Order Bound for the Realization of a Combination of Positive Filters

Béla Nagy, Máté Matolcsi, Márta Szilvási

Béla Nagy is with the Department of Analysis, Technical University of Budapest (BME), Budapest, H-1111, Hungary (e-mail: bnagy@math.bme.hu). His work was supported by OTKA-T047276.

Máté Matolcsi is with the Alfréd Rényi Institute of Mathematics, Budapest, H-1053, Hungary (e-mail matomate@renyi.hu). His work was supported by OTKA-T047276, T049301, PF64601 and by the Hungarian-Spanish Scientific and Technological Governmental Cooperation, no. E-38/04.

Márta Szilvási is with the Department of Geometry, Technical University of Budapest (BME), Budapest, H-1111, Hungary (e-mail: szilvasi@math.bme.hu). Her work was supported by OTKA-T047276.

Abstract

In a problem on the realization of digital filters, initiated by Gersho and Gopinath [8], we extend and complete a remarkable result of Benvenuti, Farina and Anderson [4] on decomposing the transfer function $t(z)$ of an *arbitrary* linear, asymptotically stable, discrete, time-invariant SISO system as a difference $t(z) = t_1(z) - t_2(z)$ of two *positive, asymptotically stable* linear systems. We give an easy-to-compute *algorithm to handle the general problem*, in particular, also the case of transfer functions $t(z)$ with *multiple poles*, which was left open in [4]. One of the appearing positive, asymptotically stable systems is always 1-dimensional, while the other has dimension depending on the order and, in the case of nonreal poles, also on the location of the poles of $t(z)$. The appearing dimension is seen to be *minimal* in some cases and it can always be calculated *before* carrying out the realization.

Keywords

Positive linear systems, charge routing networks, discrete time filtering, positive realizations

I. Introduction

Assume we are given the transfer function

$$t(z) = \frac{p_1 z^{n-1} + ... + p_n}{z^n + q_1 z^{n-1} + ... + q_n} \quad ; \quad p_j, q_j \in \mathbb{R} \;\; \text{for all } 1 \leq \text{j} \leq \text{n} \tag{1}$$

of a *discrete time-invariant linear SISO system* of McMillan degree n. In many applications (e.g., digital filters) it is very important, hence we also assume (as in [4]), that $t(z)$ is *asymptotically stable*, a.s. in short, i.e. its poles lie in the open unit disk.

The *positive realization problem* is to find, if possible, a triple $A \in \mathbb{R}^{N\times N}$, $b, c \in \mathbb{R}^N$ with nonnegative entries, such that $t(z) = c^T(zI - A)^{-1}b$ holds, the *minimality problem* is to find the minimal possible value of N (clearly, $N \geq n$.) The nonnegativity condition in applications is a consequence of *underlying physical constraints* such as in the design of charge routing networks (CRN's, [8]). Due to the nonnegativity constraint, positive filters are restricted in their achievable performance. However, as suggested in [8], and elaborated in the seminal paper [4], one can try to decompose an *arbitrary a.s.* transfer function as the *difference of two positive asymptotically stable systems*, and thus remove the performance limitations and retain the advantages offered by CRN's at the same time. Therefore, we shall be interested in decompositions of the form $t(z) = t_1(z) - t_2(z)$, where $t_1(z)$ and $t_2(z)$ are a.s. transfer functions with positive realizations of dimensions N_1 and N_2. Preferably, one would like to have an *a priori* upper bound on the numbers N_1, N_2, in terms of the location and order of the poles of $t(z)$. We emphasize that a number of technical problems in the solution stems from *the requirement of asymptotic stability of* $t_1(z)$ *and* $t_2(z)$, which is perfectly reasonable from the point of view of engineering applications (e.g. realization of digital filters).

This *positive decomposition problem* was solved in [4] for a class of transfer functions $t(z)$. Indeed, under the assumption that $t(z)$ has *exclusively simple* (but possibly complex) *poles*, it was shown that we can take $N_2 = 1$, and a reasonably good upper estimate on the value of N_1 was presented (see Theorem 8 in [4]). The case of transfer functions with *multiple poles* was left open (see the Concluding Remarks of [4]). A slight improvement on the value of N_1 was given in [10], where nonnegative simple poles with negative residues were handled in a more efficient way. Later in [12], the open case of *nonnegative multiple poles* was settled.

