Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007



# Explicit SOS Decompositions of Univariate Polynomial Matrices and the Kalman-Yakubovich-Popov Lemma

Erin M. Aylward, Sleiman M. Itani, and Pablo A. Parrilo

*Abstract*— **The purpose of this paper is twofold. The first is to make explicit the relationship between sum of squares (SOS) decompositions of univariate polynomial matrices and the Kalman-Yakubovich-Popov (KYP) lemma. The second is to present an efficient algorithm for explicitly finding an SOS decomposition of such matrices, inspired by the Hamiltonian-type methods for the solution of Riccati equations.**

## I. INTRODUCTION

Methods based on sum of squares (SOS) decomposition of multivariate polynomials have found numerous applications in systems, control, and optimization; see for instance [14] and the edited volume [6]. Of particular interest with regard to applications is the development of numerical and computational methods that exploit the structure present in SOS problems.

Besides nonnegative scalar polynomials, in many applications we are also faced with multivariate polynomial *matrices*, that must be positive semidefinite for all values of the indeterminates. Just like in the univariate case, it is possible to define a computationally convenient relaxation, using the concept of *sum of squares matrices*.

Although SOS matrices can be interpreted as a special case of the standard SOS formulation for scalar polynomials, they appear often enough in applications that a special study of their properties is warranted. In the terminology of Choi, Lam, and Reznick [1], univariate SOS matrices are equivalent to $(2, n; m, 2)$ *biforms*, which are homogeneous polynomials in two distinct sets of variables, quadratic in the first set, and bivariate in the second. SOS matrices were defined by Gatermann and Parrilo in [4] in the context of SOS problems invariant under the action of a finite group, and by Hol and Scherer [19] and Kojima [10] for relaxations of polynomial optimization problems.

An important result, rediscovered a number of times in a variety of contexts, is the fact that a univariate positive semidefinite matrix is necessarily an SOS matrix. This is a natural simultaneous generalization of the two classical cases (univariate and quadratic polynomials, respectively) where nonnegativity and sum of squares conditions are known to be equivalent. This statement was proved, in somewhat different contexts, by Djoković [2], Yakubovich [23], Popov [15], and Rosenblum-Rovnyak [18]. In particular, Choi, Lam and Reznick [1] give a constructive proof that any positive semidefinite biform of bidegree $(m, 2)$ is a sum of $2n$ squares of biforms in [1]. This "assertion that an $n$-ary quadratic form $\sum_{i,j} f_{ij}(y, z)x_i x_j$ ($f_{ij} \in H_m(\mathbb{R}^2)$) that is psd for every fixed $(y, z)$ is a sum of squares of forms that are $\mathbb{R}[y, z]$-linear in the $x_i$'s" [17]. In the signal processing and control context, this statement can be shown to be essentially equivalent to the spectral factorization theorem for vector-valued stochastic processes (see, e.g., [9]), or the factorization corollary of the Kalman-Yakubovich-Popov (KYP) lemma.

Our objective in this paper is partly expository, but also includes original research contributions. For the first part, we explore and explain some of the very interesting links between SOS matrices and the KYP lemma in detail, which based on our informal survey are not sufficiently well-known, even among experts. For the second, we develop an efficient algorithm to explicitly obtain an SOS decomposition of a positive semidefinite univariate polynomial matrix, using only standard matrix factorizations, with no need to solve an optimization problem. This method is essentially the analogue of the Hamiltonian approach to Riccati equations ([12]). To the best of our knowledge (and much to our surprise), this method does not seem to have been explicitly reported in the literature (see Section IV-A for a discussion of related work).

The explicit algorithm we present to find the sum of squares decomposition is much faster and reliable than conventional methods that use semidefinite programs. An interesting possibility, to be explored elsewhere, is whether is it possible to use this technique as an efficient subroutine for the much more complicated problems with multivariate matrix inequalities.

The paper is organized as follows: In Section II we present the basic notation and background material. In Section III we describe a condition that is equivalent to the Hamiltonian part of the KYP lemma. In Section IV we use that result to prove that a univariate polynomial matrix is SOS if and only if it is positive definite. In this section we also describe our efficient algorithm for factorizing positive definite matrix polynomials inspired by the Hamiltonian-based methods to solve Riccati equations. These Hamiltonian-based methods are often used to obtain spectral factors in the spectral factorization corollary in the KYP Lemma. In Section VI we present some examples, followed by concluding remarks and a discussion of possibilities for future work.

