6, 10].

Theorem 3.2. ([6, Lem. 3.7].) *Let $H \in \mathbb{K}(x)^{m\times m}$ be a strictly proper power series, with left and right matrix fractions descriptions of degree at most d. A denominator Q of a right irreducible description $H(x) = P(x)Q(x)^{-1}$ can be computed in $\tilde{O}\left(m^{\omega}d\right)$ arithmetic operations from the first $2d+1$ terms of the expansion of H.*

In our case, from Theorem 3.1 we will obtain the existence of appropriate fractions of degree less than $\lceil n/m \rceil$, and use Theorem 3.2 for bounding the cost of the computation of Q.

4 A BABY-STEP GIANT STEP ALGORITHM

In this section, we propose a direct adaptation of the baby steps/giant steps variant of Coppersmith's block-Wiedemann algorithm developped in [17, Sec. 4] to the case of structured matrices. In order to compute the terms of the series (1), we will assume that the input matrix T has been inverted, using [3, Theorem 6.6]. In this section we will therefore denote by T this inverse and compute the projections of its powers.

4.1 Description of the Algorithm

Let $V, W \in \mathbb{K}^{n\times m}$ be the block vectors used for the projections. Algorithm 1 performs r baby steps and s giant steps to compute the first terms of the sequence $H_k = V^{\mathsf{T}}T^{k+1}W = V^{\mathsf{T}}(T^r)^jT^{i+1}W$ for $0 \le k \le 2\lceil n/m \rceil$, $0 \le i < r$, $0 \le j < s$ and $rs \ge 2\lceil n/m \rceil + 1$.

Algorithm 1 [17] Compute $H_k = V^{\mathsf{T}}T^{k+1}W$ for $0 \le k \le 2\lceil n/m \rceil$

Input: Generators of $T \in \mathbb{K}^{n\times n}$, Toeplitz-like or Hankel-like
Input: $m, r, s \in \mathbb{N}$ s.t. $rs \ge 2\lceil n/m \rceil + 1$, r even if T is Hankel-like
Input: $V, W \in \mathbb{K}^{n\times m}$
Output: $H = \left(H_{rj+i}\right)_{j<s,i<r}$ where $H_k = V^{\mathsf{T}}T^{k+1}W$

1: $W_0 \leftarrow TW$
2: **for** $1 \le i \le r-1$ **do**
3: $\quad W_i \leftarrow TW_{i-1}$
4: $R \leftarrow T^r$
5: $V_0 \leftarrow V$
6: **for** $1 \le j \le s-1$ **do**
7: $\quad V_j^{\mathsf{T}} \leftarrow V_{j-1}^{\mathsf{T}}R$
8: $H \leftarrow \begin{pmatrix} V_0 & \dots & V_{s-1} \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} W_0 & \dots & W_{r-1} \end{pmatrix}$

This algorithm relies on three main matrix operations:

(1) The product of a structured matrix by a dense rectangular matrix, supported by Proposition 2.1 for Steps 3 and 7;
(2) The exponentiation of a structured matrix, supported by Proposition 2.4 for Step 4;
(3) The product of two dense rectangular matrices for Step 8.

4.2 Cost Analysis

Proposition 4.1. *Algorithm 1 runs in $\tilde{O}\left(n^{\omega-\frac{\omega-1}{5-\omega}}\alpha^{\frac{\omega-1}{5-\omega}}\right)$ operations in $\mathbb{K}$ for well chosen m, r and s.*

Proof. From Proposition 2.1, applying an $n \times m$ block to T can be done in $\tilde{O}\left(n\max(m,\alpha)\min(m,\alpha)^{\omega-2}\right)$ field operations. Hence

the r baby steps, Step 3, computing the $(T^iW)_{0\le i<r}$ cost overall

$$\tilde{O}\left(nr\max(m,\alpha)\min(m,\alpha)^{\omega-2}\right) \quad (5)$$

field operations.

By Proposition 2.4, the initialization of the giant steps at Step 4 is the computation of a structured representation for T^r, which can be done in

$$\tilde{O}\left(nr^{\omega-1}\alpha^{\omega-1}\right) \quad (6)$$

operations in $\mathbb{K}$.

Then each of the giant steps, at Step 7, is a product of an $m\times n$ dense matrix by an $n\times n$ matrix of displacement rank αr. From Proposition 2.1, these s steps cost

$$\tilde{O}\left(ns\max(m,\alpha r)\min(m,\alpha r)^{\omega-2}\right) \quad (7)$$

Lastly, the computation of the product resulting in H at Step 8 uses $\tilde{O}\left(n\max(mr,ms)\min(mr,ms)^{\omega-2}\right)$ or equivalently

$$\tilde{O}\left(nm^{\omega-1}\max(r,s)\min(r,s)^{\omega-2}\right) \quad (8)$$

field operations.

