Structures preserved by Schur complementation

Steven Delvaux *Marc Van Barel*

Report TW 415, December 2004

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Structures preserved by Schur complementation

Steven Delvaux   *Marc Van Barel*

Report TW 415, December 2004

Department of Computer Science, K.U.Leuven

Abstract

In this paper we investigate some matrix structures on $\mathbb{C}^{m\times n}$ that have a good behaviour under Schur complementation. The first type of structure is closely related to low displacement rank matrices. Next, we show that for a matrix having a low rank submatrix, also the Schur complement must have a low rank submatrix, which we can explicitly determine. This property holds even if the low rank submatrix contains a certain correction term, which we call the shift matrix.

Keywords : displacement structures, rank structures, lower semiseparable (plus diagonal) matrices, Schur complements.
AMS(MOS) Classification : Primary : 15A09, Secondary : 15A03,65F05.

Structures preserved by Schur complementation

Steven Delvaux *, Marc Van Barel *

24th December 2004

Abstract

In this paper we investigate some matrix structures on $\mathbb{C}^{m \times n}$ that have a good behaviour under Schur complementation. The first type of structure is closely related to low displacement rank matrices. Next, we show that for a matrix having a low rank submatrix, also the Schur complement must have a low rank submatrix, which we can explicitly determine. This property holds even if the low rank submatrix contains a certain correction term, which we call the shift matrix.

Keywords: displacement structures, rank structures, lower semiseparable (plus diagonal) matrices, Schur complements

1 Introduction

In this paper we will handle several matrix structures that are preserved by Schur complementation, as a continuation of [1] where we handled structures preserved by matrix inversion. Nevertheless, all results will be developed independently of [1].

Section 2 deals with the preservation of *displacement structures*. As in [1], the idea is to generalize the classical examples of displacement structures (Toeplitz-like, Cauchy-like, Vandermonde-like, circulant matrices etc., see [10]), by 'decoupling' the displacement equation. This means that the displacement equation is allowed to involve two variables A and B rather than only one variable A. We will then illustrate this definition by some examples.

*Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium. email: {Steven.Delvaux,Marc.VanBarel} @cs.kuleuven.ac.be.

The research was partially supported by the Research Council K.U.Leuven, project OT/00/16 (SLAP: Structured Linear Algebra Package), by the Fund for Scientific Research–Flanders (Belgium), projects G.0078.01 (SMA: Structured Matrices and their Applications), G.0176.02 (ANCILA: Asymptotic aNalysis of the Convergence behavior of Iterative methods in numerical Linear Algebra), G.0184.02 (CORFU: Constructive study of Orthogonal Functions) and G.0455.0 (RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation), and by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, project IUAP V-22 (Dynamical Systems and Control: Computation, Identification & Modelling). The scientific responsibility rests with the authors.

Section 3 handles the preservation of what we call *rank structures*. As in [2, 1], such a structure is defined as a collection of *structure blocks*: these are low rank submatrices of a given matrix $A \in \mathbb{C}^{m\times n}$, together with a certain correction term Λ, called the shift matrix. We will prove that these rank structures are preserved under Schur complementation, and we provide some examples to illustrate this.

Section 4 considers *Möbius transformations* of a matrix A. Each Möbius transformation can be realized as the Schur complement of a very special block matrix, and hence this connection can be used to translate the preservation results of Schur complements into properties of Möbius transformations.

For further reference, let us recall here some basic definitions and properties of Schur complements [5].

Definition 1 *Given $A \in \mathbb{C}^{m\times n}$, and given $k \in \mathbb{N}$. We define the k*-partitioning *of A as*

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}_k, \tag{1}$$

with $A_{1,1} \in \mathbb{C}^{k\times k}$. We define the Schur complement *induced by this k-partitioning as*

$$S_{A,k} := A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2},$$

where we supposed that $A_{1,1}$ is invertible. We denote with $A_{\mathrm{c},1}$ the first block column and with $A_{\mathrm{r},1}$ the first block row of (1).

Schur complements are related to Gaussian elimination steps on A with pivot block $A_{1,1}$, in the sense that

$$L_{Gauss} \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} R_{Gauss} = \begin{bmatrix} A_{1,1} & 0 \\ 0 & S_{A,k} \end{bmatrix}, \tag{2}$$

where

$$L_{Gauss} := \begin{bmatrix} I & 0 \\ -A_{2,1}A_{1,1}^{-1} & I \end{bmatrix}, \quad R_{Gauss} := \begin{bmatrix} I & -A_{1,1}^{-1}A_{1,2} \\ 0 & I \end{bmatrix},$$

which are unit block lower and upper triangular matrices, respectively. Hence the following lemma should not come as a surprise.

Lemma 2 *Given $L \in \mathbb{C}^{l\times m}$, $A \in \mathbb{C}^{m\times n}$ and $R \in \mathbb{C}^{n\times p}$. Suppose we can partition*

$$L = \begin{bmatrix} L_{1,1} & 0 \\ L_{1,2} & L_{2,2} \end{bmatrix}, \quad A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \text{ and } R = \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0 & R_{2,2} \end{bmatrix},$$

with $L_{1,1}$, $A_{1,1}$ and $R_{1,1}$ in $\mathbb{C}^{k\times k}$ nonsingular. Then

$$S_{LAR,k} = L_{2,2}S_{A,k}R_{2,2}.$$

PROOF. First, let us prove the property for the case $R = I$. Then we can expand the matrix LA as

$$\begin{bmatrix} L_{1,1}A_{1,1} & L_{1,1}A_{1,2} \\ L_{1,2}A_{1,1} + L_{2,2}A_{2,1} & L_{1,2}A_{1,2} + L_{2,2}A_{2,2} \end{bmatrix},$$

from which it follows that

$$\begin{aligned} S_{LA,k} = L_{1,2}A_{1,2} + L_{2,2}A_{2,2} - (L_{1,2}A_{1,1} + L_{2,2}A_{2,1})A_{1,1}^{-1}L_{1,1}^{-1}(L_{1,1}A_{1,2}) \\ = L_{2,2}[A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}] = L_{2,2}S_{A,k}, \end{aligned}$$

as we had to prove. In a similar way, one can prove the property for the case $L = I$. The general result follows then by composing these two results. □

Note that in particular, it follows from the above lemma that the left multiplication of A with a matrix $\begin{bmatrix} X & 0 \\ X & I \end{bmatrix}$, or the right multiplication with the transpose of such a matrix always preserves the Schur complement of A.

We may recall also the so-called *transitivity* of Schur complements, i.e. the fact that $S_{S_{A,k},l} = S_{A,k+l}$ whenever all involved Schur complements are defined. The underlying reason is that for a Gaussian elimination step applied on A, the same result is obtained if the Gaussian elimination step is split up in two separate steps with smaller pivot blocks. Alternatively, one can prove this property by direct computation.

2 Displacement structures

In this section we handle the preservation of displacement structure under Schur complementation. As a general reference, we can refer to [10, 9] for an overview of the many applications of displacement theory in numerical linear algebra. Some references of historical interest are [3, 8].

2.1 Sylvester type displacement

First we handle Sylvester type displacement equations. As a classical example, let A be a Hankel matrix, i.e. $A = [a_{i+j}]_{i,j=1}^{n}$. Putting $Z := [\mathbf{e}_2 \dots \mathbf{e}_n \mathbf{0}]$ with $\mathbf{e}_k$ the kth column of the identity matrix, it is easy to check that

$$AZ^T - ZA = \text{Rk } 2, \tag{3}$$

where Rk 2 denotes a matrix of rank at most 2.

Generalizing, we can come to a more general definition. The main difference with (3) is that the variable A is 'decoupled' into two variables A and B.

Definition 3 *Let A, B, F and G be rectangular matrices and let $r \in \mathbb{N}$. We say A and B to satisfy the* Sylvester type displacement equation *induced by (F, G, r) if*

$$AF - GB = \text{Rk } r, \tag{4}$$

where Rk r *denotes a matrix of rank at most* r*.*

Here we supposed (4) to be well-defined, or equivalently we supposed the block matrix

$$\begin{bmatrix} A & G \\ F^T & B^T \end{bmatrix} \tag{5}$$

to have compatible matrix dimensions. Moreover, this block representation (5) is useful in several other aspects, as will become clear soon.

Let us show that for F^T and G block lower triangular, Sylvester type displacement structure is preserved under Schur complementation. This generalizes the corresponding property for the case $A = B$ (see [10]).

Theorem 4 *(Sylvester type inheritance:) Let* $k, l \in \mathbb{N}$*, and suppose that*

$$\left[\begin{array}{c|c} A & G \\ \hline F^T & B^T \end{array}\right] = \left[\begin{array}{cc|cc} A_{1,1} & A_{1,2} & G_{1,1} & 0 \\ A_{2,1} & A_{2,2} & G_{2,1} & \tilde{G} \\ \hline F_{1,1}^T & 0 & B_{1,1}^T & B_{2,1}^T \\ F_{1,2}^T & \tilde{F}^T & B_{1,2}^T & B_{2,2}^T \end{array}\right], \tag{6}$$

with $A_{1,1} \in \mathbb{C}^{k\times k}$ *and* $B_{1,1} \in \mathbb{C}^{l\times l}$*. If*

$$AF - GB = \text{Rk } r, \tag{7}$$

where Rk r *denotes a matrix of rank at most* r*, then*

$$S_{A,k}\tilde{F} - \tilde{G}S_{B,l} = \widetilde{\text{Rk}}\ r, \tag{8}$$

where $\widetilde{\text{Rk}}\ r$ *denotes a new matrix of rank at most* r*.*

PROOF. We will prove the theorem for A and B square and nonsingular. (For the general case, see the paragraph following this proof). Multiplying (7) on the left with A^{-1} and on the right with B^{-1}, it follows that

$$FB^{-1} - A^{-1}G = \text{Rk } r, \tag{9}$$

with Rk r a new matrix of rank at most r. Now we use the fact that for any matrix A, the $(2,2)$ block element of A^{-1} is precisely the inverse of the Schur complement $S_{A,k}$. (Proof: invert both sides of (2)). Using this, and using the partitioning in (6), it follows by evaluating the $(2,2)$ block element of (9) that

$$\tilde{F}S_{B,l}^{-1} - S_{A,k}^{-1}\tilde{G} = \text{Rk } r,$$

with Rk r a new matrix of rank at most r. Hence by multiplying on the left with $S_{A,k}$ and on the right with $S_{B,l}$, we obtain the desired equation (8). □

Although the above proof of inheritance of structure is rather 'clean', it only worked for square and nonsingular matrices. (One could use a reduction to square matrices and a 'continuity argument' to remove these restrictions, but we will not do this here, due to the complexity of the argument). Furthermore, the proof of Theorem 4 gives rather complicated formulae for the new $\widetilde{\mathrm{Rk}\ r}$ matrix in the right hand side of (8).

