Structures preserved by generalized inversion and Schur complementation

Steven Delvaux   *Marc Van Barel*

Report TW478, November 2006

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

Structures preserved by generalized inversion and Schur complementation

Steven Delvaux *Marc Van Barel*

Report TW478, November 2006

Department of Computer Science, K.U.Leuven

Abstract

In this paper we investigate the inheritance of certain structures under generalized matrix inversion. These structures contain the case of rank structures, and the case of displacement structures. We do this in an intertwined way, in the sense that we develop an argument that can be used for deriving the results for displacement structures from thoses for rank structures. We pay particular attention to the Moore-Penrose generalized inverse, showing that for the cases of most interest, the ranks of the structure satisfied by the Moore-Penrose inverse can at most double with respect to the original ranks. We consider also the case of inheritance of structure by generalized Schur complements.

Keywords : generalized inversion, generalized Schur complementation, rank structure, displacement structure.
AMS(MOS) Classification : Primary : 15A09, Secondary : 15A03, 65F20.

Structures preserved by generalized inversion and Schur complementation

Steven Delvaux *, Marc Van Barel *

15th November 2006

Abstract

In this paper we investigate the inheritance of certain structures under generalized matrix inversion. These structures contain the case of rank structures, and the case of displacement structures. We do this in an intertwined way, in the sense that we develop an argument that can be used for deriving the results for displacement structures from thoses for rank structures. We pay particular attention to the Moore-Penrose generalized inverse, showing that for the cases of most interest, the ranks of the structure satisfied by the Moore-Penrose inverse can at most double with respect to the original ranks. We consider also the case of inheritance of structure by generalized Schur complements.

Keywords: generalized inversion, generalized Schur complementation, rank structure, displacement structure.

AMS subject classifications: 15A09, 15A03, 65F20.

1 Introduction

Definition 1 *Let $A \in \mathbb{C}^{m\times n}$ be a given matrix, possibly rectangular. The set of* Moore-Penrose equations *is defined as*

1. $AXA = A$
2. $XAX = X$

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

The research was partially supported by the Research Council K.U.Leuven, project OT/05/40 (Large rank structured matrix computations), Center of Excellence: Optimization in Engineering, by the Fund for Scientific Research–Flanders (Belgium), G.0455.0 (RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation), G.0423.05 (RAM: Rational modelling: optimal conditioning and stable algorithms), 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.

3. $(AX)^H = AX$

4. $(XA)^H = XA$.

Here we used the superscript ^{H} to denote the Hermitian transpose of a matrix, i.e., the complex conjugate transpose.

It was shown by Penrose [10] that the above equations always have a unique solution matrix $X \in \mathbb{C}^{n\times m}$, which is called the *Moore-Penrose inverse* of A and denoted by $A^\dagger$. In case where A is square nonsingular, $A^\dagger$ is nothing but the usual matrix inverse A^{-1}.

More generally [1], one may be interested in matrices that are a solution of only a limited number of Moore-Penrose equations.

Definition 2 *Let $A \in \mathbb{C}^{m\times n}$ be a given matrix, possibly rectangular, then X is called an S-inverse of A if it satisfies the set of Moore-Penrose equations indexed by $S \subseteq \{1, 2, 3, 4\}$.*

More specifically, X is called

- a *generalized inverse* of A if it is a $\{1\}$-inverse,
- a *reflexive generalized inverse* of A if it is a $\{1, 2\}$-inverse,
- the Moore-Penrose inverse of A if it is the (unique) $\{1, 2, 3, 4\}$-inverse.

In what follows, we will always work with a generalized inverse, i.e., a $\{1\}$-inverse of A. Note that such a matrix X must necessarily belong to $\mathbb{C}^{n\times m}$: this reflects the fact that the Moore-Penrose equations must have compatible matrix dimensions.

This paper deals with structures preserved by generalized inversion and generalized Schur complementation, hereby extending our earlier papers [3, 4]. Moreover, we will make use of the occasion to establish some additional facts and interconnections for *usual* inversion and Schur complementation as well, which have not been considered yet in [3, 4].

Section 2 deals with structures that have a good behavior under *generalized inversion*. We start with the case where A is a rank structured matrix. This means that we assume A to satisfy a collection of so-called *structure blocks*: these are low rank submatrices of A, together with a certain notion of correction term which we call *shift matrix*. We are interested in the inheritance of rank structure by the generalized inverses of A. We will see that the preservation results for nonsingular matrices [5, 3] can essentially be taken over, at the price of a possible increase of the rank. We will show how this extra term can be bounded in terms of the nullity of A, thereby generalizing a result of Bevilacqua et al. [2]. Moreover, we derive an additional bound which holds for the special case of the Moore-Penrose inverse $A^\dagger$.

We will then translate these results on rank structures to the context of displacement structures. This allows us to obtain some results in the style of Heinig and Hellinger [8, 7]. The generalization consists in the fact that we work with *decoupled* displacement structures, involving two matrices A and B rather than a single matrix A.

Section 3 handles the preservation of structure under *generalized Schur complementation*. The outline of this section is similar to the one of Section 2, in the sense that we first obtain results for rank structures, which we generalize then later so that they imply results for displacement structured matrices as well. In particular, this approach leads to a self-contained proof of the inheritance of Stein type displacement structure under (usual) Schur complementation, which does not require any embedding approach or brute-force formula of any type.

Section 4 describes some additional topics for the Moore-Penrose inversion of a full column rank matrix.

2 Generalized inversion

In this section we investigate the preservation of structure under generalized inversion, and in particular Moore-Penrose inversion. This section is organized as follows. In Subsection 1 we start by recalling some facts on the inversion of rank structures in the nonsingular case. Subsection 2 considers some generalizations of these results to the case of generalized inversion. Subsection 3 shows how these results on rank structures can be reformulated to a more general form, which allows then in Subsection 4 to derive some results on the generalized inversion of displacement structures as well.

2.1 Rank structures: the nonsingular case

We start with the generalized inversion of rank structures. The present subsection reviews some facts from the square nonsingular case.

First we recall the definition of a rank structure. We give here the definition used in [3]; a relaxation of the size restrictions occurring in this definition will be provided later.

Definition 3 *We define a* structure block $\mathcal{B}$ *on* $\mathbb{C}^{n\times n}$ *as a 4-tuple*

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

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 $\mathcal{B}$ (it is assumed here that $j-i+1 \geq 0$). We say a matrix $A \in \mathbb{C}^{n\times n}$ to satisfy the structure block if, making a partitioning

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

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

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

where Rk r *denotes 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* $\mathcal{B}$ *with the '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 structure block with shift matrix of the form* $\Lambda = 0 \oplus \infty I$*, is called* pure*, denoted* $\mathcal{B}_{\text{pure}}$*.*

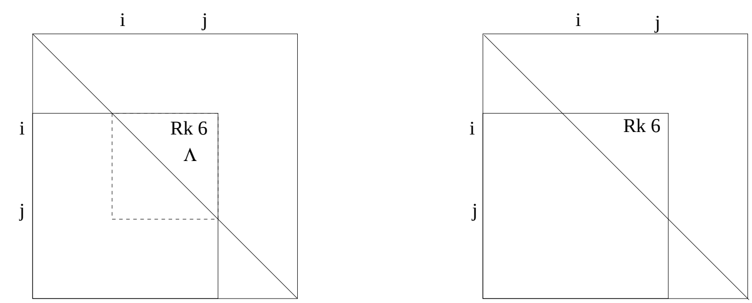

Figure 1: The structure block $\mathcal{B}$ in the left picture has the following meaning: after subtracting the shift matrix $\Lambda \in \mathbb{C}^{4\times 4}$ from the dashed square submatrix in the middle, the indicated bottom left submatrix must be of rank at most 6. The structure block $\mathcal{B}_{\text{pure}}$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

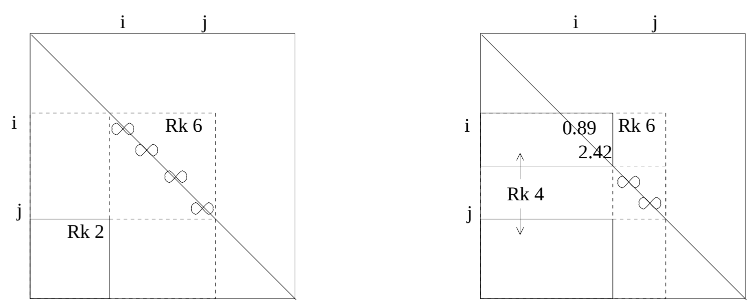

Figure 2: The structure block $\mathcal{B}_{\text{pure}}$ in the left picture has shift matrix $\Lambda = \infty I_4$. Hence by definition, it should be identified with the Rk 2 structure block in the bottom left corner. The structure block $\mathcal{B}$ in the right picture has $\Lambda = \text{diag}(0.89, 2.42, \infty, \infty)$. 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.

Theorem 4 *(see [3, Corollary 16]:) Let* $A \in \mathbb{C}^{n\times n}$ *be a nonsingular matrix satisfying the structure block* $\mathcal{B} = (i, j, r, \Lambda)$*, where* $\Lambda = \Lambda_{\text{ns}} \oplus 0 \oplus \infty I$*, with*

$\Lambda_{\rm ns}$ *nonsingular. Then the inverse matrix* A^{-1} *will satisfy the structure block* $\mathcal{B}^{-1} := (i, j, r, \Lambda^{-1})$*, with* $\Lambda^{-1} := \Lambda_{\rm ns}^{-1} \oplus \infty I \oplus 0$ *(hence using the rules* $\frac{1}{0} = \infty$ *and* $\frac{1}{\infty} = 0$*).*

As an illustration of this theorem, the reader should try to find the inverse of each of the structures shown in Figures 1 and 2.