In this paper we *solve the general problem* of transfer functions with possibly *multiple real and multiple nonreal* poles. Moreover, our approach here is *universal*, i.e. it provides a unified method for the solution of the positive decomposition problem for *any* a.s. transfer function $t(z)$ (see Theorem 4 below). In some cases we can claim minimality of the dimension N_1 (see Remark 4). At the end we illustrate on the example of a Chebyshev filter how the realization algorithm works.

For the general theory and applications of positive linear systems we refer the reader to [6], of digital filters to [5]. A thorough overview on positive realizations has recently been published in these Transactions [2]. For direct applications of the positive decomposition problem see [4] and [3]. Finally, we note that some technical ideas applied in the proofs appeared in another context in [7].

II. Notation and preliminaries

Throughout this paper we consider the transfer function $t(z)$ of a linear discrete time-invariant scalar system given in (1). We also assume that the partial fraction decomposition form of $t(z)$ is known, and $t(z)$ is asymptotically stable: all its poles lie within the open unit disk. We emphasize that the poles can have *any location and order* (apart from the obvious constraints arising from the real-valued coefficients: conjugate nonreal poles must have the same order). It is well known that such a function $t(z)$ has a *real minimal Jordan realization* (c, J, b) of order n (the McMillan degree), where the dimensions of the matrices are $1 \times n, n \times n, n \times 1$ (in that order), and we have $t(z) = c(zI - J)^{-1}b$. These matrices have exclusively real entries, and J is a *real Jordan* matrix, i.e. a direct sum (recall that $A \oplus B$ simply means the block matrix $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$) of *real Jordan blocks* of the indicated orders:

$$J = [\oplus_{i=1}^{I} J(r_i; p_i)] \oplus [\oplus_{k=1}^{K} J(x_k, y_k; q_k)].$$

Here the first I terms have the real eigenvalues r_i and are of order p_i, and the terms of the type

$$J(x, y; q) := \begin{pmatrix} x & y & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -y & x & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & x & y & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & -y & x & 0 & 1 & \dots & 0 & 0 \\ \ddots & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & x & y \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & -y & x \end{pmatrix} \tag{2}$$

have the pairs of conjugate nonreal eigenvalues $x_k + iy_k, x_k - iy_k$ and are of order $2q_k$ $(k = 1, \dots K)$. Note that, since the realization is *minimal*, to each (pair of) eigenvalue(s) there corresponds exactly 1 real Jordan block. We shall consider the corresponding partitions (direct sums) of the matrices (vectors) c, b:

$$c = [\oplus_{i=1}^{I} c^{r_i}] \oplus [\oplus_{k=1}^{K} c^k], \quad b = [\oplus_{i=1}^{I} b^{r_i}] \oplus [\oplus_{k=1}^{K} b^k],$$

and also the corresponding *direct sums of (Jordan) realizations*:

$$(c, J, b) = [\oplus_{i=1}^{I} (c^{r_i}, J(r_i, p_i), b^{r_i})] \oplus [\oplus_{k=1}^{K} (c^k, J(x_k, y_k; q_k), b^k)].$$

As stated in the Introduction, we are looking for *positive asymptotically stable* systems $t_1(z)$ and $t_2(z)$ such that $t_2(z)$ is 1-dimensional and the decomposition $t(z) = t_1(z) - t_2(z)$ holds. In our construction of a positive realization of $t_1(z)$ we will make use of the following well-known result from positive system theory (see e.g. [1]):

Lemma 1: Let $t(z)$ be a rational transfer function as in (1), and let (c, A, b) be *any minimal realization* of $t(z)$, i.e. $t(z) = c(zI - A)^{-1}b$, and the dimensions of the matrices (c, A, b) are $1 \times n, n \times n, n \times 1$, respectively. Assume that there exists a *system invariant polyhedral cone* $\mathcal{P} \subset \mathbb{R}^n$, i.e. a finitely generated cone $\mathcal{P} \subset \mathbb{R}^n$ such that $b \in \mathcal{P}$, $A\mathcal{P} \subset \mathcal{P}$ and $c \cdot p \geq 0$ for all $p \in \mathcal{P}$. If the number of extremal rays of $\mathcal{P}$ is N, then there exists a positive realization of $t(z)$ of dimension N. $\square$