This research was funded in part by AFOSR MURI subawards 2003-07688-1 and 102-1080673.

E.M. Aylward is with BBN Technologies, Cambridge MA 02139, eaylward@bbn.com

S. Itani and P.A. Parrilo are with the Laboratory for Information and Decision Systems, Cambridge, MA 02139, {ssolomon,parrilo}@mit.edu

1-4244-1498-9/07/$25.00 ©2007 IEEE.

## II. Notation and Preliminaries

In this section we describe the notation we use and present some background material.

### A. Notation

We use the following standard notation. The set of $n \times n$ real symmetric matrices is $\mathcal{S}^n$. The set of polynomials in the scalar variable $x$ with real coefficients is denoted by $\mathbb{R}[x]$. The $p \times n$ matrices with entries in $\mathbb{R}[x]$ will be denoted by $\mathbb{R}^{p\times n}[x]$.

### B. Nonnegativity and Sum of Squares

A multivariate polynomial $p(x) \in \mathbb{R}[x]$, where $x = [x_1, ..., x_m]$, is *nonnegative* or *positive semidefinite* (PSD) if

$$p(x) \geq 0 \qquad \text{for all } x \in \mathbb{R}^m. \tag{1}$$

A multivariate polynomial $p(x) \in \mathbb{R}[x]$ is a *sum of squares* (SOS) if there exist polynomials $f_1(x), ..., f_s(x) \in \mathbb{R}[x]$ such that

$$p(x) = \sum_{i=1}^{s} f_i^2(x). \tag{2}$$

The existence of an SOS representation is a sufficient condition for its global nonnegativity, i.e., equation (2) clearly implies that (1) holds. In the univariate case, it is well-known that the converse also holds, i.e., nonnegativity and SOS are equivalent.

These definitions can be naturally extended from scalar polynomials to symmetric polynomial matrices. A symmetric polynomial matrix $P(x) \in \mathbb{R}[x]^{n\times n}$ is PSD if

$$P(x) \succeq 0 \qquad \text{for all } x \in \mathbb{R}^n, \tag{3}$$

and it is positive definite (PD) if

$$P(x) \succ 0 \qquad \text{for all } x \in \mathbb{R}^n. \tag{4}$$

*Definition 1:* A symmetric polynomial matrix $P(x) \in \mathbb{R}[x]^{n\times n}$, $x \in \mathbb{R}^m$ is a *sum of squares (SOS) matrix* if the scalar polynomial $y^\top P(x) y$ is a sum of squares in $\mathbb{R}[x, y]$. Equivalently, $P(x)$ is an SOS matrix if there exists a factorization $P(x) = F^\top(x)F(x)$ where $F(x) \in \mathbb{R}[x]^{s\times n}$.

Clearly, just like in the scalar case, $P(x)$ being an SOS matrix is a sufficient condition for a polynomial matrix to be positive semidefinite for all values of the indeterminates.

### C. The KYP lemma

The Kalman-Yakubovich-Popov (KYP) lemma is a cornerstone in the analysis of control systems. It establishes the equivalence between four distinct characterizations of the nonnegativity of a rational function, each of independent interest.

For simplicity of presentation, we state the results in purely algebraic terms, without reference to their important interpretations in the context of systems and control. For this, and the corresponding proofs, we refer the reader to well-known papers such as [22], [16], or any modern textbook (e.g., [24], [3]). The version of the KYP lemma presented below corresponds to the usual case where $R \succ 0$, and $A$ has no imaginary axis eigenvalues (these conditions are somewhat restrictive, and can be weakened).

*Lemma 1 (Kalman-Yakubovich-Popov):* Consider real matrices $A, B, C, Q, S, R$, of compatible dimensions, where $A$ has no imaginary axis eigenvalues and $R \succ 0$. The following four statements are equivalent:

- Frequency domain inequality (FDI):
$$\Pi(j\omega) = \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix} \succ 0 \tag{5}$$
for all $\omega \in \mathbb{R}$.
- Linear matrix inequality (LMI). There exists a symmetric matrix $X$ such that
$$\begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} + \begin{bmatrix} A^\top X + XA & XB \\ B^\top X & 0 \end{bmatrix} \succ 0. \tag{6}$$
- Hamiltonian: The infinitesimally symplectic matrix
$$\begin{bmatrix} A - BR^{-1}S^\top & -BR^{-1}B^\top \\ -Q + SR^{-1}S^\top & -A^\top + SR^{-1}B^\top \end{bmatrix} \tag{7}$$
has no eigenvalues on the imaginary axis.
- Riccati equation. There exists a symmetric $X$ such that
$$Q + A^\top X + XA - (S + XB)R^{-1}(S + XB)^\top = 0. \tag{8}$$

