Computing specified generators of structured matrix inverses
Claude-Pierre Jeannerod, Christophe Mouilleron

To cite this version:

HAL Id: ensl-00450272
https://hal-ens-lyon.archives-ouvertes.fr/ensl-00450272
Submitted on 26 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Computing specified generators of structured matrix inverses

Claude-Pierre Jeannerod
LIP - ENS de Lyon
INRIA
claude-pierre.jeannerod@ens-lyon.fr

Christophe Mouilleron
LIP - ENS de Lyon
Université de Lyon
christophe.mouilleron@ens-lyon.org

ABSTRACT
The asymptotically fastest known divide-and-conquer methods for inverting dense structured matrices are essentially variations or extensions of the Morf/Bitmead-Anderson algorithm. Most of them must deal with the growth in length of intermediate generators, and this is done by incorporating various generator compression techniques into the algorithms. One exception is an algorithm by Cardinal, which in the particular case of Cauchy-like matrices avoids such growth by focusing on well-specified, already compressed generators of the inverse. In this paper, we extend Cardinal’s method to a broader class of structured matrices including those of Vandermonde, Hankel, and Toeplitz types. Besides, some first experimental results illustrate the practical interest of the approach.

Categories and Subject Descriptors
I.1.2 [Computing Methodologies]: Symbolic and Algebraic Manipulation—Algebraic Algorithms

General Terms
Algorithms, Theory

Keywords
Structured linear algebra, matrix inversion

1. INTRODUCTION
Since [10], a classical way of exploiting the structure of dense matrices is via the displacement rank approach: typically, \( n \times n \) matrices are represented by pairs \((G, H)\) of \( n \times \alpha\) matrices such that \( L(A) = GH^T \) for some linear operator \( L \) called a displacement. Classical choices for \( L \) are Stein’s displacement \( \Delta[M, N] : A \mapsto A - MAN \) and Sylvester’s displacement \( \nabla[M, N] : A \mapsto MA - AN \). With respect to a given displacement, \((G, H)\) is called a generator of length \( \alpha \) for \( A \), and \( A \) is considered to be structured when \( \alpha \) is “small” (in a sense that depends on the context) compared to \( n \). According to the unified treatment [19], many of the structures encountered in practice are covered by the following operator matrices: for a field \( \mathbb{K} \) and a positive integer \( n \), let

\[
M, N \in \{ D(x), Z_{m,\psi}, Z_{m,\varphi}^T \}, \quad x \in \mathbb{K}^n, \quad \varphi, \psi \in \mathbb{K}, \quad (1)
\]

with \( D(x) \) the diagonal matrix whose entry \((i, i)\) is the \( i \)th coefficient \( x_i \) of vector \( x \), and \( Z_{m, \varphi}\) the \( n \times n \) unit \( \varphi \)-circulant matrix having a \( \varphi \) in position \((1, n)\), ones in positions \((i, i)\), and zeros everywhere else.

When a structured matrix \( A \in \mathbb{K}^{n \times n} \) is invertible, its inverse \( A^{-1} \) is known to be structured too, and some asymptotically fast algorithms are available for computing length-\( \alpha \) generators for \( A^{-1} \) and linear system solutions, whose costs in terms of operations in \( \mathbb{K} \) are in \( O(\alpha^2 n) \) (see [19] and the references therein) and, since more recently, in \( O(\alpha^{n-1} n) \) (see [2, 3]). (Here and hereafter the \( O \) notation hides all logarithmic factors.) Such algorithms are essentially variations or extensions of the Morf/Bitmead-Anderson (MBA) divide-and-conquer approach [13]. In practice, they apply to important types of structures like those of (1). However, most of these algorithms must deal with the growth in length of intermediate generators, and this is done by recursively using a generator compression stage which, given matrices \( G, H \in \mathbb{K}^{n \times \beta} \) such that \( GH^T \) has rank \( \alpha \leq \beta \), computes matrices \( G_\alpha, H_\beta \) that satisfy \( G_\alpha, H_\beta \mapsto GH^T \) but now have exactly \( \alpha \) columns; see [17, 15, 16, 11, 12] and [19, §4.6].

One exception is a variant of MBA due to Cardinal [4, 5]: assuming Sylvester’s displacement equation

\[
\nabla[M, N](A) = GH^T \quad (2)
\]

and in the particular case where both \( M \) and \( N \) are diagonal (Cauchy-like structure), Cardinal’s algorithm completely avoids generator compression by directly computing

\[
Y = -A^{-1} G, \quad Z = A^{-1} H. \quad (3)
\]

As already noted in [9] —and this is readily verified by pre-and postmultiplying (2) with the inverse of \( A \)—, the matrix pair \((Y, Z)\) is a \( \nabla[N, M] \)-generator of length \( \alpha \) for \( A^{-1} \). Due to its very special form, we shall call it a specified generator for the inverse of \( A \).

The goal of this paper is to extend Cardinal’s algorithm beyond the Cauchy-like structure and to show that, in MBA and for Sylvester’s displacement, generator compression can be systematically avoided by targeting a specified generator for the inverse, rather than just an arbitrary one of length \( \alpha \). More precisely, our three main contributions can be summarized as follows:
First, we propose a recursive formula that allows to factor a specified generator of the inverse for $A$ in terms of specified generators for the inverse of its upper-left block $A_{11}$ and for the inverse of the Schur complement of $A_{11}$ in $A$.

Second, we show how to reduce the computation of specified inverse generators for the structures defined in (1) to the computation of specified inverse generators for the three basic cases below:

$$\{(M, N) \in \{(D(x), D(y)), (D(x), Z_{n,0}^T), (Z_{n,0}, Z_{n,0}^T)\}\}.$$  

(4)

For each of those three structures, which are of the Cauchy-like, Vandermonde-like, and Hankel-like types, respectively, we further give and analyze explicit algorithms for computing a specified generator of the inverse. These algorithms are compression-free and thus, in that sense, simpler to analyze and implement than traditional MBA variants. Moreover, although removing generator compression does not affect the overall asymptotic costs, it yields smaller dominant terms.

Third, we report on a first set of experiments done with our C++ implementation of MBA and of several of the new compression-free algorithms. For the Cauchy-like structure, for example, the speed-ups compared to MBA are by a factor from 4.6 to 6.7. This suggests that our extension of Cardinal’s compression-free approach may yield algorithms that are not only simpler but also significantly faster in practice.