In the space $\mathbb{R}^n$ we shall use the l^1 norm of a vector $v = (v_1, \dots, v_n)$ defined by $||v||_1 := \sum_{j=1}^{n} |v_j|$, and for an operator A (identified with its matrix of order n with respect to the canonical basis) the induced matrix norm (called *the maximal column norm* in [11]) $||A|| = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ji}|$. In a similar vein, we shall need the Euclidean l^2 vector norm on $\mathbb{R}^n$ and the induced operator norm (called *the spectral norm* in [11]) $||A||_2$.

Let $z = x + iy \in \mathbb{C} \setminus \mathbb{R}$, and consider $J(z, k) \oplus J(\bar{z}, k)$, the direct sum of two (nonreal) Jordan blocks of orders k with eigenvalues z and $\bar{z}$, respectively. Let $C \equiv C(x, y) := \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$, and let $D := \oplus_{h=1}^{k} C(x, y)$. For any real number $f \neq 0$ let $N(f)$ denote the nilpotent matrix of order $2k$ with entries $n_{j,j+2} := f$ $(j = 1, \dots, 2k-2)$ and 0 otherwise, and let

$$M(x, y; f, k) \equiv M(f) := D + N(f) = \begin{pmatrix} x & y & f & 0 & 0 & 0 & \dots & 0 & 0 \\ -y & x & 0 & f & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & x & y & f & 0 & \dots & 0 & 0 \\ 0 & 0 & -y & x & 0 & f & \dots & 0 & 0 \\ \ddots & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & x & y \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & -y & x \end{pmatrix}. \tag{3}$$

Then the matrix $M(1) = J(x, y; k)$ is *the canonical real Jordan form* of $J(z, k) \oplus J(\bar{z}, k)$. Define the diagonal matrix $d \equiv d(f)$ by $d := diag(f^{k-1}, f^{k-1}, f^{k-2}, f^{k-2}, \dots, f, f, 1, 1)$. Then $M(f)d = dM(1)$, hence $M(f)$ can also be regarded as *a certain real Jordan form* of $J(z,k) \oplus J(\bar{z},k)$. For short we shall write M for $M(f)$, and we shall fix the value of f later.

Introduce the notation $r := |z|$, and let us estimate the operator norm $||M^m||$ *induced by the* l^1 *norm* $||\cdot||_1$ in the finite dimensional real vector space $\mathbb{R}^{2k}$. We call attention to the fact that in propositions and theorems below *conditions of the type* $0 < w < 1$ *will ensure asymptotic stability of the constructed transfer function.*

Proposition 1: With the notation above for every $f, w > 0$ such that $r + f < w < 1$ we can determine $Q \equiv Q(r,f,w) \in \mathbb{N}_0$ such that $m > Q$ implies $||M(f)^m w^{-m}|| < 1$.

Proof. It is well known that for any matrix $A := (a_{ji})$ acting in $\mathbb{R}^n$ we have $||A|| = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ji}| \le \sqrt{n}||A||_2$. Here $||A||_2$ denotes the *operator norm of* A *induced by the* l^2 *norm* in $\mathbb{R}^n$. In particular, denoting by * the (conjugate) transpose of any matrix, we have

$$C^{m*}C^m = \begin{pmatrix} (x^2+y^2)^m & 0 \\ 0 & (x^2+y^2)^m \end{pmatrix}.$$

The square root of the spectral radius of this matrix is $||C^m||_2 = |z|^m = r^m$. Hence $||C^m|| \le r^m\sqrt{2}$. Since D is the direct sum of k copies of C, we obtain $||D^m|| = ||C^m|| \le r^m\sqrt{2}$.

