Principal pivot transforms of quasidefinite matrices and semidefinite Lagrangian subspaces[*]

Federico Poloni[†] *and Nataša Strabić*[‡]

Abstract

Lagrangian subspaces are linear subspaces that appear naturally in control theory applications, and especially in the context of algebraic Riccati equations. We introduce a class of *semidefinite* Lagrangian subspaces and show that these subspaces can be represented by a subset $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ and a Hermitian matrix $X \in \mathbb{C}^{n \times n}$ with the property that the submatrix $X_{\mathcal{II}}$ is negative semidefinite and the submatrix $X_{\mathcal{I}^c \mathcal{I}^c}$ is positive semidefinite. A matrix X with these definiteness properties is called $\mathcal{I}$-semidefinite and it is a generalization of a quasidefinite matrix. Under mild hypotheses which hold true in most applications, the Lagrangian subspace associated to the stabilizing solution of an algebraic Riccati equation is semidefinite, and in addition we show that there is a bijection between Hamiltonian and symplectic pencils and semidefinite Lagrangian subspaces; hence this structure is ubiquitous in control theory. The (symmetric) principal pivot transform (PPT) is a map used by Mehrmann and Poloni [*SIAM J. Matrix Anal. Appl.*, 33(2012), pp. 780–805] to convert between two different pairs $(\mathcal{I}, X)$ and $(\mathcal{J}, X')$ representing the same Lagrangian subspace. For a semidefinite Lagrangian subspace, we prove that the symmetric PPT of an $\mathcal{I}$-semidefinite matrix X is a $\mathcal{J}$-semidefinite matrix X', and we derive an implementation of the transformation $X \mapsto X'$ that both makes use of the definiteness properties of X and guarantees the definiteness of the submatrices of X' in finite arithmetic. We use the resulting formulas to obtain a semidefiniteness-preserving version of an optimization algorithm introduced by Mehrmann and Poloni to compute a pair $(\mathcal{I}_{\mathrm{opt}}, X_{\mathrm{opt}})$ with $M = \max_{i,j} |(X_{\mathrm{opt}})_{ij}|$ as small as possible. Using semidefiniteness allows one to obtain a stronger inequality on M with respect to the general case.

Key words. Lagrangian subspace, symplectic pencil, Hamiltonian pencil, principal pivot transform, quasidefinite matrix, Riccati matrix, graph matrix

AMS Subject Classifications. 65F30, 15A23, 15B99, 15A09

1 Introduction and preliminaries

The *symmetric principal pivot transform* of a matrix $X \in \mathbb{C}^{n \times n}$ with respect to an index set $\mathcal{K} \subseteq \{1, 2, \ldots, n\}$ is defined as the matrix Y such that

$$
\begin{aligned}
Y_{\mathcal{KK}} &= -X_{\mathcal{KK}}^{-1}, & Y_{\mathcal{KK}^c} &= X_{\mathcal{KK}}^{-1} X_{\mathcal{KK}^c}, \\
Y_{\mathcal{K}^c \mathcal{K}} &= X_{\mathcal{K}^c \mathcal{K}} X_{\mathcal{KK}}^{-1}, & Y_{\mathcal{K}^c \mathcal{K}^c} &= X_{\mathcal{K}^c \mathcal{K}^c} - X_{\mathcal{K}^c \mathcal{K}} X_{\mathcal{KK}}^{-1} X_{\mathcal{KK}^c},
\end{aligned}
\tag{1}
$$

---

where we denote by $X_{\mathcal{I}\mathcal{J}}$ a submatrix of X with rows and columns indexed by the sets $\mathcal{I}$ and $\mathcal{J}$, respectively (the order of the indices does not matter as long as it is chosen consistently), and $\mathcal{K}^c$ denotes the complement of $\mathcal{K}$ in $\{1, 2, \ldots, n\}$.

For instance, if $\mathcal{K} = \{1, 2, \ldots, k\}$ is the set of indices corresponding to the leading block of X, then

$$
X = \begin{array}{c} k \\ n-k \end{array}\!\! \begin{array}{cc} k & n-k \\ \left[\begin{array}{cc} X_{11} & X_{12} \\ X_{21} & X_{22} \end{array} \right], \end{array} \quad Y = \begin{array}{c} k \\ n-k \end{array}\!\! \begin{array}{cc} k & n-k \\ \left[\begin{array}{cc} -X_{11}^{-1} & X_{11}^{-1}X_{12} \\ X_{21}X_{11}^{-1} & X_{22} - X_{21}X_{11}^{-1}X_{12}. \end{array} \right]. \end{array}
$$

Note the peculiar structure of this transformation: we invert a principal submatrix of X and we perform a Schur complementation on its complement.

The map $X \mapsto Y$ defined in (1) is a symmetric variant of the *principal pivot transform* (PPT), which appears across various fields under different names. In statistics it is known as the *sweep operator* when it is used to solve least-squares regression problems [9], or as *partial inversion* in the context of linear graphical chain models [25]. Duffin, Hazony and Morrison analyze network synthesis [7] and call it *gyration*. In numerical linear algebra the PPT is often called the *exchange operator* and it is of interest since it relates computations in one structured class of matrices to another. Stewart and Stewart [22] use the exchange operator to generate J-orthogonal matrices (matrices $Q \in \mathbb{R}^{n \times n}$ such that $Q^T J Q = J$, where $J = \operatorname{diag}(\pm 1)$ is a signature matrix) from hyperbolic Householder transformations. Higham [11] further shows how to obtain a hyperbolic CS decomposition of a J-orthogonal matrix directly from the standard CS decomposition via the exchange operator. Moreover, certain important classes of matrices are invariant under this operation. Tucker [24] shows that the principal pivot transform of a P-matrix (a matrix whose principal minors are all positive) is again a P-matrix when the matrix is real. This result was extended to complex P-matrices by Tsatsomeros in [23], where further details of the history and properties of the PPT can be found. An overview by Higham [11, sec. 2] provides additional references.

Mehrmann and Poloni [18] use the symmetric PPT (1) in the context of Lagrangian subspaces, which are an essential structure in control theory applications (e.g. [1, 8, 14, 17]), to obtain their *permuted Lagrangian graph representation*, and show that it both preserves the Lagrangian structure in computations and is numerically stable to work with. In this representation a Lagrangian subspace is identified with the pair $(\mathcal{I}, X)$, where $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ and $X \in \mathbb{C}^{n \times n}$ is Hermitian. The symmetric PPT (1) is used to convert between two different representations in an optimization algorithm [18, Alg. 2] which computes a subset $\mathcal{I}_{\mathrm{opt}}$ and an associated X_{opt} whose elements are bounded by a small constant. Using this matrix X_{opt} improves numerical stability in several contexts, see [20].

In this paper we focus on a class of Lagrangian subspaces whose representation $(\mathcal{I}, X)$ has additional structure. Let the symbol $\succ$ denote the Löwner ordering: $A \succ B$ ($A \succeq B$) means that $A - B$ is positive (semi)definite. We say that a Hermitian matrix $X = X^* \in \mathbb{C}^{n \times n}$ is $\mathcal{I}$-*definite*, for $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$, if

$$
X_{\mathcal{I}\mathcal{I}} \prec 0 \quad \text{and} \quad X_{\mathcal{I}^c\mathcal{I}^c} \succ 0. \tag{2}
$$

If the previous definition holds with the symbols $\succ, \prec$ replaced by $\succeq, \preceq$ then X is $\mathcal{I}$-*semidefinite*. For $\mathcal{I} = \emptyset$ an $\mathcal{I}$-definite matrix is simply a positive definite matrix and for $\mathcal{I} = \{1, 2, \ldots, n\}$ an $\mathcal{I}$-definite matrix is negative definite. In all other cases, an $\mathcal{I}$-definite matrix is a generalization of a *quasidefinite* matrix, which is $\mathcal{I}$-definite for $\mathcal{I} = \{1, 2, \ldots, k\}$ with some $k < n$.

Identifying this class of subspaces and exploiting its properties in applications has several advantages: one can improve a bound on the elements of the matrix X_{opt} and preserve this additional structure, which is, for instance, crucial for the existence of a positive semidefinite solution X of an algebraic Riccati equation.

The rest of the paper is structured as follows. In Section 2 we present the basic definitions and concepts related to Lagrangian subspaces and their representations $(\mathcal{I}, X)$. We introduce a class of Lagrangian *(semi)definite* subspaces in Section 3 and prove that for these subspaces the Hermitian matrix X in the pair $(\mathcal{I}, X)$ which represents the Lagrangian semidefinite subspace is $\mathcal{I}$-semidefinite for all possible choices of $\mathcal{I}$. In Section 4 we link Lagrangian semidefinite subspaces to Hamiltonian and symplectic pencils appearing in control theory. In Section 5 we derive an implementation of the symmetric PPT (1) which converts between two different representations $(\mathcal{I}, X)$ and $(\mathcal{J}, X')$ of a Lagrangian semidefinite subspace. Specifically, we show how an $\mathcal{I}$-semidefinite matrix X can be converted to a $\mathcal{J}$-semidefinite matrix X' for a given index set $\mathcal{J}$ by the PPT that both makes use of the definiteness properties of X and guarantees the definiteness of the blocks of X' in finite arithmetic. The symmetric PPT in one case requires the computation of the inverse of a quasidefinite matrix with factored diagonal blocks and we also present an inversion formula which uses unitary factorizations to directly compute the factors of the diagonal blocks of the quasidefinite inverse. In Section 6 we prove that all elements of an $\mathcal{I}_{\text{opt}}$-semidefinite matrix X_{opt} associated with a semidefinite Lagrangian subspace are bounded by 1 in modulus, and present the optimization algorithm which computes an optimal representation. We test the performance of the algorithm on several numerical experiments in Section 7 and present some concluding remarks in Section 8.

2 PPTs and Lagrangian subspaces

An n-dimensional subspace $\mathcal{U}$ of $\mathbb{C}^{2n}$ is called *Lagrangian* if $u^* J_n v = 0$ for every $u, v \in \mathcal{U}$, where u^* denotes the conjugate transpose of a vector u,

$$J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix},$$

and I_n is the identity matrix. A full column rank matrix $U \in \mathbb{C}^{2n \times n}$ is a basis for a Lagrangian subspace if and only if $U^* J_n U = 0$. For $U, V \in \mathbb{C}^{2n \times n}$ of full column rank we write $U \sim V$ if $U = VM$ for a square invertible $M \in \mathbb{C}^{n \times n}$. Note that this implies that U and V have the same column space, i.e. $\text{Im}(U) = \text{Im}(V)$.

In computational practice, a subspace $\mathcal{U}$ is typically represented through a matrix U whose columns span it. A key quantity is its *condition number* $\kappa(U) = \sigma_{\max}(U)/\sigma_{\min}(U)$: the sensitivity of $\mathcal{U} = \text{Im}(U)$ as a function of U depends on it [21, p. 154], as well as the numerical stability properties of several linear algebra operations associated to it, for instance, QR factorizations and least-squares problems [10, Chap. 19 and 20]. Hence, in most applications the natural choice for a basis is a matrix U with orthonormal columns, which ensures $\kappa(U) = 1$. However, if a matrix U is partitioned as

$$U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \in \mathbb{C}^{2n \times n}, \ U_1, U_2 \in \mathbb{C}^{n \times n},$$

then it spans a Lagrangian subspace if and only if $U_1^* U_2 = U_2^* U_1$, which is a property very difficult to preserve in finite arithmetic. If the matrix U_1 is invertible, one can write

$$U = \begin{bmatrix} I_n \\ X \end{bmatrix} U_1, \quad X = U_2 U_1^{-1}, \tag{3}$$

and hence obtain a different matrix $V = \begin{bmatrix} I_n \\ X \end{bmatrix}$ whose columns span the same subspace.

Matrices of the form

$$\mathcal{G}(X) = \begin{bmatrix} I_n \\ X \end{bmatrix}, \quad X \in \mathbb{C}^{n \times n} \tag{4}$$

are called *graph matrices*, since their form resembles the definition of the graph of a function as the set of pairs $(x, f(x))$, or *Riccati matrices*, since they are related to the algebraic Riccati equations [14]. Namely, the matrix X is a solution to a continuous-time algebraic Riccati equation $Q + XA + A^*X - XGX = 0$, where $A, G, Q, X \in \mathbb{C}^{k \times k}$ and $G = G^*$, $Q = Q^*$, if and only if

$$H \begin{bmatrix} I_k \\ X \end{bmatrix} = \begin{bmatrix} I_k \\ X \end{bmatrix} (A - GX), \tag{5}$$

where the associated matrix H is Hamiltonian and given by

$$H = \begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix}.$$

Unlike previous publications, we prefer to use the name Riccati matrix here for $\mathcal{G}(X)$, since it is less likely to induce confusion with graphs as the mathematical objects with nodes and edges. From (3), since U_1 is nonsingular it follows that

$$U \sim \mathcal{G}(U_2 U_1^{-1}) \tag{6}$$

and it is easy to see from the definition that $\operatorname{Im} \mathcal{G}(X)$ is Lagrangian if and only if $X = X^*$, a condition which is trivial to ensure in numerical computation. Hence, if the object of interest is the Lagrangian subspace $\operatorname{Im} U$, one can associate it with the Hermitian matrix X and use only this matrix to store and work on the subspace. The potential difficulties with this approach come from computing $X = U_2 U_1^{-1}$ since U_1 could be ill-conditioned or even singular.