**Outline of the paper.** After some notation and preliminaries in §2, some properties of specified generators are studied in §3. Then §4 gives a compression-free algorithm for the case where $M$ and $N^T$ are lower triangular. The algorithm is specialized to the Cauchy-like and Vandermonde-like structures in §4.1 and §4.2, and then extended in §4.3 to the irregular Hankel-like case $(M, N) = (Z_{n,0}, Z_{n,0}^T)$. Experiments are reported in §5 and we conclude in §6.

### 2. NOTATION AND PRELIMINARIES

Here and hereafter, $I_n$ is the identity matrix of order $n$, $e_{n,i}$ is the $i$th unit vector of $\mathbb{K}^n$, and $J_n$ is the reflexion matrix of order $n$, whose $(i,j)$ entry is 1 if $i+j = n+1$, and 0 otherwise. For $A \in \mathbb{K}^{n \times m}$, $a_{ij}$ denotes its $(i,j)$ entry and $a_j$ its $j$th column. Also, for $\alpha \leq m$, we write $A^{\alpha \times m}$ for the matrix $[a_1, \ldots, a_{\alpha}] \in \mathbb{K}^{\alpha \times m}$.

Given positive integers $n_1$ and $n_2$ such that $n_1 + n_2 = n$, we will often partition $A$, $G$, $H$, $M$, $N$ into $n_{1 \times n_{1}}, n_{2 \times n_{2}}$ blocks as

\[
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad G = \begin{bmatrix} G_{11} \\ G_{21} \end{bmatrix}, \quad H = \begin{bmatrix} H_{11} \\ H_{21} \end{bmatrix}.
\]

(5)

\[
M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}.
\]

(5)

We shall write $\mu$ and $\nu$ for the rank of, respectively, $M_{12}$ and $N_{21}$. Consequently, those two matrices can be written

\[
M_{12} = U_1 V_2^T, \quad N_{21} = U_2 V_1^T
\]

for some full column rank matrices $U_1 \in \mathbb{K}^{n_1 \times \mu}$, $V_2 \in \mathbb{K}^{n_2 \times \nu}$, and $V_1 \in \mathbb{K}^{n_1 \times \nu}$.

The Schur complement of $A_{11}$ in $A$ is written $S$:

\[
S = A_{22} - A_{21} A_{11}^{-1} A_{12}.
\]

Recall that $S$ is nonsingular if $A_{11}$ and $A$ are nonsingular; if $A$ is strongly regular then so are $A_{11}$ and $S$. Finally, let

\[
E = \begin{bmatrix} I_{n_1} & -A_{11}^{-1} A_{12} \\ -A_{12} A_{11}^{-1} & I_{n_2} \end{bmatrix}, \quad F = \begin{bmatrix} I_{n_1} & -A_{11}^{-1} A_{12} \\ I_{n_2} & -A_{11}^{-1} A_{12} \end{bmatrix}.
\]

(7)

From $E A F = \text{diag}(A_{11}, S)$, we deduce the following classical recursive factorization of the inverse of $A$ [19, p. 157]:

\[
A^{-1} = F \begin{bmatrix} A_{11}^{-1} \\ S^{-1} \end{bmatrix} E.
\]

(8)

### 2.1 Properties of Sylvester’s displacement

The properties below show how to deduce, given a $\nabla[M, N]$-generator for $A$, a generator for various matrices related to $A$. All of them appear in/follow immediately from [18, 19].

#### Generation of the transpose

Let $A, G, H$ be as in (2). By transposing the identity $MA - AN = GH^T$, we obtain

\[
\nabla[N^T, M^T](A^T) = -HG^T,
\]

(9)

so that the pair $(-H, G)$ is a $\nabla[N^T, M^T]$-generator of $A^T$.

#### Generation of products

One has the following classical rule for generating matrix products [19, p. 10]:

\[
\nabla[M, N](AB) = \nabla[M, \ast](A) B + A \nabla[\ast, N](B),
\]

(10)

for any matrix $\ast$ of conforming dimensions. Applying this rule twice, we can straightforwardly deduce explicit formulas for generating products of three matrices:

**Lemma 1.** Let $A, G, H$ be as in (2) and, for two matrices $P_1$ and $P_2$, let

\[
\tilde{A} = P_1 A P_2.
\]

(11)

If $\nabla[M, M](P_1) = G_{P_1} H_{P_1}^T$, and $\nabla[N, N](P_2) = G_{P_2} H_{P_2}^T$ then

\[
\nabla[M, N](\tilde{A}) = \tilde{G} \nabla[H, \ast].
\]

(12)

As an example, let us mention three special cases which we will use later: assuming $M, N \in \{Z_{n,0}, Z_{n,0}^T\}$, let first

$$\nabla[M, M](P_1) = (I_n, J_n) \quad \text{and} \quad (M, N) = (M, N^T).$$

Then obviously $\nabla[M, M](P_1)$ is zero and, using the facts that $Z_{n,0}^T = I_n$ and $Z_{n,0} Z_{n,0}^T = Z_{n,0}^T$ (see [19, p. 24]), we deduce that $\nabla[N, N](P_2)$ is zero as well. Consequently, since $J_n$ is symmetric, applying (12) yields

\[
\nabla[M, N^T](A J_n) = G_{J_n} H_{J_n}^T.
\]

(13a)

Similarly, exchanging the roles of $P_1$ and $P_2$ yields

\[
\nabla[M^T, N^T](J_n A) = G_{J_n} H_{J_n}^T,
\]

(13b)

while taking $P_1 = P_2 = J_n$ gives

\[
\nabla[M^T, N^T](J_n J_n) = G_{J_n} (J_n H_{J_n})^T.
\]

(13c)

#### Generation of submatrices

From (2) and (6) and the partitioning into blocks we deduce that, for $i, j \in \{1, 2\}$, submatrix $A_{i,j}$ satisfies the following matrix equation

\[
\nabla[M_{i,j}, N_{i,j}](A_{i,j}) = G_{i,j} H_{i,j}^T,
\]

(14)

where, in particular (see for example [18, Proposition 4.4]),

\[
G_{i,1} = \begin{bmatrix} | & -U_1 \end{bmatrix} U_1 \mathbb{K}^{n_1 \times (n+\mu+\nu)}.
\]

(15a)

\[
H_{1,i} = \begin{bmatrix} H_1 \end{bmatrix} A_{12} V_2 \mathbb{K}^{n_1 \times (n+\mu+\nu)}.
\]

(15b)