To address these questions, one can proceed in a more direct way by directly computing the Schur complements.

Let us work this out. Thus we start with the equation

$$AF - GB = \mathrm{Rk}\ r =: UV, \tag{10}$$

with U having r columns and V having r rows. Let us recall the general property

$$A - A_{\mathrm{c},1}A_{1,1}^{-1}A_{\mathrm{r},1} = 0 \oplus S_{A,k}, \tag{11}$$

where we used the notations of Definition 1. Keeping in mind this property, and keeping in mind the partitioning in (6), we have

$$\begin{aligned} 0 \oplus \widetilde{\mathrm{Rk}\ r} &:= 0 \oplus (S_{A,k}\tilde{F} - \tilde{G}S_{B,l}) \\ &= (0 \oplus S_{A,k})F - G(0 \oplus S_{B,l}) \\ &= (A - A_{\mathrm{c},1}A_{1,1}^{-1}A_{\mathrm{r},1})F - G(B - B_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1}) \\ &= UV - A_{\mathrm{c},1}A_{1,1}^{-1}A_{\mathrm{r},1}F + GB_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1}, \end{aligned} \tag{12}$$

where the last transition follows from (10). Still keeping in mind (10) and the partitioning in (6), we can further work this out as

$$\begin{aligned} &= UV - A_{\mathrm{c},1}A_{1,1}^{-1}(U_{\mathrm{r},1}V + \underline{G_{1,1}B_{\mathrm{r},1}}) + (\underline{A_{\mathrm{c},1}F_{1,1}} - UV_{\mathrm{c},1})B_{1,1}^{-1}B_{\mathrm{r},1} \\ &= UV - A_{\mathrm{c},1}A_{1,1}^{-1}U_{\mathrm{r},1}V - UV_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1} + \underline{A_{\mathrm{c},1}A_{1,1}^{-1}U_{\mathrm{r},1}V_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1}} \\ &= (U - A_{\mathrm{c},1}A_{1,1}^{-1}U_{\mathrm{r},1})(V - V_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1}). \end{aligned} \tag{13}$$

We see from this that $\widetilde{\mathrm{Rk}\ r}$ is indeed a matrix of rank at most r, which we can explicitly determine. Moreover, the only condition for the above derivation to be valid was the nonsingularity of $A_{1,1}$ and $B_{1,1}$, i.e. the existence of the Schur complements $S_{A,k}$ and $S_{B,l}$.

2.2 Stein type displacement

We come to a second type of displacement structure.

Definition 5 *Let A, B, G and H be rectangular matrices and let $r \in \mathbb{N}$. We say A and B to satisfy the* Stein type displacement equation *induced by (G, H, r) if*

$$A - GBH = \mathrm{Rk}\ r, \tag{14}$$

where Rk r *denotes a matrix of rank at most r.*

Here we supposed (14) to be well-defined, or equivalently we supposed the block matrix

$$\begin{bmatrix} A & G \\ H & B^T \end{bmatrix} \tag{15}$$

to have compatible matrix dimensions. Moreover, this block representation (15) is useful in several other aspects, as will become clear soon.

As in the Sylvester case, for H^T and G block lower triangular, Stein type displacement structure will be preserved under Schur complementation. This generalizes again the corresponding property for the case $A = B$ (see [10]).

Theorem 6 *(Stein type inheritance:) Let $k, l \in \mathbb{N}$, and suppose that*

$$\left[\begin{array}{c|c} A & G \\ \hline H & B^T \end{array}\right] = \left[\begin{array}{cc|cc} A_{1,1} & A_{1,2} & G_{1,1} & 0 \\ A_{2,1} & A_{2,2} & G_{2,1} & \tilde{G} \\ \hline H_{1,1} & H_{1,2} & B_{1,1}^T & B_{2,1}^T \\ 0 & \tilde{H} & B_{1,2}^T & B_{2,2}^T \end{array}\right], \tag{16}$$

with $A_{1,1} \in \mathbb{C}^{k\times k}$ and $B_{1,1} \in \mathbb{C}^{l\times l}$. If

$$A - GBH = \mathrm{Rk}\ r, \tag{17}$$

then

$$S_{A,k} - \tilde{G} S_{B,l} \tilde{H} = \widetilde{\mathrm{Rk}}\ \tilde{r}, \tag{18}$$

where $\tilde{r} := r - k + l$.

PROOF. We will prove the theorem for $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ square and nonsingular. (For the general case, see the paragraph following this proof). Applying [1, Theorem 2], it follows from (17) that

$$B^{-1} - HA^{-1}G = \mathrm{Rk}(r + n - m). \tag{19}$$

Now we recall the fact that for any matrix A, the $(2, 2)$ block element of A^{-1} is precisely the inverse of the Schur complement $S_{A,k}$. Using this, and using the partitioning in (16), it follows by evaluating the $(2, 2)$ block element of (19) that

$$S_{B,l}^{-1} - \tilde{H} S_{A,k}^{-1} \tilde{G} = \mathrm{Rk}(r + n - m),$$

with $\mathrm{Rk}(r + n - m)$ still a matrix of rank at most $r + n - m$. Hence by applying again [1, Theorem 2], we obtain the desired equation (18), i.e.

$$S_{A,k} - \tilde{G} S_{B,l} \tilde{H} = \widetilde{\mathrm{Rk}}\,(r + n - m + (m - k) - (n - l)) =: \widetilde{\mathrm{Rk}}\ \tilde{r},$$

with $\tilde{r} := r - k + l$. □

Again, the above proof was only valid for A and B square and nonsingular. Instead of showing theoretically that these restrictions are not essential (by using a reduction to square matrices, together with a 'continuity argument' to remove the nonsingularity condition), let us indicate how to prove the theorem by a direct approach.

We start with the equation

$$A - GBH = \mathrm{Rk}\ r =: UV, \tag{20}$$

with U having r columns and V having r rows. In a similar way as the derivation of (12), we obtain

$$0 \oplus \widetilde{\mathrm{Rk}}\ r = UV - A_{\mathrm{c},1}A_{1,1}^{-1}A_{\mathrm{r},1} + GB_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1}H. \tag{21}$$

Then keeping in mind (20) and the partitioning in (16), we can further work out the right hand side of (21) as

$$\begin{aligned} &= UV - (UV_{\mathrm{c},1} + GB_{\mathrm{c},1}H_{1,1})A_{1,1}^{-1}(U_{\mathrm{r},1}V + G_{1,1}B_{\mathrm{r},1}H) \\ &\quad +GB_{\mathrm{c},1}B_{1,1}^{-1}B_{\mathrm{r},1}H \\ &= U(I - V_{\mathrm{c},1}A_{1,1}^{-1}U_{\mathrm{r},1})V - UV_{\mathrm{c},1}A_{1,1}^{-1}G_{1,1}B_{\mathrm{r},1}H \\ &\quad -GB_{\mathrm{c},1}H_{1,1}A_{1,1}^{-1}U_{\mathrm{r},1}V - GB_{\mathrm{c},1}(H_{1,1}A_{1,1}^{-1}G_{1,1} - B_{1,1}^{-1})B_{\mathrm{r},1}H. \end{aligned} \tag{22}$$

To proceed further, we will assume that

Assumption: $k = l$.

Then we claim that there exist matrices $X_1 \in \mathbb{C}^{k\times r}$, $X_3 \in \mathbb{C}^{r\times k}$, $X_2, X_4 \in \mathbb{C}^{r\times r}$ that satisfy the embedding relation

$$\begin{bmatrix} B_{1,1}^{-1} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} H_{1,1} & X_1 \\ V_{\mathrm{c},1} & X_2 \end{bmatrix} \begin{bmatrix} A_{1,1}^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} G_{1,1} & U_{\mathrm{r},1} \\ X_3 & X_4 \end{bmatrix}, \tag{23}$$

where all involved matrices are square of size $k + r$. Assuming this for the moment, then we have the componentwise equations

$$\begin{cases} B_{1,1}^{-1} &= H_{1,1}A_{1,1}^{-1}G_{1,1} + X_1X_3 \\ 0 &= H_{1,1}A_{1,1}^{-1}U_{\mathrm{r},1} + X_1X_4 \\ 0 &= V_{\mathrm{c},1}A_{1,1}^{-1}G_{1,1} + X_2X_3 \\ I &= V_{\mathrm{c},1}A_{1,1}^{-1}U_{\mathrm{r},1} + X_2X_4. \end{cases}$$

Hence (22) can be rewritten as

$$\begin{aligned} &= UX_2X_4V + UX_2X_3B_{\mathrm{r},1}H + GB_{\mathrm{c},1}X_1X_4V + GB_{\mathrm{c},1}X_1X_3B_{\mathrm{r},1}H \\ &= (UX_2 + GB_{\mathrm{c},1}X_1)(X_4V + X_3B_{\mathrm{r},1}H). \end{aligned} \tag{24}$$

We see from this that $\widetilde{\mathrm{Rk}}\ r$ is indeed a matrix of rank at most r, which we can explicitly determine in terms of X_1, X_2, X_3 and X_4.

To prove the solvability of the embedding relation (23) is beyond the scope of the paper. We may notice that it suffices to find X_i, $i = 1, \dots, 4$ which solve the equivalent embedding relation

$$\begin{bmatrix} A_{1,1} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} G_{1,1} & U_{\mathrm{r},1} \\ X_3 & X_4 \end{bmatrix} \begin{bmatrix} B_{1,1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} H_{1,1} & X_1 \\ V_{\mathrm{c},1} & X_2 \end{bmatrix}, \tag{25}$$

where the $(1, 1)$ block element of this equation is nothing but the equality $A_{1,1} = G_{1,1}B_{1,1}H_{1,1} + U_{\mathrm{r},1}V_{\mathrm{c},1}$, which is satisfied by (20). To prove that also the other block elements of this equation can be satisfied, we refer to [9] where completely similar problems are handled.