Note that the remarkable thing about the inversion theorem is that both the shape and the rank upper bound of a structure block are invariant under matrix inversion. But one should not forget that this property only holds if consistent use is made of 'shift elements ∞'. However, sometimes it is useful to get '∞-free' versions of these results. We recall that each independent shift element ∞ has the effect of decreasing the rank by one, and skipping one row and column out of the structure block; it is clear that these operations may influence the rank and shape but not the *nullity* of the structure block.

Definition 5 *The* nullity *of a matrix* $A \in \mathbb{C}^{m \times n}$ *is defined as the dimension of the right null space of* A*, i.e., the number of dependencies between the columns of* A*. We denote it by* Null A*.*

Theorem 6 *(See [3, Theorem 20]:) Let* $n \in \mathbb{N}$ *and define an index set* $N = \{1, \ldots, n\}$*. Suppose that we have partitions* $N = R \cup S \cup T = \tilde{R} \cup \tilde{S} \cup \tilde{T}$ *with* S *and* $\tilde{S}$ *having the same size. Then*

$$\mathrm{Null}(A^{-1}_{\Lambda^{-1}}(\tilde{S} \cup \tilde{T}, R \cup S)) = \mathrm{Null}(A_{\Lambda}(S \cup T, \tilde{R} \cup \tilde{S})), \tag{3}$$

where $A^{-1}_{\Lambda^{-1}}$ *is defined from* A^{-1} *by putting* $A^{-1}_{\Lambda^{-1}}(\tilde{S}, S) = A^{-1}(\tilde{S}, S) - \Lambda^{-1}$*, and similarly* A_{Λ} *is defined from* A *by putting* $A_{\Lambda}(S, \tilde{S}) = A(S, \tilde{S}) - \Lambda$*.*

An illustration of this property is given in Figure 3, where the property is illustrated for the distribution of index sets which is of main interest.

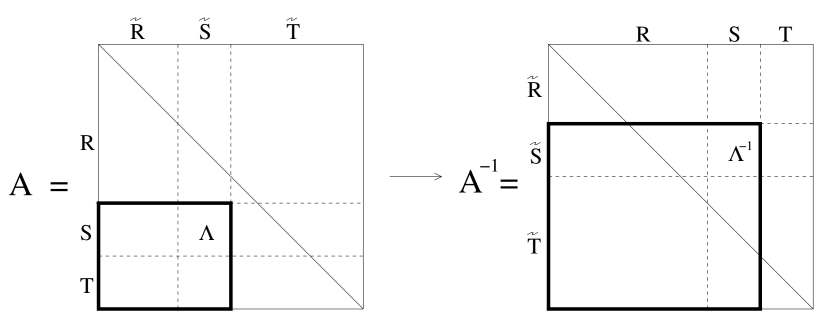

Figure 3: Inheritance of structure by the inverse matrix

The proof of this theorem follows from the above remarks about shift elements ∞ leaving the nullity invariant, combined with the use of suitable permutations to bring the structure to the lower left matrix corner [3]. Note that we recover the so-called nullity theorem of Fiedler and Markham [5] if both S and $\tilde{S}$ are empty sets: this theorem expresses that the nullity of complementary subsets of a matrix and its transposed inverse are equal.

Corollary 7 *(See [5]:) Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Then for any two index sets I and J, we have that*

$$\text{Null } A^{-1}(I,J) = \text{Null } A(N \setminus J, N \setminus I), \tag{4}$$

where $N := \{1, 2, \ldots, n\}$.

2.2 Rank structures: generalized inversion

Inspired by the results for matrix inversion, we are going to establish analogous results for the generalized inversion of a rank structured matrix $A \in \mathbb{C}^{m\times n}$. The first concern is that A may be a rectangular matrix, and hence that we should incorporate that the matrix dimensions are transposed under generalized inversion.

Still we will be able to take over the definitions for square matrices almost literally.

Definition 8 *Let $m, n \in \mathbb{N}$ and define index sets $M = \{1, \ldots, m\}$, $N = \{1, \ldots, n\}$. Suppose that we have partitions $M = R \cup S \cup T$ and $N = \tilde{R} \cup \tilde{S} \cup \tilde{T}$ with S and $\tilde{S}$ having the same size, and let Λ be square of size $|S|$ and $r \in \mathbb{N}$. Then we say A to satisfy the structure block $\mathcal{B} = (R, S, T, \tilde{R}, \tilde{S}, \tilde{T}, \Lambda, r)$ if*

$$A_\Lambda(S \cup T, \tilde{R} \cup \tilde{S}) = \text{Rk } r, \tag{5}$$

where A_Λ is defined from A by putting $A_{\Lambda^{-1}}(S, \tilde{S}) = A(S, \tilde{S}) - \Lambda$. We define the inverse structure block $\mathcal{B}^{-1} := (\tilde{R}, \tilde{S}, \tilde{T}, R, S, T, \Lambda^{-1}, \tilde{r})$, with the understanding that we consider the value of $\tilde{r}$ as unspecified: see Figure 4.

Note that we did not specify the exact rank of the inverse structure block $\mathcal{B}^{-1}$. We do this since it may depend on the actual choice of the generalized inverse matrix.

We will derive the inversion results for generalized inverses from those for square nonsingular matrices by using the well-known embedding approach. To this end we recall the following lemma.

Lemma 9 *Let $A \in \mathbb{C}^{m\times n}$ be a given matrix, and suppose that we add to this matrix extra rows and columns such that*

$$\begin{array}{c} \\ \text{Null } A \; \{ \end{array} \left[\begin{array}{cc} A & \overbrace{A(M,\tilde{U})}^{\text{Null } A^H} \\ A(U,N) & A(U,\tilde{U}) \end{array}\right] \tag{6}$$

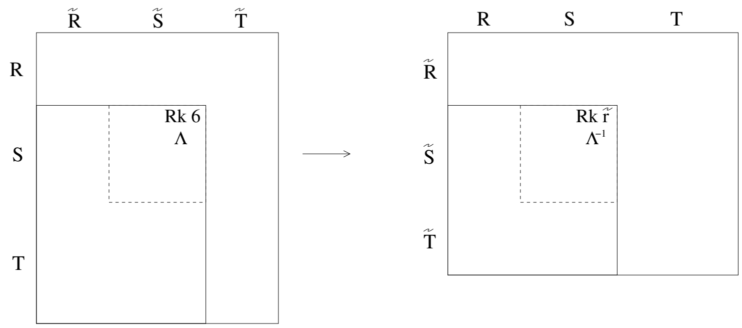

Figure 4: Given the Rk 6 structure block $\mathcal{B}$ in the left picture, the right picture shows the position of the structure block $\mathcal{B}^{-1}$. The actual value of $\tilde{r}$ depends on the choice of the generalized inverse.

is nonsingular, where we denoted index sets $M = \{1, \ldots, m\}$ *and* $N = \{1, \ldots, n\}$, *and where the index sets* U *and* $\tilde{U}$ *contain the added elements and are of the indicated dimensions. Then the inverse of this matrix admits a partition*

$$\begin{array}{c} \\ \text{Null } A^H \left\{ \vphantom{\begin{matrix} 1 \end{matrix}} \right. \end{array} \begin{bmatrix} X & \overbrace{*}^{\text{Null } A} \\ * & * \end{bmatrix} \tag{7}$$

where $X \in \mathbb{C}^{n \times m}$ *is a generalized inverse of* A.

Conversely, any generalized inverse X *of* A *can be realized by this procedure.*

The generalized inverse X *is* reflexive *(Definition 2) if and only if* $A(U, \tilde{U})$ *equals zero.*

Finally, X *equals the Moore-Penrose inverse* $A^\dagger$ *if and only if* $A(U, \tilde{U})$ *equals zero, the columns of* $A(M, \tilde{U})$ *form a basis for the right null space of* A^H, *and the columns of* $(A(U, N))^H$ *form a basis for the right null space of* A.

We will not repeat a proof of Lemma 9 here.

Now we can use Lemma 9 to obtain results for the structure preservation under generalized inversion, by reducing them to usual matrix inversion. Thus assume that A satisfies a certain structure block $\mathcal{B}$. Then consider the subdivision of A induced by this structure block, as in (1), and partition the matrices (6) and (7) correspondingly as

$$\left[\begin{array}{ccc|c} A_{1,1} & A_{1,2} & A_{1,3} & \overbrace{A_{1,4}}^{\text{Null } A^H} \\ A_{2,1} & A_{2,2} & A_{2,2} & A_{2,4} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} \\ \hline A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} \end{array} \right] \begin{array}{l} \\ \\ \\ \} \text{ Null } A \end{array} \tag{8}$$

and

$$
\begin{array}{r}
\\
\\
\\
\text{Null } A^H \{
\end{array}
\left[\begin{array}{ccc|c}
 & & & \overbrace{\quad}^{\text{Null } A} \\
X_{1,1} & X_{1,2} & X_{1,3} & * \\
X_{2,1} & X_{2,2} & X_{2,2} & * \\
X_{3,1} & X_{3,2} & X_{3,3} & * \\
\hline
* & * & * & *
\end{array}\right]. \tag{9}
$$

Theorem 10 *(Generalized inversion:) Let $A \in \mathbb{C}^{m\times n}$ satisfy a structure block $\mathcal{B}$. Then for any generalized inverse X of A we have*

$$
\text{Null}\begin{bmatrix} X_{2,1} & X_{2,2}-\Lambda^{-1} \\ X_{3,1} & X_{3,2} \end{bmatrix} = \text{Null}\left[\begin{array}{cc|c} A_{2,1} & A_{2,2}-\Lambda & A_{2,4} \\ A_{3,1} & A_{3,2} & A_{3,4} \\ \hline A_{4,1} & A_{4,2} & A_{4,4} \end{array}\right], \tag{10}
$$

where the partitions are defined as in (8), (9)*.*

PROOF. This follows immediately from Theorem 6. □

We can now generalize a result proved in [2].