The spectral factorization corollary of the KYP lemma states that if $\Pi(j\omega)$ is positive definite on the imaginary axis it can be factored as

$$\Pi(j\omega) = \Psi^*(j\omega)\Psi(j\omega), \tag{9}$$

where

$$\Psi(j\omega) = R^{\frac{1}{2}}[I + R^{-1}(S^\top + B^\top X)(j\omega I - A)^{-1}B], \tag{10}$$

and where $R^{\frac{1}{2}}$ is the matrix square-root of $R$.

The KYP lemma is also equivalent to a number of other important results in systems theory. Among them we mention the exactness of the LMI upper bound for the structured singular value $\mu$ in the case of one full and one scalar block [13]. There have also been a number of interesting extensions of the basic KYP lemma in a number of directions, such as finite frequency ranges [7] or the behavioral setting [21].

### D. Block Companion Matrices and Matrix Polynomials

We will represent our polynomial matrix $P(x) \in \mathbb{R}^{n\times n}[x]$ of degree $2d$ in the form:

$$P(x) = \sum_{k=0}^{2d} P_k x^k, \tag{11}$$

where $P_k \in \mathcal{S}^n$, $k = 0, \ldots, 2d$. For simplicity, we will assume that the matrix $P(x)$ is strictly positive definite at infinity (i.e., $P_{2d} \succ 0$). This assumption can be relaxed, at the expense of a slightly more complicated formulation (e.g., eigenvalues of matrix pencils rather than standard eigenvalues), as is usual in the Riccati case.

Under this assumption, by pre- and post-multiplying by $P_{2d}^{-1/2}$, we can always normalize the matrix polynomial so that it is *monic*, i.e., the leading term satisfies $P_{2d} = I_n$. We assume this for the rest of the paper.

The block companion matrix of $P(x)$ is the $2dn \times 2dn$ real matrix $C_P$ defined as:

$$C_P := \begin{bmatrix} 0 & I_n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_n \\ -P_0 & -P_1 & \dots & -P_{2d-1} \end{bmatrix}. \tag{12}$$

It is well-known, and easy to prove, that $\det(xI - C_P) = \det P(x)$. Since $P_{2d} = I_n$, there are $2dn$ values of $x$ (counted with multiplicities) that make $P(x)$ singular. These values $x_i$ of $x$ are called the *latent roots* of $P(x)$, and the vectors $v_i$ such that $P(x_i)v_i = 0$ are called the *right latent vectors*; see e.g. [8].

## III. Companion matrix eigenvalues of positive definite polynomials

In this section we derive a condition for the positive definiteness of a matrix $P(x)$ that is similar to the Hamiltonian condition in the KYP lemma.

*Theorem 1:* If $P(x)$ is monic, the following two statements are equivalent:

1) $P(x) \succ 0$ for all $x \in \mathbb{R}$.
2) $C_P$ has no eigenvalues on the real axis.

*Proof:* $(1 \Rightarrow 2)$: If there is an $x_0 \in \mathbb{R}$ that is an eigenvalue of $C_P$, then $\det(P(x_0)) = \det(x_0 I - C_P) = 0$.

$(2 \Rightarrow 1)$: Assume that there is an $x_0$ in $\mathbb{R}$ such that $P(x_0) \not\succ 0$. At least one of the eigenvalues of $P(x_0)$ is not positive, denote that eigenvalue $\sigma(x_0)$. If $\sigma(x_0) = 0$, then $\det(x_0 I - C_P) = 0$ and $x_0$ is a real eigenvalue of $C_P$. Otherwise $\sigma(x_0) < 0$. Since $P(x)$ is monic we have $\lim_{x\to\infty} \frac{P(x)}{x^{2d}} = I \succ 0$, and thus $\lim_{x\to\infty} \sigma(x) > 0$. Therefore, by the continuity of the eigenvalues of a polynomial matrix, there is an $x_1 > x_0$ such that $\sigma(x_1) = 0$. Therefore $\det(x_1 I - C_P) = 0$ and so $x_1$ is a real eigenvalue of $C_P$. ∎

## IV. Factorization of Univariate PSD Polynomial Matrices

In this section we describe a simple and explicit algorithm to efficiently factorize a given polynomial matrix $P(x)$. The method is inspired by the Hamiltonian-based methods to solve Riccati equations (e.g., [12]). Despite the considerable amount of earlier work in this area, we believe our method has not been previously reported in the literature.