Finally, we recall our above assumption that $k = l$. For $k \neq l$, we should additionally look at two special cases. First is when $k \neq 0$ and $l = 0$: then $B_{1,1}$ is the empty matrix, and hence the partitioning in (16) implies $G = \begin{bmatrix} 0 \\ \tilde{G} \end{bmatrix}$ and $H = \begin{bmatrix} 0 & \tilde{H} \end{bmatrix}$. Thus the displacement equation $A - GBH = \mathrm{Rk}\ r$ can be rewritten as

$$A - (0 \oplus \tilde{G}B\tilde{H}) = \mathrm{Rk}\ r. \tag{26}$$

Applying on (26) the block unit lower and upper triangular transformations L_{Gauss} and R_{Gauss} appearing in (2), we obtain

$$A_{1,1} \oplus (S_{A,k} - \tilde{G}B\tilde{H}) = \mathrm{Rk}\ r,$$

with Rk r a new matrix of rank at most r. Hence it follows that indeed $S_{A,k} - \tilde{G}B\tilde{H} = \widetilde{\mathrm{Rk}}\ \tilde{r}$ with $\tilde{r} := r - k$.

The second special case is when $l \neq 0$ and $k = 0$, and then it can be seen by a similar argument that indeed $A - \tilde{G}S_{B,l}\tilde{H} = \widetilde{\mathrm{Rk}}\ \tilde{r}$ with $\tilde{r} := r + l$.

The general case $k \neq l$ follows then by combining the results for $k = l$ together with the above two special cases, by using the 'transitivity' of Schur complements. We will not go further into this.

2.3 Stein-Sylvester hybrid displacement

The reader will have noticed a lot of similarity between the Sylvester and Stein type displacement. In fact, there exist also *Stein-Sylvester hybrid displacement structures*, in the sense described by Kailath and Sayed [9, Section 7.4].

To introduce these structures, let us start from the Sylvester type displacement equation $AF - GB = \mathrm{Rk}\ r$. Suppose we can factor

$$A := E\tilde{A}, \quad B := \tilde{B}H, \tag{27}$$

for certain block lower triangular matrices E, H^T with $E_{1,1}$ and $H_{1,1}$ square and nonsingular. Then Lemma 2 implies that

$$S_{A,k} := E_{2,2}S_{\tilde{A},k}, \quad S_{B,l} := S_{\tilde{B},l}H_{2,2},$$

and moreover it is easy to see that by substituting (27) into (13), the latter transforms into the expression

$$\widetilde{\mathrm{Rk}}\ r = (U - E\tilde{A}_{\mathrm{c},1}\tilde{A}_{1,1}^{-1}E_{1,1}^{-1}U_{\mathrm{r},1})(V^T - V_{\mathrm{c},1}^T H_{1,1}^{-1}\tilde{B}_{1,1}^{-1}\tilde{B}_{\mathrm{r},1}H). \tag{28}$$

Similarly, suppose that in the *Stein* type displacement equation $A - GBH = \mathrm{Rk}\ r$ we can factor $A := E\tilde{A}F$, for certain block lower triangular matrices E, F^T with $E_{1,1}$ and $F_{1,1}$ square and nonsingular. Then Lemma 2 implies that

$$S_{A,k} = E_{2,2}S_{\tilde{A},k}F_{2,2},$$

and moreover it is easy to see that equation (24) remains invariant; the only change is that the embedding relation (23) must be updated by substituting $A_{1,1}^{-1} := F_{1,1}^{-1}\tilde{A}_{1,1}^{-1}E_{1,1}^{-1}$.

Note that in both cases, we were led to an equation of the form

$$EAF + GBH = \mathrm{Rk}\ r.$$

In particular, the block matrix

$$\begin{bmatrix} E & G & 0 \\ A^T & 0 & F \\ 0 & B^T & H \end{bmatrix} \tag{29}$$

must have compatible matrix dimensions. We can then resume the above facts in the following theorem.

Theorem 7 *(Stein-Sylvester hybrid inheritance:) Let $k, l \in \mathbb{N}$, and consider the k-partitioning of A and the l-partitioning of B (Definition 1). Partition the block matrix (29) accordingly with A and B, and suppose that E, F^T, G and H^T are block lower triangular w.r.t. this partitioning, and such that each of the sets $\{E_{1,1}, G_{1,1}\}$ and $\{F_{1,1}, H_{1,1}\}$ contains at least one square and nonsingular matrix, i.e.*

$$\left[\begin{array}{c|c|c} E & G & 0 \\ \hline A^T & 0 & F \\ \hline 0 & B^T & H \end{array}\right] = \left[\begin{array}{cc|cc|cc} E_{1,1} & 0 & G_{1,1} & 0 & 0 & 0 \\ E_{2,1} & \tilde{E} & G_{2,1} & \tilde{G} & 0 & 0 \\ \hline A_{1,1}^T & A_{2,1}^T & 0 & 0 & F_{1,1} & F_{1,2} \\ A_{1,2}^T & A_{2,2}^T & 0 & 0 & 0 & \tilde{F} \\ \hline 0 & 0 & B_{1,1}^T & B_{2,1}^T & H_{1,1} & H_{1,2} \\ 0 & 0 & B_{1,2}^T & B_{2,2}^T & 0 & \tilde{H} \end{array}\right].$$

If

$$EAF + GBH = \mathrm{Rk}\ r, \tag{30}$$

then

$$\tilde{E}S_{A,k}\tilde{F} + \tilde{G}S_{B,l}\tilde{H} = \widetilde{\mathrm{Rk}}\ \tilde{r}, \tag{31}$$

with $\tilde{r} := r$ if either both $\{E_{1,1}, H_{1,1}\}$ or both $\{F_{1,1}, G_{1,1}\}$ are square and nonsingular, and $\tilde{r} := r - k + l$ if either both $\{E_{1,1}, F_{1,1}\}$ or both $\{G_{1,1}, H_{1,1}\}$ are square and nonsingular.

PROOF. This follows from the paragraph preceding the statement of the theorem. We even showed there how to update the explicit formulae for the new $\widetilde{\mathrm{Rk}}\ \tilde{r}$ matrix, if so desired. □

As an application of Stein-Sylvester hybrid displacement structure, we will use it to establish a converse to the reasoning in the proof of Theorem 6, i.e. we will show how the preservation of structure under matrix inversion [1, Theorem

2] is a consequence of the preservation of structure under Schur complementation. Thus let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be nonsingular matrices satisfying $A - GBH = \mathrm{Rk}\ r$, for arbitrary G and H. Hence

$$\begin{bmatrix} A - GBH & G - G \\ H - H & 0 \end{bmatrix} = \mathrm{Rk}\ r,$$

or by a small calculation

$$\begin{bmatrix} I & 0 \\ 0 & H \end{bmatrix}\begin{bmatrix} A & -I \\ I & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & G \end{bmatrix} - \begin{bmatrix} G & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} B & -I \\ I & 0 \end{bmatrix}\begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix} = \mathrm{Rk}\ r.$$

But now the Schur complements in this last equation are precisely the inverse matrices A^{-1} and B^{-1}. Hence by Theorem 7, it follows that

$$HA^{-1}G - B^{-1} = \widetilde{\mathrm{Rk}}\ \tilde{r},$$

with $\tilde{r} := r + n - m$, as we had to prove.

Remark 8 1. *We showed now how the inversion result of [1, Theorem 2] leads to Theorem 6, and conversely. This may seem a circular reasoning; but recall that we also indicated how to give a direct proof of Theorem 6, hence avoiding the use of [1, Theorem 2].*

2. *Note that in case $A = B$, the procedure of the preceding paragraph suggests an efficient way to compute A^{-1} by the* generalized Schur algorithm *[10]. A treatment of such computational aspects for $A \neq B$, will be the subject of the next subsection.*

2.4 Computational aspects

The preservation results of this section were in the first place *theoretically* oriented, in the sense that there seems to be no analog if $A \neq B$ for the so-called generalized Schur algorithm [10].

To state the problem more precisely, let us first make some assumptions. Suppose that F^T, G (for Sylvester type displacement) and H^T, G (for Stein type displacement) are not just *block* lower triangular, but completely lower triangular matrices. Then the preservation of structure holds for any choice of indices $k = l$. Moreover, by the transitivity of Schur complements we are allowed to recursively pull off rows and columns of A and B, one at a time, so that we can assume that $k = 1 = l$.

Now let us recall the explicit formulae (13) and (24) that we obtained for the new low rank matrix $\widetilde{\mathrm{Rk}}\ r$. These formulae involved information about the first row and column of the matrices A and B. (For the Stein type, this dependence appeared also in an indirect way, via the embedding relation (23)). The ideal situation would be the following: we use the given displacement equation in order to determine these first rows and columns, next update the generators of the $\widetilde{\mathrm{Rk}}\ r$ matrix, and then repeat this procedure in a recursive way on the

Schur complements $S_{A,k}$ and $S_{B,l}$ (which we do not actually compute, but only store in a 'coded' form by means of the subsequent $\widetilde{\mathrm{Rk}}\ r$ matrices). Repeating this procedure, at the end we would obtain information about the LDU decompositions of both the matrices A and B.

Unfortunately, the above scheme can impossibly work since the given displacement equation does not contain enough information to determine the first block rows and columns of *both* the matrices A and B.

The situation may be different if an additional connection is given between A and B. For example, it could be that (i) a factorization $B = LDU$ is given, and we want to compute the LDU decomposition of A; (ii) we have a relation in the style $B = A$ (leading to the generalized Schur algorithm as described in [10]), or $B = A^T$. Thus only in such cases, we can hope the above scheme to work.

3 Rank structures

In this section we handle the preservation of rank structures. The following result could already have been mentioned in the previous section. It is a special case of both the preservation of Sylvester and of Stein type displacement structure.

Corollary 9 *Let $k \in \mathbb{N}$, and let A and B be matrices for which the Schur complements $S_{A,k}$ and $S_{B,k}$ exist. If*

$$A - B = \mathrm{Rk}\ r,$$

then

$$S_{A,k} - S_{B,k} = \widetilde{\mathrm{Rk}}\ r.$$

Note that there appeared only one index k in the statement of the above corollary, rather than two indices k and l as in the previous section. This is because we have here F^T and G equal to the identity matrix, which by Theorem 4 has to be block lower triangular w.r.t. the indices k and l; hence $k = l$ is the only relevant choice.

Now let us take for B an arbitrary Hermitian matrix. By the general property $S_{B^H,k} = (S_{B,k})^H$, also the Schur complement $S_{B,k}$ must be Hermitian. Hence Corollary 9 reveals the following fact.

Corollary 10 *The property $A = \mathrm{Herm} + \mathrm{Rk}\ r$, i.e. A is Hermitian plus rank at most r, is inherited under Schur complementation.*

The rest of this section is devoted to the preservation of what we call rank structures. First we recall some definitions from [1]. We will use here the subscript *weak* to distinguish these definitions from the actual definition of rank structures, which is given later.