#### Generation of Schur complements

By combining [18, Proposition 4.5] with (6), we have the following description of the structure of the Schur complement $S$ of $A_{11}$ in $A$:

\[
\nabla[M_{22}, N_{22}](S) = G_8 H_8^T.
\]

(16)
with $G_S$ and $H_S$ the two matrices in $\mathbb{K}^{n_2 \times (\alpha + \mu + \nu)}$ given by

\[
G_S = \begin{bmatrix} G_2 & A_2 A_T^{-1} G_1 & A_2 A_T^{-1} U_1 - SU_2 \end{bmatrix}, \quad (17a)
\]
\[
H_S = \begin{bmatrix} H_2 - A_T^{-1} A_1^T H_1 [S_T V_2 A_T^{-1} A_1^T V_1] \end{bmatrix}. \quad (17b)
\]

When the operator matrices $M$ and $N^*$ are lower triangular, one has $\mu = \nu = 0$ and the above formulas for generating the Schur complement can thus be simplified as follows (see [8, Theorem 2.3], [14, Lemma 3.1], [19, §5.4]):

\[
G_S = G_2 - A_2 A_T^{-1} G_1, \quad H_S = H_2 - A_T^{-1} A_1^T H_1. \quad (18)
\]

### 2.2 Computing with basic structures

We compute our preliminaries by reviewing three basic invertible displacement operators that we shall repeatedly use in the sequel, as well as some associated cost functions. Here we assume that $(2)$ holds in the rectangular case, that is, for $A \in \mathbb{K}^{n \times n}$, $G \in \mathbb{K}^{n \times c}$, and $H \in \mathbb{K}^{n \times c}$; this assumption will allow us to handle off-diagonal blocks in Section 4. Recall also that $\nabla [M, N]$ is invertible if and only if the spectra of $M$ and $N$ are disjoint [19, p. 123].

#### Cauchy-like structure.

For $x \in \mathbb{K}^n$ and $y \in \mathbb{K}^m$, assume

\[
M = D(x), \quad N = D(y), \quad x_i \neq y_j \text{ for all } (i, j). \quad (19a)
\]

Then $\nabla [M, N]$ is invertible and it is known [7] (see also [19, p. 8] and [20, Lemma 2.1]) that $(2)$ is equivalent to

\[
A = \sum_{i,j} D(g_{i,j}) C(x, y) D(h_{i,j}), \quad (19b)
\]

with $C(x, y)$ the $n$ by $m$ Cauchy matrix $[1/(x_i - y_j)]_{i,j}$.

#### Vandermonde-like structure.

For $x \in \mathbb{K}^n$, assume now

\[
M = D(x), \quad N = Z_{m,0}, \quad x_i \neq 0 \text{ for all } i. \quad (20a)
\]

Then $\nabla [M, N]$ is invertible and, in this case, $A$ can be recovered as follows (see [19, Example 4.4.6(d)]):

\[
A = \sum_{n=1}^\infty D(x^{-1}, g_n) V(x^{-1}, m) U(h_n), \quad (20b)
\]

where, for $x \in \mathbb{K}^n$, $V(x, m)$ is the $n$ by $m$ Vandermonde matrix whose $(i, j)$ entry equals $x_i^{-j}$, and $U(h_n)$ is the $m$ by $m$ upper triangular Toeplitz matrix whose first row is $h_n^T$.

#### Hankel-like structure.

Assume finally that

\[
M = Z_{n,1}, \quad N = Z_{m,0}. \quad (21a)
\]

Since $Z_{n,1}$ and $Z_{m,0}$ have disjoint spectra, $\nabla [M, N]$ is invertible. In addition, we can recover $A$ as follows:

\[
A = \sum_{j=1}^n T^{n \times m}(g_j) L(h_j) L_m, \quad (21b)
\]

where, for $x \in \mathbb{K}^n$, $T^{n \times m}(x)$ is the $n$ by $m$ Toeplitz matrix $[x_{i+j-1} \text{ mod } n]_{i,j}$, and where $L(h_j)$ is the $m$ by $m$ lower triangular Toeplitz matrix whose first column is $h_j$. (A proof of (21b) is given in Appendix A.)

#### Cost functions.

Our algorithms in the next sections will essentially require the ability to efficiently evaluate products of the form $Av$ and $A^TVv$, where $A$ has one of the three basic structures above, $m$ is of the same order of $n$, and $v$ consists of one or several vectors.

In order to relate the costs of our algorithms in Section $4$ to the costs of such products, we introduce the following functions. For the Cauchy-like structure (19a), let \(\text{MM}_C : \mathbb{N}_{\geq 0} \times \mathbb{N}_{\geq 0} \to \mathbb{N}_{\geq 0}\) be such that, for $A \in \mathbb{K}^{n \times n}$ given by the right-hand-side of (19b) and $v \in \mathbb{K}^{n \times \beta}$, the products $Av$ and $A^TVv$ can be computed using $\text{MM}_C(n, \beta)$ operations in $K$. We define the functions $\text{MM}_V$ and $\text{MM}_H$ in a similar way for, respectively, the Vandermonde-like and Hankel-like structures. Also, when $\beta = \alpha$ we shall simply write $\text{MM}_{\sigma}(\alpha, n)$, for $\sigma = C, V, H$.

Following [6, p. 242], we write $M(n)$ for the cost of multiplying two polynomials of degree less than $n$ over $\mathbb{K}[x]$, and we assume that $M(n)$ is “superlinear”, that is, $M(n)/n$ is nondecreasing.

It is known (see for example [19]) that $C(x, y) = -C(y, x)$ and that multiplying $C(x, y)$, $V(x, n)$, or $V(x, n)^T$ by a vector can be done in time $O(M(n)\log(n))$ via (transposed) multipoint evaluation. Hence by a straightforward application of the summation formulas (19b), (20b), and (21b), one has

\[
\text{MM}_V(\alpha, n, 1) \in O(\alpha M(n)\log(n)) \quad \text{for } \sigma = C, V,
\]

\[
\text{MM}_H(\alpha, n, 1) \in O(\alpha M(n)).
\]

We shall also use the three basic properties given below:

#### Lemma 2. Let $k, \ell \in \Omega(1)$. Then

\[
\text{MM}_V(\alpha + k, n, \alpha) \in \text{MM}_V(\alpha, n) + O(\alpha M(n)\log(n)),
\]

and, for $\sigma = C, V, H$,

\[
\text{MM}_H(\alpha + k, n, \alpha) \in \text{MM}_H(\alpha, n) + O(\alpha M(n)).
\]

#### Proof. To get (22) note that, for all $\sigma$, $\text{MM}_V(\alpha + k, n, \alpha)$ is in $\text{MM}_V(\alpha, n, \alpha) + \text{MM}_V(k, n, \alpha) + O(\alpha n)$. Indeed, one can evaluate our sum of $\alpha + k$ products by adding the first $\alpha$ terms and the last $k$ terms separately, and then combining the two intermediate results. Since moreover $\text{MM}_V(k, n, \alpha) \leq \text{MM}_V(k, n, 1)$, (22a) and (22b) follow from the complexities of $\text{MM}_V(n, m)$ and $\text{MM}_H(n, m)$ mentioned above. To establish (23), notice that a sum of $k$ terms for $\ell$ vectors can be evaluated via $k$ sums of $\alpha$ terms for $\alpha$ vectors plus a final sum in $O(\alpha n)$, repeated $\ell$ times.