Corollary 11 *Let A satisfy a structure block $\mathcal{B}$. Then any generalized inverse X of A will satisfy the structure block $\mathcal{B}^{-1}$, with nullity lying between the extremal values* $\text{Null } \mathcal{B} - \text{Null } A$ *and* $\text{Null } \mathcal{B} + \text{Null } A^H$.

We can also incorporate permutations.

Theorem 12 *Let $A \in \mathbb{C}^{m\times n}$ and define index sets $M = \{1,\dots,m\}$, $N = \{1,\dots,n\}$. Suppose that we have partitions $M = R\cup S\cup T$ and $N = \tilde{R}\cup\tilde{S}\cup\tilde{T}$ with S and $\tilde{S}$ having the same size. Then*

$$
\text{Null}(X_{\Lambda^{-1}}(\tilde{S}\cup\tilde{T}, R\cup S)) = \text{Null}\begin{bmatrix} A_\Lambda(S\cup T, \tilde{R}\cup\tilde{S}) & A(S\cup T,\tilde{U}) \\ A(U,\tilde{R}\cup\tilde{S}) & A(U,\tilde{U}) \end{bmatrix}, \tag{11}
$$

where A_Λ and $A^{-1}_{\Lambda^{-1}}$ are defined as in Theorem 6, and where U and $\tilde{U}$ denote the indices of adjoined rows and columns constructed in Lemma 9.

Indeed, this result is nothing but a restatement of Theorem 10.

Now we return to Theorem 10. This result can be refined for the case of the Moore-Penrose inverse $A^\dagger$, by virtue of the more precise information about the elements standing in the index sets U and $\tilde{U}$ in Lemma 9.

Rather than stating such a refinement for the Moore-Penrose inverse, we will present here a more transparent and easily applicable bound. The following result will be stated only for *pure* structure blocks.

Theorem 13 *(Moore-Penrose inversion:) Let $A \in \mathbb{C}^{m\times n}$ satisfy the pure structure block $\mathcal{B}$ specified by $A(I,J) = \text{Rk } r$ with $r \in \mathbb{N}$. Then if $\mathcal{B}$ is 'complemented' by another pure structure block, in the sense that $A(M\setminus I, N\setminus J) = \text{Rk } s$ for $s \in \mathbb{N}$, we have that $A^\dagger$ satisfies the structure block $\mathcal{B}^{-1}$ with rank at most equal to $r+s$, i.e.,* $\text{Rank}(A^\dagger(N\setminus J, M\setminus I)) \le r+s$*: see Figure 5.*

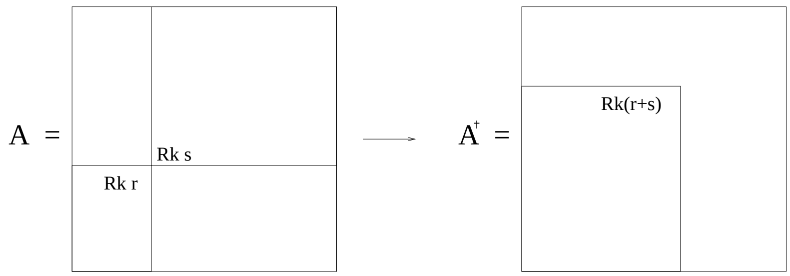

Figure 5: Illustration of the bound for Moore-Penrose inversion of a Rk r structure block which is complemented by a Rk s structure block.

PROOF. By suitable permutations, it may be assumed that the given Rk r structure block is situated in the bottom left corner of A, and hence that the Rk s structure block is situated in the top right corner of A: see the left part of Figure 5.

Now let us transform A into a new matrix

$$\tilde{A} := (I \oplus U)A(V \oplus I), \tag{12}$$

where U, V are unitary transformations acting inside I, J, respectively, chosen such that the rank-s block maximally expands, i.e., such that $\tilde{A}(M \setminus \tilde{I}, N \setminus \tilde{J}) =$ Rk s, where $\tilde{I} \subseteq I$ and $\tilde{J} \subseteq J$ are minimal. Obviously, we have then also $\tilde{A}(\tilde{I}, \tilde{J}) =$ Rk r. Suppose then by induction that we could prove from these assumptions that $\text{Rank}(\tilde{A}^\dagger(N \setminus \tilde{J}, M \setminus \tilde{I})) \leq r + s$. It would follow then a fortiori that $\text{Rank}(\tilde{A}^\dagger(N \setminus J, M \setminus I)) \leq r + s$. Now by the compatibility of Moore-Penrose inversion with the multiplication with unitary transformations, we have from (12) that $\tilde{A}^\dagger = (V^{-1} \oplus I)A^\dagger(I \oplus U^{-1})^*$, where the two unitary operations act *outside* the structure block $\mathcal{B}^{-1}$ and hence do not influence its structure block rank. We obtain then the desired bound $\text{Rank}(A^\dagger(N \setminus J, M \setminus I)) \leq r + s$.

We conclude from the previous paragraph that it is sufficient to prove the theorem under the assumption that no row of I depends on previous rows, and no column of J depends on later columns. (Since the existence of such a dependency would allow us to apply a reduction of the form (12).) Thus we can assume that all row and column dependencies of A must be strictly inside the index sets $M \setminus I$, $N \setminus J$, respectively.

To prove this remaining case, we apply a tranformation of A into a new matrix

$$\tilde{A} := (U \oplus I)A(I \oplus V), \tag{13}$$

where now U, V are unitary transformations acting inside $M \setminus I$, $N \setminus J$, respectively, chosen such that all linear dependencies inside these index sets are

*The fact that $(UAV)^\dagger = V^{-1}A^\dagger U^{-1}$ whenever U and V are unitary can be easily verified by means of the Moore-Penrose equations of Definition 1.

transformed into zero rows on top and zero columns on the right. Thus the matrix $\tilde{A}$ takes the form

$$\tilde{A} = \begin{bmatrix} 0 & 0_{a\times b} \\ X & 0 \end{bmatrix}, \tag{14}$$

with X nonsingular, and for suitable $a, b \in \mathbb{N}$. The Moore-Penrose inverse of this matrix is given by

$$\tilde{A}^{\dagger} = \begin{bmatrix} 0 & X^{-1} \\ 0_{b\times a} & 0 \end{bmatrix}. \tag{15}$$

Observe now that the submatrix X in (14) must be such that $X(I,J) = \text{Rk } r$ and $X(\tilde{M}\setminus I, \tilde{N}\setminus J) = \text{Rk } s$, where $\tilde{M} \subseteq M$, $\tilde{N} \subseteq N$ are the index sets obtained by removing the indices of the zero rows and columns in (14). Suppose then by induction that we could prove from these assumptions that $X^{-1}(\tilde{N}\setminus J, \tilde{M}\setminus I) \leq r+s$. Then the presence of the zeros in (15) allows us to refine this conclusion to $\tilde{A}^{\dagger}(N\setminus J, M\setminus I) \leq r+s$. From this result on $\tilde{A}^{\dagger}$, we can then again derive that exactly the same result must hold for the matrix $A^{\dagger}$ (without the tilde). Indeed, this follows since the row and column operations $I \oplus V^{-1}$, $U^{-1} \oplus I$ obtained from inverting (13) are both nonsingular operations acting completely *inside* the structure block $\mathcal{B}^{-1}$, and hence not influencing its structure block rank.

We conclude from the previous paragraph that it is sufficient to prove the theorem for the case where A is square nonsingular. Let us prove now this remaining case. From the assumption that $A(I,J) = \text{Rk } r$, it follows from Corollary 7 that $A^{-1}(N\setminus J, N\setminus I) = \text{Rk } \tilde{r}$, where $\tilde{r} = r + n - |I| - |J|$. (Indeed, this value of $\tilde{r}$ is such that the *nullity* of the structure blocks $\mathcal{B}$ and $\mathcal{B}^{-1}$ is the same.) It will then be sufficient to prove that $n - |I| - |J| \leq s$. To this end, we recall the assumption $A(N\setminus I, N\setminus J) = \text{Rk } s$ to obtain the bound $A = \text{Rk } \tilde{s}$ with $\tilde{s} := s + |I| + |J|$. Since by assumption A is nonsingular, it follows that $s + |I| + |J| \geq n$, which was to be demonstrated. □

Remark 14 *A trivial case where the above theorem is satisfied is when* $r = s = 0$ *(the block diagonal case). It might then be tempting to conjecture that the result for general* r *and* s *can be derived from this special case. Indeed, note first that the matrix with general* r *and* s *can be written as a block diagonal matrix, plus a correction term of rank at most* $r+s$*. It would then suffice to prove that the Moore-Penrose of a rank-*k *correction of a matrix* C*, equals a rank-*k *correction of the Moore-Penrose* $C^{\dagger}$*. Unfortunately, although this property is true in the nonsingular case, it fails for the case of Moore-Penrose inversion, even when* C *is square nonsingular. This is related to the fact that the so-called Sherman-Morrison formula [6] does not have a straightforward analogue for the case of Moore-Penrose inversion.*

Note that Theorem 13 deviates from the earlier results in this section in the sense that it bounds the *rank* of the structure block $\mathcal{B}^{-1}$, rather than its nullity. Therefore even in the square nonsingular case, this bound is not completely straightforward (cf. the last paragraph of the proof of Theorem 13).