Mehrmann and Poloni [18] consider a slightly more general form instead. For each subset $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$, the *symplectic swap matrix* associated with $\mathcal{I}$ is defined as

$$\Pi_{\mathcal{I}} = \begin{bmatrix} I_n - D & D \\ -D & I_n - D \end{bmatrix} \in \mathbb{R}^{2n \times 2n}, \tag{7}$$

where D is the diagonal matrix such that

$$D_{ii} = \begin{cases} 1, & i \in \mathcal{I}, \\ 0, & i \notin \mathcal{I}. \end{cases}$$

The matrices $\Pi_{\mathcal{I}}$ are symplectic ($\Pi_{\mathcal{I}}^T J_n \Pi_{\mathcal{I}} = J_n$) and orthogonal ($\Pi_{\mathcal{I}}^T \Pi_{\mathcal{I}} = I_{2n}$), and the multiplication with $\Pi_{\mathcal{I}}$ permutes (up to a sign change) the elements of a $2n$-length vector, with the limitation that the ith entry can only be exchanged with the $(n+i)$th, for each $i = 1, 2, \ldots, n$.

Example 1. When $n = 2$, the four symplectic swap matrices are

$$\Pi_{\emptyset} = I_4, \qquad \Pi_{\{1\}} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \Pi_{\{2\}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \qquad \Pi_{\{1,2\}} = J_2.$$

Given a full column rank matrix $U \in \mathbb{C}^{2n \times n}$ such that $\operatorname{Im} U$ is Lagrangian and a set $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$, define the symplectic swap $\Pi_{\mathcal{I}}$ as in (7) and partition

$$\Pi_{\mathcal{I}} U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \quad U_1, U_2 \in \mathbb{C}^{n \times n}. \tag{8}$$

If the top $n \times n$ block U_1 is invertible then

$$U \sim \mathcal{G}_{\mathcal{I}}(U_2 U_1^{-1}), \tag{9}$$

where

$$\mathcal{G}_{\mathcal{I}}(X) = \Pi_{\mathcal{I}}^T \begin{bmatrix} I_n \\ X \end{bmatrix}, \quad X \in \mathbb{C}^{n \times n}. \tag{10}$$

Note that (10) generalizes the notion of a Riccati matrix (4) by not requiring that the identity matrix is contained in the top block but that it can be pieced together (modulo signs) from a subset of rows of the matrix $\mathcal{G}_{\mathcal{I}}(X)$. Clearly, the pair $(\mathcal{I}, X)$, with $X = U_2 U_1^{-1}$, identifies $\operatorname{Im} U$ uniquely.

The representation (9) is called the *permuted Lagrangian graph representation* in [18] and it generalizes the representation (6), while keeping the property that $\operatorname{Im} \mathcal{G}_{\mathcal{I}}(X)$ is Lagrangian if and only if X is Hermitian. We use the name *permuted Riccati representation* (or *basis*) here.

Theorem 2.1 ([18, sec. 3]). *Let $U \in \mathbb{C}^{2n \times n}$. The following properties are equivalent.*

1. *$\operatorname{Im} U$ is Lagrangian.*

2. *For a particular choice of $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ we have $U \sim \mathcal{G}_{\mathcal{I}}(X)$ and it holds that $X = X^*$.*

3. *For all choices of $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ such that $U \sim \mathcal{G}_{\mathcal{I}}(X)$, it holds that $X = X^*$.*

Moreover, for each U satisfying the above properties there exists at least one $\mathcal{I}_{\text{opt}} \subseteq \{1, 2, \ldots, n\}$ such that $U \sim \mathcal{G}_{\mathcal{I}_{\text{opt}}}(X_{\text{opt}})$ and $X_{\text{opt}} = X_{\text{opt}}^*$ satisfies

$$|(X_{\text{opt}})_{ij}| \leq \begin{cases} 1, & \text{if } i = j, \\ \sqrt{2}, & \text{otherwise.} \end{cases} \tag{11}$$

As with the Riccati matrix representation, we can use any of the matrices X such that $U \sim \mathcal{G}_{\mathcal{I}}(X)$ to store the Lagrangian subspace $\operatorname{Im} U$ on a computer and operate on it, since the property that X must be Hermitian can be easily enforced. The choice with $\mathcal{I}_{\text{opt}}$ is particularly convenient from a numerical point of view: using (11), one can prove that $\kappa(\mathcal{G}_{\mathcal{I}}(X))$ cannot be too large [18, Thm. 8.2].

Example 2. For the matrix

$$U = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 5 & 3 \\ 8 & 5 \end{bmatrix},$$

whose column space $\operatorname{Im} U$ is Lagrangian we have

$$U \sim \mathcal{G}_{\emptyset} \left(\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \right), \qquad\qquad U \sim \mathcal{G}_{\{1\}} \left(\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \right),$$

$$U \sim \mathcal{G}_{\{2\}} \left(\begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix} \right), \qquad U \sim \mathcal{G}_{\{1,2\}} \left(\begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix} \right).$$

All the matrices X in $\mathcal{G}_{\mathcal{I}}(X)$ are Hermitian. For $\mathcal{I}_{\text{opt}} = \{2\}$, the condition (11) is satisfied.

Example 3. For the matrix

$$U = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 6 & 4 \\ 6 & 4 \end{bmatrix},$$

whose column space $\operatorname{Im} U$ is Lagrangian we have

$$U \sim \mathcal{G}_{\emptyset} \left(\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \right),$$

and for both

$$U \sim \mathcal{G}_{\{1\}} \left(\begin{bmatrix} -1/2 & 1 \\ 1 & 0 \end{bmatrix} \right) \quad \text{and} \quad U \sim \mathcal{G}_{\{2\}} \left(\begin{bmatrix} 0 & 1 \\ 1 & -1/2 \end{bmatrix} \right)$$

the condition (11) is satisfied. The top 2×2 block of $\Pi_{\{1,2\}} U$ is singular, hence the permuted Riccati representation (9) does not exist for $\mathcal{I} = \{1, 2\}$.

The following result shows how the symmetric PPT (1) converts between two different permuted Riccati representations, which is then used in the optimization algorithm [18, Alg. 2] to compute $(\mathcal{I}_{\mathrm{opt}}, X_{\mathrm{opt}})$.

Lemma 2.2 ([18, Lem. 5.1]). *Let $\mathcal{I}, \mathcal{J} \subseteq \{1, 2, \ldots, n\}$, and let $U \in \mathbb{C}^{2n \times n}$ be a matrix whose column space is Lagrangian and such that $U \sim \mathcal{G}_{\mathcal{I}}(X)$. Let $\mathcal{K}$ be the symmetric difference set*

$$\mathcal{K} = \{i \in \{1, 2, \ldots, n\} : i \text{ is contained in exactly one among } \mathcal{I} \text{ and } \mathcal{J}\}.$$

Then, $U \sim \mathcal{G}_{\mathcal{J}}(X')$ if and only if $X_{\mathcal{K}\mathcal{K}}$ is invertible, and in this case $X' = DYD$, where Y is the symmetric PPT of X defined in (1) for the index set $\mathcal{K}$, and D is the diagonal matrix such that

$$D_{ii} = \begin{cases} -1, & i \in \mathcal{I} \setminus \mathcal{J}, \\ 1, & \text{otherwise.} \end{cases}$$

Informally speaking, when we wish to transform the matrix X such that $U \sim \mathcal{G}_{\mathcal{I}}(X)$ into the matrix X' so that $U \sim \mathcal{G}_{\mathcal{J}}(X')$ for a new set $\mathcal{J}$, we have to perform a symmetric PPT (1) with respect to the indices that we wish to add to or remove from $\mathcal{I}$, and then flip the signs in the rows and columns with the indices that we remove from $\mathcal{I}$.

Example 4. Take $\mathcal{I} = \{1\}$ and the matrix U from Example 2 so that $U \sim \mathcal{G}_{\{1\}}(X)$ with

$$X = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}.$$

Applying Lemma 2.2 transforms between the remaining three representations as follows. For $\mathcal{J} = \emptyset$ Lemma 2.2 defines $\mathcal{K} = \{1\}$ and $D = \mathrm{diag}(-1, 1)$. Applying (1) to X gives

$$Y = \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix}, \quad X' = DYD = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

Therefore, $U \sim \mathcal{G}_{\emptyset}(X')$ holds. For $\mathcal{J} = \{2\}$ we have $\mathcal{K} = \{1, 2\}$ and $D = \mathrm{diag}(-1, 1)$. In this case

$$Y = - \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} -1/3 & -2/3 \\ -2/3 & -1/3 \end{bmatrix}, \quad X' = DYD = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix},$$

leading to the representation $U \sim \mathcal{G}_{\{2\}}(X')$. Finally, for $\mathcal{J} = \{1, 2\}$ we have $\mathcal{K} = \{2\}$ and $D = I_2$. It follows that $U \sim \mathcal{G}_{\{1,2\}}(X')$ for

$$X' = Y = \begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix}.$$

3   Semidefinite Lagrangian subspaces

The equation (5) shows that solving a continuous-time algebraic Riccati equation $Q + XA + A^*X - XGX = 0$ is equivalent to solving an invariant subspace problem for the associated Hamiltonian matrix H, if we impose that the subspace is represented via a Riccati basis $\mathcal{G}(X)$. If the matrices Q and G are positive semidefinite, under standard conditions (see, e.g. [5, 17, 19]) the Riccati equation has a unique positive semidefinite solution, and this is the solution that is usually of interest. A common approach to computing it is to determine a basis for the *stable invariant subspace* of H, i.e., the one corresponding to the eigenvalues of H in the open left half plane (e.g. [2, 15, 17]). This subspace is Lagrangian and if $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ is its basis then the matrix U_1 is invertible and $X = U_2 U_1^{-1}$ is the positive semidefinite solution to the Riccati equation [19]. Specifically, $U \sim \mathcal{G}(X)$ and the matrix $X = X^* \geq 0$.

In this section we take a closer look at Lagrangian subspaces which have a Riccati basis with this property. We call a Lagrangian subspace *definite* if it can be written as

$$\mathcal{U} = \operatorname{Im} \mathcal{G}(X), \quad X = X^* \succ 0, \quad X \in \mathbb{C}^{n \times n},$$

where $\mathcal{G}(X)$ is defined in (4). The following result relates $\mathcal{I}$-definite matrices defined by the property (2) and definite Lagrangian subspaces.

Theorem 3.1. *Let $U \in \mathbb{C}^{2n \times n}$ have full column rank. The following properties are equivalent.*

1. *$\mathcal{U} = \operatorname{Im} U$ is Lagrangian definite.*

2. *For some $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ we have $U \sim \mathcal{G}_{\mathcal{I}}(X)$, where X is $\mathcal{I}$-definite and $\mathcal{G}_{\mathcal{I}}(X)$ is defined in (10).*

3. *For all $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ we have $U \sim \mathcal{G}_{\mathcal{I}}(X)$ and X is $\mathcal{I}$-definite.*

Proof. Let $U \sim \mathcal{G}(X)$ and $X \succ 0$. From the definition of a symplectic swap matrix (7) it follows that $\Pi_\emptyset = I_{2n}$ and hence $\mathcal{G}_\emptyset(X) = \mathcal{G}(X)$. Therefore, the definition of a Lagrangian definite subspace can be reformulated as stating $U \sim \mathcal{G}_{\mathcal{I}}(X)$, where X is $\mathcal{I}$-definite for $\mathcal{I} = \emptyset$. If this holds, then for each $\mathcal{J} \subseteq \{1, 2, \ldots, n\}$ Lemma 2.2 defines $\mathcal{K} = \mathcal{J}$ and $D = I_n$. Since $X \succ 0$ every principal submatrix $X_{\mathcal{K}\mathcal{K}}$ is also positive definite and therefore $U \sim \mathcal{G}_{\mathcal{J}}(X')$ for every $\mathcal{J}$, where X' is the symmetric PPT (1) of X. It is clear from the formulas (1) and the properties of Schur complements [12, sec. 12.3] that X' is $\mathcal{J}$-definite, as required.