Finally, we assume as for $M(n)$ that the function $M(\cdot, n)$ is “superlinear,” that is, $M(\cdot, n)/n$ is nondecreasing. This assumption will allow us to simplify the cost bounds of the algorithms of Section 4 and can be easily supported by “naive” implementations in $O(n^2)$ as those used in Section 5.

### 3. PROPERTIES OF SPECIFIED GENERATORS OF THE MATRIX INVERSE

#### 3.1 Recovery after matrix transformations

We recalled in Section 2 some formulas for generating the matrix $A \in \{A^T, P_1 AP_2\}$ from some generators of the matrix $A$. Conversely, we give in the theorem below some formulas for recovering specified generators of the inverse of $A$ from specified generators of the inverse of $\bar{A}$.

#### Theorem 1. Let $A \in \mathbb{K}^{n \times n}$ be invertible and let $G, H \in \mathbb{K}^{n \times \alpha}$ and $Y, Z \in \mathbb{K}^{n \times \beta}$ be as in (2) and (3). Let $\bar{A} \in \mathbb{K}^{n \times \beta}$ be invertible and, for $\bar{G}, \bar{H} \in \mathbb{K}^{n \times \beta}$, $\beta \geq \alpha$, define

\[
\bar{Y} = -\bar{A}^{-1} \bar{G}, \quad \bar{Z} = \bar{A}^{-T} \bar{H}.
\]

Then

- for $\bar{A} = A^T$ and $(\bar{G}, \bar{H}) = (-H, G)$, one has

\[
Y = -\bar{Z}, \quad Z = \bar{Y};
\]

- for $\bar{A} = A$ and $(\bar{G}, \bar{H}) = (-G, H)$, one has

\[
X = -\bar{Y}, \quad \bar{X} = \bar{Z};
\]

- for $\bar{A} = A^T$ and $(\bar{G}, \bar{H}) = (G, H)$, one has

\[
X = Y, \quad \bar{X} = \bar{Z};
\]

- for $\bar{A} = A$ and $(\bar{G}, \bar{H}) = (-G, H)$, one has

\[
X = -\bar{Y}, \quad \bar{X} = \bar{Z}.
\]
Theorem 2. Let $A \in \mathbb{K}^{n \times n}$ be nonsingular and generated by $G$ and $H$ as in (2). Assume that $A_{11}$ is nonsingular as well, that it is generated by $G_{11}$ and $H_{11}$ as in (15), and let

$$Y_{11} = -A_{11}^{-1}G_{11}, \quad Z_{11} = A_{11}^{-1}H_{11}.$$  

Assume further that the Schur complement $S$ of $A_{11}$ in $A$ is generated by $G_{5}$ and $H_{5}$ as in (17), and let

$$Y_{5} = -A_{5}^{-1}G_{5}, \quad Z_{5} = A_{5}^{-1}H_{5}.$$  

Then the matrices $Y$ and $Z$ in (3) satisfy

$$Y = F \left[ \begin{array}{c} Y_{11}^\alpha \\ Y_{5}^\alpha \end{array} \right], \quad Z = E^{T} \left[ \begin{array}{c} Z_{11}^\alpha \\ Z_{5}^\alpha \end{array} \right],$$

where $E$ and $F$ are the elimination matrices defined in (7).

Proof. Using (5) and (8), we obtain

$$-A^{-1}G = F \left[ \begin{array}{c} -A_{11}^{-1}G_{1} \\ -S^{-1}(G_{2} - A_{21}^{-1}G_{1}) \end{array} \right].$$

It follows from (15a) and (17a) that $G_{1} = G_{11}^{\alpha}$ and that $G_{2} - A_{21}^{-1}G_{1} = G_{5}^{\alpha}$. The expression claimed for $Y = -A^{-1}G$ then follows from applying the rule $A(B^{-\alpha}) = (AB)^{-\alpha}$ twice, and from the definitions of $Y_{11}$ and $Y_{5}$. The expression for $Z$ can be obtained in a similar way, using (15b) and (17b).

A first consequence of this theorem is a “compressed” analogue of the classical recursive factorization formula (8):

$$YZ^{T} = F \left[ \begin{array}{c} Y_{11}^\alpha \\ Y_{5}^\alpha \end{array} \right] \left[ \begin{array}{c} Z_{11}^\alpha \\ Z_{5}^\alpha \end{array} \right]^{T} E.$$
• Recursively, compute \((Y_S, Z_S) = (-S^{-1}G_S, S^{-T}H_S)\);

• Compute \((-A^{-1}G, -A^{-T}H)\) from the first \(\alpha\) columns of \(Y_{11}, Y_S, Z_{11}, Z_S\), using Theorem 2.

4. ALGORITHMS FOR LOWER TRIANGULAR OPERATOR MATRICES \(M\) AND \(N^T\)

In order to cover simultaneously the three displacements in \((4)\) to which we have previously reduced, we assume in this section that both operator matrices \(M\) and \(N^T\) are lower triangular.

This assumption implies in particular that the blocks \(M_{12}\) and \(N_{21}\) in \((6)\) are zero, so that their respective ranks \(\mu\) and \(\nu\) satisfy \(\mu = \nu = 0\). From \((15)\) it then follows that the submatrix \(A_{11}\) satisfies

\[
\nabla[M_{11}, N_{11}](A_{11}) = G_1H_1^T.
\]

Thus, some generators of length at most \(\alpha\) for \(A_{11}\) can be read off the first \(n_1\) rows of some generators of length at most \(\alpha\) for \(A\).