*Algorithm 1:* Given a polynomial matrix $P(x)$, of the form (11) with $P_{2d} = I_n$:

1) Form the companion matrix $C_P \in \mathbb{R}^{2dn\times 2dn}$, as defined in (12).
2) If $C_P$ has any purely real eigenvalues, then stop: there exists a real $x \in \mathbb{R}$, $v \in \mathbb{R}^n$ such that $P(x)v = 0$, and thus $P(x) \not\succ 0$.
3) Construct orthogonal bases of the invariant subspaces of $C_P$ (e.g., using a complex Schur factorization),
$$C_P = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} \Lambda_1 & \Gamma \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}^* \tag{13}$$
where $\Lambda_1$ and $\Lambda_2$ have conjugate spectra (i.e., $\sigma(\Lambda_1) = \sigma(\Lambda_2)^*$).
4) Let $Q := VU_{11}^{-1}$, where $V$ is the submatrix of $U_{21}$ corresponding to its first $n$ rows. This matrix has dimensions $n \times dn$. Let $Q_r$ and $Q_i$ be the real and imaginary parts of $Q$, respectively.
5) The factor $F(x)$ is then given by:
$$F(x) = \begin{bmatrix} -Q_r & I_n \\ -Q_i & 0_n \end{bmatrix} \begin{bmatrix} I_n \\ xI_n \\ \vdots \\ x^d I_n \end{bmatrix}.$$

A proof that this algorithm gives a correct SOS factorization of $P(x)$ is given in the Appendix.

### A. Related work

Besides the references already cited, there have been other approaches to the spectral factorization of polynomial matrices. Generally speaking, these have focused on the parahermitian case (i.e., $P(s) = P^\top(-s)$, where $s$ is a complex variable), although it is not too difficult to adapt some of the results from the imaginary axis to the real line. In particular, we highlight the results of Kwakernaak and Šebek [11] and Trentelman and Rapisarda [20]. In [11], the authors present several algorithms for $J$-spectral factorization. Of these, their "interpolation" based method most closely resembles ours, except that the last three steps of their method are replaced by a simple linear system solving. The technique in [20] also has some similarities, except they rely on the solution of LMIs rather than a direct Hamiltonian approach.

## V. SOS matrices and KYP

The results described above allow us to formulate the following theorem, which is the exact SOS polynomial analogue of the KYP lemma.

*Theorem 2:* Consider a monic symmetric polynomial matrix $P(x) \in \mathbb{R}[x]^{n\times n}$ of degree $2d$, of the form (11). The following statements are equivalent:

- Positive definiteness:
$$P(x) \succeq 0, \qquad \forall x \in \mathbb{R} \tag{14}$$
- SOS factorization with bound on the number of squares:
$$P(x) = F(x)^\top F(x) \tag{15}$$
for some $F(x) \in \mathbb{R}[x]^{s\times n}$, with $s \le 2n$.
- "Gram matrix" SOS LMI (kernel version). There exists a matrix $G \in \mathcal{S}^{(d+1)n}$ such that $G \succeq 0$ and
$$P_k = \sum_{i+j=k} G_{ij}, \qquad k = 0, \dots, 2d,$$
where $G$ is partitioned in $(d+1)^2$ square blocks of size $n \times n$, and the indices range from $0$ to $d$.
- SOS LMI (image version). There exists a skew-symmetric matrix $S \in \mathbb{R}^{dn\times dn}$ ($S = -S^\top$) such that
$$Q + \begin{bmatrix} 0_{dn\times n} & S \\ 0_{n\times n} & 0_{n\times dn} \end{bmatrix} + \begin{bmatrix} 0_{n\times dn} & 0_{n\times n} \\ S & 0_{dn\times n} \end{bmatrix} \succeq 0, \tag{16}$$

where

$$Q := \frac{1}{2}\begin{bmatrix} 2P_0 & P_1 & \cdots & P_{d-1} & P_d \\ P_1 & 0 & \cdots & 0 & P_{d+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_{d-1} & 0 & \cdots & 0 & P_{2d-1} \\ P_d & P_{d+1} & \cdots & P_{2d-1} & 2P_{2d} \end{bmatrix}$$

Furthermore, $P(x)$ is strictly positive definite if and only if the matrix $C_P$ defined in (12) has no real eigenvalues.