Definition 11 *(See [1]:) We define a* weak structure block $\mathcal{B}_{\text{weak}}$ *on* $\mathbb{C}^{n\times n}$ *as a 4-tuple*

$$\mathcal{B}_{\text{weak}} = (i, j, r, \Lambda),$$

where i is the row index, j the column index, r the rank upper bound and $\Lambda \in \mathbb{C}^{(j-i+1)\times(j-i+1)}$ is called the shift matrix of $\mathcal{B}_{\text{weak}}$ (it is assumed here that $j - i + 1 \geq 0$). We say a matrix $A \in \mathbb{C}^{n\times n}$ to satisfy the weak structure block if, making a partitioning

$$A =: \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix}, \tag{32}$$

where $A_{2,2}$ is square and containing rows and columns $i, \ldots, j$, we have

$$\begin{bmatrix} A_{2,1} & A_{2,2} - \Lambda \\ A_{3,1} & A_{3,2} \end{bmatrix} = \text{Rk } r, \tag{33}$$

i.e. a matrix of rank at most r: see Figure 1.

As an extension, we can allow shift matrices $\Lambda = \Lambda_{\text{fin}} \oplus \infty I$, with Λ_{fin} having only finite entries. In this case we identify $\mathcal{B}_{\text{weak}}$ with the 'weak structure block' obtained by dropping all rows and columns involving ∞, and with the rank upper bound r decreased by the number of these dropped rows: see Figure 2. A weak structure block with shift matrix of the form $\Lambda = 0 \oplus \infty I$, is called pure, *denoted* $\mathcal{B}_{\text{weak,pure}}$.

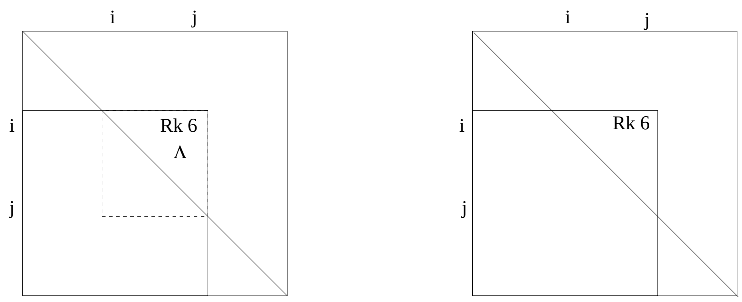

Figure 1: The structure block $\mathcal{B}_{\text{weak}}$ in the left picture has the following meaning: after subtracting the shift matrix $\Lambda \in \mathbb{C}^{4\times 4}$ from the dashed square submatrix in the middle, the indicated bottom left submatrix must be of rank at most 6. The structure block $\mathcal{B}_{\text{weak,pure}}$ in the right picture is a special case of this, with $\Lambda = 0$.

Theorem 12 *(see [1, Corollary 16]:) Let $A \in \mathbb{C}^{n\times n}$ be a nonsingular matrix satisfying the structure block $\mathcal{B}_{\text{weak}} = (i, j, r, \Lambda)$, where $\Lambda = \Lambda_{\text{ns}} \oplus 0 \oplus \infty I$, with Λ_{ns} nonsingular. Then the inverse matrix A^{-1} will satisfy the structure block $\mathcal{B}_{\text{weak}}^{-1} := (i, j, r, \Lambda^{-1})$, with $\Lambda^{-1} := \Lambda_{\text{ns}}^{-1} \oplus \infty I \oplus 0$ (hence using the rules $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$).*

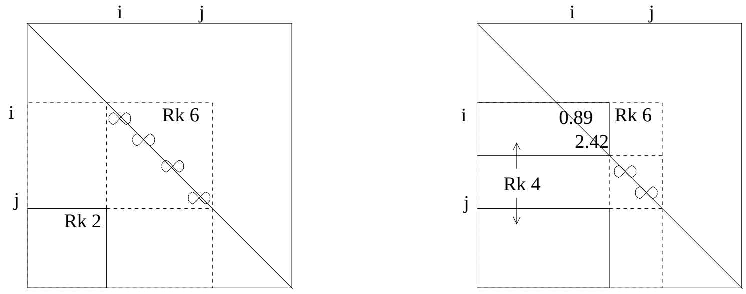

Figure 2: The structure block $\mathcal{B}_{\text{weak,pure}}$ in the left picture has shift matrix $\Lambda = \infty I_4$. Hence by definition, it should be identified with the Rk 2 structure block in the bottom left corner. The structure block $\mathcal{B}_{\text{weak}}$ in the right picture has $\Lambda = \text{diag}(0.89, 2.42, \infty, \infty)$. Hence it should be identified with the smaller Rk 4 structure block, consisting of two pieces. Note that the shift submatrix $\Lambda_{\text{fin}} := \text{diag}(0.89, 2.42)$ is inherited.

Let us recall also that by absorbing permutation matrices into the structure, structure blocks can be moved to any matrix position, not necessarily situated in the bottom left matrix corner anymore [1].

Now we come to the actual definition of structure blocks in the context of Schur complements. Such structure blocks will be denoted just as $\mathcal{B}$, hence dropping the subscript *weak*.

Definition 13 *Given a matrix $A \in \mathbb{C}^{m\times n}$, $k \in \mathbb{N}$ and consider the k-partitioning of A (Definition 1). We define a structure block w.r.t. this k-partitioning as a collection*

$$\mathcal{B} = (I, J, I_\Lambda, J_\Lambda, r, \Lambda),$$

where $I_\Lambda \subseteq I$ and $J_\Lambda \subseteq J$ are the index sets, r is the rank upper bound and $\Lambda \in \mathbb{C}^{|I_\Lambda|\times|J_\Lambda|}$ is called the shift matrix of the structure block. We define a partition $I = I_1 \cup I_2$, with $I_1 := I \cap \{1, \ldots, k\}$ and $I_2 := I \cap \{k+1, \ldots, n\}$, and we define similar partitions for the other index sets. We also partition

$$\Lambda = \begin{bmatrix} \Lambda_{1,1} & \Lambda_{1,2} \\ \Lambda_{2,1} & \Lambda_{2,2} \end{bmatrix},$$

with $\Lambda_{1,1}$ having dimension $|I_{\Lambda,1}|$ by $|J_{\Lambda,1}|$. Here we assume that

Condition: $\Lambda_{1,1}$ is square of size $|J_1| - |I_1| + 1$.

We say the matrix A to satisfy the structure block $\mathcal{B}$ if

$$\tilde{A}(I, J) = \text{Rk } r,$$

where $\tilde{A}$ has been defined from A by $\tilde{A}(I_\Lambda, J_\Lambda) = A(I_\Lambda, J_\Lambda) - \Lambda$: see Figure 3.

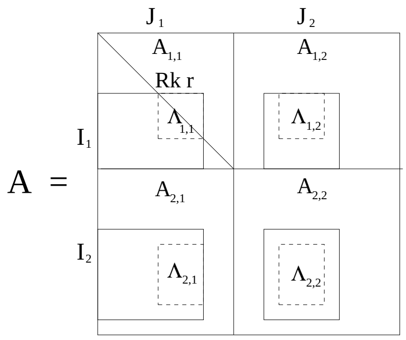

Figure 3: Example of a structure block $\mathcal{B}$. The meaning is that after subtracting the shift matrix $\Lambda \in \mathbb{C}^{7\times 6}$ from the dashed matrix positions, the indicated submatrix $A(I,J)$ (consisting of four parts) must be of rank at most r. Note that in general, there are no restrictions on the size of $\Lambda_{2,2}$ while on the other hand $\Lambda_{1,1}$ must be square and such that, up to permutation, the restriction $\mathcal{B}|_{A_{1,1}}$ is a *weak* structure block $\mathcal{B}_{\text{weak}}$.

Now we come to the preservation of structure blocks under Schur complementation.

Theorem 14 *Suppose given a matrix $A \in \mathbb{C}^{m\times n}$, a k-partitioning of A and a structure block $\mathcal{B}$ w.r.t. this k-partitioning. Using the notations of Definition 13, let us suppose that $\Lambda_{1,1}$ is square and nonsingular. Then the Schur complement $S_{A,k}$ satisfies the structure block $S_{\mathcal{B}} := (I_2, J_2, I_{\Lambda,2}, J_{\Lambda,2}, r, S_\Lambda)$ with $S_\Lambda := \Lambda_{2,2} - \Lambda_{2,1}\Lambda_{1,1}^{-1}\Lambda_{1,2}$: see Figure 4.*

PROOF. By definition of structure block, there exists a matrix B having the form

$$B = P \left[\begin{array}{ccc|ccc} X & X & X & X & X & X \\ 0 & \Lambda_{1,1} & X & 0 & \Lambda_{1,2} & X \\ 0 & 0 & X & 0 & 0 & X \\ \hline X & X & X & X & X & X \\ 0 & \Lambda_{2,1} & X & 0 & \Lambda_{2,2} & X \\ 0 & 0 & X & 0 & 0 & X \end{array}\right] \tilde{P},$$

for certain permutation matrices $P = P_1 \oplus P_2$ and $\tilde{P} = \tilde{P}_1 \oplus \tilde{P}_2$, such that $A - B = \text{Rk}\ r$. By Corollary 9, it follows that $S_{A,k} - S_{B,k} = \text{Rk}\ r$. But by the form of B, it is easy to see that its Schur complement satisfies

$$S_{B,k} = P_2 \begin{bmatrix} X & X & X \\ 0 & S_\Lambda & X \\ 0 & 0 & X \end{bmatrix} \tilde{P}_2,$$

where $S_\Lambda = \Lambda_{2,2} - \Lambda_{2,1}\Lambda_{1,1}^{-1}\Lambda_{1,2}$. It follows that $S_{A,k}$ satisfies the structure block $S_{\mathcal{B}} = (I_2, J_2, I_{\Lambda,2}, J_{\Lambda,2}, r, S_\Lambda)$, as we had to prove. □

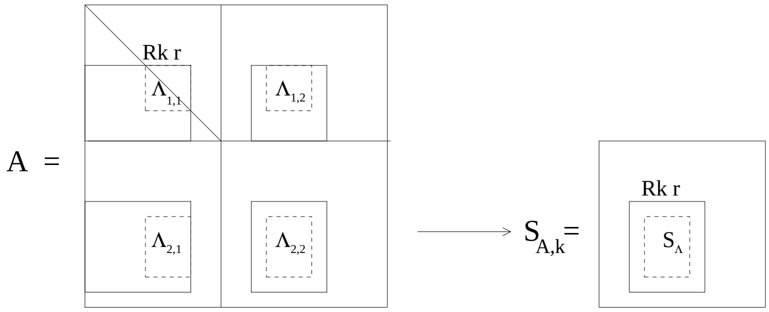

Figure 4: Given the matrix in the left hand side, satisfying the huge structure block $\mathcal{B}$, consisting of four parts. Then the Schur complement $S_{A,k} = A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}$ inherits this structure block, with new shift matrix given by $S_\Lambda := \Lambda_{2,2} - \Lambda_{2,1}\Lambda_{1,1}^{-1}\Lambda_{1,2}$.