2.3 Pure rank structures: reformulation of the results

In this subsection we reconsider the results on *pure* rank structures of the previous two subsections, and tranlate them to a slightly more general form involving orthogonal projectors.

We start with the nonsingular case.

Theorem 15 *(The nonsingular case:) Let* $A \in \mathbb{C}^{n\times n}$ *be nonsingular, and let* $S \in \mathbb{C}^{n\times s}, T \in \mathbb{C}^{n\times t}$ *be full rank matrices,* $\{s,t\} < n$*. Then we have*

$$\text{Null } (S^{\perp}A^{-1}T) = \text{Null } (T^{\perp}AS), \tag{16}$$

where $S^{\perp} \in \mathbb{C}^{(n-s)\times n}, T^{\perp} \in \mathbb{C}^{(n-t)\times n}$ *denote full rank matrices for which* $S^{\perp}S = 0$ *and* $T^{\perp}T = 0$*.*

PROOF. We start by completing each of S, T to a square nonsingular matrix by adding extra columns to it. This leads to

$$S = U\begin{bmatrix} I_s \\ 0 \end{bmatrix}, \quad T = V\begin{bmatrix} I_t \\ 0 \end{bmatrix}, \tag{17}$$

where U and V are nonsingular matrices. We can then identify

$$S^{\perp} := \begin{bmatrix} 0 & I_{n-s} \end{bmatrix} U^{-1}, \quad T^{\perp} := \begin{bmatrix} 0 & I_{n-t} \end{bmatrix} V^{-1}. \tag{18}$$

Indeed, these definitions are such that the required equations $S^{\perp}S = 0$ and $T^{\perp}T = 0$ are satisfied. It follows now from the above definitions that

$$\text{Null } (S^{\perp}A^{-1}T) = \text{Null } (\begin{bmatrix} 0 & I_{n-s} \end{bmatrix} U^{-1}A^{-1}V \begin{bmatrix} I_t \\ 0 \end{bmatrix}). \tag{19}$$

This is a structure block for the matrix $U^{-1}A^{-1}V$. By Corollary 7, it follows that its nullity equals the nullity of

$$\text{Null } (\begin{bmatrix} 0 & I_{n-t} \end{bmatrix} V^{-1}AU \begin{bmatrix} I_s \\ 0 \end{bmatrix}) =: \text{Null } (T^{\perp}AS),$$

which is the desired result (16). □

Remark 16 *By a QR-factorization of* S, T *(with square nonsingular R-factor) we may assume without loss of generality that these matrices are normalized such that* $S^H S = I_s$ *and* $T^H T = I_t$*. Let us in addition normalize* $S^{\perp}, T^{\perp}$ *by means of an RQ-factorization. The pair* $\{S, S^{\perp}\}$ *can then be interpreted as a pair of* orthogonal projectors*. It is in this terminology that the equivalent of Theorem 15 in Bevilacqua et al. [2] was stated. However, the current formulation of Theorem 15 will be more suited for deriving the results for displacement structures in the next subsection.*

In a similar way one can reformulate the theorems for singular or even rectangular A to the setting of orthogonal projectors. Recall that for $A \in \mathbb{C}^{m \times n}$, Lemma 9 asserts that we can realize any generalized inverse X of A as the top left block of the inverse of an embedded matrix

$$\begin{bmatrix} A & A_{c,4} \\ A_{r,4} & A_{4,4} \end{bmatrix}, \tag{20}$$

where $A_{r,4}, A_{c,4}, A_{4,4}$ denote the added elements. (The provenance of the subscript '4' follows from the partition in (8), although we will have no shift matrix Λ involved here).

We have the following result.

Theorem 17 *(Generalized inversion:) Let $A \in \mathbb{C}^{m \times n}$, and let X be a generalized inverse of A. Define the corresponding embedded matrix of A as in* (20)*. Then we have*

$$\text{Null } (S^{\perp} X T) = \text{Null } \begin{bmatrix} T^{\perp} A S & T^{\perp} A_{c,4} \\ A_{r,4} S & A_{4,4} \end{bmatrix}, \tag{21}$$

where $S, S^{\perp}, T, T^{\perp}$ are defined as in Theorem 15 (except that some occurrences of n should now be replaced by m since we have now A rectangular of size m by n; this will be clear from the context).

PROOF. Let us again identify $S, S^{\perp}, T, T^{\perp}$ as in (17), (18) for suitable invertible matrices U, V. It follows then that

$$\text{Null } (S^{\perp} X T) = \text{Null } (\begin{bmatrix} 0 & I_{n-s} \end{bmatrix} U^{-1} X V \begin{bmatrix} I_t \\ 0 \end{bmatrix}). \tag{22}$$

This is a structure block for the matrix $U^{-1} X V$. Now observe that the latter matrix is a generalized inverse of $V^{-1} A U$. More precisely, we can realize $U^{-1} X V$ as the top left block of the inverse of the embedded matrix

$$\begin{bmatrix} V^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & A_{c,4} \\ A_{r,4} & A_{4,4} \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix}. \tag{23}$$

By Theorem 10, the structure block in the right hand side of (22) is then in correspondence with a structure block of this embedded matrix (23):

$$\begin{aligned} & \text{Null } (\begin{bmatrix} 0 & I_{n-s} \end{bmatrix} U^{-1} X V \begin{bmatrix} I_t \\ 0 \end{bmatrix}) \\ = \; & \text{Null } (\begin{bmatrix} 0 & I_{m-t} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} V^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & A_{c,4} \\ A_{r,4} & A_{4,4} \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I_s & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix}) \\ = \; & \text{Null } (\begin{bmatrix} T^{\perp} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & A_{c,4} \\ A_{r,4} & A_{4,4} \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & I \end{bmatrix}) \\ = \; & \text{Null } (\begin{bmatrix} T^{\perp} A S & T^{\perp} A_{c,4} \\ A_{r,4} S & A_{4,4} \end{bmatrix}), \end{aligned}$$

which is the desired result (21). $\square$

Finally, one can translate the result for Moore-Penrose inversion to the setting of orthogonal projectors.

Theorem 18 *(Moore-Penrose inversion:) Let $A \in \mathbb{C}^{m\times n}$, then we have the following bound for the Moore-Penrose inverse $A^\dagger$:*

$$\mathrm{Rank}(S^\perp A^\dagger T) \leq \mathrm{Rank}(T^\perp AS) + \mathrm{Rank}(S^\perp A^H T), \tag{24}$$

where $S, S^\perp, T, T^\perp$ are defined as in Theorem 15.

PROOF. By a QR-factorization of S, T we may assume without loss of generality that these matrices are normalized such that $S^H S = I_s$ and $T^H T = I_t$. We can then identify $S, S^\perp, T, T^\perp$ as in (17), (18), where now U and V can be chosen to be *unitary* matrices. It follows that

$$\mathrm{Rank}(S^\perp A^\dagger T) = \mathrm{Rank}(\begin{bmatrix} 0 & I_{n-s} \end{bmatrix} U^{-1} A^\dagger V \begin{bmatrix} I_t \\ 0 \end{bmatrix}).$$

This is a structure block for the matrix $U^{-1}A^\dagger V = (V^{-1}AU)^\dagger$. By Theorem 13, the rank of this structure block can be bounded by the sum of the following two ranks

$$\mathrm{Rank}(\begin{bmatrix} 0 & I_{m-t} \end{bmatrix} (V^{-1}AU) \begin{bmatrix} I_s \\ 0 \end{bmatrix})$$

and

$$\mathrm{Rank}(\begin{bmatrix} 0 & I_{n-s} \end{bmatrix} (V^{-1}AU)^H \begin{bmatrix} I_t \\ 0 \end{bmatrix})$$

(note the Hermitian transpose sign ^{H}. We could in fact freely interchange this by a usual transposition sign ^{T} since the complex conjugation of a matrix does not change its rank). The desired equation (24) now follows. $\square$

2.4 Displacement structures

In this subsection we apply the results of the previous subsection to obtain results on the generalized inversion of *displacement structures.* The trick will be to rewrite the displacement structure in a block matrix form.

First we recall the definition of displacement structured matrices. We distinguish between two types.

Definition 19 *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*

$$AF - GB = \mathrm{Rk}\ r, \tag{25}$$

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

Definition 20 *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*

$$A - GBH = \text{Rk } r, \tag{26}$$

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

Note that we assume all matrix dimensions in the above definitions to be compatible.

In practical applications, the case where $A = B$ and F, G, H have a simple form is of particular interest [9]. With the aim of structure preservation under Schur complementation, the matrices F, G, H must in addition be block upper or lower triangular (see Section 3).