Let $m=\left\lceil n^{\frac{\omega-3}{\omega-5}}\alpha^{\frac{2}{5-\omega}}\right\rceil$ and set $r=s=\left\lceil\sqrt{2n/m}\right\rceil$. Note that $\alpha\le m\le\alpha r$, therefore the bound of Eq. (5) is dominated by the one of Eq. (8). Moreover the bound of Eq. (7) can be rewritten as $\tilde{O}\left(n^2m^{\omega-3}\alpha\right)$, and from Eq. (8) we have $\tilde{O}\left(n^{\frac{\omega+1}{2}}m^{\frac{\omega-1}{2}}\right)$, and these two quantities are

$$\tilde{O}\left(n^{\omega-\frac{\omega-1}{5-\omega}}\alpha^{\frac{\omega-1}{5-\omega}}\right).$$

Finally, the bound of Eq. (6) can be rewritten as $\tilde{O}\left(n^{\frac{\omega+1}{2}}(\frac{\alpha^2}{m})^{\frac{\omega-1}{2}}\right)$, which is dominated by the one of Eq. (8). □

When the displacement rank α is constant, and with the best known estimate $\omega=2.373$ [1, 20] the cost bound given in Proposition 4.1 becomes $\tilde{O}\left(n^{1.851}\right)$, while it is $\tilde{O}\left(n^2\right)$ for $\omega=3$.

Let us now suppose that the entries of V and W are sampled uniformly and independently from a finite subset $S\subseteq\mathbb{K}$, we then have the following.

Theorem 4.2. *The minimal polynomial of an $n\times n$ Toeplitz-like or Hankel-like matrix with displacement rank α can be computed by a randomized Monte Carlo algorithm using*

$$\tilde{O}\left(n^{\omega-\frac{\omega-1}{5-\omega}}\alpha^{\frac{\omega-1}{5-\omega}}\right)$$

field operations, with probability of success at least $1-(n^2+3n+2n^{5/3}\alpha^2)/|S|$.

Proof. The first step is to compute the inverse of T, using [3, Theorem 6.6] in $\tilde{O}\left(n\alpha^{\omega-1}\right)$ operations in $\mathbb{K}$. Then running Algorithm 1 on T^{-1} costs $\tilde{O}\left(n^{\omega-\frac{\omega-1}{5-\omega}}\alpha^{\frac{\omega-1}{5-\omega}}\right)$ which dominates $\tilde{O}\left(n\alpha^{\omega-1}\right)$ since $\alpha\le n$. From the sequence of matrices $(H_k)_{0\le k\le 2n/m}$, one can compute a minimal denominator Q for $H(x)=V^{\mathsf{T}}(xI_n-T)^{-1}W\in\mathbb{K}[x]^{m\times m}$ in $\tilde{O}\left(nm^{\omega-1}\right)$ field operations, by Theorem 3.2.

Using Theorem 3.1, the minimal polynomial is then obtained as the first invariant factor in the Smith form of Q, computed by [22, Proposition 41]. This step also costs $\tilde{O}\left(nm^{\omega-1}\right)$ field operations and since $m\le n$ we have

$$nm^{\omega-1}\le n^{\frac{\omega+1}{2}}m^{\frac{\omega-1}{2}}$$

which shows that the cost of these last two computations will always be dominated by the cost of the product in Eq. (8). The probability of failure for the computation of T^{-1} is at most $n(n+1)/|S|$ by [3, Lemma 6.2]. For the computation of the minimal polynomial it is at most $2m^2n\le 2n^{5/3}\alpha^2$, from [22, Concl.] and [15, Thm 3.3]. A union bound combining this probability and the failure probablity of Theorem 3.1 yields a probability of failure of $(n^2+3n+2n^{5/3}\alpha^2)/|S|$. □

Note that this result carries over to the computation of the characteristic polynomial of any Toeplitz-like or Hankel-like matrix T having fewer than m invariant factors in its Frobenius normal form.

5 USING STRUCTURED INVERSION

In this section we develop a new approach for computing the characteristic polynomial of generic structured polynomial matrices $T\in\mathbb{K}^{n\times n}$ with displacement rank α. Following [25, Sec. 7], in the Toeplitz-like case the idea is to exploit the structure of the ΣLU representation [12]. For Hankel-like matrices (see the discussion after Theorem 5.4), we generalize the approach using both generators and irregularity sets that has been introduced in Section 2.2 [21].

Principle of the approach. Here, rather than using successive applications and powering of T^{-1} as in Section 4, the first terms of the sequence sequence $\{H_k\}_k=\{V^{\mathsf{T}}T^{-k-1}W\}_k$ are obtained as the matrix coefficients of the series expansion of $V^{\mathsf{T}}M(x)^{-1}W$. Since $2\lceil n/m\rceil+1$ terms are required, and with the special choice $V=W=X=\begin{pmatrix}I_m & 0\end{pmatrix}^{\mathsf{T}}\in\mathbb{K}^{n\times m}$, this boils down to computing a dense representation of the $m\times m$ leading principal submatrix of $M(x)^{-1}\bmod x^{2\lceil n/m\rceil+1}$. The outline of the algorithm is as follows.