On the other hand, if $U \sim \mathcal{G}_{\mathcal{I}}(X)$ and X is $\mathcal{I}$-definite for some $\mathcal{I}$, then X is Hermitian by definition and hence U spans a Lagrangian subspace. We prove that the subspace is Lagrangian definite by applying Lemma 2.2 with $\mathcal{J} = \emptyset$ to X. It follows that $\mathcal{K} = \mathcal{I}$ and since $X_{\mathcal{K}\mathcal{K}} = X_{\mathcal{I}\mathcal{I}} \prec 0$, we have $U \sim \mathcal{G}_\emptyset(X') = \mathcal{G}(X')$, with $X' = DYD$ as in Lemma 2.2. Since X' is defined via congruence it is sufficient to prove that Y, the symmetric PPT of an $\mathcal{I}$-definite matrix with respect to the index set $\mathcal{I}$, is positive definite. This follows from (1) due to the definiteness properties of the blocks of X: both $Y_{\mathcal{K}\mathcal{K}} = -X_{\mathcal{I}\mathcal{I}}^{-1}$ and its Schur complement $X_{\mathcal{I}^c\mathcal{I}^c} - X_{\mathcal{I}^c\mathcal{I}}X_{\mathcal{I}\mathcal{I}}^{-1}X_{\mathcal{I}\mathcal{I}^c} = Y_{\mathcal{K}^c\mathcal{K}^c}$ are positive definite, so again by the properties of Schur complements Y is positive definite and the proof is complete. $\qquad\square$

More interesting is the corresponding semidefinite case, in which existence of the permuted Riccati representation is not guaranteed for all $\mathcal{I}$, cf. Example 3.

Theorem 3.2. *Let $U \in \mathbb{C}^{2n \times n}$ have full column rank. The following properties are equivalent.*

1. *For some $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ we have $U \sim \mathcal{G}_{\mathcal{I}}(X)$, where X is $\mathcal{I}$-semidefinite and $\mathcal{G}_{\mathcal{I}}(X)$ is defined in (10).*

2. *For all $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ such that the permuted Riccati representation exists, i.e., $U \sim \mathcal{G}_{\mathcal{I}}(X)$, the matrix X is $\mathcal{I}$-semidefinite.*

When these properties hold, we call the subspace *Lagrangian semidefinite*.

Proof. Let $\mathcal{I}$ be such that $U \sim \mathcal{G}_{\mathcal{I}}(X)$ and X is $\mathcal{I}$-semidefinite. Consider the matrix Y obtained by perturbing the diagonal entries of X so that the blocks $X_{\mathcal{I}\mathcal{I}}$ and $X_{\mathcal{I}^c\mathcal{I}^c}$ become strictly definite, that is, for some $\varepsilon > 0$,

$$
\begin{aligned}
Y_{\mathcal{I}\mathcal{I}} &= X_{\mathcal{I}\mathcal{I}} - \varepsilon I \prec 0, & Y_{\mathcal{I}\mathcal{I}^c} &= X_{\mathcal{I}\mathcal{I}^c}, \\
Y_{\mathcal{I}^c\mathcal{I}} &= X_{\mathcal{I}^c\mathcal{I}}, & Y_{\mathcal{I}^c\mathcal{I}^c} &= X_{\mathcal{I}^c\mathcal{I}^c} + \varepsilon I \succ 0.
\end{aligned}
$$

Then the subspace $\operatorname{Im} U_\varepsilon$, with $U_\varepsilon \sim \mathcal{G}_{\mathcal{I}}(Y)$, is Lagrangian definite, and by Theorem 3.1, the permuted Riccati representations $U_\varepsilon \sim \mathcal{G}_{\mathcal{I}}(Z)$ exist for every $\mathcal{I}$ with Z having the required definiteness properties. By passing to the limit $\varepsilon \to 0$, we get the semidefiniteness of the blocks of Z (whenever the representation exists). $\qquad\square$

Example 5. Consider the subspace in Example 3. We have $U \sim \mathcal{G}_{\emptyset}(X)$ for X positive semidefinite, so U is Lagrangian semidefinite. Other choices of the index set for which the permuted Riccati representations exist are $\mathcal{I} = \{1\}$ and $\mathcal{I} = \{2\}$ and the corresponding matrices X are $\{1\}$-semidefinite and $\{2\}$-semidefinite, respectively.

4 Semidefinite Lagrangian subspaces associated with control-theory pencils

Section 6 of [18] introduces a method to map matrix pencils with special structures to Lagrangian subspaces. The main reason why this kind of bijection is used is that changing a basis in the subspace is equivalent to premultiplying the pencil by a nonsingular matrix, which preserves eigenvalues and right eigenvectors of regular pencils. This makes it possible to apply several techniques based on PPTs to pencils as well. Specifically, we write

$$
M_1 - xN_1 \sim M_2 - xN_2,
$$

and say that the two pencils are *left equivalent*, if there exists a nonsingular square matrix S such that $M_1 = SM_2$ and $N_1 = SN_2$. It follows that $M_1 - xN_1 \sim M_2 - xN_2$ if and only if $\begin{bmatrix} M_1 & N_1 \end{bmatrix}^* \sim \begin{bmatrix} M_2 & N_2 \end{bmatrix}^*$. Hence, if we are interested in the eigenvalues and right eigenvectors of a regular pencil we may instead work with any regular pencil left equivalent to it.

We construct here a simple variation of the map from [18] which sends the pencils appearing in most applications in control theory to semidefinite Lagrangian subspaces.

A *Hamiltonian pencil* is a matrix pencil $M - xN \in \mathbb{C}^{2k \times 2k}[x]$ such that $MJ_kN^* + NJ_kM^* = 0$. In several problems in control theory, e.g. [16, 17], one deals with Hamiltonian pencils in the form

$$
\begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix} - xI_{2k}, \quad A, G, Q \in \mathbb{C}^{k \times k}, \quad G = G^* \succeq 0, \quad Q = Q^* \succeq 0; \tag{12}
$$

moreover, factorizations $G = BB^*$ and $Q = C^*C$ (with $B \in \mathbb{C}^{k \times t}$, $C \in \mathbb{C}^{r \times k}$, $r, t \leq k$) are known in advance. In the following theorem, we show that this kind of structure is mapped to a semidefinite Lagrangian subspace by a special bijection between pencils and $4k \times 2k$ matrices.

Theorem 4.1. *Let*

$$
M - xN = \begin{bmatrix} M_1 & M_2 \end{bmatrix} - x \begin{bmatrix} N_1 & N_2 \end{bmatrix}, \quad M_1, M_2, N_1, N_2 \in \mathbb{C}^{2k \times k}
$$

be a matrix pencil without common row kernel (i.e., there exists no vector $v \neq 0$ such that $v^* M = v^* N = 0$). Construct the matrix

$$U = \begin{bmatrix} M_1 & -N_1 & -N_2 & M_2 \end{bmatrix}^*. \tag{13}$$

Then,

1. $M - xN$ is Hamiltonian if and only if $\operatorname{Im} U$ is Lagrangian.

2. If $M - xN$ is in the form (12), then $\operatorname{Im} U$ is Lagrangian semidefinite.

Proof. The first claim is proved by expanding the relation $U^* J_{2k} U = 0$ into blocks. This leads to the expression $-M_1 N_2^* - N_1 M_2^* + N_2 M_1^* + M_2 N_1^* = 0$, which we can recombine to get $M J_k N^* + N J_k M^* = 0$.

For the second claim, take $M - xN$ as in (12), and $\mathcal{I} = \{1, 2, \ldots, k\}$. We have

$$\Pi_{\mathcal{I}} U = \begin{bmatrix} 0 & 0 & I_k & 0 \\ 0 & I_k & 0 & 0 \\ -I_k & 0 & 0 & 0 \\ 0 & 0 & 0 & I_k \end{bmatrix} \begin{bmatrix} A^* & -Q \\ -I_k & 0 \\ 0 & -I_k \\ -G & -A \end{bmatrix} = \begin{bmatrix} 0 & -I_k \\ -I_k & 0 \\ -A^* & Q \\ -G & -A \end{bmatrix} \sim \begin{bmatrix} I_k & 0 \\ 0 & I_k \\ -Q & A^* \\ A & G \end{bmatrix}. \tag{14}$$

Hence, $U \sim \Pi_{\mathcal{I}}^T \mathcal{G}(X) = \mathcal{G}_{\mathcal{I}}(X)$, with $X = \begin{bmatrix} -Q & A^* \\ A & G \end{bmatrix}$, which is $\mathcal{I}$-semidefinite. Thus, by Theorem 3.2, the subspace $\operatorname{Im} U$ is Lagrangian semidefinite. $\qquad\square$

Equation (6.4) in [18] gives a matrix U in a form similar to (13), which satisfies only the first part of the theorem.

Similarly, a *symplectic pencil* is a matrix pencil $M - xN \in \mathbb{C}^{2k \times 2k}[x]$ such that $M J_k M^* = N J_k N^*$. In several problems in discrete-time control theory, e.g. [8, 16, 17], one deals with symplectic pencils in the form

$$\begin{bmatrix} A & 0 \\ -Q & I_k \end{bmatrix} - x \begin{bmatrix} I_k & G \\ 0 & A^T \end{bmatrix}, \quad A, G, Q \in \mathbb{C}^{k \times k}, \quad G = G^* \succeq 0, \quad Q = Q^* \succeq 0; \tag{15}$$

again, factorizations $G = BB^*$, $Q = C^*C$ as above are often available. Similarly to the Hamiltonian case, there is a bijection which maps this structure into a semidefinite Lagrangian subspace.

Theorem 4.2. *Let*

$$M - xN = \begin{bmatrix} M_1 & M_2 \end{bmatrix} - x \begin{bmatrix} N_1 & N_2 \end{bmatrix}, \quad M_1, M_2, N_1, N_2 \in \mathbb{C}^{2k \times k}$$

be a matrix pencil without common row kernel (i.e., there exists no vector $v \neq 0$ such that $v^* M = v^* N = 0$). Construct the matrix

$$U = \begin{bmatrix} M_1 & -N_1 & -M_2 & -N_2 \end{bmatrix}^*. \tag{16}$$

Then,

1. $M - xN$ is symplectic if and only if $\operatorname{Im} U$ is Lagrangian.

2. If $M - xN$ is in the form (15), then $\operatorname{Im} U$ is Lagrangian semidefinite.

Proof. The proof of both claims is analogous to the proof of Theorem 4.1. Specifically, the Lagrangian semidefinite subspace U from (16) is also associated to the quasidefinite matrix $X = \begin{bmatrix} -Q & A^* \\ A & G \end{bmatrix}$. $\qquad\square$

Once again, a construction given in Equation (6.2) in [18] provides an analogous bijection that satisfies only the first part of the theorem. The main use for these bijections is producing left-equivalent pencils with better numerical properties. We show it in a simple case.

Example 6. Consider $k = 1$, $A = 1$, $G = 10^5$, $Q = 0.1$. The Hamiltonian pencil $M - xN$ obtained as in (12) has the condition number $\kappa(\left[\begin{smallmatrix} M \\ N \end{smallmatrix}\right]) \approx 10^5$, that is, a perturbation of relative magnitude 10^{-5} can turn it into a singular pencil (one in which M, N have a common kernel). If we construct the matrix U in (13) associated with it and apply Algorithm 1 described in Section 6 to obtain an equivalent permuted Riccati representation of U with smaller entries, we obtain a left-equivalent pencil

$$\hat{M} - x\hat{N} = \begin{bmatrix} 0.1 & 0 \\ -10^{-5} & 1 \end{bmatrix} - x \begin{bmatrix} 10^{-5} & -1 \\ -10^{-5} & 0 \end{bmatrix},$$

with $\kappa(\left[\begin{smallmatrix} \hat{M} \\ \hat{N} \end{smallmatrix}\right]) \approx 14$, a considerably lower value. The two pencils are Hamiltonian and have the same eigenvalues and right eigenvectors, so they are completely equivalent from a numerical perspective.

The optimization algorithm [18, Alg. 2] uses the PPT formulas (1) to compute the optimal permuted Riccati representation of a Lagrangian subspace and it can be used to normalize pencils [18, sec. 6]. If a PPT is applied to a Lagrangian semidefinite subspace $\mathrm{Im}\, U$, where U is for example given in (13) or (16), the definiteness properties of the blocks G and Q are not exploited. Furthermore, due to Theorem 3.2, for the computed optimal representation $(\mathcal{I}_{\mathrm{opt}}, X_{\mathrm{opt}})$ the matrix X_{opt} must be $\mathcal{I}_{\mathrm{opt}}$-semidefinite but the definiteness properties of its submatrices are not guaranteed due to possible numerical errors. Note the structure of the matrix X appearing in the proof of the second part of Theorem 4.1 and Theorem 4.2: when the factors B and C are known for representations (12) and (15), the quasidefinite matrix X is

$$X = \begin{bmatrix} -Q & A^* \\ A & G \end{bmatrix} = \begin{bmatrix} -C^*C & A^* \\ A & BB^* \end{bmatrix}.$$

In the next section we describe the structure preserving implementation of the symmetric PPT (1) for $\mathcal{I}$-semidefinite matrices X in factored form, which resolves the issues described above and leads to the structured version of the optimization algorithm presented in section 6.