Since the matrices D and $N \equiv N(f)$ commute, for $m \in \mathbb{N}_0$ we have

$$||M^m|| = ||(D+N)^m|| \le \sum_{j=0}^{m} \binom{m}{j} ||D^{m-j}|| \cdot ||N||^j \le \sqrt{2} \sum_{j=0}^{m} \binom{m}{j} r^{m-j} f^j = \sqrt{2}(r+f)^m. \tag{4}$$

Denoting by $[q]$ the integer part of any real number q, define

$$Q(r,f,w) := [\frac{\log\sqrt{2}}{\log w - \log(r+f)}]. \tag{5}$$

Then, for $m > Q(r,f,w)$ we have $\left(\frac{r+f}{w}\right)^m < \frac{1}{\sqrt{2}}$, and hence $||M(f)^m w^{-m}|| \le \sqrt{2}\left(\frac{r+f}{w}\right)^m < 1$ by (4). $\square$

Remark 1. Formula (5) shows that, as expected, increasing the value of w will (or, at least, may) decrease, whereas increasing the value of $r = |z|$ or of f will (or may) increase the (needed) value of Q. $\square$

III. Decomposition results

With the help of Proposition 1 we can now solve the positive decomposition problem for transfer functions with complex multiple poles. As a preparation, we will first deal with the case of *one pair* of complex conjugate poles of higher order. The most general case (including poles of any location and order) is treated in Theorem 4 below.

Theorem 1: Assume that the transfer function $t(v)$ with real coefficients has exclusively the two non-real poles z and $\bar{z}$ such that $|z| < 1$. Then the orders of the poles are identical, say k. For every $f, w > 0$ such that $|z| = r < w < 1$ and $r + f < w$ define the nonnegative integer $Q = Q(r,f,w)$ as in formula (5). Then, for some appropriately large $p > 0$ the function $t_1(v) := t(v) + \frac{p}{v-w}$ has a nonnegative realization of dimension $N_1 \le 4k(Q+1)$.

Proof. Consider any *real* minimal realization (c_1, A_1, b_1). There is a *real* matrix S establishing system similarity of this realization with *a real Jordan* realization (c, A, b), which is clearly also minimal. Hence A has exactly the two (complex) elementary divisors $(v-z)^k$ and $(v-\bar{z})^k$. Consequently, A can be any of the matrices $M = M(f)$ defined in (3). Pick the numbers w and f as in the statement of the Theorem, and define $Q = Q(r, f, w)$ as in formula (5).

Let c_0, b_0 be positive numbers, let I_1 denote the 1×1 identity matrix, and consider the triple

$$[\begin{pmatrix} c_0 & c \end{pmatrix}, wI_1 \oplus M, \begin{pmatrix} b_0 \\ b \end{pmatrix}]. \tag{6}$$

Consider the following vectors $u_j \in \mathbb{R}^{2k+1}$ $(j = 1, 2, \ldots, 4k)$, (where ^{T} denotes transpose):
$u_1 := (1, 1, 0, 0, \ldots, 0)^T,\ u_2 := (1, -1, 0, 0, \ldots, 0)^T,\ u_3 := (1, 0, 1, 0, \ldots, 0)^T,\ u_4 := (1, 0, -1, 0, \ldots, 0)^T \ldots$
$\ldots,\ u_{4k} := (1, 0, 0, 0, \ldots, -1)^T$.

Consider now the polyhedral cone $\mathcal{K}_u$ generated by these vectors: $\mathcal{K}_u := cone[u_1, u_2, \ldots, u_{4k}]$. Clearly, any vector $s := (s_0, s_1, \ldots, s_{2k})^T$ is in $\mathcal{K}_u$ if and only if $\sum_{j=1}^{2k} |s_j| \le s_0$. Applying Proposition 1, we see that the matrix $\hat{A} := wI_1 \oplus M$ has the property that

$$m > Q(r, f, w) \text{ implies } \hat{A}^m \mathcal{K}_u \subset \mathcal{K}_u.$$

Consider now the polyhedral cone $\mathcal{K}$ generated by the following vectors:

$$\mathcal{K} := \{\hat{A}^j u_h :\ h = 1, 2, \ldots, 4k,\ j = 0, 1, 2, \ldots, Q\}.$$

It follows that $\hat{A}\mathcal{K} \subset \mathcal{K}$, i.e. the polyhedral cone $\mathcal{K}$ is $\hat{A}$-invariant. We can check, e.g., that $\hat{A}\hat{A}^Q u_h \in \mathcal{K}_u \subset \mathcal{K}$.