Now we come to the generalized inversion of displacement structures of Stein type. We start by rewriting the expression $A - GBH$ as

$$\begin{bmatrix} I & G \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I \\ -H \end{bmatrix}. \tag{27}$$

Inspired by this equation, we define

$$S = \begin{bmatrix} I \\ -H \end{bmatrix}, \quad T = \begin{bmatrix} G \\ -I \end{bmatrix}, \tag{28}$$

$$S^{\perp} := \begin{bmatrix} H & I \end{bmatrix}, \quad T^{\perp} := \begin{bmatrix} I & G \end{bmatrix}. \tag{29}$$

Note that these definitions are indeed such that $S^{\perp}S = 0$ and $T^{\perp}T = 0$. (Let us stress here that both G and H may be rectangular of any size, with the only size restriction being that $A - GBH$ is well-defined. Hence, each of the four identity matrices I occurring in (28) and (29) can be of different size too.)

It follows now from Theorem 15 and the identifications (28), (29) that in the case where both A and B are nonsingular, the nullity of (27) equals the nullity of

$$\begin{bmatrix} H & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} G \\ -I \end{bmatrix},$$

which is the nullity of $HA^{-1}G - B^{-1}$, i.e.,

$$\text{Null } (A - GBH) = \text{Null } (HA^{-1}G - B^{-1}). \tag{30}$$

This retrieves the result for the inversion of displacement structures of Stein type in the nonsingular case.

Now for the singular case. Let X, Y be any pair of generalized inverses of A, B, respectively, and define the corresponding embedded versions of A and B as in (20).

We can then 'glue' these two embedded matrices together to an embedding of the matrix $A \oplus B$, defined by

$$\left[\begin{array}{cc|cc} A & 0 & A_{c,4} & 0 \\ 0 & B & 0 & B_{c,4} \\ \hline A_{r,4} & 0 & A_{4,4} & 0 \\ 0 & B_{r,4} & 0 & B_{4,4} \end{array}\right].$$

Indeed: it is easily checked that the top left block of the inverse of this matrix equals precisely $X \oplus Y$.

Defining $S, S^{\perp}, T, T^{\perp}$ again by means of (28), (29), it follows now from Theorem 17 that

$$\text{Null }(HXG - Y) = \text{Null} \begin{bmatrix} A - GBH & A_{c,4} & GB_{c,4} \\ A_{r,4} & A_{4,4} & 0 \\ -B_{r,4}H & 0 & B_{4,4} \end{bmatrix}. \tag{31}$$

Finally, let us specify to the Moore-Penrose inverse. We can then apply Theorem 18 to obtain

$$\text{Rank}(HA^{\dagger}G - B^{\dagger}) \leq \text{Rank}(A - GBH) + \text{Rank}(HA^{H}G - B^{H}). \tag{32}$$

(Again, let us stress that this result is valid for *any* rectangular A, B, G and H; the only restriction is that $A - GBH$ is well-defined.)

It follows from (32) that in case G and H are *unitary* matrices, the displacement rank of the Moore-Penrose inverses can at most *double* with respect to the original displacement rank. A direct argument reveals that in the case $A = B$, this doubling result holds also for each of G and H either unitary or Hermitian: this was shown by Heinig and Hellinger [8]. Unfortunately we can not retrieve the latter result as a consequence of Theorem 18.

In a similar way as the results for the Stein displacement equation described in the paragraphs above, one can also obtain bounds for the generalized inversion of the *Sylvester* displacement equation $AF - GB = \text{Rk } r$. We start then by rewriting the expression $AF - GB$ as

$$\begin{bmatrix} I & G \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} F \\ -I \end{bmatrix},$$

from which we define

$$S = \begin{bmatrix} F \\ -I \end{bmatrix}, \quad T = \begin{bmatrix} G \\ -I \end{bmatrix},$$

$$S^{\perp} := \begin{bmatrix} I & F \end{bmatrix}, \quad T^{\perp} := \begin{bmatrix} I & G \end{bmatrix}.$$

The results for the Sylvester displacement structure are then quite similar to the results for the Stein displacement structure (30), (31) and (32) and will be stated now without derivation:

$$\text{Null }(A^{-1}G - FB^{-1}) = \text{Null }(AF - GB). \tag{33}$$

$$\text{Null }(XG - FY) = \text{Null} \begin{bmatrix} AF - GB & A_{c,4} & GB_{c,4} \\ A_{r,4}F & A_{4,4} & 0 \\ -B_{r,4} & 0 & B_{4,4} \end{bmatrix}. \tag{34}$$

$$\text{Rank}(A^{\dagger}G - FB^{\dagger}) \leq \text{Rank}(AF - GB) + \text{Rank}(A^{H}G - FB^{H}). \tag{35}$$

3 Generalized Schur complementation

In this section we will consider the generalized Schur complementation of rank and displacement structured matrices. The outline of this section is quite similar to Section 2. In particular, the first two subsections deal with the generalized Schur complementation of rank structures. Subsection 3.3 shows how these results on rank structures can be reformulated to a more general form, which allows then in Subsection 3.4 to derive some results on the generalized Schur complementation of displacement structures as well.

3.1 Rank structures: Schur complementation

In the present subsection, we recall some of the results for the (usual) Schur complementation of rank structures. These results will then be extended in the next subsection to obtain bounds for *generalized* Schur complementation.

Let us start with some elementary definitions.

Definition 21 *Given $A \in \mathbb{C}^{m\times n}$, and given an integer k. We define the* k-partitioning *of A as*

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}_k, \tag{36}$$

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.

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

$$L_{Gauss} \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} R_{Gauss} = \begin{bmatrix} A_{1,1} & 0 \\ 0 & S_{A,k} \end{bmatrix}, \tag{37}$$

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 invertible block lower and upper triangular matrices, respectively.

In relation to this, the following result is well-known.

Lemma 22 *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

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

Let us now give the definition of structure blocks in the context of Schur complements. We will do this in a slightly more general way than in [4]. It could be useful for the reader to have first a glimpse at Figure 6 before embarking on the following definition.

Definition 23 *Let k, m, n be integers, $k \leq \{m, n\}$, and define corresponding index sets $K = \{1, \ldots, k\}$, $M = \{k+1, \ldots, m\}$ and $N = \{k+1, \ldots, n\}$. Suppose that we have partitions $K = R \cup S \cup T = \tilde{R} \cup \tilde{S} \cup \tilde{T}$, $M = R' \cup S' \cup T'$, $N = \tilde{R}' \cup \tilde{S}' \cup \tilde{T}'$ such that S and $\tilde{S}$ have the same size. Suppose also given a matrix Λ of size $|S| + |S'|$ by $|\tilde{S}| + |\tilde{S}'|$, and let*

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

be the natural partitioning of this matrix, thus with $\Lambda_{1,1}$ of size $|S|$ by $|\tilde{S}|$, $\Lambda_{1,2}$ of size $|S|$ by $|\tilde{S}'|$, and so on. (By the restriction that S and $\tilde{S}$ have the same size, we have actually that $\Lambda_{1,1}$ is square.) Suppose also given an integer r.

We say the matrix $A \in \mathbb{C}^{m \times n}$ to satisfy the structure block $\mathcal{B}$ *induced by the above data if*

$$\tilde{A}(S \cup T \cup S' \cup T', \tilde{R} \cup \tilde{S} \cup \tilde{R}' \cup \tilde{S}') = \text{Rk } r,$$

where $\tilde{A}$ has been defined from A by $\tilde{A}(S \cup S', \tilde{S} \cup \tilde{S}') = A(S \cup S', \tilde{S} \cup \tilde{S}') - \Lambda$: see Figure 6.

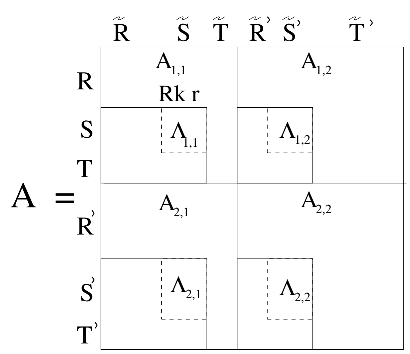

Figure 6: Given a matrix A together with a k-partitioning of A, which is visualized by the horizontal and vertical line in the figure. The figure shows an example of a structure block $\mathcal{B}$ satisfied by this matrix. The meaning is that after subtracting the shift matrix Λ (consisting of four parts) from the dashed matrix positions, the indicated submatrix of A (also consisting of four parts) must be of rank at most r.

Figure 6 illustrates that a given structure block $\mathcal{B}$ can be considered as a collection of four individual parts w.r.t. the given k-partitioning. Moreover, note that the size restriction in Definition 23 expresses precisely that the top left part of the structure block has a *square* shift matrix $\Lambda_{1,1}$. This reflects the fact that the Schur complement is defined as $A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}$, with $A_{1,1}$ appearing in *inversed* form.

Theorem 24 *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 23, let $\Lambda_{1,1}$ be square and nonsingular. Then the Schur complement $S_{A,k}$ satisfies the structure block*

$$\tilde{S}_{A,k}(S' \cup T', \tilde{R}' \cup \tilde{S}') = \text{Rk } \tilde{r},$$

where $\tilde{S}_{A,k}$ is defined from $S_{A,k}$ by setting $\tilde{S}_{A,k}(S', \tilde{S}') = S_{A,k}(S', \tilde{S}') - S_\Lambda$, with $S_\Lambda := \Lambda_{2,2} - \Lambda_{2,1}\Lambda_{1,1}^{-1}\Lambda_{1,2}$, and with $\tilde{r} := r + |R| - |\tilde{R}|$: see Figure 7.