5   Applying a PPT to a factored representation of an $\mathcal{I}$-semidefinite matrix

Let $X \in \mathbb{C}^{n \times n}$ be $\mathcal{I}$-semidefinite and $k = \mathrm{card}(\mathcal{I})$, where $\mathrm{card}(\mathcal{I})$ denotes the number of elements of the set $\mathcal{I}$. Due to the definiteness properties there exist matrices $A \in \mathbb{C}^{(n-k)\times k}$, $B \in \mathbb{C}^{(n-k)\times t}$ and $C \in \mathbb{C}^{r \times k}$ such that

$$X_{\mathcal{I}\mathcal{I}} = -C^*C \in \mathbb{C}^{k \times k}, \quad X_{\mathcal{I}^c\mathcal{I}} = A, \quad X_{\mathcal{I}\mathcal{I}^c} = A^*, \quad X_{\mathcal{I}^c\mathcal{I}^c} = BB^* \in \mathbb{C}^{(n-k)\times(n-k)}. \tag{17}$$

Any A, B and C satisfying (17) are called the *factors* of the $\mathcal{I}$-semidefinite matrix X. We also introduce the following compact form of (17):

$$X = \mathcal{C}_{\mathcal{I}}\left(\begin{bmatrix} C & 0 \\ A & B \end{bmatrix}\right),$$

and say that the map $\mathcal{C}_{\mathcal{I}}$ converts between any factor representation of the $\mathcal{I}$-semidefinite matrix X and the real matrix. Clearly, the factors B and C are not unique as for any unitary matrices H and U of conformal size we have

$$X = \mathcal{C}_{\mathcal{I}}\left(\begin{bmatrix} C & 0 \\ A & B \end{bmatrix}\right) = \mathcal{C}_{\mathcal{I}}\left(\begin{bmatrix} HC & 0 \\ A & BU \end{bmatrix}\right).$$

Given an $\mathcal{I}$-semidefinite matrix X in a factored form (17) and an index set $\mathcal{J}$, our goal in this section is to derive formulas for the symmetric PPT (1) needed in Lemma 2.2 to compute a $\mathcal{J}$-semidefinite matrix X' so $\mathcal{G}_{\mathcal{I}}(X) \sim \mathcal{G}_{\mathcal{J}}(X')$ where

$$
X' = \mathcal{C}_{\mathcal{J}}\left(\begin{bmatrix} C' & 0 \\ A' & B' \end{bmatrix}\right),
$$

and the factors A', B' and C' are computed *directly* from A, B and C. We distinguish three cases for the index set $\mathcal{J}$ we are converting to:

Case 1: $\mathcal{J} \subseteq \mathcal{I}$ (the negative semidefinite block shrinks, the positive semidefinite block expands), in which case $\mathcal{K} = \mathcal{I} \setminus \mathcal{J}$,

Case 2: $\mathcal{J} \supseteq \mathcal{I}$ (the negative semidefinite block expands, the positive semidefinite block shrinks), where $\mathcal{K} = \mathcal{J} \setminus \mathcal{I}$, and

Case 3: $\mathcal{I} \setminus \mathcal{J} \neq \emptyset$ and $\mathcal{J} \setminus \mathcal{I} \neq \emptyset$, in which case $\mathcal{K} = (\mathcal{I} \setminus \mathcal{J}) \cup (\mathcal{J} \setminus \mathcal{I})$.

We now derive the formulas for A', B' and C' in each case. For simplicity, so that we may use a simpler matrix form instead of working with a generic block partition (17), take $\mathcal{I} = \{1, 2 \ldots, k\}$ so that X is

$$
X = \mathcal{C}_{\mathcal{I}}\left(\begin{bmatrix} C & 0 \\ A & B \end{bmatrix}\right) = \begin{array}{c} k \\ n-k \end{array} \overset{\begin{array}{cc} k & n-k \end{array}}{\begin{bmatrix} -C^*C & A^* \\ A & BB^* \end{bmatrix}}. \tag{18}
$$

5.1 Case 1.

Recall that we have $A \in \mathbb{C}^{(n-k)\times k}$, $B \in \mathbb{C}^{(n-k)\times t}$, $C \in \mathbb{C}^{r\times k}$ as factors of X. Again for simplicity, we take $\mathcal{J} = \{1, 2, \ldots, k-l\}$ for some l with $1 \leq l \leq k$. Let H be a unitary matrix such that

$$
A = \begin{array}{c} n-k \end{array} \overset{\begin{array}{cc} k-l & l \end{array}}{\begin{bmatrix} A_1 & A_2 \end{bmatrix}}, \quad HC = \begin{array}{c} r-l \\ l \end{array} \overset{\begin{array}{cc} k-l & l \end{array}}{\begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix}}. \tag{19}
$$

Then the factor representation (18) of X is

$$
X = \mathcal{C}_{\mathcal{I}}\left(\left[\begin{array}{c|c} HC & 0 \\ \hline A & B \end{array}\right]\right) = \left[\begin{array}{cc|c} -C_{11}^*C_{11} - C_{21}^*C_{21} & -C_{21}^*C_{22} & A_1^* \\ -C_{22}^*C_{21} & -C_{22}^*C_{22} & A_2^* \\ \hline A_1 & A_2 & BB^* \end{array}\right].
$$

Now we use Lemma 2.2. The pivot index set is $\mathcal{K} = \{k-l+1, \ldots, k\}$. Note that for this PPT to exist it must hold $r \geq l$, otherwise the matrix $-C^*C$ has rank $r < l$ and hence $X_{\mathcal{K}\mathcal{K}}$, which is a submatrix of $-C^*C$, cannot be invertible. For the pivot submatrix $X_{\mathcal{K}\mathcal{K}} = -C_{22}^*C_{22}$ to be nonsingular, the square matrix C_{22} must be invertible. The diagonal sign change matrix D is the block diagonal matrix $D = \mathrm{diag}(I_{k-l}, -I_l, I_{n-k})$, and applying (1) to X to compute Y we get

$$
X' = DYD = \left[\begin{array}{c|cc} -C_{11}^*C_{11} & -C_{21}^*C_{22}^{-*} & A_1^* - C_{21}^*C_{22}^{-*}A_2^* \\ \hline -C_{22}^{-1}C_{21} & (C_{22}^*C_{22})^{-1} & (C_{22}^*C_{22})^{-1}A_2^* \\ A_1 - A_2C_{22}^{-1}C_{21} & A_2(C_{22}^*C_{22})^{-1} & BB^* + A_2(C_{22}^*C_{22})^{-1}A_2^* \end{array}\right].
$$

The matrix X' is $\mathcal{J}$-semidefinite (as follows by Theorem 3.2) and it is easy to check that it can be represented as $X' = \mathcal{C}_{\mathcal{J}}\left(\left[\begin{array}{c|c} C' & 0 \\ \hline A' & B' \end{array}\right]\right)$ for $A' \in \mathbb{C}^{(l+n-k)\times(k-l)}$, $B' \in \mathbb{C}^{(l+n-k)\times(l+t)}$, $C' \in \mathbb{C}^{(r-l)\times(k-l)}$ given by

$$A' = \begin{bmatrix} -C_{22}^{-1}C_{21} \\ A_1 - A_2C_{22}^{-1}C_{21} \end{bmatrix}, \quad B' = \begin{bmatrix} C_{22}^{-1} & 0 \\ A_2C_{22}^{-1} & B \end{bmatrix}, \quad C' = C_{11}. \tag{20}$$

5.2 Case 2.

Case 2 is very similar to Case 1. We again start from $A \in \mathbb{C}^{(n-k)\times k}$, $B \in \mathbb{C}^{(n-k)\times t}$, $C \in \mathbb{C}^{r\times k}$ and now take $1 \le m \le n-k$, with $m \le t$, to apply Lemma 2.2 to X from (18) for $\mathcal{J} = \{1, 2, \ldots, k, k+1, \ldots, k+m\}$, for simplicity. Let U be a unitary matrix such that

$$A = \begin{array}{c} \\ m \\ n-k-m \end{array}\overset{\displaystyle k}{\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}}, \quad BU = \begin{array}{c} \\ m \\ n-k-m \end{array}\overset{\displaystyle m \quad t-m}{\begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix}}. \tag{21}$$

The factor representation (18) expands to

$$X = \mathcal{C}_{\mathcal{I}}\left(\left[\begin{array}{c|c} C & 0 \\ \hline A & BU \end{array}\right]\right) = \left[\begin{array}{c|cc} -C^*C & A_1^* & A_2^* \\ \hline A_1 & B_{11}B_{11}^* & B_{11}B_{21}^* \\ A_2 & B_{21}B_{11}^* & B_{21}B_{21}^* + B_{22}B_{22}^* \end{array}\right].$$

From Lemma 2.2 we have $\mathcal{K} = \{k+1, \ldots, k+m\}$ and $D = I_n$. The pivot submatrix is $X_{\mathcal{K}\mathcal{K}} = B_{11}B_{11}^*$ and B_{11} must be invertible for this PPT operation to be defined. If this is the case, we have

$$X' = DYD = \left[\begin{array}{cc|c} -C^*C - A_1^*(B_{11}B_{11}^*)^{-1}A_1 & A_1^*(B_{11}B_{11}^*)^{-1} & A_2^* - A_1^*B_{11}^{-*}B_{21}^* \\ (B_{11}B_{11}^*)^{-1}A_1 & -(B_{11}B_{11}^*)^{-1} & B_{11}^{-*}B_{21}^* \\ \hline A_2 - B_{21}B_{11}^{-1}A_1 & B_{21}B_{11}^{-1} & B_{22}B_{22}^* \end{array}\right]$$

$$= \mathcal{C}_{\mathcal{J}}\left(\left[\begin{array}{c|c} C' & 0 \\ \hline A' & B' \end{array}\right]\right),$$

where $A' \in \mathbb{C}^{(n-k-m)\times(k+m)}$, $B' \in \mathbb{C}^{(n-k-m)\times(t-m)}$ and $C' \in \mathbb{C}^{(r+m)\times(k+m)}$ are given by

$$A' = \begin{bmatrix} A_2 - B_{21}B_{11}^{-1}A_1 & B_{21}B_{11}^{-1} \end{bmatrix}, \quad B' = B_{22}, \quad C' = \begin{bmatrix} C & 0 \\ -B_{11}^{-1}A_1 & B_{11}^{-1} \end{bmatrix}. \tag{22}$$

5.3 Case 3

Case 3 is somewhat more complicated. We start from $A \in \mathbb{C}^{(n-k)\times k}$, $B \in \mathbb{C}^{(n-k)\times t}$, $C \in \mathbb{C}^{r\times k}$ and take $1 \le l \le k$ and $1 \le m \le n-k$, such that $l \le r$ and $m \le t$. For simplicity, we look at $\mathcal{J} = \{1, 2, \ldots, k-l\} \cup \{k+1, \ldots, k+m\}$. Let H and U be unitary matrices such that

$$A = \begin{array}{c} \\ m \\ n-k-m \end{array}\overset{\displaystyle k-l \quad l}{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}, \quad BU = \begin{array}{c} \\ m \\ n-k-m \end{array}\overset{\displaystyle m \quad t-m}{\begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix}}, \quad HC = \begin{array}{c} \\ r-l \\ l \end{array}\overset{\displaystyle k-l \quad l}{\begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix}}. \tag{23}$$

In this case (18) is

$$
\begin{aligned}
X &= \mathcal{C}_{\mathcal{I}}\left(\left[\begin{array}{c|c} HC & 0 \\ \hline A & BU \end{array}\right]\right) \\
&= \left[\begin{array}{cc|cc}
-C_{11}^*C_{11} - C_{21}^*C_{21} & -C_{21}^*C_{22} & A_{11}^* & A_{21}^* \\
-C_{22}^*C_{21} & -C_{22}^*C_{22} & A_{12}^* & A_{22}^* \\
\hline
A_{11} & A_{12} & B_{11}B_{11}^* & B_{11}B_{21}^* \\
A_{21} & A_{22} & B_{21}B_{11}^* & B_{21}B_{21}^* + B_{22}B_{22}^*
\end{array}\right].
\end{aligned} \tag{24}
$$

From Lemma 2.2 we have $\mathcal{K} = \{k - l + 1, \ldots, k, k + 1, \ldots, k + m\}$ and the pivot submatrix whose inverse is required is the quasidefinite matrix

$$
X_{\mathcal{KK}} = \begin{bmatrix} -C_{22}^*C_{22} & A_{12}^* \\ A_{12} & B_{11}B_{11}^* \end{bmatrix}. \tag{25}
$$

5.3.1   An inversion formula for quasidefinite matrices

It is not difficult to see that whenever a quasidefinite matrix is invertible, its inverse is quasidefinite, too. Hence, given A, B, C of conformal sizes, one can write

$$
\begin{bmatrix} -C^*C & A^* \\ A & BB^* \end{bmatrix}^{-1} = \begin{bmatrix} -NN^* & K \\ K^* & L^*L \end{bmatrix} \tag{26}
$$

for suitable matrices K, L, N. In this section, we describe a method to compute K, L, N directly from A, B, C. In principle, one can assemble the matrix in (26), invert it, and then find the Cholesky factors of its diagonal blocks. However, this does not appear sound from a numerical point of view, since it means forming BB^* and C^*C and then factoring the corresponding blocks in the computed inverse (which may not be semidefinite due to numerical errors). It is a problem similar to the infamous normal equations for least-squares problems [10, sec. 20.4]. The only condition appearing in Lemma 2.2 is that the pivot submatrix (25) is invertible and we wish to keep only that assumption for the existence of the PPT. Hence, formulas which rely on Schur complements [13, sec. 0.7.3] cannot be used, since BB^* and C^*C are not guaranteed to have full rank (consider, e.g. the case $A = 1$, $B = C = 0$).