Assuming that \(A_{11}\) is invertible, consider now the associated specified generators of \(A_{11}^{-1}\), that is,

\[
Y_{11} = -A_{11}^{-1}G_1, \quad Z_{11} = A_{11}^{-T}H_1.
\]

Combining the two identities in \((28)\) with the explicit Schur complement generation formulas in \((17)\) and \((18)\) yields

\[
\nabla[M_{22}, N_{22}](S) = \begin{pmatrix} G_2 + A_{21}Y_{11} \end{pmatrix}(H_2 - A_{22}Z_{11})^T.
\]

In other words, the precise specification of the above generators of the inverse of \(A_{11}\) can be exploited to simplify even further the generators of the Schur complement. In \([4, Proposition 1]\), Cardinal had already noted this formula but only for the Cauchy-like structure \((M\) and \(N\) diagonal).

Now, if we assume further that \(A\) is strongly regular (which, if randomization is allowed, makes sense in view of the probabilistic reductions to strong regularity shown in Section 3.1), we obtain the following general algorithm:

**Theorem 3.** Algorithm GenInvLT is correct.

**Proof.** When \(\alpha = 1\), the assumption on \(M\) and \(N\) implies that \(A\) is the scalar \((\sum_{i=1}^{\alpha} g_i h_i)/(m_{11} - m_{11})\). Correctness then follows immediately in this case. Assume now that \(n > 1\) and, in order to proceed by induction, assume correctness for \(n' < n\). The matrix \(A_{11}\) is strongly regular (since \(A\) is) and it satisfies \((27)\), where, by assumption \(M_{11}\) and \(N_{11}\) are both lower triangular and with disjoint diagonals. Since \(n_1 < n\), the induction assumption then implies that the pair \((Y_{11}, Z_{11})\) returned by the first recursive call is precisely \((-A_{11}^{-1}G_1, A_{11}^{-T}H_1)\). Therefore, the computed pair \((G_2, H_2)\) satisfies \((29)\), where, by assumption, \(S\) is strongly regular (since \(A\) is) and where \(M_{22}\) and \(N_{22}\) are both lower triangular and have disjoint diagonals. Since \(n_2 < n\), the induction assumption implies that the pair \((Y_S, Z_S)\) returned by the second recursive call is exactly \((-S^{-1}G_S, S^{-T}H_S)\). The conclusion then follows from Theorem 2. \(\square\)

To implement Algorithm GenInvLT and bound its cost, all we need is to be able to evaluate the four matrix products

\[
A_{21}Y_{11}, \quad A_{21}^{-1}Z_{11}, \quad A_{11}^{-1}A_{12}Y_S, \quad A_{11}^{-T}A_{21}^{-1}Z_S.
\]

(30)

In the next subsections, we study the evaluation of those expressions for each of three basic structures of the Cauchy, Vandermonde, and Hankel types. That requires in each case a detailed analysis of the structure of the matrices \(A_{11}^{-1}, A_{12}, A_{21}\), and their transposes. Since in \((30)\) there are two ways of parenthesizing the products of three matrices, we will also study the structure of \(A_{11}^{-1}A_{12}\) and \((A_{21}A_{21}^{-1})^T\). The parenthesizations \((A_{11}^{-1}A_{12})Y_S\) and \((A_{21}A_{21}^{-1})^T Z_S\) will be referred to as “Cardinal’s trick” later on, as they have been initially used in \([4]\) for the Cauchy-like case.

4.1 Application to Cauchy-like matrices

We consider here the specialization of Algorithm GenInvLT to the Cauchy-like structure defined in \((19a)\). Partitioning the two vectors \(x\) and \(y\) conformally with \(A\) yields

\[
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad x_1, y_1 \in \mathbb{R}^{n_1}, \quad x_2, y_2 \in \mathbb{R}^{n_2}.
\]

**Lemma 3.** Let the matrices \(A, G, H, Y_{11}, Z_{11}, G_S, H_S\) be as in Algorithm GenInvLT. Then

• \(\nabla[D(x_i), D(y)](A_{ij}) = G_{ij}H_{ij}^T\) for \(1 \leq i, j \leq 2\),

• \(\nabla[D(y_1), D(x_1)](A_{11}^{-1}) = Y_{11}Z_{11}^T\),

• \(\nabla[D(y_1), D(y_2)](A_{11}^{-1}A_{12}) = -Y_{11}H_{12}^T\),

• \(\nabla[D(x_2), D(x_1)](A_{21}A_{11}^{-1}) = G_SZ_{11}^T\).

**Proof.** Since \(D(x)\) and \(D(y)\) are diagonal, all their off-diagonal blocks are zero, and the first identity follows from \((2)\). To get the second identity, it suffices to pre- and postmultiply by \(A_{11}^{-1}\) both sides of the first identity for \((i, j) = (1, 1)\), and then to use the specification of \(Y_{11}\) and \(Z_{11}\). Using the multiplication rule \((10)\), we deduce further from the first identity for \((i, j) = (1, 2)\) and from the second one that

\[
\nabla[D(y_1), D(y_2)](A_{11}^{-1}A_{12}) = Y_{11}Z_{11}^T A_{12} + A_{11}^{-1}G_SH_{12}^T = Y_{11}(Z_{11}^TA_{12} - H_{12}^T),
\]

which by definition of \(H_S\) equals \(-Y_{11}H_{12}^T\). Similarly,

\[
\nabla[D(x_2), D(x_1)](A_{21}A_{11}^{-1}) = G_SH_{11}^TA_{11}^{-1} + A_{21}Y_{11}Z_{11}^T = (G_S + A_{21}Y_{11})Z_{11}^T,
\]

which by definition of \(G_S\) equals \(G_SZ_{11}^T\). \(\square\)
4.2 Application to Vandermonde-like matrices

Let us now focus on the cost of Algorithm GenInvLT when M and N correspond to the Vandermonde-like structure (20a). We assume x to be partitioned as in the previous section.

**Lemma 4.** Let the matrices A, G, H, Y, Z, A₁₁, G₁₁, H₁₁, Z₁₁ be as in Algorithm GenInvLT. Let also w₁₁ be the last column of A₁₁ and v₁₁ be the first row of A₁₁A₂₂. Then
- \( \nabla [D(x₁), Z₁₁^T, 0]|(A₁₂) = G₁₁H₁₁^T + w₁₁e₁₁^T₁₁ \),
- \( \nabla [D(x₂), Z₂₂^T]|(A₂₁) = G₂₂H₂₂^T \),
- \( \nabla [Z₁₁^T, G, D(x₁)]|(A₁₁^T) = Y₁₁Z₁₁^T \),
- \( \nabla [D(x₂), D(x₁)]|(A₁₂A₂₁) = G₁₁Z₂₂ \),
- \( \nabla [Z₂₂^T]|(A₂₂^T)|A₁₁A₂₂) = (−Y₁₁H₁₁^T + e₁₁,e₁₁^T₁₁ + v₁₁^T) \).