*Proof:* The proofs of these statements are omitted for space reasons. They will appear in the full version of this paper. ■

## VI. Examples

In this section we present two examples demonstrating the application of our algorithm. The first one considers a $2 \times 2$ polynomial matrix, while the second example illustrates the case of roots with higher multiplicities.

*Example 1:* The first example we present is the strictly positive definite polynomial matrix

$$P(x) = \begin{bmatrix} x^2 - 2x + 2 & x \\ x & x^2 + 1 \end{bmatrix}.$$

The corresponding companion matrix is

$$C_P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 0 & 2 & -1 \\ 0 & -1 & -1 & 0 \end{bmatrix}.$$

The eigenvalues of $C_P$ are

$$\begin{array}{rcl} 1.1898 & + & 0.6028i, \\ 1.1898 & - & 0.6028i \\ -0.1898 & + & 1.0432i \\ -0.1898 & - & 1.0432i \end{array}$$

none of which are purely real. The matrix $Q$, calculated according to the algorithm in Section IV, is

$$Q = \begin{bmatrix} 1.0000 + 0.7585i & -0.3938 + 0.2393i \\ -0.6062 + 0.2393i & -0.0000 + 0.8875i \end{bmatrix}.$$

From this, we construct the corresponding factor $F(x)$ as:

$$F(x) = \begin{bmatrix} x - 1 & 0.3938 \\ 0.6062 & x \\ -0.7585 & -0.2393 \\ -0.2393 & -0.8875 \end{bmatrix},$$

and

$$P(x) = F^\top(x)F(x).$$

*Example 2:* Our second example is the scalar univariate positive definite polynomial given by

$$p(x) = x^4 + 2x^2 + 1.$$

The associated companion matrix is

$$C_p = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & -2 & 0 \end{bmatrix}.$$

The eigenvalues of the companion matrix $C_p$, or equivalently roots of the polynomial $p(x)$ are

$$-i, \quad -i, \quad +i, \quad +i.$$

Notice that we can group the eigenvalues in two different ways to satisfy the condition that if $x_i \in \Lambda_1$, then $x_i^* \in \Lambda_2$. Specifically, we can group them as $-i, -i \in \Lambda_1$ and $+i, +i \in \Lambda_2$, or we can group them as $-i, +i \in \Lambda_1$ and $+i, -i \in \Lambda_2$. We will see that these different groupings give rise to different factorizations. For the first case, we have that the matrix $Q$ is

$$Q = \begin{bmatrix} 1 & 2i \end{bmatrix},$$

and thus we obtain the factor

$$F_1(x) = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = \begin{bmatrix} x^2 - 1 \\ -2x \end{bmatrix}.$$

For the second grouping, we have

$$Q = \begin{bmatrix} -1 & 0 \end{bmatrix},$$

and

$$F_2(x) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = \begin{bmatrix} x^2 + 1 \\ 0 \end{bmatrix}.$$

We have then two different SOS decompositions

$$p(x) = F_1^\top(x)F_1(x) = F_2^\top(x)F_2(x),$$

or equivalently,

$$x^4 + 2x^2 + 1 = (x^2 - 1)^2 + (-2x)^2 = (x^2 + 1)^2.$$

## VII. CONCLUSIONS AND FUTURE WORK

We introduced a theorem for polynomial matrices that is the analogue of the KYP lemma, and presented an efficient algorithm to find an explicit SOS decomposition of univariate positive definite matrices.

One area for possible future work is to explore the possibility of using our procedure as an efficient subroutine in optimization problems involving multivariate polynomial matrix inequalities. Also of interest would be to explore the possibilities of computing the central solution of the inequality (the analytic center of the feasible set), an issue originally suggested by Genin, Nesterov, and van Dooren in the context of Riccati inequalities in [5].

## References


[1] M.D. Choi, T.Y Lam, and B. Reznick. Real zeros of positive semidefinite forms I. *Mathematische Zeitschrift*, 171:1–26, 1980.

[2] D.Z. Djoković. Hermitian matrices over polynomial rings. *Journal of Algebra*, 43:359–374, 1976.

[3] G.E. Dullerud and F. Paganini. *A Course in Robust Control Theory: A Convex Approach*. Springer Verlag, 1999.

[4] K. Gatermann and P.A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95–128, 2004.

[5] Y. Genin, Y. Nesterov, and P. van Dooren. The analytic center of LMIs and Riccati equations. In *Proceedings European Control Conference*, 1999.

[6] D. Henrion and A. Garulli, editors. *Positive polynomials in Control*, volume 312 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 2005.