In the following, we present an alternative expression that relies heavily on unitary factorizations.

Theorem 5.1. *Let*

$$
P = \begin{bmatrix} -C_{22}^*C_{22} & A_{12}^* \\ A_{12} & B_{11}B_{11}^* \end{bmatrix}, \quad A_{12} \in \mathbb{C}^{m \times l}, \ B_{11} \in \mathbb{C}^{m \times m}, \ C_{22} \in \mathbb{C}^{l \times l}
$$

be an invertible matrix and let Q and H be unitary matrices such that

$$
\begin{bmatrix} B_{11}^* \\ A_{12}^* \end{bmatrix} = Q \begin{bmatrix} R^* \\ 0 \end{bmatrix}, \quad Q = \begin{array}{c} m \\ l \end{array} \begin{bmatrix} \overset{m}{Q_{11}} & \overset{l}{Q_{12}} \\ Q_{21} & Q_{22} \end{bmatrix} \tag{27}
$$

and

$$
M = \begin{bmatrix} I_m & 0 \\ 0 & C_{22} \end{bmatrix} Q = H \begin{bmatrix} M_{11} & 0 \\ M_{21} & M_{22} \end{bmatrix}, \quad H = \begin{array}{c} m \\ l \end{array} \begin{bmatrix} \overset{m}{H_{11}} & \overset{l}{H_{12}} \\ H_{21} & H_{22} \end{bmatrix}. \tag{28}
$$

Then,

1. R and M_{22} are invertible.

2. We have
$$P^{-1} = \begin{bmatrix} -NN^* & K \\ K^* & L^*L \end{bmatrix},$$

   with
$$N = Q_{22}M_{22}^{-1}, \quad K = (Q_{21} - Q_{22}M_{22}^{-1}M_{21})R^{-1}, \quad L = M_{11}R^{-1}.$$

3. The following relations hold:
$$C_{22}N = H_{22}, \qquad\qquad C_{22}K = H_{21}L, \qquad\qquad (29)$$
$$LB_{11} = H_{11}^*, \qquad\qquad KB_{11} = -NH_{12}^*. \qquad\qquad (30)$$

Proof. We use a few manipulations of quasidefinite matrices which are standard in the context of preconditioners for saddle-point matrices; see for instance [4, sec. 10.4].

Note that P is the Schur complement of $-I_m$ in
$$T = \begin{bmatrix} -I_m & 0 & B_{11}^* \\ 0 & -C_{22}^*C_{22} & A_{12}^* \\ B_{11} & A_{12} & 0 \end{bmatrix},$$

so by the standard results on Schur complements T is nonsingular and
$$P^{-1} = \begin{bmatrix} 0 & I_l & 0 \\ 0 & 0 & I_m \end{bmatrix} T^{-1} \begin{bmatrix} 0 & 0 \\ I_l & 0 \\ 0 & I_m \end{bmatrix}.$$

Inserting factors $\hat{Q} = \mathrm{diag}(Q, I)$ and its inverse, we get
$$P^{-1} = \begin{bmatrix} 0 & I_l & 0 \\ 0 & 0 & I_m \end{bmatrix} \hat{Q} \left(\hat{Q}^* \begin{bmatrix} -I_m & 0 & B_{11}^* \\ 0 & -C_{22}^*C_{22} & A_{12}^* \\ B_{11} & A_{12} & 0 \end{bmatrix} \hat{Q} \right)^{-1} \hat{Q}^* \begin{bmatrix} 0 & 0 \\ I_l & 0 \\ 0 & I_m \end{bmatrix}$$
$$= \begin{bmatrix} Q_{21} & Q_{22} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} * & * & R^* \\ * & * & 0 \\ R & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} Q_{21}^* & 0 \\ Q_{22}^* & 0 \\ 0 & I \end{bmatrix}.$$

The top–left 2×2 block which we have marked with asterisks is $-M^*M$, with M as in (28), so we can write it also as
$$P^{-1} = \begin{bmatrix} Q_{21} & Q_{22} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} -M_{11}^*M_{11} - M_{21}^*M_{21} & -M_{21}^*M_{22} & R^* \\ -M_{22}^*M_{21} & -M_{22}^*M_{22} & 0 \\ R & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} Q_{21}^* & 0 \\ Q_{22}^* & 0 \\ 0 & I \end{bmatrix}. \qquad (31)$$

The middle matrix in (31) is equal to $\hat{Q}^*T\hat{Q}$, which is invertible. Hence R and M_{22} must be invertible, too, which proves our first statement. The inverse of this block antitriangular matrix can be computed explicitly as
$$P^{-1} = \begin{bmatrix} Q_{21} & Q_{22} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} 0 & 0 & R^{-1} \\ 0 & -M_{22}^{-1}M_{22}^{-*} & -M_{22}^{-1}M_{21}R^{-1} \\ R^{-*} & -R^{-*}M_{21}^*M_{22}^{-*} & R^{-*}M_{11}^*M_{11}R^{-1} \end{bmatrix} \begin{bmatrix} Q_{21}^* & 0 \\ Q_{22}^* & 0 \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} -Q_{22}M_{22}^{-1}M_{22}^{-*}Q_{22}^* & (Q_{21} - Q_{22}M_{22}^{-1}M_{21})R^{-1} \\ R^{-*}(Q_{21}^* - M_{21}^*M_{22}^{-*}Q_{22}^*) & R^{-*}M_{11}^*M_{11}R^{-1} \end{bmatrix} = \begin{bmatrix} -NN^* & K \\ K^* & L^*L \end{bmatrix},$$

which proves the second claim.

Expanding the multiplications in the second block column of (28), we get $C_{22}Q_{22} = H_{22}M_{22}$ and $C_{22}Q_{21} = H_{21}M_{11} + H_{22}M_{21}$, from which the two equations (29) follow easily. From the first block row of (27) we get $B_{11}^* = Q_{11}R^*$, and again from (28) we get

$$H^* \begin{bmatrix} I & 0 \\ 0 & C_{22} \end{bmatrix} = \begin{bmatrix} M_{11} & 0 \\ M_{21} & M_{22} \end{bmatrix} Q^*,$$

whose first block column reads $H_{11}^* = M_{11}Q_{11}^*$, $H_{12}^* = M_{21}Q_{11}^* + M_{22}Q_{12}^*$. Putting together these relations, (30) follows. $\qquad\square$

We now continue computing the factored version of the PPT in Case 3. Assuming that the matrix $X_{\mathcal{K}\mathcal{K}}$ from (25) is nonsingular, the symmetric principal pivot transform Y of X from (24) exists and we partition it as

$$Y = \left[\begin{array}{c|cc|c} Y_{11} & Y_{21}^* & Y_{31}^* & Y_{41}^* \\ \hline Y_{21} & -Y_{22} & -Y_{32}^* & Y_{42}^* \\ Y_{31} & -Y_{32} & -Y_{33} & Y_{43}^* \\ \hline Y_{41} & Y_{42} & Y_{43} & Y_{44} \end{array} \right]. \tag{32}$$

The middle block is $-X_{\mathcal{K}\mathcal{K}}^{-1}$ and from Theorem 5.1 defining K, L, N we have

$$X_{\mathcal{K}\mathcal{K}}^{-1} = \begin{bmatrix} -C_{22}^*C_{22} & A_{12}^* \\ A_{12} & B_{11}B_{11}^* \end{bmatrix}^{-1} = \begin{bmatrix} -NN^* & K \\ K^* & L^*L \end{bmatrix} = \begin{bmatrix} Y_{22} & Y_{32}^* \\ Y_{32} & Y_{33} \end{bmatrix}. \tag{33}$$

Lemma 5.2. *The remaining blocks of Y from* (32) *are given by*

$$\begin{aligned}
Y_{11} &= -C_{11}^*C_{11} - C_{21}^*C_{21} + C_{21}^*C_{22}NN^*C_{22}^*C_{21} + A_{11}^*K^*C_{22}^*C_{21} \\
&\quad + C_{21}^*C_{22}KA_{11} - A_{11}^*L^*LA_{11}, \\
Y_{21} &= NN^*C_{22}^*C_{21} + KA_{11}, \\
Y_{31} &= -K^*C_{22}^*C_{21} + L^*LA_{11}, \\
Y_{41} &= A_{21} - A_{22}NN^*C_{22}^*C_{21} - A_{22}KA_{11} + B_{21}B_{11}^*K^*C_{22}^*C_{21} - B_{21}B_{11}^*L^*LA_{11}, \\
Y_{42} &= -A_{22}NN^* + B_{21}B_{11}^*K^*, \\
Y_{43} &= A_{22}K + B_{21}B_{11}^*L^*L, \\
Y_{44} &= B_{21}B_{21}^* + B_{22}B_{22}^* + A_{22}NN^*A_{22}^* - B_{21}B_{11}^*K^*A_{22}^* \\
&\quad - A_{22}KB_{11}B_{21}^* - B_{21}B_{11}^*L^*LB_{11}B_{21}^*.
\end{aligned}$$

Proof. We get the above formulas after some tedious but straightforward algebra from the PPT formulas (1), the expression (24) for X and (33). $\qquad\square$

The sign change matrix D from Lemma 2.2 is $D = \mathrm{diag}(I_{k-l}, -I_l, I_m, I_{n-k-m})$, and we finally have

$$X' = DYD = \left[\begin{array}{c|cc|c} Y_{11} & -Y_{21}^* & Y_{31}^* & Y_{41}^* \\ \hline -Y_{21} & -Y_{22} & Y_{32}^* & -Y_{42}^* \\ Y_{31} & Y_{32} & -Y_{33} & Y_{43}^* \\ \hline Y_{41} & -Y_{42} & Y_{43} & Y_{44} \end{array} \right],$$

where the blocks are defined in (33) and Lemma 5.2. What remains is to show that $X' = \mathcal{C}_{\mathcal{J}}\left(\begin{bmatrix} C' & 0 \\ A' & B' \end{bmatrix} \right)$, where $\mathcal{J} = \{1, 2, \ldots, k-l\} \cup \{k, k+1, \ldots, k+m\}$, by finding the factors $B' \in$

$\mathbb{C}^{(n-k+l-m)\times(t-m+l)}$ and $C' \in \mathbb{C}^{(r-l+m)\times(k-l+m)}$ such that

$$\begin{bmatrix} -Y_{22} & -Y_{42}^* \\ -Y_{42} & Y_{44} \end{bmatrix} = B'(B')^* \quad \text{and} \quad \begin{bmatrix} -Y_{11} & -Y_{31}^* \\ -Y_{31} & Y_{33} \end{bmatrix} = (C')^*C'. \tag{34}$$

The factor $A' \in \mathbb{C}^{(n-k+l-m)\times(k+m-l)}$ is given by

$$A' = \begin{bmatrix} -Y_{21} & Y_{32}^* \\ Y_{41} & Y_{43} \end{bmatrix}. \tag{35}$$

Lemma 5.3. *The equalities* (34) *hold with*

$$B' = \begin{bmatrix} N & 0 \\ B_{21}H_{12} + A_{22}N & B_{22} \end{bmatrix} \quad \text{and} \quad C' = \begin{bmatrix} C_{11} & 0 \\ H_{21}^*C_{21} - LA_{11} & L \end{bmatrix}, \tag{36}$$

where H_{12}, H_{21}, L and N are defined in Theorem 5.1.

Proof. Define $Z = B_{21}H_{12} + A_{22}N$. Then

$$B'(B')^* = \begin{bmatrix} NN^* & NZ^* \\ ZN^* & ZZ^* + B_{22}B_{22}^* \end{bmatrix}$$

and we only need to check that these blocks match the blocks specified in (34). From (33) we have $NN^* = -Y_{22}$ and from (30) we get

$$ZN^* = B_{21}H_{12}N^* + A_{22}NN^* = -B_{21}B_{11}^*K^* + A_{22}NN^* = -Y_{42},$$

where we have used the formula for Y_{42} from Lemma 5.2 for the last equality. What remains is to show that

$$ZZ^* + B_{22}B_{22}^* = Y_{44}.$$

Multiplying out the left hand side and using (30) we see that the above equality holds if and only if

$$B_{21}H_{12}H_{12}^*B_{21}^* = B_{21}B_{21}^* - B_{21}H_{11}H_{11}^*B_{21}^*,$$

which is true because H from (28) is a unitary matrix and so $H_{11}H_{11}^* + H_{12}H_{12}^* = I$.