As an illustrative example, suppose that

$$A_{i,j} = \begin{bmatrix} \times & \times & \times \\ 1 & 1+\lambda_{i,j} & \times \\ 1 & 1 & \times \end{bmatrix},$$

for $i, j = 1, 2$. Then we claim that $A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}$ (if $A_{1,1}^{-1}$ exists) will satisfy the structure block $\mathcal{B}_{\text{weak}} : (i, j, r, \lambda) = (2, 2, 1, S_\lambda)$, with new shift element defined by $S_\lambda := \lambda_{2,2} - \lambda_{2,1}\lambda_{1,1}^{-1}\lambda_{1,2}$ (if $\lambda_{1,1}^{-1}$ exists). Indeed: the proof follows immediately from Theorem 14 by working with the embedded matrix

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix},$$

and observing that the given data can be translated in terms of a huge structure block $\mathcal{B}$ on A.

Note that in this last example, it was necessary that the low rank blocks $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ of the several matrices $A_{i,j}$ were 'compatible' with each other. If this were not the case, it could be that the rank upper bound of the huge structure block $\mathcal{B}$ (and hence of the Schur complement $S_{A,k}$) must be increased.

A way to avoid the latter problem is to choose several of the low rank blocks equal to zero. Suppose for example that $A_{1,1} := T$ is a given matrix, satisfying a given structure block $\mathcal{B}_{\text{weak}}$. Suppose that we choose $A_{1,2}$, $A_{2,2}$ and $A_{2,1}$ with sparse bottom left parts as illustrated in Figure 5. Then it is clear that the structure block $\mathcal{B}_{\text{weak}}$ can *always* be extended to a huge structure block $\mathcal{B}$ in the matrix A, with new shift matrix

$$\Lambda = \begin{bmatrix} \Lambda_{1,1} & -I \\ I & 0 \end{bmatrix}.$$

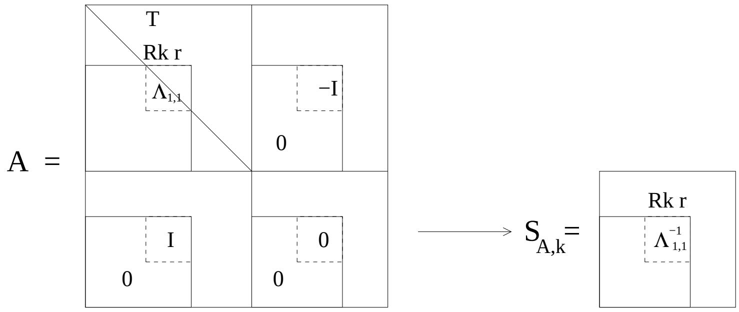

Figure 5: Given the matrix $A_{1,1} := T$ in the top left position, satisfying a structure block $\mathcal{B}_{\text{weak}}$. Then the data in the figure imply a huge structure block $\mathcal{B}$ in the matrix A. Hence the Schur complement $S_{A,k}$ will inherit the structure block $S_{\mathcal{B}}$, with new shift matrix $S_\Lambda \equiv \Lambda_{1,1}^{-1}$.

See figure 5.

We can derive two things from this example. First, consider the special case where $A_{1,2} = -I$, $A_{2,2} = 0$ and $A_{2,1} = I$. Then Figure 5 shows that $S_A \equiv T^{-1}$ satisfies the structure block $S_{\mathcal{B}} \equiv \mathcal{B}_{\text{weak}}^{-1}$, which is precisely the structure block inversion result of [1, Theorem 11]. Second, we can interpret Figure 5 in the following way: it can be used to generate matrices $A_{1,2}$, $A_{2,2}$ and $A_{2,1}$ such that $A_{2,2} - A_{2,1}T^{-1}A_{1,2}$ inherits the structure of T^{-1}. We can interpret this as a set of *structure preserving transformations* for T^{-1}.

To conclude this section, we want to relax the nonsingularity condition in Theorem 14. At the same time we want to introduce shift elements equal to ∞, in the sense of Definition 11. Here we will restrict ourselves to the case where

Assumption: $\Lambda_{1,1} = \Lambda_{\text{ns}} \oplus \infty I \oplus 0_l$,

where Λ_{ns} is square and nonsingular, and with 0_l being the zero matrix of size l by l. The other parts $\Lambda_{1,2}$, $\Lambda_{2,1}$ and $\Lambda_{2,2}$ are not allowed to contain elements equal to ∞.

Now let us write

$$\Lambda = \left[\begin{array}{c|c} \Lambda_{1,1} & \Lambda_{1,2} \\ \hline \Lambda_{2,1} & \Lambda_{2,2} \end{array}\right] = \left[\begin{array}{cc|c} \Lambda_{1,1}^{TL} & 0 & \Lambda_{1,2}^{T} \\ 0 & 0_l & \Lambda_{1,2}^{B} \\ \hline \Lambda_{2,1}^{L} & \Lambda_{2,1}^{R} & \Lambda_{2,2} \end{array}\right], \tag{34}$$

where $\Lambda_{1,1}^{TL} := \Lambda_{\text{ns}} \oplus \infty I$, and where the superscripts T, B, L and R denote the top, bottom, left and right parts of the corresponding matrices. It is easy to see that the Schur complement of (34) can be written as a 'dyadic decomposition'

$$S_\Lambda = S_{\text{fin}} + S_\infty, \tag{35}$$

where S_{fin} and S_∞ are the Schur complements of the respective matrices

$$\left[\begin{array}{cc} \Lambda_{1,1}^{TL} & \Lambda_{1,2}^{T} \\ \Lambda_{2,1}^{L} & \Lambda_{2,2} \end{array}\right], \quad \left[\begin{array}{cc} 0_l & \Lambda_{1,2}^{B} \\ \Lambda_{2,1}^{R} & 0 \end{array}\right]. \tag{36}$$

Here S_{fin} contains only finite elements, and hence this will just be a *finite* correction term to the structure of $S_{A,k}$. The problem is instead to determine the meaning of S_∞.

To achieve this, we will suppose that operations have been applied on the second block row and column of A, such that

$$\left[\begin{array}{c|c} 0_l & \Lambda_{1,2}^{B} \\ \hline \Lambda_{2,1}^{R} & 0 \end{array}\right] = \left[\begin{array}{c|cc} 0_l & 0 & \Lambda_{\text{ind col}} \\ \hline 0 & 0 & 0 \\ \Lambda_{\text{ind row}} & 0 & 0 \end{array}\right], \tag{37}$$

where $\Lambda_{\text{ind col}}$ contains independent columns and $\Lambda_{\text{ind row}}$ contains independent rows. (Here the row and column operations which we applied on A to achieve (37), have a well-determined effect on the Schur complement $S_{A,k}$ by virtue of Lemma 2).

Then we have the following result.

Theorem 15 *Suppose given a matrix $A \in \mathbb{C}^{m\times n}$, a k-partitioning of A and a structure block $\mathcal{B}$ w.r.t. this k-partitioning. Suppose that $\Lambda_{1,1} = \Lambda_{\text{ns}} \oplus \infty I \oplus 0_l$, that S_{fin} and S_∞ are defined as in (35), and that (37) holds. Then if we update*

$$S_{A,k}(I_{\Lambda,2}, J_{\Lambda,2}) := S_{A,k}(I_{\Lambda,2}, J_{\Lambda,2}) - S_{\text{fin}},$$

and if we drop in this updated matrix the rows of $I_{\Lambda,2}$ and columns of $J_{\Lambda,2}$ which are nonzero in (37), the resulting part of $S_{A,k}$ will have rank at most $r - l$: see Figure 6.

PROOF. If necessary, we can virtually add extra rows and columns to A until the blocks $\Lambda_{\text{ind row}}$ and $\Lambda_{\text{ind col}}$ in (37) become square and nonsingular (of size l by l). Then since S_∞ was defined as the Schur complement of (37), and by approximating 0_l as $0_l = \lim_{\epsilon\to 0} \epsilon I$, we obtain

$$S_\infty = \left[\begin{array}{cc} 0 & 0 \\ 0 & \Lambda_{\text{ind row}}(\infty I_l)\Lambda_{\text{ind col}} \end{array}\right].$$

But by our knowledge of the meaning of shift elements ∞, this means that we should drop in $S_{A,k}$ all l rows and columns where ∞ is standing, and decrease the rank upper bound r by this same number l. The theorem now follows. $\square$

As an illustrative example, suppose that

$$A_{i,j} = \left[\begin{array}{ccc} \times & \times & \times \\ 1 & 1 & \times \\ 1 & 1 & \times \end{array}\right],$$

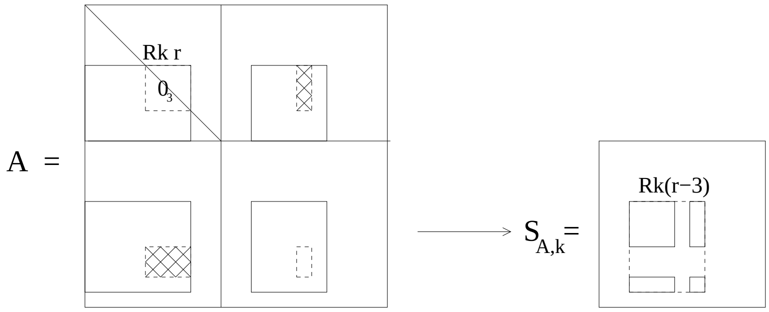

Figure 6: Given the matrix in the left hand side, satisfying the huge structure block $\mathcal{B}$ with $\Lambda_{1,1} = 0_3$. For the other parts of the shift matrix, the places where the nonzero elements act are indicated with a cross. (We assume here for simplicity of illustration that the finite correction term S_{fin} in (35) is equal to zero). Then the Schur complement $S_{A,k}$ satisfies a $\text{Rk}(r-3)$ structure block consisting of four parts, as indicated.

for $i, j = 1, 2$. Then we claim that $A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}$ (if $A_{1,1}^{-1}$ exists) will satisfy the structure block $\mathcal{B}_{\text{weak}} : (i, j, r) = (2, 2, 0)$, i.e. that

$$A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2} = \begin{bmatrix} \times & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \times \end{bmatrix}.$$

Indeed, this follows from Theorem 15 by working with the embedded matrix A (as usual), and by observing that the given data can be translated in terms of a huge structure block $\mathcal{B}$ with shift matrix

$$\begin{bmatrix} \lambda_{1,1} & \lambda_{1,2} \\ \lambda_{2,1} & \lambda_{2,2} \end{bmatrix} \equiv \begin{bmatrix} 0_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus indeed the rank upper bound r decreases by the value $l = 1$.