**Proof.** In this case, the upper-right block of N satisfies \( N₁₂ = e₁₁,e₁₁^T₁₁ \). Hence we deduce from (2) that \( \nabla [D(x₁), Z₁₁^T]|(A₁₂) = G₁₁H₁₁^T + A₁₁e₁₁,e₁₁^T₁₁ \), and the first identity follows from the definition of vector w₁₁. The second to fourth identities are obtained in the same way as in the proof of Lemma 3. Let us now verify the last identity, which displays the structure of the product \( A₁₁^T A₁₂ \). First, applying the techniques of Lemma 3, we deduce that \( \nabla [Z₁₁^T, 0]|A₁ₙ₁^T A₁₂ = −Y₁₁H₁₁^T + e₁₁,e₁₁^T₁₁ \).

Then, using (26) with \((φ, n) = (1, r₁)\) together with the definition of v₁₂ yields the announced expression. □

**Theorem 4.** Let \( n \) be a power of two and \( M, N \in \mathbb{K}^{n \times n} \) be as in (19a). Then Algorithm GenInvLT requires at most
\[ 3\log(n) MM_C(\alpha, n) + O(\alpha n \log(n)) \]
field operations. If, in addition, the set \( \{x₁, \ldots, x_n, y₁, \ldots, y_n\} \) has cardinality \( 2n \) this bound drops to
\[ 2\log(n) MM_C(\alpha, n) + O(\alpha n \log(n)) \].

**Proof.** When \( n = 1 \), \( A = (\sum_{i=1}^{n} g_i h_i)/(m₁₁ − m₁₁) \). Hence \( A⁻¹ \) can be computed using \( 2a + 1 \) operations in \( \mathbb{K} \), and the cost for \( n = 1 \) is \( C(a, 1) := 4\alpha + 1 \). Consider now the case \( n ≥ 2 \). Using Lemma 3 together with (9), we see that the matrices \( A₁₁^T, A₁₂, A₂₁, \) and their transposes are all of the Cauchy-like structure defined in (19a). Furthermore, for each of them a generator of length at most \( \alpha \) can deduced in time \( O(\alpha n) \) from the quantities computed by Algorithm GenInvLT. Consequently, one can compute \( A₁₁ Y₁₁, A₁₂ Z₁₁, A₂₁ Y₂₁ \), and \( A₁₁^T A₂₂ \) via six applications, in dimension \( n/2 \), of the reconstruction formula (19b) to \( 2n \) vectors in \( \mathbb{K}^{n/2} \). Finally, Algorithm GenInvLT uses \( 2n \) additions to deduce \( G₁₁, H₁₁, \) and the upper parts of \( Y \) and \( Z \). Overall, the cost \( n > 2 \) thus satisfies
\[ C(\alpha, n) ≤ 2C(\alpha, n/2) + 6MM_C(\alpha, n/2) + k \alpha n \]
for some constant \( k \). The superlinearity of \( MM_C(\cdot, n) \) then yields our first bound.

Assume now that the \( x_i \) and \( y_i \) are \( 2n \) pairwise distinct values. From Lemma 3 the reconstruction formula (19b) can then be applied directly to \( A₁₁ A₂₂ \) and to the transpose of \( A₁₁ A₁₂ \), in order to compute \( (A₁₁ A₂₂) Y₁₂ \) and \( (A₁₂ A₁₁) Z₂₂ \). This reduces the number of reconstructions from six to four, whence the second cost bound. □

**4.3 Extension to Hankel-like matrices**

Finally, let us consider the Hankel-like structure defined by \( M = Zₙ₀ \) and \( N = Zₙ₀^T \). Although \( M \) and \( N \) are lower triangular, Algorithm GenInvLT cannot be used directly in this case as the operator \( \nabla [Zₙ₀^T, Zₙ₀] \) is not invertible. Covering such a structure, however, is interesting in particular as it yields an immediate extension to some Toeplitz-like matrices (see [19, Remark 5.1.4] and our Section 3.1).

To cope with the singularity of the displacement operator, some additional data, called irregularity set in [19, p. 136], are needed, which typically consist in “a few” entries of \( A \). An irregularity set for \( \nabla [Zₙ₀^T, Zₙ₀] \) is given by the last row of \( A \). Indeed, for \( u^T = eₙ,n \) \( A \) we see that (2) and (26) imply
\[ \nabla [Zₙ₁, Zₙ₂]^T(A) = [G | eₙ₁] [H | u]^T, \]
so that the matrix \( A \) is Hankel-like in the sense of (21a), with displacement rank \( \alpha + 1 \). Consequently, the reconstruction formula (21b) can be used.

We need to exhibit an irregularity set for \( \nabla [Zₙ₀^T, Zₙ₀] \) too, because we shall multiply with inverses of Hankel-like matrices. A suitable choice here is \( v^T = eₙ₁ A^{-1} \), the first row of the inverse of \( A \): indeed, if \( \nabla [Zₙ₁^T, Zₙ₀^T](A^{-1}) = YZ^T \), then, recalling (13c), we may check that \( \tilde{z}_n A^{-1} \tilde{z}_n \) satisfies an identity similar to (31); it is thus fully determined by: up to reflections, \( Y, Z, \) and its last row \( v^T \).
The resulting adaptation of Algorithm GenInvHL to Hankel-like operator $\nabla[Z_{n,0}, Z_{n,0}^T]$ is as follows:

**GenInvHL(G, H, u)**

**Input:** $G, H \in \mathbb{K}^{n \times n}$ such that $\nabla[Z_{n,0}, Z_{n,0}^T](A) = GH^T$, and $u = A^T e_n, n$ (the last row of $A$).