[7] T. Iwasaki and S. Hara. Generalized KYP lemma: unified frequency domain inequalities with design applications. *IEEE Trans. Automat. Control*, 50(1):41–59, 2005.

[8] J.E. Dennis Jr., J.F. Traub, and R.P. Weber. The algebraic theory of matrix polynomials. *SIAM J. Numer. Anal.*, 13(6):831–845, 1976.

[9] T. Kailath, A.H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice-Hall, 2000.

[10] M. Kojima. Sums of squares relaxations of polynomial semidefinite programs. Research report B-397, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2003.

[11] H. Kwakernaak and M. Šebek. Polynomial $J$-spectral factorization. *IEEE Trans. Automat. Control*, 39(2):315–328, 1994.

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

[13] A. Packard and J. C. Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[14] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, May 2000.

[15] V.M. Popov. *Hyperstability of Control Systems*. Berlin-Heidelberg-New York: Springer, 1973.

[16] A. Rantzer. On the Kalman-Yakubovich-Popov lemma. *Systems & Control Letters*, 28:7–10, 1996.

[17] B. Reznick. Some concrete aspects of Hilbert's 17th problem. In *Contemporary Mathematics*, volume 253, pages 251–272. American Mathematical Society, 2000.

[18] M. Rosenblum and J. Rovnyak. The factorization problem for nonnegative operator valued functions. *Bull. Amer. Math. Soc.*, 77:287–318, 1971.

[19] C.W. Scherer and C.W.J. Hol. Matrix sum-of-squares relaxations for robust semi-definite programs. *Math. Program.*, 107(1-2, Ser. B):189–211, 2006.

[20] H.L. Trentelman and P. Rapisarda. New algorithms for polynomial $J$-spectral factorization. *Math. Control Signals Systems*, 12(1):24–61, 1999.

[21] R. van der Geest and H. Trentelman. The Kalman-Yakubovich-Popov lemma in a behavioural framework. *Systems Control Lett.*, 32(5):283–290, 1997.

[22] J.C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Transactions on Automatic Control*, 16(6):621–634, 1971.

[23] V.A. Yakubovich. Factorization of symmetric matrix polynomials. *Dok. Akad. Nauk. SSSR*, 194:532–535, 1970.

[24] K. Zhou. *Essentials of Robust Control*. Prentice Hall, Upper Saddle River, NJ, 1998.

## APPENDIX

### A. Proof of Factorization Algorithm in Section IV

In this appendix we will prove that the factorization algorithm given in Section IV works. For simplicity of notation, we assume that the companion matrix $C_P$ is diagonalizable, although this proof can be extended to the non-diagonalizable case. The algorithm in the diagonalizable case is as follows:

*Algorithm 2:* Given a monic positive definite polynomial matrix $P(x)$:

1) Diagonalize the block companion matrix $C_P$ into the form:

$$\begin{bmatrix} B_1 & B_2 \\ B_1\Lambda^d & B_2(\Lambda^*)^d \end{bmatrix} \underbrace{\begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda^* \end{bmatrix}}_{D} \begin{bmatrix} B_1 & B_2 \\ B_1\Lambda^d & B_2(\Lambda^*)^d \end{bmatrix}^{-1} \tag{17}$$

where $\Lambda$ and $\Lambda^*$ are diagonal matrices containing the eigenvalues of $C_P$. If

$$\Lambda = \operatorname{diag}(\begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_{nd} \end{bmatrix}),$$

then

$$B_1 = \begin{bmatrix} b_1 & b_2 & \cdots & b_{nd} \\ \lambda_1 b_1 & \lambda_2 b_2 & \cdots & \lambda_{nd} b_{nd} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{d-1} b_1 & \lambda_2^{d-1} b_2 & \cdots & \lambda_{nd}^{d-1} b_{nd} \end{bmatrix}, \tag{18}$$

$$B_2 = \bar{B}_1, \tag{19}$$

where $b_i \in \mathbb{R}^{n\times 1}$, and $\bar{B}_1$ denotes the conjugation of the entries of $B_1$.

2) Define $C_Q := B_1 \Lambda B_1^{-1}$, and Let $Q$ denote the last block row $(n \times dn)$ of $C_Q$, and let $Q_r$ and $Q_i$ be the real and imaginary parts of $Q$, respectively. Note that this is equivalent to the $VU_{11}^{-1}$ in the formulation in section IV.