(1) Compute the inverse $M(x)^{-1}\bmod x^{2\lceil n/m\rceil+1}$ in a compressed representation;
(2) Crop this representation to form a representation of the $m\times m$ leading principal submatrix;
(3) Extract $S^{(m)}(x)=X^{\mathsf{T}}M(x)^{-1}X\bmod x^{2\lceil n/m\rceil+1}$ in dense form.

Below we specialize the approach for the two classes of interest. Our algorithms in Theorems 5.2 and 5.4 are correct for generic matrices T (in the Zariski sense), see Assumptions (A1) and (A2) in Section 6 to which the discussion on genericity is deferred.

5.1 Generic Toeplitz-like Matrices

If T is Toeplitz-like, so it is $M(x)=xI_n-T$. If $M(x)$ is represented in ΣLU form by generators $G,H\in\mathbb{K}[x]^{n\times\alpha}$ such that $M(x)=\sum_{i=0}^{\alpha-1}L(G_{*,i})L(H_{*,i})^T$, where $L(v)$ is the lower triangular Toeplitz matrix with v as its first column [12, 13]. The $m\times m$ leading principal submatrix of any product $L(v)L(w)^{\mathsf{T}}$ is the product of the $m\times m$ leading principal submatrix of these factors, which in turn is $L(v_{0..m-1})L(w_{0..m-1})^{\mathsf{T}}$. Algorithm 2 relies on this property to produce $S^{(m)}$ from the m first rows of the generators of M^{-1}.

Proposition 5.1. *Algorithm 2 is correct for $M(x)=xI_n-T$; if T has generic rank profile it uses*

$$\tilde{O}\left(\frac{n^2}{m}\alpha^{\omega-1}+nm\alpha\right)$$

operations in $\mathbb{K}$.

Algorithm 2 Compute $S^{(m)}$: Toeplitz-like case

Input: (G, H) generators of $M \in \mathbb{K}[x]^{n\times n}$, a Toeplitz-like matrix of displacement rank α
Output: $S^{(m)} = X^{\mathsf{T}}M^{-1}X \bmod x^{2\lceil n/m\rceil+1}$ in dense form
1: $(E, F) \leftarrow$ generators for $M^{-1} \bmod x^{2\lceil n/m\rceil+1}$
2: $E' \leftarrow X^{\mathsf{T}}E$; $F' \leftarrow FX$
3: $S^{(m)} \leftarrow \sum_{i=0}^{\alpha-1} L(E'_{*,i})L(F'_{*,i})^{\mathsf{T}} \mod x^{2\lceil n/m\rceil+1}$

Proof. From the discussion at the beginning of the section, $E' = E_{0..m-1,*}$ and $F' = F_{0..m-1,*}$ are generators for $S^{(m)} = X^{\mathsf{T}}M^{-1}X$. We use the algorithm of [4, Prop. 5] for computing the generators of the inverse. Note that no division by x in the ring $\mathbb{K}[x]/\langle x^{2\lceil n/m\rceil+1}\rangle$ will occur in Step 1 since $M(0) = T$ has generic rank profile, and consequently all leading principal minors of $M(x)$ are not divisible by x which shows the correctness.

By [4, Prop. 5], computing the generators of M^{-1} at Step 1 can be done in $\tilde{O}\left(n\alpha^{\omega-1}\right)$ operations over $\mathbb{K}[x]/\langle x^{2\lceil n/m\rceil+1}\rangle$ which in turn is

$$\tilde{O}\left(\frac{n^2}{m}\alpha^{\omega-1}\right) \tag{9}$$

operations in $\mathbb{K}$.

The dense reconstruction of $S^{(m)}$ in Step 3 is achieved by α products of an $m \times m$ Toeplitz matrix $L(E'_{*,i})$ by an $m \times m$ dense matrix $L(F'_{*,i})^{\mathsf{T}}$ for a total cost of

$$\tilde{O}\left(nm\alpha\right) \tag{10}$$

operations in $\mathbb{K}$. □

From the efficient computation of the first terms of the expansion of $X^{\mathsf{T}}M(x)^{-1}X$ and using fraction reconstruction, the characteristic polynomial of T is obtained.