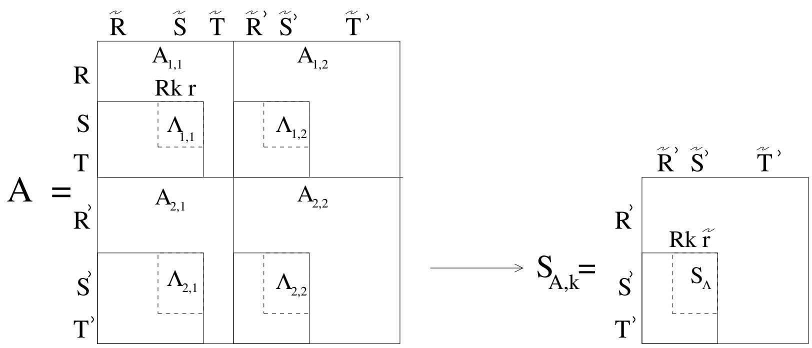

Figure 7: Given the matrix in the left hand side, satisfying the huge structure block $\mathcal{B}$, consisting of four parts. Then the Schur complement $S_{A,k} = A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}$ essentially inherits this structure block, with new shift matrix given by $S_\Lambda := \Lambda_{2,2} - \Lambda_{2,1}\Lambda_{1,1}^{-1}\Lambda_{1,2}$ and new rank given by $\tilde{r} := r + |R| - |\tilde{R}|$.

For a proof of Theorem 24, we refer to [4]. (In fact the latter proof works only when $|R| = |\tilde{R}|$, but the general case can be easily established from this by using 'shift elements ∞'; we omit the details here.)

3.2 Rank structures: generalized Schur complementation

Inspired by the results for Schur complementation, we move now to generalized Schur complementation. This means that we allow $A_{1,1}$ to be singular, or even rectangular.

More precisely, let there be given a matrix $A \in \mathbb{C}^{m\times n}$ and integers $k, \tilde{k}$. We define the $\{k, \tilde{k}\}$*-partitioning* of A as

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}_{k,\tilde{k}}, \tag{38}$$

with $A_{1,1} \in \mathbb{C}^{k\times\tilde{k}}$. We define a *generalized Schur complement* of A induced by this $\{k,\tilde{k}\}$-partitioning as an expression of the form $A_{2,2} - A_{2,1}XA_{1,2}$, where X is some generalized inverse of $A_{1,1}$.

Completing $A_{1,1}$ according to the generalized inverse X by adding extra index sets $U, \tilde{U}$, as described in Lemma 9, we can define the embedded matrix

$$\left[\begin{array}{cc|c} A_{1,1} & A_{1,1}(K,\tilde{U}) & A_{1,2} \\ A_{1,1}(U,\tilde{K}) & A_{1,1}(U,\tilde{U}) & 0 \\ \hline A_{2,1} & 0 & A_{2,2} \end{array}\right], \tag{39}$$

and note that the Schur complement of (39) is precisely the required generalized Schur complement $A_{2,2} - A_{2,1}XA_{1,2}$.

It is then easy to obtain the following analogue of Theorem 10: the generalized Schur complement $A_{2,2} - A_{2,1}XA_{1,2}$ inherits the given structure block $\mathcal{B}$ from A, with new rank bounded by

$$\tilde{r} := \operatorname{Rank}(\tilde{A}(S\cup T\cup U\cup S'\cup T', \tilde{R}\cup\tilde{S}\cup\tilde{U}\cup\tilde{R}'\cup\tilde{S}')) + |R| - |\tilde{R}| - |\tilde{U}|.$$

Here we define $\tilde{A}$ from A as usual by putting $A_\Lambda(S\cup S', \tilde{S}\cup\tilde{S}') = A(S\cup S', \tilde{S}\cup\tilde{S}') - \Lambda$, where we denote with A the embedded matrix (39).

We will now pay some attention to the generalized Schur complement formed by means of the *Moore-Penrose inverse* $A_{1,1}^\dagger$ of $A_{1,1}$, i.e., the generalized Schur complement $S_A^\dagger := A_{2,2} - A_{2,1}A_{1,1}^\dagger A_{1,2}$. It turns out that Theorem 13 can be extended almost literally to this case. The reader might wish to have first a glimpse at Figure 8 before embarking on the following theorem.

Theorem 25 *(Generalized Schur complementation, the Moore-Penrose case:) Given a matrix $A \in \mathbb{C}^{m\times n}$ together with a $\{k,\tilde{k}\}$-partitioning of A. Let A satisfy the pure structure block $\mathcal{B}$ specified by $A(T\cup T', \tilde{R}\cup\tilde{R}') = \operatorname{Rk} r$ with $r \in \mathbb{N}$; here we supposed index sets $K := \{1,\ldots,k\}$, $\tilde{K} := \{1,\ldots,\tilde{k}\}$, $M := \{k+1,\ldots,m\}$, $N := \{\tilde{k}+1,\ldots,n\}$ and partitions $K = R\cup T$, $\tilde{K} = \tilde{R}\cup\tilde{T}$, $M = R'\cup T'$, $N = \tilde{R}'\cup\tilde{T}'$. Assume that $\mathcal{B}$ is 'complemented' by another pure structure block inside $A_{1,1}$, in the sense that $A(R,\tilde{T}) = \operatorname{Rk} s$ for $s \in \mathbb{N}$. Then we have that $S_A^\dagger := A_{2,2} - A_{2,1}A_{1,1}^\dagger A_{1,2}$ inherits the structure block $\mathcal{B}$ with rank at most $r+s$, i.e., $\operatorname{Rank}(S_A^\dagger(T',\tilde{R}')) \le r+s$: see Figure 8.*

PROOF. The proof is very similar to the proof of Theorem 13. We start by applying unitary row and column operations U, V to the index sets T, $\tilde{R}$, respectively, chosen such that the rank-s block maximally expands, i.e., such that $A(R_{\text{new}}, \tilde{T}_{\text{new}}) = \operatorname{Rk} s$, where $R_{\text{new}} \supseteq R$ and $\tilde{T}_{\text{new}} \supseteq \tilde{T}$ are maximal. We can then redefine $T_{\text{new}} := K \setminus R_{\text{new}} \subseteq T$, $\tilde{R}_{\text{new}} := \tilde{K} \setminus \tilde{T}_{\text{new}} \subseteq \tilde{R}$. Since $S_A^\dagger = A_{2,2} - A_{2,1}A_{1,1}^\dagger A_{1,2}$ is invariant under the applied unitary operations, it will obviously suffice to show the result for this new matrix.

We conclude from the previous paragraph that it is sufficient to prove the theorem under the assumption that no row of T depends on previous rows of K, and no column of $\tilde{R}$ depends on later columns of $\tilde{K}$. (Since the existence

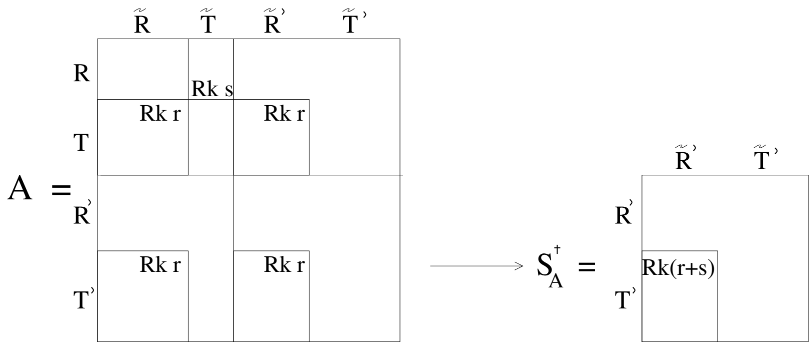

Figure 8: Given the matrix in the left hand side, satisfying the pure structure block $\mathcal{B}$, consisting of four parts, and having rank at most r. Suppose that this structure block is 'complemented' in $A_{1,1}$ by a block of rank at most s. Then the generalized Schur complement $S_A^{\dagger} = A_{2,2} - A_{2,1}A_{1,1}^{\dagger}A_{1,2}$ inherits the structure block $\mathcal{B}$, with rank bounded by $r+s$.

of such a dependency would allow us to apply a reduction as described.) Thus we can assume that all row and column dependencies of $A_{1,1}$ must be strictly inside the index sets R, $\tilde{T}$, respectively.

To prove this remaining case, we apply unitary row and column operations U, V to the index sets R, $\tilde{T}$, respectively, chosen to transform all linear dependencies inside these index sets into zero rows on top and zero columns on the right of $A_{1,1}$. Thus after this reduction, the matrix $A_{1,1}$ takes the form

$$A_{1,1} = \begin{bmatrix} 0 & 0_{a\times b} \\ X & 0 \end{bmatrix}, \tag{40}$$

with X nonsingular, and for suitable $a, b \in \mathbb{N}$. The Moore-Penrose inverse of this matrix is given by

$$A_{1,1}^{\dagger} = \begin{bmatrix} 0 & X^{-1} \\ 0_{b\times a} & 0 \end{bmatrix}. \tag{41}$$

Observe now that in the expression $S_A^{\dagger} = A_{2,2} - A_{2,1}A_{1,1}^{\dagger}A_{1,2}$, the first a rows of $A_{1,2}$ and the last b columns of $A_{2,1}$ are cancelled out by virtue of the zero pattern in (41). It follows that the generalized Schur complement $S_A^{\dagger}$ equals the Schur complement of the smaller matrix $A(K_{\text{new}} \cup M, \tilde{K}_{\text{new}} \cup N)$, where $K_{\text{new}} \subsetneq K$ is obtained by removing the first a indices from K, and where $\tilde{K}_{\text{new}} \subseteq \tilde{K}$ is obtained by removing the last b indices from $\tilde{K}$.