Note that in this last example, it was again necessary that the low rank blocks $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ of the several matrices $A_{i,j}$ were compatible with each other.

A way to avoid the latter problem is to choose several of the low rank blocks equal to zero. Suppose for example that $A_{1,1} := T$ is a given matrix, satisfying a given structure block $\mathcal{B}_{\text{weak,pure}}$ with $\Lambda = 0_3$. Suppose that we choose $A_{1,2}$, $A_{2,2}$ and $A_{2,1}$ with zero bottom left parts as illustrated in Figure 7. Then it is clear that the structure block $\mathcal{B}_{\text{weak,pure}}$ can *always* be extended to a huge structure block $\mathcal{B}_{\text{pure}}$ in the matrix A, with new shift matrix

$$\begin{bmatrix} 0_3 & 0 \\ 0 & 0 \end{bmatrix}.$$

See Figure 7.

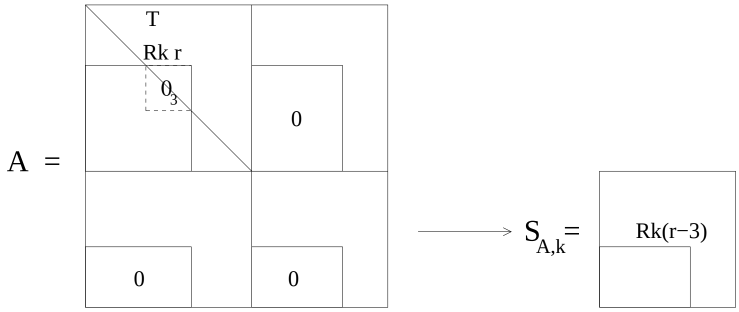

Figure 7: Specification of Figure 5 in the case of zero shift matrices.

We can derive two things from this example. First, consider the special case where $A_{1,2} = -I$, $A_{2,2} = 0$ and $A_{2,1} = I$. Then Figure 7 shows that $S_A \equiv T^{-1}$ satisfies the structure block $S_{\mathcal{B}} \equiv \mathcal{B}^{-1}_{\text{weak,pure}}$, which is precisely the structure block inversion result of [1, Corollary 16] (concerning shift matrices $\Lambda_{\text{ns}} \oplus 0 \oplus \infty I$). Second, we can interpret Figure 7 in the following way: it can be used to generate matrices $A_{1,2}$, $A_{2,2}$ and $A_{2,1}$ such that $A_{2,2} - A_{2,1}T^{-1}A_{1,2}$ inherits the structure of T^{-1}. We can interpret this as a set of *structure preserving transformations* for T^{-1}. Other examples of structure preserving transformations will be given in the next section, in the context of Möbius transformations.

We conclude this section with a final remark.

Remark 16

1. *The structure block $\mathcal{B}$ can sometimes be brought in easier form by virtue of Lemma 2. For example, it follows from this lemma that the left multiplication of A with a matrix $\begin{bmatrix} X & 0 \\ X & I \end{bmatrix}$, or the right multiplication with the transpose of such a matrix always preserves the Schur complement. Such transformations can be used for example to transform the dependent rows and columns of $\Lambda_{1,1}$ into zeros. In some cases, it may be even possible to restore in this way the size restrictions on $\Lambda_{1,1}$ occurring in Definition 13, even if these were not satisfied initially.*

2. *In many examples where structure blocks occur (for example lower semiseparable or lower semiseparable plus diagonal related matrices, see [1]), we have $I_1 = \emptyset$, $J_1 = \{1, \ldots, k\}$ and $\Lambda_{2,1} = 0$, as in the left picture of Figure 8. Then the size restrictions on $\Lambda_{1,1}$ occurring in Definition 13 are not satisfied. But they can be restored by just 'enlarging' the structure block, as illustrated in Figure 8.*

4 Möbius and Cayley transformations

In this section we will focus on Möbius transformations, as an illustration of the results on Schur complements in the previous section. Möbius transformations

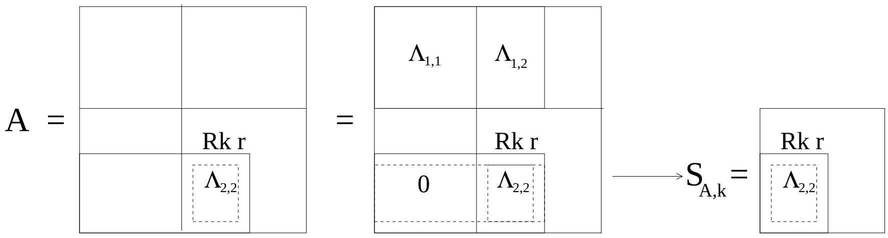

Figure 8: Given the structure block in the leftmost picture. Then we can 'enlarge' this structure block by redefining $I_1 := \{1, \ldots, k\}$, $\Lambda_{1,1} := A_{1,1}$ and $\Lambda_{1,2} := A_{1,2}|_{I_1 \times J_2}$, as illustrated in the middle picture. In this way the resulting structure block will still be of rank at most r, and the size restrictions on $\Lambda_{1,1}$ occurring in Definition 13 are restored. But then it follows that $S_{A,k}$ satisfies the new structure block $S_{\mathcal{B}} = (I_2, J_2, I_{\Lambda_2}, J_{\Lambda_2}, r, \Lambda_{2,2})$.

appear also under the name of *rational linear transformations*. As a general reference, we can refer to [7] for the treatment of Möbius transformations with scalar coefficients, and to [11, 12] for the general case of matrix-valued coefficients. Most of the results which we state without proof can be found there.

We start with a definition.

Definition 17 *Given fixed coefficient matrices $P, Q, R, S \in \mathbb{C}^{n\times n}$, then we define the* Möbius transformation *on $\mathbb{C}^{n\times n}$ to be the map*

$$\mathcal{M} : A \mapsto (PA + Q)(RA + S)^{-1}.$$

Similarly, we define the dual Möbius transformation *to be the map*

$$A \mapsto (AP + R)^{-1}(AQ + S).$$

Finally, we define

$$\begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

to be the matrix associated with $\mathcal{M}$, and we say $\mathcal{M}$ to be invertible *if its associated matrix is nonsingular.*

Unless explicitly mentioned, we will always work with usual Möbius transformations, rather than with their dual versions.

Note that the Möbius transformation is only defined on the domain $\mathcal{D} := \{A \in \mathbb{C}^{n\times n} |\ \det(RA + S) \neq 0\}$. Since the domain $\mathcal{D}$ is defined by the non-vanishing of an algebraic equation, it is either empty (a case we exclude) or a dense subset of $\mathbb{C}^{n\times n}$.

The use of the matrix associated with $\mathcal{M}$ follows by rewriting $\mathcal{M}(A) = N_{\mathcal{M}(A)} D^{-1}_{\mathcal{M}(A)}$ with

$$\begin{bmatrix} N_{\mathcal{M}(A)} \\ D_{\mathcal{M}(A)} \end{bmatrix} := \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \begin{bmatrix} A \\ I \end{bmatrix}. \tag{38}$$

This matrix representation is useful in several aspects. For example, it can be checked that for given Möbius transformations $\mathcal{M}_1$ and $\mathcal{M}_2$, the composed map $A \mapsto \mathcal{M}_2(\mathcal{M}_1(A))$ is again a Möbius transformation, with associated matrix

$$\begin{bmatrix} P_2 & Q_2 \\ R_2 & S_2 \end{bmatrix} \begin{bmatrix} P_1 & Q_1 \\ R_1 & S_1 \end{bmatrix}.$$

Since the identity map $A \mapsto A$ is a special case of a Möbius transformation, with associated matrix $\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$, it follows that the *inverse Möbius transformation* $\mathcal{M}^{-1}$ will have as its associated matrix precisely $\begin{bmatrix} P & Q \\ R & S \end{bmatrix}^{-1}$. Moreover, denoting with $\mathcal{D}$ the domain of $\mathcal{M}$, then it can be shown that $\mathcal{M}^{-1}$ has as its domain $\mathcal{M}(\mathcal{D})$ and as its range $\mathcal{D}$.

We may note here that $\mathcal{M}^{-1}$ can also be obtained by directly solving for A in terms of $\mathcal{M}(A)$ in Definition 17. This yields $\mathcal{M}^{-1} : B \mapsto -(BR-P)^{-1}(BS-Q)$, which is not a Möbius transformation anymore, but rather a dual Möbius transformation in the sense of Definition 17. In particular, it follows that every invertible Möbius transformation can be expressed as a dual Möbius transformation too.

Now note that $\mathcal{M}(A)$ can be realized as the Schur complement of

$$\begin{bmatrix} RA+S & -I \\ PA+Q & 0 \end{bmatrix}. \tag{39}$$

In particular, we can prove the following result.

Theorem 18 *Let $\mathcal{M}$ be a Möbius transformation with domain $\mathcal{D}$. Suppose $A \in \mathcal{D}$, $A + \text{Rk } r \in \mathcal{D}$ where $\text{Rk } r$ is a matrix of rank at most r. Then*

$$\mathcal{M}(A + \text{Rk } r) = \mathcal{M}(A) + \widetilde{\text{Rk}}\ r,$$

with $\widetilde{\text{Rk}}\ r$ a new matrix of rank at most r.

PROOF. Let us write $\text{Rk } r = UV^H$ with $U, V \in \mathbb{C}^{n\times r}$. Note that $\mathcal{M}(A + UV^H)$ can be realized as the Schur complement of (39), to which is added now a correction term $\begin{bmatrix} RU \\ PU \end{bmatrix} \begin{bmatrix} V^H & 0 \end{bmatrix}$. Since this correction term is still of rank at most r, the result follows by Corollary 9. □

We come to a second topic.

Definition 19 *Given fixed coefficient matrices $E, F, G \in \mathbb{C}^{n\times n}$, with E and G Hermitian, then we define the* quadratic transformation *on $\mathbb{C}^{n\times n}$ to be the map*

$$\mathcal{Q} : A \mapsto A^H EA + A^H F^H + FA + G.$$

We define

$$\begin{bmatrix} E & F^H \\ F & G \end{bmatrix} \tag{40}$$

to be the matrix associated with the quadratic transformation.