Theorem 5.2. *The characteristic polynomial of a generic $n \times n$ Toeplitz-like matrix with displacement rank α (assumptions (A1) and (A2) in Section 6) can be computed in $\tilde{O}\left(n^{2-\frac{1}{\omega}}\alpha^{\frac{(\omega-1)^2}{\omega}}\right)$ operations in $\mathbb{K}$ when $\alpha = O\left(n^{\frac{\omega-2}{-\omega^2+4\omega-2}}\right)$, and $\tilde{O}\left(n^{\frac{3}{2}}\alpha^{\frac{\omega}{2}}\right)$ otherwise.*

Proof. From Lemma 6.1 (genericity assumption (A2)), irreducible left and right fractions descriptions of $X^{\mathsf{T}}M^{-1}X$ have degree at most $\lceil n/m\rceil$. Thus Theorem 3.2 ensures that a denominator Q of a right description can be computed from $S^{(m)}(x) = X^{\mathsf{T}}M(x)^{-1}X \bmod x^{2\lceil n/m\rceil+1}$. By Lemma 6.1 again, the determinant of Q gives the characteristic polynomial of T.

Besides the computation of $S^{(m)}$ by Proposition 5.1 (genericity assumption (A1)), the computation of the denominator Q of its irreducible right fraction description costs

$$\tilde{O}\left(nm^{\omega-1}\right) \tag{11}$$

operations by Theorem 3.2. Computing the determinant of Q has same cost using the algorithm in [19]. The total cost depends on α.

Case 1: $\alpha = O\left(n^{\frac{\omega-2}{-\omega^2+4\omega-2}}\right)$. We set $m = n^{\frac{1}{\omega}}\alpha^{\frac{\omega-1}{\omega}}$ so that $\alpha = O\left(m^{\omega-2}\right)$ and the term in Eq. (10) is dominated by the one in Eq. (11). For the chosen value of m the terms in Eq. (9) (decreasing in m) and Eq. (11) (increasing in m) are equal, leading to a full cost of $\tilde{O}\left(n^{2-\frac{1}{\omega}}\alpha^{\frac{(\omega-1)^2}{\omega}}\right)$ operations in $\mathbb{K}$.

Case 2: $\alpha = \Omega\left(n^{\frac{\omega-2}{-\omega^2+4\omega-2}}\right)$. We set $m = n^{\frac{1}{2}}\alpha^{\frac{\omega-2}{2}}$ so that $\alpha = \Omega\left(m^{\omega-2}\right)$. In this case the term in Eq. (11) is dominated by the one in Eq. (10) and for this value of m we have equality between the terms in Eq. (9) and Eq. (10), leading to a full cost of $\tilde{O}\left(n^{\frac{3}{2}}\alpha^{\frac{\omega}{2}}\right)$ operations in $\mathbb{K}$. □

The exponent in Theorem 5.2 is $O\left(n^{1.579}\right)$ (resp. $O\left(n^{1.667}\right)$) for α constant and $\omega = 2.373$ (resp. $\omega = 3$). When $\alpha = \Theta\left(n^{\frac{\omega-2}{-\omega^2+4\omega-2}}\right)$ and taking $\omega = 2.373$ (resp. $\omega = 3$), both expressions become $\tilde{O}\left(n^{1.74}\right)$ (resp. $\tilde{O}\left(n^3\right)$). The complexity bound when α is small can also be written as

$$\tilde{O}\left(n^{\omega-f(\omega)}\alpha^{f(\omega)}\right),$$

similarly as in Proposition 4.1, which can be interpreted as a transfer of part of the exponent from n to α by using the structure of the matrix.

5.2 Generic Toeplitz+Hankel-like Matrices

We now adapt the previous approach to more general structures. If T is Hankel-like then $M(x) = xI_n - T$ is Toeplitz+Hankel-like. In this section we consider generic matrices with such a structure.

Compared to the Toeplitz case in Section 5.1, only the computation of the truncated expansion of $X^{\mathsf{T}}M(x)^{-1}X$ is modified. Computing the characteristic polynomial from there does not depend on the structure of M or T (dense matrix polynomial operations).

In addition to the generators one has to consider an irregularity set for M^{-1}. These data are computed by Algorithm 3 at Step 1 using the recursive matrix decomposition in [21, Ch. 5]. The irregularity set we consider is the first column. The dense form of $S^{(m)}(x) = X^{\mathsf{T}}M(x)^{-1}X \bmod x^{2\lceil n/m\rceil+1}$ is then recovered from its compressed representation using Proposition 2.9.

Algorithm 3 Compute $S^{(m)}$: Toeplitz+Hankel-like case

Input: (G, H, v) generators and irregularity set of $M \in \mathbb{K}[x]^{n\times n}$, a Toeplitz+Hankel-like matrix of displacement rank α.
Output: $S^{(m)} = X^{\mathsf{T}}M^{-1}X \bmod x^{2\lceil n/m\rceil+1}$ in dense form
1: $(E, F, c) \leftarrow$ generators and irregularity set for M^{-1}, the irregularity set is the first column ($Mc = e_0^n$)
2: $c_0 \leftarrow \begin{pmatrix} I_{2m-1} & 0 \end{pmatrix} c$
3: $c_1 \leftarrow U_{2m-1}c_0 - \sum_{j=0}^{\alpha-1} E_{0,j}F_{0\ldots2m-2,j}$
4: **for** $1 \le k \le m-2$ **do**
5: $\quad c_{k+1} \leftarrow U_{2m-1}c_k - c_{k-1} - \sum_{j=0}^{\alpha-1} F_{k,j}E_{0\ldots2m-2,j}$
6: $S^{(m)}(x) = \begin{pmatrix} I_m & 0 \end{pmatrix} (c_0 \cdots c_{m-1})$

Proposition 5.3. *Algorithm 3 is correct for $M(x) = xI_n - T$; if T has generic rank profile it uses*

$$\tilde{O}\left(\frac{n^2}{m}\alpha^2 + mn\alpha\right)$$

operations in $\mathbb{K}$.