The proof involving the matrix C' uses (29) and is identical to the above. $\qquad\square$

To summarize the results for Case 3, we have an $\mathcal{I}$-semidefinite matrix $X = \mathcal{C}_{\mathcal{I}}\left(\begin{bmatrix} HC & 0 \\ A & BU \end{bmatrix}\right)$ for $\mathcal{I} = \{1,\ldots,k\}$ and the factors A, BU and HC as in (23), and we wish to transform into a $\mathcal{J}$-semidefinite matrix X' for $\mathcal{J} = \{1,\ldots k-l\} \cup \{k+1,\ldots,k+m\}$. Providing that the matrix $X_{\mathcal{K}\mathcal{K}}$ from (25) is invertible and its inverse defined by (33), X' can be represented as $X' = \mathcal{C}_{\mathcal{J}}\left(\begin{bmatrix} C' & 0 \\ A' & B' \end{bmatrix}\right)$, where the factor A' is given by (35) and the factors B' and C' are defined in Lemma 5.3.

5.4   General formulas

The index sets $\mathcal{I}$ and $\mathcal{J}$ from the previous sections were chosen for simplicity. Nothing essentially changes if we use general index sets but the notation becomes much more obscured. To illustrate this, we briefly sketch the general version of Case 1.

We start from an $\mathcal{I}$-semidefinite matrix $X \in \mathbb{C}^{n\times n}$, where $\mathcal{I}$ is now any subset of $\{1,2,\ldots,n\}$ and $\mathcal{I}^c = \{1,2,\ldots,n\} \setminus \mathcal{I}$. Let $k = \text{card}(\mathcal{I})$ and $A \in \mathbb{C}^{(n-k)\times k}$, $B \in \mathbb{C}^{(n-k)\times t}$ and $C \in \mathbb{C}^{r\times k}$ such that

$$X_{\mathcal{I}\mathcal{I}} = -C^*C, \quad X_{\mathcal{I}^c\mathcal{I}} = A, \quad X_{\mathcal{I}\mathcal{I}^c} = A^*, \quad X_{\mathcal{I}^c\mathcal{I}^c} = BB^*.$$

Take $\mathcal{J} \subseteq \mathcal{I}$ and let $\hat{\mathcal{J}} \subseteq \{1, \ldots, k\}$ be the local index set for the columns of the matrix C corresponding to the global $\mathcal{J}$. Let $\hat{\mathcal{J}}^c = \{1, 2, \ldots, k\} \setminus \hat{\mathcal{J}}$. From Lemma 2.2 we have $\mathcal{K} = \mathcal{I} \setminus \mathcal{J}$.

Define the unitary matrix H which introduces zeros into each column c_j of C, $j \in \hat{\mathcal{J}}^c$, except in the positions c_{ij} for $i \in \hat{\mathcal{L}}$, where $\hat{\mathcal{L}}$ is a subset of $\{1, 2, \ldots, r\}$ corresponding to the local row index set of the matrix C, and such that $\mathrm{card}(\hat{\mathcal{L}}) = \mathrm{card}(\hat{\mathcal{J}}^c)$. Then we have

$$HC = \begin{bmatrix} (HC)_{\hat{\mathcal{L}}^c \hat{\mathcal{J}}} & (HC)_{\hat{\mathcal{L}}^c \hat{\mathcal{J}}^c} \\ (HC)_{\hat{\mathcal{L}} \hat{\mathcal{J}}} & (HC)_{\hat{\mathcal{L}} \hat{\mathcal{J}}^c} \end{bmatrix} = \begin{matrix} \mathrm{card}(\hat{\mathcal{L}}^c) \\ \mathrm{card}(\hat{\mathcal{L}}) \end{matrix} \overset{\overset{\mathrm{card}(\hat{\mathcal{J}}) \quad \mathrm{card}(\hat{\mathcal{J}}^c)}{}}{\begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix}}.$$

Note that $\mathcal{J} \cup \mathcal{K} = \mathcal{I}$ and hence we can partition A as

$$A = \begin{bmatrix} X_{\mathcal{I}^c \mathcal{J}} & X_{\mathcal{I}^c \mathcal{K}} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \end{bmatrix}.$$

Then the negative semidefinite block in $X = \mathcal{C}_{\mathcal{I}}\left(\begin{bmatrix} HC & 0 \\ A & B \end{bmatrix}\right)$ is now given by

$$X_{\mathcal{I}\mathcal{I}} = Z, \quad \text{where} \quad Z = \begin{bmatrix} X_{\mathcal{J}\mathcal{J}} & X_{\mathcal{J}\mathcal{K}} \\ X_{\mathcal{K}\mathcal{J}} & X_{\mathcal{K}\mathcal{K}} \end{bmatrix} = \begin{bmatrix} -C_{11}^* C_{11} - C_{21}^* C_{21} & -C_{21}^* C_{22} \\ -C_{22}^* C_{21} & -C_{22}^* C_{22} \end{bmatrix}.$$

We apply the PPT (1) with $X_{\mathcal{K}\mathcal{K}}$ as the pivot matrix, provided that the square matrix C_{22} is nonsingular, to get the following X':

$$\begin{aligned} X' = DYD &= \begin{bmatrix} X'_{\mathcal{J}\mathcal{J}} & X'_{\mathcal{J}\mathcal{K}} & X'_{\mathcal{J}\mathcal{I}^c} \\ X'_{\mathcal{K}\mathcal{J}} & X'_{\mathcal{K}\mathcal{K}} & X'_{\mathcal{K}\mathcal{I}^c} \\ X'_{\mathcal{I}^c \mathcal{J}} & X'_{\mathcal{I}^c \mathcal{K}} & X'_{\mathcal{I}^c \mathcal{I}^c} \end{bmatrix} \\ &= \begin{bmatrix} -C_{11}^* C_{11} & -C_{21}^* C_{22}^{-*} & A_1^* - C_{21}^* C_{22}^{-*} A_2^* \\ -C_{22}^{-1} C_{21} & (C_{22}^* C_{22})^{-1} & (C_{22}^* C_{22})^{-1} A_2^* \\ A_1 - A_2 C_{22}^{-1} C_{21} & A_2 (C_{22}^* C_{22})^{-1} & BB^* + A_2 (C_{22}^* C_{22})^{-1} A_2^* \end{bmatrix}, \end{aligned}$$

where the sign change matrix D is partitioned conformally so is the same as in Case 1 (Section 5.1). It follows that the matrix $X' = \mathcal{C}_{\mathcal{J}}\left(\begin{bmatrix} C' & 0 \\ A' & B' \end{bmatrix}\right)$ is $\mathcal{J}$-semidefinite for A', B' and C' given by (20).

Note that the above could have also been achieved by defining permutation matrices P_1 and P_2 and symmetrically permuting the rows and columns of X and X' so that the leading diagonal blocks of the matrices $P_1 X P_1^T$ and $P_2 X' P_2^T$ are $X_{\mathcal{I}\mathcal{I}}$ and $X'_{\mathcal{J}\mathcal{J}}$, respectively.

6 PPTs with bounded elements

We now use the factor-based formulas for the PPT derived in Section 5 to compute an optimal permuted Riccati basis for a Lagrangian semidefinite subspace. From [18, Thm. 3.4], which is here stated as the final part of Theorem 2.1, we know that for a Lagrangian subspace $\mathrm{Im}\, U$ there exists at least one optimal permuted Riccati representation with X_{opt} satisfying

$$|(X_{\mathrm{opt}})_{ij}| \leq \begin{cases} 1, & \text{if } i = j, \\ \sqrt{2}, & \text{otherwise.} \end{cases}$$

The above inequality is sharp, as can be seen from the example [18, sec. 3] where $U = \begin{bmatrix} I_2 \\ X \end{bmatrix}$, $X = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix}$. However, a stronger version can be obtained under the additional hypothesis that $\mathrm{Im}\, U$ is Lagrangian semidefinite (instead of merely Lagrangian).

Theorem 6.1. *Let $U \in \mathbb{C}^{2n \times n}$ be such that $\operatorname{Im} U$ is Lagrangian semidefinite. Then, there exists $\mathcal{I}_{\mathrm{opt}} \subseteq \{1, 2, \ldots, n\}$ such that $U \sim \mathcal{G}_{\mathcal{I}_{\mathrm{opt}}}(X)$ and*

$$|x_{ij}| \leq 1 \quad \forall i, j. \tag{37}$$

Proof. Since $\operatorname{Im} U$ is Lagrangian, from the proof of [18, Thm. 3.4] it follows that there exists an index set $\mathcal{I}$ defining the symplectic swap $\Pi_{\mathcal{I}}$ such that $U \sim \mathcal{G}_{\mathcal{I}}(X)$, $X \in \mathbb{C}^{n \times n}$ is Hermitian and

$$|x_{ii}| \leq 1, \quad \left| \det \begin{bmatrix} x_{ii} & x_{ij} \\ x_{ji} & x_{jj} \end{bmatrix} \right| \leq 1, \quad i, j = 1, 2, \ldots, n. \tag{38}$$

In addition, since $\operatorname{Im} U$ is Lagrangian semidefinite, from Theorem 3.2 it follows that X is $\mathcal{I}$-semidefinite. We prove that the choice $\mathcal{I}_{\mathrm{opt}} = \mathcal{I}$ satisfies (37).

For $i = j$ this trivially follows from (38). When $i \neq j$, we distinguish four cases.

Case A: $i, j \in \mathcal{I}$ The block $X_{\mathcal{I}\mathcal{I}}$ is negative semidefinite, hence its submatrix $(X)_{\{i,j\}\{i,j\}}$ is negative semidefinite, too, and this implies

$$|x_{ij}|^2 = x_{ij} x_{ji} \leq |x_{ii}||x_{jj}| \leq 1. \tag{39}$$

Case B: $i, j \notin \mathcal{I}$ The proof is analogous to the previous one, since $X_{\mathcal{I}^c \mathcal{I}^c}$ is positive semidefinite.

Case C: $i \in \mathcal{I}, j \notin \mathcal{I}$ By semidefiniteness it follows that $-1 \leq x_{ii} \leq 0$ and $0 \leq x_{jj} \leq 1$. Moreover, by the 2×2 case from (38) we get

$$|x_{ii} x_{jj} - |x_{ij}|^2| = |x_{ii}|x_{jj} + |x_{ij}|^2 \leq 1,$$

and hence $|x_{ij}| \leq 1$.

Case D: $i \notin \mathcal{I}, j \in \mathcal{I}$ The proof is analogous to the previous one by swapping i and j. $\qquad \square$

6.1 The optimization algorithm

Algorithm 1, which is a modified version of [18, Alg. 2], can be used to compute $\mathcal{I}_{\mathrm{opt}}$ such that (37) holds. The algorithm is based on the following observation: in a Lagrangian semidefinite matrix X with factors A, B, C the entry of maximum modulus cannot occur in an off-diagonal entry of $X_{\mathcal{I}\mathcal{I}}$ or $X_{\mathcal{I}^c \mathcal{I}^c}$, because those blocks are semidefinite and thus the first inequality in (39) holds. So there can be three possible locations for the entry of maximum modulus:

- in the block $C^* C$: then it must be on the diagonal and equal to the squared norm $\|C_{:,j}\|^2$ of a column of C;

- in the block BB^*: then it must be on the diagonal and equal to the squared norm $\|B_{i,:}\|^2$ of a row of B;

- in the block A, and then nothing more specific can be said on its position.

When this maximum modulus exceeds a given threshold $\tau \geq 1$, in each of the three cases we can perform a PPT that strictly reduces this maximal entry. We repeat this procedure until all entries are smaller than τ. The algorithm terminates since each PPT uses a pivot matrix with determinant at least τ and hence $|\det X|$ is reduced by a factor at least τ^2 at each step. This argument is similar to, but slightly different from, the one used in [18, Thm. 5.2], where a determinant argument is applied to U_1 in (8) rather than X.

Algorithm 1: Computing a bounded permuted Riccati basis of a semidefinite Lagrangian subspace.

Input: $\mathcal{I}_{\text{in}} \subseteq \{1, 2, \ldots, n\}$, and factors $A_{\text{in}}, B_{\text{in}}, C_{\text{in}}$ of an $\mathcal{I}_{\text{in}}$-semidefinite matrix $X_{\text{in}} = X_{\text{in}}^* \in \mathbb{C}^{n \times n}$ as defined by (17); a threshold $\tau \geq 1$.