**Assumption:** $A$ strongly regular.

**Output:** $Y = -A^{-1} G$, $Z = A^{-1} H$, and $v = A^{-1} e_{n,1}$ (the first row of $A^{-1}$).

if $n = 1$ then

$Y := -u^\top G$, $Z := u^\top H$, $v := u^{-1}$;

else

$n_1 := \lfloor n/2 \rfloor$, $n_2 := \lfloor n/2 \rfloor$;

$(W_{21}, W_{23}) := A^T e_{n,1, n}$;

$(Y_{11}, Z_{11}) := \text{GenInvHL}(G_1, H_1, u_{11})$;

$G_2 := G_2 - A_{21} Y_{11}$; $H_2 := H_2 - A_{21} Z_{11}$;

$u_{21} := u_{21} - A_{21} Y_{11}$;

$(Y_{5,5}, Z_{5,5}) := \text{GenInvHL}(G_{5,5}, H_{5,5}, u_{5,5})$;

$(i, j, v) := Y_{11} - A_{11}^T Z_{11}; Z := Z_{11} - A_{11}^T Z_{11}$;

$w := -S^T A_{12} v_{11}; v := [v_{11} - A_{12}^T Z_{11}]$;

fi;

return $(Y, Z, v)$.

**Theorem 6.** Algorithm GenInvHL is correct.

**Proof.** When $n = 1$, both $A$ and $v$ are reduced to the scalar $a_{11}$ and correctness is then straightforward. Assume now that $n > 1$ and, in order to proceed by induction, assume correctness for $n' < n$. The vector $u$ is split into $u_{21}$ in $\mathbb{K}^{n' \times 1}$ and $u_{22}$ in $\mathbb{K}^{2 \times 1}$. Similarly, the vector of coefficients of row $n_1$ of $A$ is split into $u_{11} \in \mathbb{K}^{n' \times 1}$ and $u_{12} \in \mathbb{K}^{2 \times 1}$. Hence $u_{11}$ equals $A_{11} e_{n,1, n}$ (that is, the vector of coefficients of the last row of $A_{11}$), $u_{21} = A_{21} e_{n,2, n-2}$, and $u_{22} = A_{22} e_{n,2, n-2}$. Recalling that $S = A_{22} + A_{21} A_{11}$, we deduce that the vector $w$ computed by Algorithm GenInvHL satisfies $w = S^T e_{2, n-2}$ and thus is the vector of coefficients of the last row of $S$. Since the computation of $Y$ and $Z$ is unchanged in comparison to Algorithm GenInvLT, we still have $Y = -A^{-1} G$ and $Z = A^{-1} H$. All that remains is to prove that $v$ is actually the vector of coefficients of the first row of $A^{-1}$. By induction, $v_{11}$ and $v_{5,5}$ correspond to the first rows of $A_{11}^{-1}$ and $S^{-1}$, respectively. Using the definition of $A_{11}$ seen in (8) and letting $w_{1} = -v_{11} A_{12} S^{-1}$, we get:

$e_{n,1}^T A^{-1} = \begin{bmatrix} \epsilon_{n,1}^T - e_{n,1}^T A_{11}^{-1} A_{12} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} A_{12} S^{-1} \end{bmatrix} E$

$\epsilon_{11} = \begin{bmatrix} v_{11}^T \end{bmatrix} E$

$\epsilon_{11} = \begin{bmatrix} v_{11}^T - w^T A_{12} A_{11}^{-1} \end{bmatrix} w^T$,

which is exactly the way vector $v$ is computed.

**Lemma 5.** Let $A, G, H, Y_{11}, Z_{11}, G_5, H_5, u_{11}$ be as in Algorithm GenInvHL. Recall that $u_{11}$ is the last row of the matrix $A_{11}$ and let $w_{11}$ be its last column. Then

- $\nabla[Z_{n,0}, Z_{n,0}^T](A_{12}) = G_1 H_2^T + w_{11} e_{n,1,1}$,
- $\nabla[Z_{n,0}, Z_{n,0}^T](A_{21}) = G_2 H_1^T - e_{n,2,1} u_{11}^T$,
- $\nabla[Z_{n,0}, Z_{n,0}^T](A_{11}^T) = Y_{11} Z_{11}^T$,
- $\nabla[Z_{n,0}, Z_{n,0}^T](A_{12}^T A_{12}) = -Y_1 H_2^T + e_{n,1,1} T_{n,1,1,1}$.

**Theorem 7.** Let $n$ be a power of two and $M, \alpha, n \in \mathbb{K}^{n \times n}$ be as in (21a). Then Algorithm GenInvHL requires at most $2 \log(n) M M_M(\alpha, n) + O(\alpha M(n) \log(n))$ field operations.

**Proof.** Proceed as for Lemma 3 and Lemma 4.

**5. EXPERIMENTAL RESULTS**

We have implemented the two variants of GenInvLT (with and without Cardinal’s trick) as well as the MBA algorithm for Sylvester’s displacement. Moreover, we have developed some code to handle Cauchy-like and Hankel-like structures.

For our experiments, we take $\mathbb{K} = \mathbb{F}_p$ with $p = 999999937$, which lets us measure the algebraic costs. Basic operations in $\mathbb{K}$ are provided by NTL, and we also use some code for fast polynomial arithmetic. All the computations are carried out on a desktop machine with an Intel® Core™ i7-2600 K processor at 3.40 GHz. Finally, generators $(G,H)$ are picked randomly, while operator matrices $D(x), D(y)$ are chosen in order to satisfy all the assumptions made on the algorithms.

Figure 1 shows computing times for inverting Cauchy-like matrices of displacement rank $\alpha = 10$ when $n$ is increasing. It appears that the computing time is quasi-linear with respect to $n$ for each method, and that the compression steps in MBA have negligible cost. Thus, the main difference explaining the various performances lies in the number of products “Cauchy-like matrix $\times$ vectors.” We have already seen in Theorem 4 that the choice in the parentheses leads to one variant in $3 \log(n)$ $\mathbb{M}_M(\alpha, n)$ and, up to stronger conditions on the input, to another variant in $2 \log (n) M M_M(\alpha, n)$. Let us now estimate this cost for our implementation of the MBA algorithm. Generators for the Schur complement and the inverse of $A$ before the compression steps are computed using (10) according to the following parenthesisization: $X_1 = A_{11}^{-1} A_{12}$, $S = A_{22} - A_{21} X_1$.

1\text{http://www.math.uvsq.fr/~lecerf/software/tellegen/}

2\text{http://www.shoup.net/ntl/}
superlinearity of additional time due to subblocks having displacement rank
of stopping recursive calls (and reconstructing...

cost dominated by $14 \log(n)$ MM$_C(\alpha,n)$ in the recurrence equation for the cost of MBA, which gives a total cost dominated by $14 \log(n)$ MM$_C(\alpha,n)$. In Figure 1, we observe a speed-up around $4.6 \approx 14/3$ between MBA and our first variant (GenInvLT), and around $6.7 \approx 14/2$ between MBA and the second variant (GenInvLT + Cardinal’s trick), which is in agreement with our analysis above.