We conclude from the previous paragraph that it is sufficient to prove the theorem for the case where $A_{1,1}$ is nonsingular. Let us prove this remaining case. From the nonsingularity assumption, we can invoke Theorem 24 to obtain that $S_{A,k}(T', \tilde{R}') = \text{Rk } \tilde{r}$, where $\tilde{r} = r + |R| - |\tilde{R}|$. It will then be sufficient to prove that $|R| - |\tilde{R}| \leq s$. To this end, we recall the assumption $A_{1,1}(R, \tilde{T}) = \text{Rk } s$ to obtain the bound $A_{1,1} = \mathrm{Rk}\ \tilde{s}$ with $\tilde{s} := s + |\tilde{R}| + |T|$. Since by assumption $A_{1,1}$ is nonsingular, it follows that $s + |\tilde{R}| + |T| \geq k$, or equivalently $s + |\tilde{R}| \geq |R|$, which was to be demonstrated. □

3.3 Pure rank structures: reformulation of the results

In this subsection we reconsider the results on the generalized Schur complementation of pure rank structures, and tranlate them to a more general form. These results will then be used in the next subsection to derive results for *displacement* structures as well.

For brevity, we will explain the main ideas of this subsection only for the case of (usual) Schur complementation. The case of *generalized* Schur complementation will then be briefly stated without derivation, further in this subsection.

Theorem 26 *Suppose that*

$$\mathrm{Rank}(\begin{bmatrix} S_{1,1} & 0 \\ S_{2,1} & S_{2,2} \end{bmatrix} A \begin{bmatrix} T_{1,1} & T_{1,2} \\ 0 & T_{2,2} \end{bmatrix}) = r, \tag{42}$$

where $S_{1,1}$ is of size s_1 by k, $s_1 \leq k$, having full rank s_1, and where $T_{1,1}$ is of size k by t_1, $t_1 \leq k$, having full rank t_1. Then for the Schur complement $S_{A,k}$ of A we have that

$$\mathrm{Rank}(S_{2,2} S_{A,k} T_{2,2}) = \tilde{r}, \tag{43}$$

with $\tilde{r} := r + (k - s_1 - t_1)$.

PROOF. Consider the matrix factorization

$$\begin{bmatrix} S_{1,1} & 0 \\ S_{2,1} & S_{2,2} \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & I_{s_1} & 0 & 0 \\ \hline 0 & 0 & 0 & I_{s_2} \end{array}\right] \left[\begin{array}{c|c} K & 0 \\ S_{1,1} & 0 \\ \hline X & X \\ S_{2,1} & S_{2,2} \end{array}\right], \tag{44}$$

where $K \in \mathbb{C}^{(k-s_1)\times s_1}$ can be chosen such that $\begin{bmatrix} K \\ S_{1,1} \end{bmatrix}$ is a square nonsingular matrix, where each I denotes the identity matrix of the indicated size, and where the value of the blocks marked with X is irrelevant. We can write down a decomposition of exactly the same form for the matrix T^H. Let us then denote the rightmost matrix in (44) by L_S, and the similar factor in the decomposition of T^H by L_T. We can then reformulate (42) as

$$\mathrm{Rank}(\begin{bmatrix} 0 & I_{s_1} & 0 & 0 \\ 0 & 0 & 0 & I_{s_2} \end{bmatrix} L_S A L_T^H \begin{bmatrix} 0 & 0 \\ I_{t_1} & 0 \\ 0 & 0 \\ 0 & I_{t_2} \end{bmatrix}) = r.$$

But this states now precisely that the matrix $L_S A L_T^H$ must satisfy a structure block in the sense of Definition 23 (no shift matrix Λ is involved here). It follows then from Theorem 24 that

$$\text{Rank}(\begin{bmatrix} 0 & I_{s_2} \end{bmatrix} S_{L_S A L_T^H,k} \begin{bmatrix} 0 \\ I_{t_2} \end{bmatrix}) = \tilde{r},$$

where $\tilde{r} := r + (k - s_1 - t_1)$. By virtue of Lemma 22, we have now

$$S_{L_S A L_T^H,k} = L_S(2,2) S_{A,k} L_T^H(2,2),$$

and the desired result (43) follows then easily by using the definition of L_S and L_T. □

Remark 27 1. *One could devise also an alternative proof of Theorem 26. The idea is to complete $S_{1,1}$ and $T_{1,1}$ by extra rows and columns, respectively, so that they become square nonsingular. Then the proof follows immediately from the relation between Schur complements and Gaussian elimination, cf. Lemma 22 and Eq.* (37). *We omit the details.*

2. *Theorem 26 is a generalization of Theorem 15. Indeed, we have that the inverse matrix A^{-1} can be realized as the Schur complement of the block matrix $\begin{bmatrix} A & I \\ -I & 0 \end{bmatrix}$. Then if $S \in \mathbb{C}^{n \times s}, T \in \mathbb{C}^{n \times t}$ and if $S^\perp$, $T^\perp$ are maximal matrices such that $S^\perp S = 0$ and $T^\perp T = 0$, we have that*

$$\text{Rank}(\begin{bmatrix} T^\perp & 0 \\ 0 & S^\perp \end{bmatrix} \begin{bmatrix} A & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & T \end{bmatrix}) = \text{Rank}(T^\perp A S).$$

It follows then easily from Theorem 26 that Null $(S^\perp A^{-1} T) =$ Null $(T^\perp A S)$, *which is Theorem 15.*

Let us now briefly consider the case of *generalized* Schur complementation.

Theorem 28 *Given a generalized inverse X of $A_{1,1}$, with $A_{1,1}$ of size k by $\tilde{k}$, and a corresponding embedding of A as in* (39)*. We have the following bound:*

$$\text{Rank}(S_{2,2} S_A^X T_{2,2}) \leq$$
$$\text{Rank}(\left[\begin{array}{cc|c} S_{1,1} & 0 & 0 \\ 0 & I_a & 0 \\ \hline S_{2,1} & 0 & S_{2,2} \end{array}\right] \left[\begin{array}{cc|c} A_{1,1} & A_{1,1}(K, \tilde{U}) & A_{1,2} \\ A_{1,1}(U, \tilde{K}) & A_{1,1}(U, \tilde{U}) & 0 \\ \hline A_{2,1} & 0 & A_{2,2} \end{array}\right] \left[\begin{array}{cc|c} T_{1,1} & 0 & T_{1,2} \\ 0 & I_b & 0 \\ \hline 0 & 0 & T_{2,2} \end{array}\right])$$
$$+ k - s_1 - t_1 - \text{Null } A^H,$$

where I_a, I_b denote the identity matrix of size $a :=$ Null A and $b :=$ Null A^H, respectively, where S_A^X denotes the generalized Schur complement $A_{2,2} - A_{2,1} X A_{1,2}$, and where we used the same notations as in Theorem 26 (except that now $T_{1,1}$ is of size $\tilde{k}$ by t_1, $t_1 \leq \tilde{k}$).

PROOF. This follows from Theorem 26. $\square$

Finally, we state the result for generalized Schur complementation formed by means of the Moore-Penrose inverse.

Theorem 29 *We have the following bound:*

$$\mathrm{Rank}(S_{2,2}S_A^{\dagger}T_{2,2}) \leq \mathrm{Rank}(\begin{bmatrix} S_{1,1} & 0 \\ S_{2,1} & S_{2,2} \end{bmatrix} A \begin{bmatrix} T_{1,1} & T_{1,2} \\ 0 & T_{2,2} \end{bmatrix}) + \mathrm{Rank}(T_{1,1}^{\perp} A_{1,1}^{T} S_{1,1}^{\perp}),$$

where $S_A^{\dagger}$ denotes the generalized Schur complement $A_{2,2} - A_{2,1}A_{1,1}^{\dagger}A_{1,2}$, and where we used the same notations as in Theorem 26 (except that now $T_{1,1}$ is of size $\tilde{k}$ by t_1, $t_1 \leq \tilde{k}$).

PROOF. This follows from Theorem 25. $\square$

3.4 Displacement structures

In the present subsection we will apply the results of the previous subsection to obtain results for the generalized Schur complementation of displacement structured matrices. For example, the Stein displacement $A - GBH$ can again be realized as

$$\begin{bmatrix} I & G \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I \\ -H \end{bmatrix},$$

or equivalently

$$\begin{bmatrix} I & 0 & G_{1,1} & 0 \\ 0 & I & G_{2,1} & G_{2,2} \end{bmatrix} \begin{bmatrix} A_{1,1} & A_{1,2} & 0 & 0 \\ A_{2,1} & A_{2,2} & 0 & 0 \\ 0 & 0 & B_{1,1} & B_{1,2} \\ 0 & 0 & B_{2,1} & B_{2,2} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \\ -H_{1,1} & -H_{1,2} \\ 0 & -H_{2,2} \end{bmatrix}.$$

Note that we assumed here that G and H are block lower and upper triangular matrices, respectively, with position of the zeros compatible with the dimensions of A and B. Although this assumption was not needed in the case of matrix inversion, it will be required in order to have structure preservation in the context of *Schur complements*.