Proof. As discussed in the proof of Proposition 5.1, no division by x occur in the ring $\mathbb{K}[x]/\langle x^{2\lceil n/m\rceil+1}\rangle$ since since $M(0) = T$ has generic rank profile. Step 1 can be performed in $\tilde{O}(\alpha^2 n)$ operations on truncated power series, so $\tilde{O}\left(\frac{n^2}{m}\alpha^2\right)$ operations in $\mathbb{K}$ [21, Corollary 5.3.3]. Each step of the for loop consists of a number of polynomial operations modulo $x^{2\lceil n/m\rceil+1}$ linear in $m\alpha$ as U_{2m-1} has only two non-zero entries on each row. Lines 2 to 5 can be performed in $\tilde{O}(m^2\alpha)$ power series operations, so $\tilde{O}(nm\alpha)$ operations in $\mathbb{K}$. By Proposition 2.9, if the first $2m-k$ coefficients of c_{k-1} are equal to the ones of column $k-1$ of M^{-1}, then the first $2m-k-1$ coefficients of c_k are equal to the ones of column k of M^{-1}. Since c_0 gives the $2m-1$ first coefficients of column 0 of M^{-1}, Step 6 outputs $S^{(m)}(x)$. □

The characteristic polynomial is then obtained following our general strategy.

Theorem 5.4. *The characteristic polynomial of a generic $n \times n$ Toeplitz+Hankel-like matrix with displacement rank α (assumptions (A1) and (A2) in Section 6) can be computed in $\tilde{O}\left(n^{2-\frac{1}{\omega}}\alpha^{\frac{2(\omega-1)}{\omega}}\right)$ field operations when $\alpha = O\left(n^{\frac{\omega-2}{4-\omega}}\right)$, and $\tilde{O}\left(n^{\frac{3}{2}}\alpha^{\frac{3}{2}}\right)$ otherwise.*

Proof. The arguments are similar to those of the proof of Theorem 5.2, we do not repeat them here. We have only have to discuss the slighlty different cost bound. The overall cost is that for computing the matrix denominator Q and its determinant in $\tilde{O}\left(nm^{\omega-1}\right)$ operations in $\mathbb{K}$, plus the cost of computing the sequence $\{H_k\}_k$. We distinguish two cases:

If $\alpha = O\left(n^{\frac{\omega-2}{4-\omega}}\right)$: we take $m = n^{\frac{1}{\omega}}\alpha^{\frac{2}{\omega}}$ so that $\alpha = O\left(m^{\omega-2}\right)$, with overall cost bound $\tilde{O}\left(n^{2-\frac{1}{\omega}}\alpha^{\frac{2(\omega-1)}{\omega}}\right)$.

If $\alpha = \Omega\left(n^{\frac{\omega-2}{4-\omega}}\right)$: we take $m = n^{\frac{1}{2}}\alpha^{\frac{1}{2}}$ so that $\alpha = \Omega\left(m^{\omega-2}\right)$, with overall cost bound $\tilde{O}\left(n^{\frac{3}{2}}\alpha^{\frac{3}{2}}\right)$. □

Given $\nabla_{Z_n,Z_{n,1}^{\mathsf{T}}}$-generators of length α for a Hankel-like matrix T, ∇_{U_n,U_n}-generators of length $O(\alpha)$ can be computed in time $\tilde{O}(n\alpha)$. T can be written as a sum of α terms of the form LUJ_n, where L and U are Toeplitz and $J_n = \left(\delta_{i,n-1-j}\right)_{0\le i,j\le n-1}$ [21, Example 4.4.4]. Constant-length $\nabla_{Z_n,Z_n^{\mathsf{T}}}$- and $\nabla_{Z_n^{\mathsf{T}},Z_n}$-generators for each of the α terms can then be derived from ∇_{Z_n,Z_n}- and $\nabla_{Z_n^{\mathsf{T}},Z_n^{\mathsf{T}}}$-generators for the products LU using [21, Theorem 1.5.4] and the fact that J_n is in the kernel of $\nabla_{Z_n,Z_n^{\mathsf{T}}}$ and $\nabla_{Z_n^{\mathsf{T}},Z_n}$. Concatenation of the obtained generators yields the result.