Output: $\mathcal{I}_{\text{out}} \subseteq \{1, 2, \ldots, n\}$, and factors $A_{\text{out}}, B_{\text{out}}, C_{\text{out}}$ of an $\mathcal{I}_{\text{out}}$-semidefinite matrix $X_{\text{out}} = X_{\text{out}}^* \in \mathbb{C}^{n \times n}$ such that $\mathcal{G}_{\mathcal{I}_{\text{in}}}(X_{\text{in}}) \sim \mathcal{G}_{\mathcal{I}_{\text{out}}}(X_{\text{out}})$ and $|(X_{\text{out}})_{ij}| \leq \tau$ for each i, j.

$A = A_{\text{in}},\ B = B_{\text{in}},\ C = C_{\text{in}},\ \mathcal{I} = \mathcal{I}_{\text{in}}$;

for it $= 1, 2, \ldots, \text{max\_iterations}$ **do**

    $\hat{\jmath} = \arg\max \|C_{:,j}\|^2$;

    if $\|C_{:,\hat{\jmath}}\|^2 > \tau$ **then**

        use the formulas in Case 1 in Section 5.1, with $\mathcal{J} = \mathcal{I} \setminus \{\hat{\jmath}\}$, to update $(A, B, C, \mathcal{I}) \leftarrow (A', B', C', \mathcal{J})$;

        continue;

    end

    $\hat{\imath} = \arg\max \|B_{i,:}\|^2$;

    if $\|B_{\hat{\imath},:}\|^2 > \tau$ **then**

        use the formulas in Case 2 in Section 5.2, with $\mathcal{J} = \mathcal{I} \cup \{\hat{\imath}\}$, to update $(A, B, C, \mathcal{I}) \leftarrow (A', B', C', \mathcal{J})$;

        continue;

    end

    $\hat{\imath}, \hat{\jmath} = \arg\max |A_{i,j}|$;

    if $|A_{\hat{\imath},\hat{\jmath}}| > \tau$ **then**

        use the formulas in Case 3 in Section 5.2, with $\mathcal{J} = (\mathcal{I} \setminus \{\hat{\jmath}\}) \cup \{\hat{\imath}\}$, to update $(A, B, C, \mathcal{I}) \leftarrow (A', B', C', \mathcal{J})$;

        continue;

    end

    break;

end

$A_{\text{out}} = A,\ B_{\text{out}} = B,\ C_{\text{out}} = C,\ \mathcal{I}_{\text{out}} = \mathcal{I}$;

6.2 Special formulas for the scalar cases $l = 1$, $m = 1$

The pivot sets used in the algorithm have at most 2 elements, so some simplifications can be done to the general formulas (20), (22), (35) and (36).

For Case 1, the factor partition (19) is (up to the ordering of indices)

$$
A = \begin{array}{c} \\ n-k \end{array} \begin{array}{cc} {\scriptstyle k-1} & {\scriptstyle 1} \\ \left[\begin{array}{cc} A_1 & a \end{array} \right] \end{array}, \quad
HC = \begin{array}{c} \\ r-1 \\ 1 \end{array} \begin{array}{cc} {\scriptstyle k-1} & {\scriptstyle 1} \\ \left[\begin{array}{cc} C_{11} & 0 \\ c^* & \gamma \end{array} \right] \end{array},
$$

and the PPT that gives the updated factors in (20) for $\mathcal{J} = \{1, \ldots, k-1\}$ is

$$
\left[\begin{array}{c|c} HC & 0 \\ \hline A & B \end{array} \right] =
\left[\begin{array}{cc|c} C_{11} & 0 & 0 \\ c^* & \gamma & 0 \\ \hline A_1 & a & B \end{array} \right] \mapsto
\left[\begin{array}{c|cc} C_{11} & 0 & 0 \\ \hline -\gamma^{-1}c^* & \gamma^{-1} & 0 \\ A_1 - \gamma^{-1}ac^* & \gamma^{-1}a & B \end{array} \right] =
\left[\begin{array}{c|c} C' & 0 \\ \hline A' & B' \end{array} \right].
$$

Similarly, for Case 2 the starting factors (21) are now partitioned as

$$
A = \begin{array}{c} 1 \\ n-k-1 \end{array} \begin{array}{c} {\scriptstyle k} \\ \left[\begin{array}{c} a^* \\ A_2 \end{array} \right] \end{array}, \quad
BU = \begin{array}{c} 1 \\ n-k-1 \end{array} \begin{array}{cc} {\scriptstyle 1} & {\scriptstyle t-1} \\ \left[\begin{array}{cc} \beta & 0 \\ b & B_{22} \end{array} \right] \end{array},
$$

and the updated factors (22) for $\mathcal{J} = \{1, \ldots, k, k+1\}$ corresponds to the PPT

$$
\left[\begin{array}{c|c} C & 0 \\ \hline A & BU \end{array} \right] =
\left[\begin{array}{c|cc} C & 0 & 0 \\ \hline a^* & \beta & 0 \\ A_2 & b & B_{22} \end{array} \right] \mapsto
\left[\begin{array}{cc|c} C & 0 & 0 \\ -\beta^{-1}a^* & \beta^{-1} & 0 \\ \hline A_2 - \beta^{-1}ba^* & \beta^{-1}b & B_{22} \end{array} \right] =
\left[\begin{array}{c|c} C' & 0 \\ \hline A' & B' \end{array} \right].
$$

For Case 3 the initial partition of factors (23) is

$$
A = \begin{array}{c} 1 \\ n-k-1 \end{array} \begin{array}{cc} {\scriptstyle k-1} & {\scriptstyle 1} \\ \left[\begin{array}{cc} a^* & \alpha \\ A_{21} & d \end{array} \right] \end{array}, \quad
BU = \begin{array}{c} 1 \\ n-k-1 \end{array} \begin{array}{cc} {\scriptstyle 1} & {\scriptstyle t-1} \\ \left[\begin{array}{cc} \beta & 0 \\ b & B_{22} \end{array} \right] \end{array}, \quad
HC = \begin{array}{c} r-1 \\ 1 \end{array} \begin{array}{cc} {\scriptstyle k-1} & {\scriptstyle 1} \\ \left[\begin{array}{cc} C_{11} & 0 \\ c^* & \gamma \end{array} \right] \end{array}.
$$

The PPT for the updated factors for $\mathcal{J} = \{1, \ldots k-1\} \cup \{k+1\}$ has the pivot set $\mathcal{K} = \{k, k+1\}$ and it requires the inverse of the 2×2 matrix $X_{\mathcal{KK}} = \begin{bmatrix} -|\gamma|^2 & \overline{\alpha} \\ \alpha & |\beta|^2 \end{bmatrix}$, which can be computed explicitly as

$$
X_{\mathcal{KK}}^{-1} = \frac{1}{\Delta^2} \begin{bmatrix} -|\beta|^2 & \overline{\alpha} \\ \alpha & |\gamma|^2 \end{bmatrix}, \quad \Delta = \sqrt{|\alpha|^2 + |\beta\gamma|^2}.
$$

Therefore, we can write $X_{\mathcal{KK}}^{-1} = \begin{bmatrix} -NN^* & K \\ K^* & L^*L \end{bmatrix}$ for $N = \beta/\Delta$, $K = \overline{\alpha}/\Delta^2$ and $L = \gamma/\Delta$. Lemma 5.3 gives

$$
B' = \begin{bmatrix} \beta/\Delta & 0 \\ (\beta d - \alpha b)/\Delta & B_{22} \end{bmatrix} \quad \text{and} \quad
C' = \begin{bmatrix} C_{11} & 0 \\ (\alpha c^* - \gamma a^*)/\Delta & \gamma/\Delta \end{bmatrix}.
$$

Finally, the factor update formula for is

$$
\left[\begin{array}{c|c} HC & 0 \\ \hline A & BU \end{array}\right] = \left[\begin{array}{cc|cc} C_{11} & 0 & 0 & 0 \\ c^* & \gamma & 0 & 0 \\ \hline a^* & \alpha & \beta & 0 \\ A_{21} & d & b & B_{22} \end{array}\right]
$$

$$
\mapsto \left[\begin{array}{cc|cc} C_{11} & 0 & 0 & 0 \\ (\alpha c^* - \gamma a^*)/\Delta & \gamma/\Delta & 0 & 0 \\ -Y_{21} & Y_{32}^* & \beta/\Delta & 0 \\ Y_{41} & Y_{43} & (\beta d - \alpha b)/\Delta & B_{22} \end{array}\right] = \left[\begin{array}{c|c} C' & 0 \\ \hline A' & B' \end{array}\right],
$$

where

$$
Y_{21} = (\overline{\gamma}|\beta|^2 c^* + \overline{\alpha} a^*)/\Delta^2,
$$

$$
Y_{41} = A_{21} - \frac{1}{\Delta^2} \begin{bmatrix} d & b \end{bmatrix} \begin{bmatrix} \overline{\gamma}|\beta|^2 & \overline{\alpha} \\ -\alpha\overline{\beta}\overline{\gamma} & \overline{\beta}|\gamma|^2 \end{bmatrix} \begin{bmatrix} c^* \\ a^* \end{bmatrix},
$$

$$
Y_{32} = \overline{\alpha}/\Delta^2,
$$

$$
Y_{43} = (\overline{\alpha} d + |\gamma|^2 \overline{\beta} b)/\Delta^2.
$$

7   Numerical experiments

We have implemented a MATLAB version of Algorithm 1 and carried out the tests in MAT-LAB R2014a on a machine with an Intel Core i7-4910MQ 2.90GHz processor and 16GB RAM.

In our first experiment we use `randn` to generate random factors $C \in \mathbb{R}^{14 \times 14}$, $A \in \mathbb{R}^{16 \times 14}$, $B \in \mathbb{R}^{16 \times 16}$ and a random index set $\mathcal{I}$ with card$(\mathcal{I}) = 14$ defining the $\mathcal{I}$-semidefinite matrix X of order 30. The threshold parameter for the optimization algorithm is $\tau = 1.5$. In Figure 1 we display a color plot of the matrix $|X|$, where $(|X|)_{ij} = |x_{ij}|$ at the start of the optimization procedure, after 10 and 20 iterations, and the final matrix. The algorithm took 31 iterations and produced the matrix X with $\max |x_{ij}| = 0.46$ and card$(\mathcal{I}_{\mathrm{opt}}) = 16$. The effect of semidefinite blocks on the reduction can be seen in plots (b) and (c) of Figure 1, where the dark red stripes that appear are due to the fact that whenever a diagonal pivot is chosen (Case 1 or 2), all elements in the corresponding row and column of the matrices $-C^*C$ or BB^* are also reduced below τ.

For the same example, Figure 2 displays $\max_{i,j} |x_{ij}|$ and $|\det X|$ during the iterations. The quantity $\max_{i,j} |x_{ij}|$ is not guaranteed to decrease with iterations and we see this behaviour on the left plot but $|\det X|$ must decrease with each iteration as we explain in Section 6.1, and this is evident in the log-lin graph on the right.

We next use the matrices from the examples in the benchmark test set [6] to construct a quasidefinite matrix X to which we then apply Algorithm 1 with the threshold $\tau = 1.5$. The test set [6] contains 33 problems, which are taken from the standard `carex` test suite [3] for the numerical solution of the continuous-time algebraic Riccati equation, and in addition some examples use different parameters chosen to make the problems more challenging. Each example contains factors (cf. Section 4) A, $G = G^T$, $Q = Q^T \in \mathbb{R}^{k \times k}$, $B \in \mathbb{R}^{k \times t}$, $C \in \mathbb{R}^{r \times k}$, $R = R^T \in \mathbb{R}^{t \times t}$ and $\widetilde{Q} = \widetilde{Q}^T \in \mathbb{R}^{r \times r}$, with $r, t \leq k$, which define the Hamiltonian matrix

$$
\mathcal{H} = \begin{bmatrix} A & -G \\ -Q & -A^T \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ -C^T\widetilde{Q}C & -A^T \end{bmatrix}.
$$

From these factors we construct the quasidefinite matrix

$$
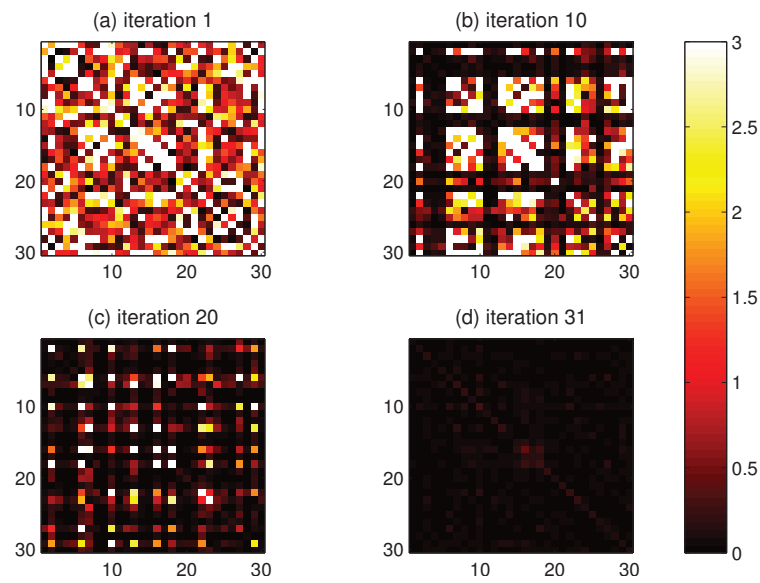X = \begin{bmatrix} -Q & A^T \\ A & G \end{bmatrix} = \begin{bmatrix} -C_f^T C_f & A^T \\ A & B_f B_f^T \end{bmatrix},
$$