3) The factor $F(x)$ such that $P(x) = F^\top(x)F(x)$ is then given by:

$$F(x) = \begin{bmatrix} -Q_r & I_n \\ -Q_i & 0_n \end{bmatrix} \begin{bmatrix} I_n \\ xI_n \\ \vdots \\ x^d I_n \end{bmatrix}. \tag{20}$$

*Proof:* The proof is divided in two parts. In the first, we will show that the (complex) polynomial matrix $Q(s) := \det(sI - C_Q)$ satisfies $P(s) = Q^*(s)Q(s)$. We will do this by proving $P^{-1}(s) = Q^{-1}(s)Q^{-*}(s)$. In the second part we will show that the real factor $F(x)$ satisfies $P(x) = F^\top(x)F(x)$.

To prove $P^{-1}(s) = Q^{-1}(s)Q^{-*}(s)$, consider the following two state-space representations:

$$\begin{aligned} A &= C_P \\ B &= \begin{bmatrix} 0 & \cdots & 0 & I_n \end{bmatrix}^\top \\ C &= \begin{bmatrix} I_n & \cdots & 0 & 0 \end{bmatrix}, \end{aligned} \tag{21}$$

and

$$\begin{aligned} \tilde{A} &= \begin{bmatrix} C_Q & E \\ 0 & C_{Q^*} \end{bmatrix} \\ \tilde{B} &= \begin{bmatrix} 0 & \cdots & 0 & I_n \end{bmatrix}^\top \\ \tilde{C} &= \begin{bmatrix} I_n & \cdots & 0 & 0 \end{bmatrix}, \end{aligned} \tag{22}$$

where $E$ is a matrix with all zeros except the lower left $n \times n$ block which is the identity matrix. The matrix $C_Q$ is the companion matrix defined in Step 2 of the algorithm above, and $C_{Q^*}$ is the companion matrix of $Q^*(s)$.

It can be easily verified that the $n \times n$ transfer function associated to the system $(A, B, C)$ is $P^{-1}(s)$ and that the transfer function associated to $(\tilde{A}, \tilde{B}, \tilde{C})$ is $Q^{-1}(s)Q^{-*}(s)$. We will show next that the two state-space representations are equivalent up to a linear state transformation, and thus they have the same transfer function.

Let

$$T := T_2 T_1^{-1},$$

where

$$T_1 = \begin{bmatrix} B_1 & B_2 \\ B_1\Lambda^d & B_2(\Lambda^*)^d \end{bmatrix}, \tag{23}$$

$$T_2 = \begin{bmatrix} B_1 & B_2 \\ 0 & B_3 \end{bmatrix}, \tag{24}$$

and

$$B_3 = \begin{bmatrix} v_1 & v_2 & \cdots & v_{nd} \\ \lambda_1^* v_1 & \lambda_2^* v_2 & \cdots & \lambda_{nd}^* v_{nd} \\ \vdots & \vdots & \vdots & \vdots \\ (\lambda_1^*)^{d-1} v_1 & (\lambda_2^*)^{d-1} v_2 & \cdots & (\lambda_{nd}^*)^{d-1} v_{nd} \end{bmatrix} \tag{25}$$

with $v_i := Q(\lambda_i^*)\bar{b}_i$. Notice that the vectors $v_i$ satisfy $Q^*(\lambda_i^*)v_i = 0$.

To prove that $T$ is a similarity transformation between systems (21) and (22) we need to show that $\tilde{A} = TAT^{-1}$, $T^{-1}\tilde{B} = B$ and $\tilde{C} = CT^{-1}$. Notice first that:

$$\begin{aligned} \tilde{A}T_2 &= \begin{bmatrix} C_Q & E \\ 0 & C_{Q^*} \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ 0 & B_3 \end{bmatrix} \\ &= \begin{bmatrix} C_Q B_1 & C_Q B_2 + E B_3 \\ 0 & C_{Q^*} B_3 \end{bmatrix} \\ &= \begin{bmatrix} B_1 \Lambda & B_2 \Lambda^* \\ 0 & B_3 \Lambda^* \end{bmatrix} \\ &= T_2 D. \end{aligned} \tag{26}$$