Note that the complexity bound in n in Theorem 5.4 is the same as in the Toeplitz-like case (Theorem 5.2), we obtain however a stronger dependence in α. Indeed, we have used a Toeplitz+Hankel-like inversion in $O(n\alpha^2)$ [21], a better cost bound in $O(n\alpha^{\omega-1})$ would require to generalize the results of [3, 4] to partly regular operators.

6 SPECIAL MATRICES FOR GENERICITY

In order to identify the matrices T for which the algorithms of Section 5 output the characteristic polynomial (Theorems 5.2 and 5.4), we use the rank of the block Hankel matrix [17]

$$\mathrm{Hk}_{m,\lceil n/m\rceil} = \left(X^{\mathsf{T}}T^{i+j}X\right)_{0\le i,j\le\lceil n/m\rceil-1}.$$

We indeed have the following.

Lemma 6.1. *Let $T \in \mathbb{K}^{n\times n}$. If rank $\mathrm{Hk}_{m,\lceil n/m\rceil} = n$ then the irreducible left and right fractions descriptions of $X^{\mathsf{T}}(xI_n - T)^{-1}X$ have degree at most $\lceil n/m\rceil$. Furthermore, the determinant (made monic) of the denominator $Q \in \mathbb{K}[x]^{m\times m}$ of such a right irreducible description is the characteristic polynomial of T.*

Proof. The determinant of a denominator Q of an irreducible right fraction description of $X^{\mathsf{T}}(xI_n - T)^{-1}X$ is a divisor of the characteristic polynomial of T [17, Thm 2.12], hence has degree at most n. The claims then follow from [25, Lem. 2.4] since $\mathrm{Hk}_{m,\lceil n/m\rceil}$ has maximal rank n. □

Genericity Assumptions. To apply Theorems 5.2 and 5.4, a matrix T is "sufficiently" generic if it satisfies the following assumptions:

(A1) T has generic rank profile, so that the truncated generators of $M(x)^{-1}$ can be computed fast [4, 21];
(A2) there exists an $n \times n$ submatrix $\mathrm{Hk}^{(n)}$ of $\mathrm{Hk}_{m,\lceil n/m\rceil}$ whose determinant is nonzero, so that Lemma 6.1 can be applied.

The genericity in the Zariski sense can be expressed either based on the coefficients of T or on its generators. Indeed, the determinant of an $n \times n$ submatrix $\mathrm{Hk}^{(n)}$ of $\mathrm{Hk}_{m,\lceil n/m\rceil}$ is a polynomial in the coefficients of T. Toeplitz and Hankel matrices have $2n-1$ independant coefficients. With non-singular displacement operators, the coefficients of a Toeplitz-like or Hankel-like matrix of displacement rank α are themselves polynomials in the coefficients of its generators, so $\det \mathrm{Hk}^{(n)}$ is by composition a polynomial on the $2n\alpha$ coefficients of the $n \times \alpha$ generators of T.

In Sections 6.1 and 6.2, we show that we can construct an $n \times n$ submatrix $\mathrm{Hk}^{(n)}$ of $\mathrm{Hk}_{m,\lceil n/m\rceil}$ such that $\det \mathrm{Hk}^{(n)}$ is not uniformly zero on the space of Toeplitz (resp. Hankel) matrices, by finding one Toeplitz (resp. Hankel) matrix for which $\mathrm{Hk}_{m,\lceil n/m\rceil}$ has rank n. This establishes that assumption (A2) is satisfied for all matrices of each class except for those with coefficients in a certain hypersurface of $\mathbb{K}^{2n-1}$. As the displacement rank of the matrices we show is at most 2, they are Toeplitz-like (resp. Hankel-like) and can be represented with larger generators (padded with zeros). (A2) is thus also satisfied for matrices with displacement rank $\alpha \ge 2$ whose generators' coefficients are not in a certain hypersurface of $\mathbb{K}^{2n\alpha}$. The special matrices we construct are also Toeplitz+Hankel and Toeplitz+Hankel-like so the same reasoning shows that (A2) is satisfied for all Toeplitz+Hankel matrices except for those with coefficients in a certain hypersurface of $\mathbb{K}^{4n-2}$ and all Toeplitz+Hankel-like matrices with displacement rank $\alpha \ge 4$ except for those on a certain hypersurface of $\mathbb{K}^{2n\alpha}$. Using the fact that in the Toeplitz+Hankel-like case the operator is partly regular [21,

Sec. 4.5], the hypersurface can also be defined by considering the coefficients of the generators together with the irregularity set.

The generic rank profile condition (A1) can be handled similarly by considering the product Δ of the principal minors of T, though we omit details. This polynomial in the coefficients of T is non-zero for $T = I_n$ in the Toeplitz case. For the Hankel case, the determinant of an $n \times n$ Hankel matrix H defined by $h_0, \ldots, h_{2n-2}$ such that $H = (h_{i+j})_{i,j}$ has a unique term in h_{n-1}^n, hence is a non-zero polynomial in the h_i's; the same holds for Δ.