The use of the matrix associated with the quadratic transformation, follows by rewriting

$$\mathcal{Q}(A) = \begin{bmatrix} A^H & I \end{bmatrix} \begin{bmatrix} E & F^H \\ F & G \end{bmatrix} \begin{bmatrix} A \\ I \end{bmatrix}. \tag{41}$$

Moreover, note that by our assumption that E and G are Hermitian, both the middle matrix and the right hand side of (41) must be Hermitian. In particular, it makes sense to speak about the *inertia* of $\mathcal{Q}(A)$.

To establish some inertia results, we will first prove the following lemma. Here we agree to label the eigenvalues of a matrix A in non-increasing order as $\lambda_{A,k}$, $k = 1, \ldots, n$.

Lemma 20 *Let* Herm *be Hermitian and let T be arbitrary. Define $H := T(\text{Herm})T^H$ and $\tilde{H} = (T + \text{Rk } r)\text{Herm}(T + \text{Rk } r)^H$, where* Rk r *is a matrix of rank at most r. Then we have the interlacing property*

$$\lambda_{H,k+r} \leq \lambda_{\tilde{H},k} \leq \lambda_{H,k-r}.$$

PROOF. We will prove the interlacing property under the slightly weaker condition that H and $\tilde{H}$ are both Hermitian and

$$\tilde{H} = H + UV^H + \tilde{V}U^H, \tag{42}$$

with U, V, $\tilde{V}$ in $\mathbb{C}^{n \times r}$. We recall that for any Hermitian matrix Herm, the eigenvalues can be determined by the Courant-Fisher characterization

$$\lambda_{\text{Herm},k} = \max_{\dim \mathcal{V}=k} \min_{x \in \mathcal{V}} \frac{x^H \text{Herm } x}{x^H x}, \tag{43}$$

with $\mathcal{V}$ running over all k-dimensional linear subspaces of $\mathbb{C}^n$. Taking such a fixed subspace $\mathcal{V}$ and taking Herm $= \tilde{H}$, we obtain by (42) that $\min_{x \in \mathcal{V}} \frac{x^H \tilde{H} x}{x^H x} \leq \min_{x \in \mathcal{V} \cap \mathcal{U}} \frac{x^H H x}{x^H x}$, where $\mathcal{U}$ denotes the $(n-r)$-dimensional linear subspace of $\mathbb{C}^n$ containing all vectors for which $U^H x = 0$. Then since $\dim(\mathcal{V} \cap \mathcal{U}) \geq k - r$, we derive by (43) that

$$\min_{x \in \mathcal{V}} \frac{x^H \tilde{H} x}{x^H x} \leq \lambda_{H,k-r}.$$

By taking the maximum over all $\mathcal{V}$, it follows that $\lambda_{\tilde{H},k} \leq \lambda_{H,k-r}$, as we had to prove. The other inequality follows by symmetry. □

Remark 21 *An alternative proof of Lemma 20 is the following. It can be checked that the inertia (π, ν, ζ) of the Hermitian term $UV^H + \tilde{V}U^H$ in (42) must necessarily be 'equally distributed' in the sense that* $\max\{\pi, \nu\} \leq r$. *Lemma 20 follows then as a consequence of the interlacing properties of eigenvalues of a Hermitian, small rank correction to a Hermitian matrix as stated in [6]. Here we may remark that the proof in [6] is essentially based on a similar argument involving the Courant-Fisher characterization of eigenvalues.*

Corollary 22

1. *Inertia*$(\mathcal{Q}(A+\text{Rk } r))-$*Inertia*$(\mathcal{Q}(A)) = (\Delta\pi, \Delta\nu, \Delta\zeta)$ *with* $\max\{|\Delta\pi|, |\Delta\nu|\} \leq r$.

2. *If the matrix associated with the quadratic transformation (40) has inertia* (π, ν, ζ) *with* $\pi = n+\delta$, $\delta \geq 1$, *then* $\mathcal{Q}(A)$ *has at least* δ *positive eigenvalues, independent of the choice of* A.

PROOF.

1. Note that in (41), going over to $\mathcal{Q}(A + \text{Rk } r)$ corresponds to a rank r correction of the factor $\begin{bmatrix} A \\ I \end{bmatrix}$. Hence the result follows from the previous lemma.

2. We will use again (41). First, by adding n extra columns, the matrix $\begin{bmatrix} A \\ I \end{bmatrix}$ can always be completed to a nonsingular $2n \times 2n$ matrix. Replacing $\begin{bmatrix} A \\ I \end{bmatrix}$ by this completed version, and replacing $\begin{bmatrix} A^H & I \end{bmatrix}$ by the Hermitian transpose of it, Sylvester's law of inertia implies that the right hand side of (41) must still have inertia (π, ν, ζ) with $\pi = n+\delta$. The result follows then by again removing the n added columns (which is a rank n perturbation) and applying the previous lemma.

□

Since both Möbius and quadratic transformations allow what we called a matrix representation, we can expect that these transformations have a good behaviour w.r.t. each other. Let us first introduce

Definition 23 *Given a quadratic transformation* $\mathcal{Q}$ *on* $\mathbb{C}^{n\times n}$ *and given a positive integer* r, *then we say* $A \in \mathbb{C}^{n\times n}$ *to satisfy the* quadratic relation *induced by* $(\mathcal{Q}, r)$ *if*

$$\mathcal{Q}(A) = \text{Rk } r,$$

for a certain matrix Rk r *of rank at most* r.

Since $\mathcal{Q}(A)$ is always Hermitian, we can in fact add any inertia condition on Rk r if so desired.

Now let $\mathcal{M}$ be a Möbius transformation with domain $\mathcal{D}$, let $A \in \mathcal{D}$ and let

$B := N_B D_B^{-1} = \mathcal{M}(A)$. Then

$$\begin{aligned}
& \mathcal{Q}(B) = \text{Rk } r \\
\Leftrightarrow \quad & \begin{bmatrix} B^H & I \end{bmatrix} \begin{bmatrix} E & F^H \\ F & G \end{bmatrix} \begin{bmatrix} B \\ I \end{bmatrix} = \text{Rk } r \\
\Leftrightarrow \quad & \begin{bmatrix} N_B^H & D_B^H \end{bmatrix} \begin{bmatrix} E & F^H \\ F & G \end{bmatrix} \begin{bmatrix} N_B \\ D_B \end{bmatrix} = \widetilde{\text{Rk}}\, r \\
\overset{\text{by (38)}}{\Leftrightarrow} \quad & \begin{bmatrix} A^H & I \end{bmatrix} \begin{bmatrix} P & Q \\ R & S \end{bmatrix}^H \begin{bmatrix} E & F^H \\ F & G \end{bmatrix} \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \begin{bmatrix} A \\ I \end{bmatrix} = \widetilde{\text{Rk}}\, r \\
\Leftrightarrow \quad & \begin{bmatrix} A^H & I \end{bmatrix} \begin{bmatrix} \tilde{E} & \tilde{F}^H \\ \tilde{F} & \tilde{G} \end{bmatrix} \begin{bmatrix} A \\ I \end{bmatrix} = \widetilde{\text{Rk}}\, r \\
\Leftrightarrow \quad & \widetilde{\mathcal{Q}}(A) = \widetilde{\text{Rk}}\, r,
\end{aligned}$$

where we defined $\widetilde{\text{Rk}}\, r := D_B^H(\text{Rk } r)D_B$ (being a matrix with same rank and inertia as Rk r), and where we defined the quadratic transformation $\widetilde{\mathcal{Q}}$ by its associated matrix

$$\begin{bmatrix} \tilde{E} & \tilde{F}^H \\ \tilde{F} & \tilde{G} \end{bmatrix} := \begin{bmatrix} P & Q \\ R & S \end{bmatrix}^H \begin{bmatrix} E & F^H \\ F & G \end{bmatrix} \begin{bmatrix} P & Q \\ R & S \end{bmatrix}. \tag{44}$$

Let us give some illustrations to the above series of equivalences. First consider the class of Hermitian matrices, defined by $\mathcal{Q}(A) := -iA + iA^H =$ Rk 0. Note that the matrix associated with $\mathcal{Q}$ is given by $\begin{bmatrix} 0 & iI \\ -iI & 0 \end{bmatrix}$. Then consider the class of J-unitary matrices, defined by $\mathcal{Q}(A) := A^H JA - J =$ Rk 0, where $J = I_r \oplus -I_s$ is a fixed signature matrix. The matrix associated with $\mathcal{Q}$ is given by $\begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix}$. Now we may observe that these two matrices obtained for Hermitian and J-unitary quadratic relations have the same inertia, namely $(n, n, 0)$. Hence by Sylvester's law of inertia, there exists a nonsingular congruence transformation which maps these matrices into each other. Indeed: one can check that

$$\begin{bmatrix} 0 & iI \\ -iI & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -iJ & iJ \end{bmatrix} \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix} \begin{bmatrix} I & iJ \\ I & -iJ \end{bmatrix} \frac{1}{\sqrt{2}}.$$

Comparing this with (44), we conclude that

$$\mathcal{M} : A \mapsto (A + iJ)(A - iJ)^{-1} \tag{45}$$

is a Möbius transformation mapping the class of Hermitian into the class of J-unitary matrices (at least for every Hermitian matrix belonging to the domain $\mathcal{D}$ of $\mathcal{M}$). Moreover, one can check that

$$\mathcal{M}^{-1} : A \mapsto i(A + I)(JA - J)^{-1}.$$

Remark 24 *Composing (45) with an arbitrary Möbius transformation mapping the class of J-unitary matrices into itself, one can obtain many other Möbius transformations* $\tilde{\mathcal{M}}$ *mapping Hermitian into J-unitary matrices. In particular, due to the cancellation of numerators and denominators, the domain* $\mathcal{D}$ *will heavily depend on the resulting Möbius transformation* $\tilde{\mathcal{M}}$. *It can be proven that any Hermitian matrix belongs to such a domain.*

For the rest of this section, we will restrict ourselves to Möbius transformations with scalar coefficients, i.e. with P, Q, R, S being scalar multiples of the identity matrix. We will denote these multiples as pI, qI, rI and sI.