Now we apply permutations to obtain

$$\begin{bmatrix} I & G_{1,1} & 0 & 0 \\ 0 & G_{2,1} & I & G_{2,2} \end{bmatrix} \left[\begin{array}{cc|cc} A_{1,1} & 0 & A_{1,2} & 0 \\ 0 & B_{1,1} & 0 & B_{1,2} \\ \hline A_{2,1} & 0 & A_{2,2} & 0 \\ 0 & B_{2,1} & 0 & B_{2,2} \end{array}\right] \begin{bmatrix} I & 0 \\ -H_{1,1} & -H_{1,2} \\ 0 & I \\ 0 & -H_{2,2} \end{bmatrix}.$$

Now provided both matrices $A_{1,1}$ and $B_{1,1}$ are square nonsingular of size k by k, it follows that the Schur complement of the middle matrix in the above equation equals precisely the block matrix $S_{A,k} \oplus S_{B,k}$. It follows then from Theorem 26 that

$$\begin{bmatrix} I & G_{2,2} \end{bmatrix} \begin{bmatrix} S_{A,k} & 0 \\ 0 & S_{B,k} \end{bmatrix} \begin{bmatrix} I \\ -H_{2,2} \end{bmatrix} = \text{Rk } \tilde{r},$$

or equivalently

$$S_{A,k} - G_{2,2} S_{B,k} H_{2,2} = \text{Rk } \tilde{r},$$

where $\tilde{r} := r + (k - s_1 - t_1)$, with $r := \text{Rank}(A - GBH)$, where s_1 denotes the number of rows of $G_{1,1}$ and t_1 denotes the number of columns of $H_{1,1}$. This retrieves the result for the (usual) Schur complementation of displacement structures of Stein type. Moreover, note that the above proof of structure preservation does not need any technical completion lemmas, as is usually done in the literature [9, 4].

One could now do similar things for the case of *generalized* Schur complementation as well. For brevity, we will restrict ourselves here to the generalized Schur complement formed by means of the Moore-Penrose inverses $A^\dagger, B^\dagger$. In this case, we can apply Theorem 29 to obtain that

$$\text{Rank}(S^\dagger_{A,k} - G_{2,2} S^\dagger_{B,k} H_{2,2}) \leq \text{Rank}(A - GBH) + \text{Rank}(H_{1,1} A_{1,1} G_{1,1} - B_{1,1}),$$

where $S^\dagger_{A,k}$ denotes the generalized Schur complement $A_{2,2} - A_{2,1} A^\dagger_{1,1} A_{1,2}$, and similarly for $S^\dagger_{B,k}$.

It is clear that similar results can be obtained for displacement structures of *Sylvester* type as well, but we will not pursue this anymore.

4 Moore-Penrose inversion of a full column rank matrix

In this final section we return to the problem of generalized inversion, more precisely we consider the problem of Moore-Penrose inversion of a *full column rank* matrix $A \in \mathbb{C}^{m \times n}$, with $m \geq n$. The Moore-Penrose inverse is then given by the formula $A^\dagger = (A^H A)^{-1} A^H$. One can realize this expression as the Schur complement of the block matrix

$$\left[\begin{array}{cc|c} 0 & A^H & A^H \\ A & I & 0 \\ \hline I & 0 & 0 \end{array}\right], \tag{45}$$

with respect to the $(m+n)$-partitioning of this matrix, as indicated by the horizontal and vertical line. This approach allows then to derive the results for Moore-Penrose inversion from those for Schur complementation.

Let us do this here for the case of rank structures. We recall Theorem 13, where we had a rank-r structure block complemented by a rank-s block, leading to $A^\dagger$ having a structure block of rank at most $r + s$. In fact, one can obtain here an alternative derivation of this result by constructing a huge structure block of rank at most $r + s$ for the block matrix (45). See Figure 9.

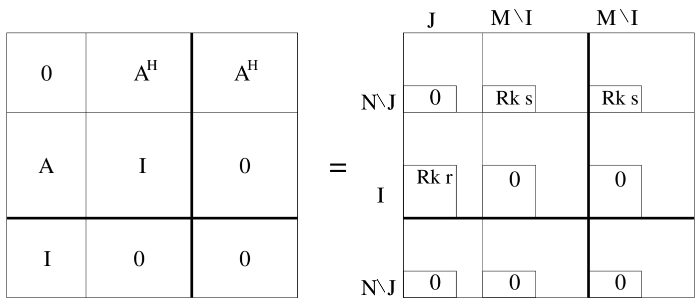

Figure 9: Under the assumptions of Theorem 13, one can construct a structure block of rank $r+s$ for the matrix (45), as indicated in the right part of the figure. The Schur complement induced by the $|M \cup N|$-partitioning of this matrix (indicated by the highlighted horizontal and vertical line in the figure) inherits then this structure block with rank at most $r+s+|J|+|M\backslash I|-|J|-|M\backslash I| = r+s$. This retrieves the result of Theorem 13.

We note that similar results have been obtained for displacement structures $AF-GA = \mathrm{Rk}\ r$ or $A-GAH = \mathrm{Rk}\ r$ with A having full column rank, provided G is either unitary or Hermitian.

Let us briefly review these results here. For example, let us focus on the displacement structure of Sylvester type $AF - GA = \mathrm{Rk}\ r =: UV^H$, with U, V both having r columns. It was essentially observed in [11] that when G is Hermitian, one can write down

$$\left[\begin{array}{cc|c} 0 & A^H & A^H \\ A & I & 0 \\ \hline I & 0 & 0 \end{array}\right] \left[\begin{array}{ccc} F & 0 & 0 \\ 0 & G & 0 \\ 0 & 0 & G \end{array}\right] - \left[\begin{array}{ccc} F^H & 0 & 0 \\ 0 & G & 0 \\ 0 & 0 & F \end{array}\right] \left[\begin{array}{cc|c} 0 & A^H & A^H \\ A & I & 0 \\ \hline I & 0 & 0 \end{array}\right]$$
$$= \left[\begin{array}{cc|c} 0 & -VU^H & -VU^H \\ UV^H & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right],$$

which is a displacement equation of Sylvester type for the matrix (45), having displacement rank at most $2r$. Hence the Moore-Penrose inverse $A^\dagger$, which is the Schur complement of the matrix (45), must inherit this displacement structure, namely $A^\dagger G - FA^\dagger = \mathrm{Rk}(2r)$.

Similarly, when G is unitary we have

$$\left[\begin{array}{cc|c} 0 & A^H & A^H \\ A & I & 0 \\ \hline I & 0 & 0 \end{array}\right] \left[\begin{array}{ccc} F & 0 & 0 \\ 0 & G & 0 \\ 0 & 0 & G \end{array}\right] - \left[\begin{array}{ccc} F^{-H} & 0 & 0 \\ 0 & G & 0 \\ 0 & 0 & F \end{array}\right] \left[\begin{array}{cc|c} 0 & A^H & A^H \\ A & I & 0 \\ \hline I & 0 & 0 \end{array}\right]$$
$$= \left[\begin{array}{cc|c} 0 & \tilde{V}\tilde{U}^H & \tilde{V}\tilde{U}^H \\ UV^H & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right],$$

where $\tilde{U} := G^H U$ and $\tilde{V} := F^{-H} V$. Hence this is again of rank at most $2r$. Again, it follows then that the Moore-Penrose inverse $A^\dagger$, which is the Schur complement of the matrix (45), must be such that $A^\dagger G - F A^\dagger = \mathrm{Rk}(2r)$.

Moreover, the above approach suggests a constructive way for computing $A^\dagger$: once there are given the low rank generators for the displacement structure of the matrix (45), one can make use of the so-called *generalized Schur algorithm* [9] to compute the Moore-Penrose inverse $A^\dagger = (A^H A)^{-1} A^H$ in an efficient way. This was essentially what was done in [11] for the case of the displacement structures defining the classes of Toeplitz-like and Cauchy-like matrices.

References

[1] A. Ben-Israel and T. N. E. Greville. *Generalized inverses: theory and applications.* CMS Books in Mathematics. Springer, second edition, 2003.

[2] R. Bevilacqua, E. Bozzo, G. M. Del Corso, and D. Fasino. Rank structure of generalized inverses of rectangular banded matrices. *Calcolo*, 42(3-4):157–169, 2005.

[3] S. Delvaux and M. Van Barel. Structures preserved by matrix inversion. *SIAM Journal on Matrix Analysis and its Applications*, 28(1):213–228, 2006.

[4] S. Delvaux and M. Van Barel. Structures preserved by Schur complementation. *SIAM Journal on Matrix Analysis and its Applications*, 28(1):229–252, 2006.

[5] M. Fiedler and T. L. Markham. Completing a matrix when certain entries of its inverse are specified. *Linear Algebra and its Applications*, 74:225–237, 1986.

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

[7] G. Heinig and F. Hellinger. Displacement structure of generalized inverse matrices. *Linear Algebra and its Applications*, 211:67–83, 1994.

[8] G. Heinig and F. Hellinger. Displacement structure of pseudoinverses. *Linear Algebra and its Applications*, 197-198:623–649, 1994.

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

[10] R. Penrose. A generalized inverse for matrices. *Proc. Cambridge Phil. Soc.*, 51:406–413, 1955.

[11] G. Rodriguez. Fast solution of Toeplitz- and Cauchy-like least squares problems. *SIAM Journal on Matrix Analysis and its Applications*, 28(3):724–748, 2006.