From the polynomial $(\det \mathrm{Hk}^{(n)}) \cdot \Delta$ in the entries of T, one can then define the general hypersurfaces outside of which our algorithms are correct.

6.1 A Toeplitz Point

Let

$$\mathcal{T} = \begin{pmatrix} 0 & I_m \\ I_{n-m} & 0 \end{pmatrix}$$

and $M(x) = xI_n - \mathcal{T}$. Let $P(x) \in \mathbb{K}[x]^{n \times m}$ defined by:

$$P_{n-m+k,k} = 1, \qquad \text{for } 0 \le k < m;$$
$$P_{i,k} = xP_{i+m,k}, \qquad \text{for } 0 \le k \le m, 0 \le i \le n-m-1.$$

With

$$D(x) = \begin{pmatrix} 0 & x^{\lfloor n/m \rfloor} I_{n \bmod m} \\ x^{\lfloor n/m \rfloor - 1} I_{-n \bmod m} & 0 \end{pmatrix}$$

we can write $P(x) = \begin{pmatrix} D(x)^\mathsf{T} & R(x) & I_m \end{pmatrix}^\mathsf{T}$ for some polynomial matrix R. From there we have $M(x)P(x) = \begin{pmatrix} xD(x)^\mathsf{T} - I_m & 0 \end{pmatrix}^\mathsf{T}$ and thus

$$X^\mathsf{T} M^{-1}(x) X = X^\mathsf{T} P(x) \left(xD(x) - I_m\right)^{-1}.$$

That is $X^\mathsf{T} M^{-1}(x) X = D(x) Q^{-1}(x)$ (we have used the form of P) with $Q(x) = xD(x) - I_m$. As $(xI_m) \cdot D(x) - I_m \cdot Q(x) = I_m$, the fraction DQ^{-1} is irreducible and

$$\det Q(x) = \pm x^{(\lfloor n/m \rfloor + 1)(n \bmod m) + \lfloor n/m \rfloor (-n \bmod m)} - 1$$

from which we get $\deg \det Q = n$. By [25, Lemma 2.4], $\mathrm{Hk}_{m,\lceil n/m \rceil}$ has rank n.

6.2 A Hankel Point

Consider the $n \times n$ Hankel matrix $\mathcal{H} = (I_n + Z_n^m) J_n$ (with $J_n = \left(\delta_{i,n-1-j}\right)_{0 \le i,j \le n-1}$). For j such that $2j \le \lceil n/m \rceil - 1$, rows jm to $(j+1)m - 1$ of $\mathcal{H}^{2j} X$ are I_m and the following rows are 0. This can be seen by recursively applying the band matrix $\mathcal{H}^2 = Z_n^m + I_n + Z_n^m Z_n^{m\mathsf{T}} + Z_n^{m\mathsf{T}}$ to X. By applying $\mathcal{H}$ to $\mathcal{H}^{2j} X$ we get that the rows $n - (j+1)m$ to $n - jm - 1$ of $\mathcal{H}^{2j+1} X$ are J_m, and the preceding rows are 0.

Let K_r be the first n columns of $\left(X | \mathcal{H}X | \ \ldots | \mathcal{H}^{\lceil n/m \rceil - 1} X\right)$. This matrix K_r is non-singular, as its columns can be permuted to get a matrix of the form

$$\begin{pmatrix} L_1^\mathsf{T} & 0 \\ 0 & L_2 \end{pmatrix}$$

where L_1 and L_2 are lower triangular with ones on the diagonal. Since $\mathcal{H}$ is symmetric, the $n \times n$ principal submatrix of $\mathrm{Hk}_{m,\lceil n/m \rceil}$ is $K_r^\mathsf{T} K_r$, hence $\mathrm{Hk}_{m,\lceil n/m \rceil}$ has rank n.

REFERENCES

[1] J. Alman and V. Vassilevska Williams. 2020. A Refined Laser Method and Faster Matrix Multiplication. arXiv:cs.DS/2010.05846

[2] B. Beckermann and G. Labahn. 1994. A uniform approach for the fast computation of matrix-type Padé approximants. *SIAM J. Matrix Analysis and Applications* 15, 3 (1994), 804–823. https://doi.org/10.1137/S0895479892230031

[3] A. Bostan, C.-P. Jeannerod, C. Mouilleron, and É. Schost. 2017. On matrices with displacement structure: generalized operators and faster algorithms. *SIAM J. on Matrix Analysis and Applications* 38, 3 (2017), 733–775. https://doi.org/10.1137/16M1062855

[4] A. Bostan, C.-P. Jeannerod, and É. Schost. 2008. Solving structured linear systems with large displacement rank. *Theoretical Computer Science* 407, 1-3 (2008), 155–181. https://doi.org/10.1016/j.tcs.2008.05.014