As in the general case, note that $\mathcal{M}(A)$ can be realized as the Schur complement of the embedded matrix

$$\begin{bmatrix} rA + sI & -I \\ pA + qI & 0 \end{bmatrix}. \tag{46}$$

Now assume that A satisfies a weak structure block $\mathcal{B}_{\text{weak}} = (i, j, r, \Lambda)$. We want to show that this weak structure block can be extended to a huge structure block in (46), with same rank upper bound r. For this, let $A_{\text{pure}} := (A - (0 \oplus \Lambda \oplus 0))|_{\mathcal{B}_{\text{weak}}}$. Note that $\operatorname{Rank} \begin{bmatrix} rA_{\text{pure}} \\ pA_{\text{pure}} \end{bmatrix} = \operatorname{Rank} A_{\text{pure}}$. This means that $\mathcal{B}_{\text{weak}}$ can indeed be extended to a huge structure block, denoted $\mathcal{B}_{\text{huge}}$, and with corresponding shift matrix

$$\Lambda_{\text{huge}} := \begin{bmatrix} p\Lambda + qI & -I \\ r\Lambda + sI & 0 \end{bmatrix}.$$

(Here we did not show all the zero blocks of Λ_{huge}; see Figure 9 for a more accurate picture).

Now by Theorem 14, the Schur complement of (46) must inherit $\mathcal{B}_{\text{huge}}$, with new shift matrix being the Schur complement of Λ_{huge}. Reformulating this in terms of the original data, we obtain

Theorem 25 *Given a matrix* $A \in \mathbb{C}^{n \times n}$ *which satisfies a weak structure block* $\mathcal{B}_{\text{weak}}$. *Then the scalar Möbius transformation* $\mathcal{M}(A)$ *will inherit the weak structure block* $\mathcal{B}_{\text{weak}}$, *with new shift matrix being* $\mathcal{M}(\Lambda)$, *i.e. precisely the scalar Möbius transformation of the original shift matrix* Λ.

We may mention that Theorem 25 was already shown in [13] for the case of lower semiseparable plus diagonal matrices. See also [4].

For this same example, note that for a *couple* of matrices A, B satisfying the same weak structure block $\mathcal{B}_{\text{weak}} = (i, j, r, \Lambda)$, the 'decoupled scalar Möbius transformation' $(pA + qI)(rB + sI)^{-1}$ will essentially inherit the weak structure block $\mathcal{B}_{\text{weak}}$, but now with rank upper bound only bounded by $2r$. The reason for this is the identity $\operatorname{Rank} \begin{bmatrix} rA_{\text{pure}} \\ pB_{\text{pure}} \end{bmatrix} \leq \operatorname{Rank} A_{\text{pure}} + \operatorname{Rank} B_{\text{pure}}$, which is much weaker than in the case $A_{\text{pure}} = B_{\text{pure}}$.

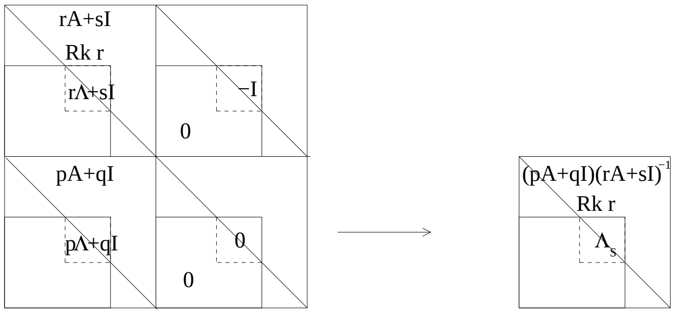

Figure 9: Given a matrix A satisfying a Rk r weak structure block with shift matrix Λ. Let $p, q, r, s \in \mathbb{C}$ be arbitrary numbers. Then the Möbius transformation $(pA + qI)(rA + sI)^{-1}$ can be obtained in the form of a Schur complement of an embedded matrix A, as illustrated. Hence it will inherit the Rk r weak structure block, with new shift matrix $(p\Lambda + qI)(r\Lambda + sI)^{-1}$.

An interesting case of a scalar Möbius transformation is by taking $J = I$ in (45). Then (45) reduces to the well-known *Cayley transformation* $\mathcal{C} : A \mapsto (A+iI)(A-iI)^{-1}$, mapping Hermitian into unitary matrices. Unlike the general situation, the domain $\mathcal{D}$ of the Cayley transformation contains the *entire* class of Hermitian matrices.

The Cayley transformation can be used to derive several properties of unitary matrices. We already remarked in [1] that the weak structure blocks of such a matrix always come in pairs, i.e. that the presence of one such weak structure block always implies the presence of a second weak structure block, which we can easily determine. The underlying reason was the fact that $\text{Uni}^{-1} = \text{Uni}^H$ for any unitary matrix Uni, together with the inversion theorem for weak structure blocks. Another way to see this is by using the Cayley transformation. This transformation can be used to establish the property that the structure blocks of a unitary matrix always come in pairs, from the corresponding property that the structure blocks of a *Hermitian* matrix always come in pairs (for obvious reasons). We will not go further into this.

The Cayley transformation can also be used as a tool to prove a similar result to the following theorem from [1].

Theorem 26 *Let $r \in \mathbb{N}$. The following are equivalent:*

(i) $A = \text{Herm} + \text{Rk } r$*, i.e.* A *is Hermitian plus rank at most* r*;*

(ii) $\mathcal{Q}(A) := i(A - A^H) = \text{Rk } 2r$*, where* $\text{Rk } 2r$ *is a matrix of rank at most* $2r$ *such that* $\text{Inertia}(\text{Rk } 2r) = (\pi, \nu, \zeta)$ *with* $\max\{\pi, \nu\} \leq r$*.*

We obtain the following, similar formulation.

Theorem 27 *Let $r \in \mathbb{N}$. The following are equivalent:*

(i) $A = \mathrm{Uni} + \mathrm{Rk}\ r$*, i.e. A is unitary plus rank at most r;*

(ii) $\mathcal{Q}(A) := A^H A - I = \mathrm{Rk}\ 2r$*, where* $\mathrm{Rk}\ 2r$ *is a matrix of rank at most* $2r$ *such that* $\mathrm{Inertia}(\mathrm{Rk}\ 2r) = (\pi, \nu, \zeta)$ *with* $\max\{\pi, \nu\} \leq r$.

PROOF. The implication (i) $\Rightarrow$ (ii) is a special case of Corollary 22.1. For the implication (ii) $\Rightarrow$ (i), suppose that A is such that $A^H A - I = \mathrm{Rk}\ 2r$ with $\mathrm{Inertia}(\mathrm{Rk}\ 2r) = (\pi, \nu, \zeta)$ with $\max\{\pi, \nu\} \leq r$. Denoting by $\mathcal{D} \subseteq \mathbb{C}^{n\times n}$ the domain of the Cayley transformation $\mathcal{C}$, then we will suppose that $A \in \mathcal{C}(\mathcal{D})$, the domain of the inverse Cayley transformation $\mathcal{C}^{-1}$. (This can always be realized by multiplying with a suitable number $e^{i\theta}$, $\theta \in \mathbb{R}$). Then we claim that $B := \mathcal{C}^{-1}(A)$ will satisfy $iB - iB^H = \widetilde{\mathrm{Rk}}\ 2r$, with $\widetilde{\mathrm{Rk}}\ 2r$ having the same inertia as $\mathrm{Rk}\ 2r$. Indeed, this follows from the series of equivalences preceding (44). Hence by Theorem 26, we can factorize $B = \mathrm{Herm} + \mathrm{Rk}\ r$. The result follows then by applying $\mathcal{C}$ on both sides of this equation. (Here it is essential that for the application of Theorem 18, $\mathcal{C}(\mathrm{Herm})$ is always defined, independent of the precise form of this Hermitian component Herm!) $\square$

Corollary 28 *Let $r \in \mathbb{N}$. Then the class of unitary plus rank at most r matrices is topologically closed.*

By composing Theorems 26, 27 with an affine transformation $A \mapsto pA + qI$, these theorems can in fact be generalized to normal matrices with eigenvalues lying on a fixed *generalized circle* in $\mathbb{C}$, i.e. either a straight line or a circle in the complex plane. We must then work with a quadratic transformion on $\mathbb{C}^{n\times n}$ of the form $\mathcal{Q}(A) := eA^H A + \bar{f}A^H + fA + gI = 0$, for suitable numbers $e, g \in \mathbb{R}$ and $f \in \mathbb{C}$.

We may note here that these theorems can *not* be generalized to arbitrary, non-scalar quadratic relations. For example, it can be shown that the class of J-unitary plus rank at most 1 matrices, is *not* topologically closed.

5 Conclusion

In this paper, we investigated some structures that have a good behaviour under Schur complementation. We handled two classes of them: displacement and rank structures. For displacement structures, we derived in a direct way the preservation of structure, leading to formulae which extend the classical displacement tools. For the case of rank structures, we showed how the preservation results could be used as a general framework to specify structure-preserving operations. In particular, we considered the Möbius transformion of a matrix and derived several structure preservation results.

References

[1] S. Delvaux and M. Van Barel. Structures preserved by matrix inversion. Report TW 414, Department of Computer Science, K.U.Leuven, Leuven, Belgium, December 2004.

[2] S. Delvaux and M. Van Barel. Structures preserved by the QR-algorithm. Report TW 399, Department of Computer Science, K.U.Leuven, Leuven, Belgium, August 2004.

[3] G. Heinig and K. Rost. *Algebraic Methods for Toeplitz-like Matrices and Operators*. Mathematical Research. Akademie Verlag-Berlin, 1984.

[4] L. Gemignani. A unitary Hessenberg QR-based algorithm via semiseparable matrices. url: fibonacci.dm.unipi.it/~gemignan/papers/unihess.pdf.

[5] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins university Press, third edition, 1996.

[6] D. A. Gregory, B. Heyink, and K. N. Vander Meulen. Inertia and biclique decompositions of joins of graphs. *J. Combin. Theory*, (Series B 88):135–151, 2003.

[7] P. Henrici. *Applied and Computational Complex Analysis*. Wiley, 1974.

[8] T. Kailath, S. Kung, and M. Morf. Displacement ranks of matrices and linear equations. *Journal of Mathematical Analysis and its Applications*, 68:395–407, 1979.

[9] T. Kailath and A. H. Sayed. Displacement structure: theory and applications. *SIAM Review*, 37:297–386, 1995.

[10] T. Kailath and A. H. Sayed, editors. *Fast reliable algorithms for matrices with structure*. SIAM, Philadelphia, PA, USA, May 1999. ISBN0-89871-431-1.

[11] V. P. Potapov. The multiplicative structure of J-contractive matrix functions. *American Mathematical Society Translations, Series 2*, 15:131–243, 1960.

[12] V. P. Potapov. Linear fractional transformations of matrices. *American Mathematical Society Translations, Series 2*, 138:21–35, 1988.

[13] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. Orthogonal rational functions and structured matrices. *SIAM Journal on Matrix Analysis and its Applications*, 2004. Accepted for publication.