Moreover, we experimented with Hankel-like matrices in order to estimate the cost of row reconstruction and the additional time due to subblocks having displacement rank $\alpha + 1$ instead of $\alpha$ like in the Cauchy-like case. Timings are summarized in Table 1, where it appears that these costs become negligible when $\alpha$ is large enough. Indeed, they are linear in $\alpha$ as expected (see (22b)) and our first variant (GenInvLT), and around 6 when

\[
\begin{align*}
X_2 &= A_{12}A_{11}^{-1}, \\
A^{-1} &= \begin{bmatrix} A^{-1}_{\alpha-1} + (x_i 5^{-1}) x_i & -x_i 5^{-1} \\ -5^{-1} x_i & 5^{-1} \end{bmatrix}.
\end{align*}
\]

Counting the costs of all these products using (23) and the superlinearity of MM$_C(\cdot,n)$ leads to a bound of $14$ MM$_C(\alpha,n)$ in the recurrence equation for the cost of MBA, which gives a total cost dominated by $14 \log(n)$ MM$_C(\alpha,n)$. In Figure 1, we observe a speed-up around $4.6 \approx 14/3$ between MBA and our first variant (GenInvLT), and around $6.7 \approx 14/2$ between MBA and the second variant (GenInvLT + Cardinal’s trick), which is in agreement with our analysis above.

Moreover, we experimented with Hankel-like matrices in order to estimate the cost of row reconstruction and the additional time due to subblocks having displacement rank $\alpha + 1$ instead of $\alpha$ like in the Cauchy-like case. Timings are summarized in Table 1, where it appears that these costs become negligible when $\alpha$ is large enough. Indeed, they are linear in $\alpha$ as expected (see (22b)) and Theorem 5) while the total cost seems quadratic in $\alpha$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>10</th>
<th>30</th>
<th>50</th>
<th>70</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cost</td>
<td>4.7</td>
<td>34.7</td>
<td>92.1</td>
<td>177.0</td>
<td>290.3</td>
</tr>
<tr>
<td>Irregularity related cost</td>
<td>0.8</td>
<td>2.5</td>
<td>3.9</td>
<td>5.3</td>
<td>7.1</td>
</tr>
<tr>
<td>Rank increase related cost</td>
<td>0.5</td>
<td>1.6</td>
<td>2.6</td>
<td>3.5</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Table 1: Cost (in seconds) of Hankel-like matrix inversion for $n = 200$ and increasing values of $\alpha$.

6. CONCLUSIONS

In this paper, we have extended Cardinal’s compression-free algorithm to a broader class of structured matrices, including not only the Cauchy-like type but also the Vandermonde, Hankel, and Toeplitz-like types. Our main conclusion is that this approach yields variants of the MBA algorithm that are simpler to analyze and implement, and, according to our first experiments, significantly faster in practice. However, this study calls for a number of extensions:

On the practical side, we should first study the impact of stopping recursive calls (and reconstructing $A^{-1}$ explicitly) via fast dense linear algebra) when $n \approx \alpha$. It would also be interesting to measure the memory gains brought by Cardinal’s extended approach over MBA.

On the algorithmic side, although we have focused only on $O'(\alpha^2 n)$ versions of MBA, it would be interesting to incorporate the matrix multiplication techniques of [3]. We should also study the impact of multiplicities in $x$ and $y$ on the cost bounds and adapt our work to structures like those of the Toeplitz+Hankel-like type.

7. ACKNOWLEDGMENTS

We thank Benoit Lacelle for providing a framework for structured matrices and the generator-compression subroutine used in our implementation of MBA, and Eric Schost for pointing out the code used for multipoint evaluation.

8. REFERENCES

APPENDIX

A. RECONSTRUCTION OF RECTANGULAR HANKEL-LIKE MATRICES

Theorem 8. Equation (21b) is correct.

Proof. The Hankel-like matrix $A \in \mathbb{K}^{n \times m}$ satisfies

$$Z_{n,1}A - AZ_{m,0}^T = GH^T.$$  

By left-multiplying with $Z_{n,1}^T$, we deduce the following recurrence formula:

$$A = Z_{n,1}^T A Z_{m,0} + Z_{n,1}^T G H^T.$$  

We can now apply Theorem 4.3.6 from [19] to obtain that, for all integer $k$,

$$A = (Z_{n,1})^k A (Z_{m,0})^k + \sum_{l=0}^{k-1} (Z_{n,1})^{l+1} G H^T (Z_{m,0})^l.$$  

As $(Z_{m,0})^l$ is the null matrix as soon as $l \geq m$, we can simplify the above equation:

$$A = \sum_{l=0}^{m-1} (Z_{n,1})^{l+1} G H^T (Z_{m,0})^l$$

$$= \sum_{l=0}^{m-1} (Z_{n,1})^{l+1} \left( \sum_{j=1}^\alpha g_j h_j^T \right) (Z_{m,0})^l$$

$$= \sum_{j=1}^\alpha \sum_{l=0}^{m-1} (Z_{n,1})^{l+1} g_j h_j^T (Z_{m,0})^l.$$  

The inner sum is a sum of $m$ outer products. This can be rewritten as a product of two matrices $B \in \mathbb{K}^{n \times m}$ and $C \in \mathbb{K}^{m \times m}$ by taking $(Z_{n,1})^{l+1} g_j$ for column $m - l$ of $B$ and $h_j^T (Z_{m,0})^l$ for row $m - l$ of $C$. Thus, the $i$th row of $C$ is exactly $[0 \ldots 0 h_{j,i} \ldots h_{j,m}]$ so that $C = \mathbb{L}(h_j) \mathbb{J}_m$.

Finally, the last column of $B$ is $[g_{j,2} \ldots g_{j,n}, g_{j,1}]^T$ and we can deduce its column $l$ by applying $Z_{n,1}^T$ to its column $l + 1$, that is by moving the elements of column $l + 1$ one location up (except for the first element which goes at bottom). This defines a Toeplitz matrix whose coefficient $(k,l)$ is $g_{j,1+(k-l+m)} \mod n$, which is $\mathbb{T}^{n \times m}(g_j)$ by definition. □