Fig. 1: Snapshots of $|X|$ for the starting matrix, iterations 10 and 20, and the final matrix for Algorithm 1 applied to a random matrix X of order 30 with the factors $C \in \mathbb{R}^{14 \times 14}$, $A \in \mathbb{R}^{16 \times 14}$ and $B \in \mathbb{R}^{16 \times 16}$.
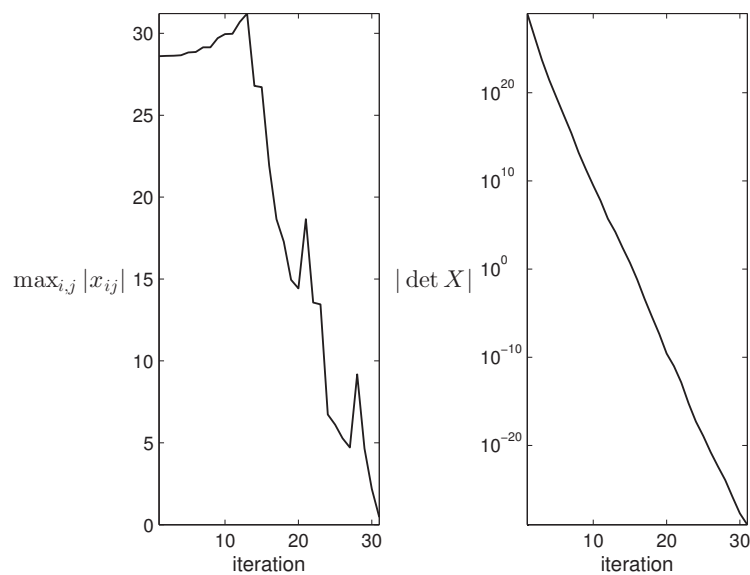

Fig. 2: The changes in $\max_{i,j} |x_{ij}|$ and $|\det X|$ for Algorithm 1 applied to a random matrix X of order 30 with the factors $C \in \mathbb{R}^{14 \times 14}$, $A \in \mathbb{R}^{16 \times 14}$ and $B \in \mathbb{R}^{16 \times 16}$. In the figure on the right, a base-10 logarithmic scale is used for the y-axis.

Tab. 1: Algorithm 1 applied to the matrices defining the test examples from [6].

Ex.	k	t	r	$\kappa(\mathcal{G}_\mathcal{I}(X))$	Subspace distance	$\max |x_{ij}|$	$\max |(X_\text{opt})_{ij}|$	it
1	2	1	2	2.41	4.71e-16	2.000	1.000	2
2	2	1	2	1.62e1	1.31e-15	9.000	9.231e-1	3
5	9	3	9	2.16e2	5.60e-15	1.472e2	7.961e-1	13
6	30	3	5	1.44e8	3.47e-13	1.440e8	1.377	29
7	2	1	1	2.03	7.67e-16	2.000	1.000	2
8	2	1	1	2.72	7.28e-16	2.000	1.000	2
9	2	2	1	1.01e4	1.99e-13	1.000e4	1.000e-1	2
10	2	2	1	1.01e6	4.07e-11	1.000e6	1.000e-1	3
11	2	1	2	1.62	4.71e-16	1.000	1.000	1
12	2	1	2	7.07e5	5.08e-16	1.000e6	1.000	2
13	2	1	2	1.41	4.71e-16	1.000	1.000	1
14	2	2	2	1.91	7.25e-16	2.000	4.000e-1	3
15	2	2	2	2.75	1.04e-15	1.000	1.000	1
16	2	2	2	2.75	1.86e-15	1.000	1.000	1
19	3	3	3	1.91	1.22e-15	2.333	1.486	3
20	3	3	3	3.54	1.12e-15	2.333e6	7.037e-7	7
21	4	1	2	1.91	1.60e-15	1.000	1.000	1
22	4	1	2	1.00e12	1.16e-10	1.000e12	1.000	3
23	4	1	1	4.16	1.09e-15	1.000	1.000	1
24	4	1	1	4.24	9.11e-16	1.000	1.000	1
25	77	39	38	1.00e1	6.44e-15	1.000e1	1.000	39
26	237	119	118	1.00e1	9.88e-15	1.000e1	1.000	119
27	397	199	198	1.00e1	1.31e-14	1.000e1	1.000	199
28	8	8	8	3.00	1.68e-15	2.000	1.400	5
29	64	64	64	3.00	6.11e-15	2.000	1.400	33
30	21	1	1	1.00	6.76e-16	1.000	1.000	1
31	21	1	1	1.00e2	6.76e-16	1.000e2	1.000	2
32	100	1	1	1.22e3	1.84e-13	4.481e2	9.985e-1	200
33	60	2	60	2.41	6.75e-15	1.000	1.000	1

where $B_f = BR_R^{-1} \in \mathbb{R}^{k \times t}$, $C_f = R_{\widetilde{Q}}C \in \mathbb{R}^{r \times k}$ and R_R and $R_{\widetilde{Q}}$ are the upper triangular Cholesky factors of the matrices R and $\widetilde{Q}$, respectively. Moreover, the matrix $\widetilde{Q} = \left[\begin{smallmatrix} 9 & 6 \\ 6 & 4 \end{smallmatrix}\right]$ in Example 2 in [6] is singular positive semidefinite and therefore we use $R_{\widetilde{Q}} = \left[\begin{smallmatrix} 3 & 2 \\ 0 & 0 \end{smallmatrix}\right]$ as its Cholesky factor. The pair $(\mathcal{I}, X)$, with $\mathcal{I} = \{1, 2, \dots, k\}$, identifies a Lagrangian subspace of $\mathbb{C}^{4k}$ which is associated with the Hamiltonian pencil (12), as described in Section 4.

This construction eliminates Examples 3, 4, 17 and 18 from [6] due to the indefiniteness of the matrix $\widetilde{Q}$, and consequently the matrix Q, since we cannot form a quasidefinite matrix X from these factors.

In Table 1, for each of the remaining examples we present the dimensions k, t and r defining the factors C_f, A and B_f of the matrix X, the number of iterations it the optimization took, the 2-norm condition number κ of the starting matrix $\mathcal{G}_\mathcal{I}(X)$, the maximum modulus of the elements in X and the computed optimal reduced matrix X_opt, and the subspace distance between $\mathcal{G}_\mathcal{I}(X)$ and $\mathcal{G}_{\mathcal{I}_\text{opt}}(X_\text{opt})$ computed by MATLAB's subspace.

Small values for the subspaces distance indicate that the algorithm produced a representation of the same subspace associated with $\mathcal{G}_\mathcal{I}(X)$, which happens in almost all examples. The largest value for this quantity corresponds to the Example 22 where the starting representation

$\mathcal{G}_\mathcal{I}(X)$ is ill-conditioned. Several examples already had X with elements bounded in modulus by 1 but we include them for completeness. In all other examples, the algorithm achieved the goal of reducing the modulus of all elements in X_{opt} below the threshold τ and the number of iterations required to do this was in general not large.

We also note that the factors B_f and C_f could have been formed from the matrices G and Q, e.g. by taking their Cholesky factors or semidefinite square roots. We chose not to do this not only because B and C are readily available, but also since in most examples G and/or Q are singular, and moreover such B_f and C_f would be square, while those computed from B and C are rectangular, often with very small number of columns and rows, respectively.

The MATLAB code used for the experiments is available at `https://bitbucket.org/fph/pgdoubling-quad`.

8   Concluding remarks

The main motivation for this work was the fact that the definiteness structure possessed by most matrices to which the PPT is applied in [18] is not used or enforced in general formulas (1), which means that it could be lost in computation due to numerical errors. These matrices of interest, which generalize the quasidefinite structure, are called $\mathcal{I}$-semidefinite.

We have shown that $\mathcal{I}$-semidefinite matrices define Lagrangian semidefinite subspaces which are associated with the standard form of Hamiltonian and symplectic pencils appearing in control theory, and this makes $\mathcal{I}$-semidefinite matrices ubiquitous in the field. We also proved that the elementwise bound on the entries of an optimal permuted Riccati representation can be improved for the case of a Lagrangian semidefinite subspace.

The central part of the paper was dedicated to deriving factored versions of the general PPT formulas used in the optimization algorithm for computing this optimal representation. These formulas now exploit the structure of an $\mathcal{I}$-semidefinite matrix X by working on the (not necessarily square) factors defining the semidefinite blocks and guarantee the definiteness properties of the resulting matrix by construction. Working directly with the factors of X is additionally appealing in view of the fact that the factors B and C are often available a priori in control theory. Furthermore, in this way we avoid forming the Gram matrices C^*C and BB^* where a possible loss of accuracy might occur.

9   References

[1] Andrei A. Agrachev and Yuri L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[2] Gregory Ammar and Volker Mehrmann. On Hamiltonian and symplectic Hessenberg forms. *Linear Algebra Appl.*, 149:55 – 72, 1991.

[3] Peter Benner, Alan J. Laub, and Volker Mehrmann. A collection of benchmark examples for the numerical solution of algebraic Riccati equations I: the continuous-time case. Technical Report SPC 95-22, Forschergruppe 'Scientific Parallel Computing', Fakultät für Mathematik, TU Chemnitz-Zwickau, 1995. Version dated February 28, 1996.

[4] Michele Benzi, Gene H. Golub, and Jörg Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, 2005.

[5] Dario A. Bini, Bruno Iannazzo, and Beatrice Meini. *Numerical Solution of Algebraic Riccati Equations*. Society for Industrial and Applied Mathematics, 2011.

[6] Delin Chu, Xinmin Liu, and Volker Mehrmann. A numerical method for computing the Hamiltonian Schur form. *Numer. Math.*, 105(3):375–412, 2007.

[7] Richard J. Duffin, Dov Hazony, and Norman Morrison. Network synthesis through hybrid matrices. *SIAM J. Appl. Math.*, 14(2):390–413, 1966.

[8] Heike Fassbender. *Symplectic methods for the symplectic eigenproblem*. Kluwer Academic/Plenum Publishers, New York, 2000.

[9] James H. Goodnight. A tutorial on the SWEEP operator. *The American Statistician*, 33(3):149–158, 1979.

[10] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second edition, 2002.

[11] Nicholas J. Higham. *J*-orthogonal matrices: Properties and generation. *SIAM Rev.*, 45(3):504–519, September 2003.

[12] Leslie Hogben, editor. *Handbook of linear algebra*. Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, second edition, 2014.

[13] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, second edition, 2013.

[14] Peter Lancaster and Leiba Rodman. *Algebraic Riccati equations*. Oxford University Press, Oxford, 1995.

[15] A.J. Laub. A Schur method for solving algebraic Riccati equations. *IEEE Trans. Automat. Control*, 24(6):913–921, 1979.

[16] Wen-Wei Lin, Volker Mehrmann, and Hongguo Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302/303:469–533, 1999.

[17] Volker Mehrmann. *The autonomous linear quadratic control problem*, volume 163 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1991. Theory and numerical solution.

[18] Volker Mehrmann and Federico Poloni. Doubling algorithms with permuted Lagrangian graph bases. *SIAM J. Matrix Anal. Appl.*, 33(3):780–805, 2012.

[19] Chris Paige and Charles Van Loan. A Schur decomposition for Hamiltonian matrices. *Linear Algebra Appl.*, 41:11 – 32, 1981.

[20] Federico Poloni. Permuted graph matrices and their applications. In Peter Benner, Matthias Bollhöfer, Daniel Kressner, Christian Mehl, and Tatjana Stykel, editors, *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 107–129. Springer International Publishing, 2015.

[21] G. W. Stewart and Ji Guang Sun. *Matrix perturbation theory.* Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1990.

[22] Michael Stewart and G. W. Stewart. On hyperbolic triangularization: Stability and pivoting. *SIAM J. Matrix Anal. Appl.*, 19(4):847–860, 1998.

[23] Michael J. Tsatsomeros. Principal pivot transforms: properties and applications. *Linear Algebra Appl.*, 307(1-3):151–165, 2000.

[24] Albert W. Tucker. Principal pivotal transforms of square matrices. *SIAM Rev.*, 5(3):305, 1963.

[25] Nanny Wermuth, Michael Wiedenbeck, and David R. Cox. Partial inversion for linear systems and partial closure of independence graphs. *BIT*, 46(4):883–901, 2006.