[5] W. Eberly, M. Giesbrecht, P. Giorgi, A. Storjohann, and G. Villard. 2007. Faster inversion and other black box matrix computation using efficient block projections. In *Proc. ISSAC*. ACM Press, 143–150. https://doi.org/10.1145/1277548.1277569

[6] P. Giorgi, C. Jeannerod, and G. Villard. 2003. On the complexity of polynomial matrix computations. In *Proc. ISSAC*. ACM Press, 135–142. https://doi.org/10.1145/860854.860889

[7] G. Heinig, P. Jankowski, and K. Rost. 1988. Fast inversion algorithms of Toeplitz-plus-Hankel matrices. *Numer. Math.* 52, 6 (1988), 665–682. https://doi.org/10.1007/BF01395817

[8] G. Heinig and K. Rost. 1984. *Algebraic Methods for Toeplitz-like Matrices and Operator*. Springer, Birkhäuser Basel. https://doi.org/10.1007%2F978-3-0348-6241-7

[9] G. Heinig and K. Rost. 2004. New fast algorithms for Toeplitz-plus-Hankel matrices. *SIAM J. Matrix Analysis and Applications* 25, 3 (2004), 842–857. https://doi.org/10.1137/S0895479802410074

[10] C.-P. Jeannerod, V. Neiger, and G. Villard. 2020. Fast computation of approximant bases in canonical form. *J. Symbolic Computation* 98 (2020), 192–224. https://doi.org/10.1016/j.jsc.2019.07.011

[11] T. Kailath. 1980. *Linear Systems*. Prentice-Hall.

[12] T. Kailath, S.Y. Kung, and M. Morf. 1979. Displacement ranks of matrices and linear equations. *J. Mathematical Analysis and Applications* 68, 2 (1979), 395–407. https://doi.org/10.1016/0022-247X(79)90124-0

[13] E. Kaltofen. 1994. Asymptotically fast solution of Toeplitz-like singular linear systems. In *Proc. ISSAC*. ACM Press, 297–304. https://doi.org/10.1145/190347.190431

[14] E. Kaltofen. 2000. Challenges of symbolic computation: my favorite open problems. *J. Symbolic Computation* 29, 6 (2000), 891–919. https://doi.org/10.1006/jsco.2000.0370

[15] E. Kaltofen, M.S. Krishnamoorthy, and B.D Saunders. 1990. Parallel algorithms for matrix normal forms. *Linear Algebra Appl.* 136 (1990), 189–208. https://doi.org/10.1016/0024-3795(90)90028-B

[16] E. Kaltofen and B.D. Saunders. 1991. On Wiedemann's method of solving sparse linear systems. In *Proc. AAECC-9 (LNCS 539, Springer Verlag)*. 29–38. https://doi.org/10.1007/3-540-54522-0_93

[17] E. Kaltofen and G. Villard. 2005. On the complexity of computing determinants. *Comput. Complex.* 13, 3 (2005), 91–130. https://doi.org/10.1007/s00037-004-0185-3

[18] E. Kaltofen and G. Yuhasz. 2013. On the matrix Berlekamp-Massey algorithm. *ACM Trans. Algorithms* 9, 4 (2013), 33:1–33:24. https://doi.org/10.1145/2500122

[19] G. Labahn, V. Neiger, and W. Zhou. 2017. Fast, deterministic computation of the Hermite normal form and determinant of a polynomial matrix. *J. Complexity* 42 (2017), 44–71. https://doi.org/10.1016/j.jco.2017.03.003

[20] F. Le Gall. 2014. Powers of Tensors and Fast Matrix Multiplication. In *Proc. ISSAC*. ACM Press, 296–303. https://doi.org/10.1145/2608628.2608664

[21] V. Y. Pan. 2001. *Structured Matrices and Polynomials: Unified Superfast Algorithms*. Springer-Verlag, Berlin, Heidelberg. https://doi.org/10.1007/978-1-4612-0129-8

[22] A. Storjohann. 2003. High-order lifting and integrality certification. *J. Symbolic Computation* 36, 3-4 (2003), 613–648. https://doi.org/10.1016/S0747-7171(03)00097-X

[23] M. Van Barel and A. Bultheel. 1992. A general module theoretic framework for vector M-Padé and matrix rational interpolation. *Numerical Algorithms* 3 (1992), 451–462. https://doi.org/10.1007/BF02141952

[24] G. Villard. 1997. *A study of Coppersmith's block Wiedemann algorithm using matrix polynomials*. RR 975 IM IMAG. http://perso.ens-lyon.fr/gilles.villard/BIBLIOGRAPHIE/PDF/rr0497.pdf

[25] G. Villard. 2018. On computing the resultant of generic bivariate polynomials. In *Proc. ISSAC*. ACM Press, 391–398. https://doi.org/10.1145/3208976.3209020