Equality holds for the top-left block of equation (26) since we constructed $C_Q$ as $C_Q = B_1 \Lambda B_1^{-1}$. It is also straightforward to show the lower-right block equality (since $Q^*(\lambda_i^*)v_i = 0$). For the top-right block we have:

$$C_Q B_2 = \begin{bmatrix} \lambda_1^* \bar{b}_1 & \cdots & \lambda_{nd}^* \bar{b}_{nd} \\ \vdots & \ddots & \vdots \\ (\lambda_1^*)^{d-1} \bar{b}_1 & \cdots & (\lambda_{nd}^*)^{d-1} \bar{b}_{nd} \\ -Q(\lambda_1^*)\bar{b}_1 + (\lambda_1^*)^d \bar{b}_1 & \cdots & -Q(\lambda_{nd}^*)\bar{b}_{nd} + (\lambda_{nd}^*)^d \bar{b}_{nd} \end{bmatrix}$$

$$E B_3 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ v_1 & \cdots & v_{nd} \end{bmatrix}. \tag{27}$$

Using the fact that $v_i = Q(\lambda_i^*)\bar{b}_i$,

$$\begin{aligned} C_Q B_2 + E B_3 &= \begin{bmatrix} \lambda_1^* \bar{b}_1 & \cdots & \lambda_{nd}^* \bar{b}_{nd} \\ \vdots & \ddots & \vdots \\ (\lambda_1^*)^{d-1} \bar{b}_1 & \cdots & (\lambda_{nd}^*)^{d-1} \bar{b}_{nd} \\ (\lambda_1^*)^{d} \bar{b}_1 & \cdots & (\lambda_{nd}^*)^{d} \bar{b}_{nd} \end{bmatrix} \\ &= B_2 \Lambda^*. \end{aligned}$$

Since $\tilde{A}T_2 = T_2 D$, we have

$$\tilde{A} = T_2 D T_2^{-1} = T_2 T_1^{-1} T_1 D T_1^{-1} T_1 T_2^{-1} = TAT^{-1}.$$

We now show that $T$ also satisfies $T^{-1}\tilde{B} = B$ and $\tilde{C} = CT^{-1}$:

$$\begin{aligned} T^{-1} &= T_1 T_2^{-1} \\ &= \begin{bmatrix} B_1 & B_2 \\ B_1\Lambda^d & B_2(\Lambda^*)^d \end{bmatrix} \begin{bmatrix} B_1^{-1} & -B_1^{-1} B_2 B_3^{-1} \\ 0 & B_3^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ B_1 \Lambda^d B_1^{-1} & L \end{bmatrix}, \end{aligned} \tag{28}$$

where $L$ is given by

$$L = -B_1 \Lambda^d B_1^{-1} B_2 B_3^{-1} + B_2(\Lambda^*)^d B_3^{-1}. \tag{29}$$

It can be shown that $L$ is a lower triangular matrix with $n \times n$ block identity matrices on the diagonal. Hence $T^{-1}\tilde{B} = B$ and $\tilde{C} = CT^{-1}$, since

$$\begin{aligned} \tilde{B} = B &= \begin{bmatrix} 0 & \cdots & 0 & I_n \end{bmatrix}^\top \\ \tilde{C} = C &= \begin{bmatrix} I_n & \cdots & 0 & 0 \end{bmatrix}. \end{aligned}$$

With this we have shown that a similarity transformation exists between systems (21) and (22), and therefore that $P^{-1}(s) = Q^{-1}(s)Q^{-*}(s)$. It follows immediately that $P(s) = Q^*(s)Q(s)$.

We now prove that a factor $F(x)$ with real coefficients and satisfying $P(x) = F^\top(x)F(x)$ can be constructed from $Q(x)$ according to equation (20). For this, first note that if $Q(x) = Q_r(x) + iQ_i(x)$, and if we construct $F(x)$ as

$$F(x) = \begin{bmatrix} Q_r(x) \\ Q_i(x) \end{bmatrix}, \tag{30}$$

then we have

$$\begin{aligned} P(x) &= Q_r^\top(x)Q_r(x) + Q_i^\top(x)Q_i(x) \\ &\quad + i[Q_r^\top(x)Q_i(x) - Q_i^\top(x)Q_r(x)] \\ &= F^\top(x)F(x). \end{aligned}$$

Since the last block row of the companion matrix $C_Q$ is the matrix $Q$, we have

$$Q(x) = \begin{bmatrix} -Q & I_n \end{bmatrix} \begin{bmatrix} I_n \\ xI_n \\ \vdots \\ x^d I_n \end{bmatrix}.$$

From equation (30), this is equivalent to

$$F(x) = \begin{bmatrix} -Q_r & I_n \\ -Q_i & 0_n \end{bmatrix} \begin{bmatrix} I_n \\ xI_n \\ \vdots \\ x^d I_n \end{bmatrix}.$$

■