We shall show that we can choose the positive numbers c_0, b_0 so that $\begin{pmatrix} b_0 \\ b \end{pmatrix} \in \mathcal{K}, \qquad \begin{pmatrix} c_0 & c \end{pmatrix} \mathcal{K} \ge 0$ will hold. Indeed, if we take any $b_0 \ge ||b||_1 \equiv \sum_{j=1}^{2k} |b_j|$, where b_j denote the components of the vector b, then we have $\begin{pmatrix} b_0 \\ b \end{pmatrix} \in \mathcal{K}_u \subset \mathcal{K}$. On the other hand, we want to satisfy for every $h = 1, 2, \ldots, 4k,\ j = 0, 1, 2, \ldots, Q$ the inequality $\begin{pmatrix} c_0 & c \end{pmatrix} [w^j I_1 \oplus M^j] u_h = \begin{pmatrix} c_0 & c \end{pmatrix} \hat{A}^j u_h \ge 0$. Denoting the projection onto the subspace of the last $2k$ coordinates (parallel to that of the first one) by P, we obtain the sufficient conditions

$$c_0 w^j + cM^j P u_h \ge c_0 w^Q + cM^j P u_h \ge 0.$$

These are a finite number of conditions on c_0, which are satisfied if we choose

$$c_0 \ge -w^{-Q} c M^j P u_h \qquad (h = 1, 2, \ldots, 4k,\ j = 0, 1, 2, \ldots, Q).$$

With such a choice of b_0 and c_0 the polyhedral cone $\mathcal{K}$ will therefore be system invariant, and can apply Lemma 1 to conclude the existence of a positive realization of $t_1(v)$ of the desired dimension. However, we prefer to conclude the proof by showing the actual construction of how the positive realization of $t_1(v)$ is obtained.

The number s of the extreme rays of the polyhedral cone $\mathcal{K}$ clearly satisfies $s \leq 4k(Q+1)$. There exists a real $(2k+1) \times s$ matrix S (having as columns the extreme rays of $\mathcal{K}$) satisfying $\mathcal{K} = S\mathbb{R}_+^s$. The proved properties of the cone $\mathcal{K}$ imply the existence of a nonnegative $s \times s$ matrix A_+ such that $\hat{A}S = SA_+$, of a nonnegative $s \times 1$ vector b_+ such that $\begin{pmatrix} b_0 \\ b \end{pmatrix} = Sb_+$, and that the $1 \times s$ vector defined by $c_+ := \begin{pmatrix} c_0 & c \end{pmatrix} S$ is nonnegative. Finally, define $p := c_0 b_0 > 0$. It is easy to check that (6) is a realization of $t_1(v)$, and $c_+(vI - A_+)^{-1} b_+ = t(v) + \frac{p}{v-w} = t_1(v)$. $\square$

Remark 2. From the proof it is clear that the statement of the Theorem holds for any w_1 such that $w < w_1 < 1$ with the same values of Q and p. However, taking the value of w closer and closer to 1 will *not* decrease the the value of Q (and hence the dimension of the realization) below a certain threshold. $\square$

Remark 3. Note that in our construction the dimension N_1 depends on the location and order of the poles. The appearance of the pole order k as a factor in N_1 should not be unexpected (see Remark 4 below for a possible explanation). Note also, that our other essential factor Q is 'circular', i.e. is a function of $|z| = r$. In fact, if we choose f very close to 0, and w very close to 1 then the value of Q will be approximately $Q \approx [\frac{\log \sqrt{2}}{-\log |z|}]$. Therefore, our dimension N_1 will be reasonably low *as long as the complex poles of $t(v)$ do not lie very close to the boundary of the unit disk, i.e. to the "boundary of asymptotic stability"*. As a comparison, in [4], in the case of one pair of conjugate complex simple poles z and $\bar{z}$, the dimension of a positive realization of $t_1(v)$ was given as the smallest integer m such that z lies in the interior of the regular polygon $\mathcal{P}_m$ with m edges inscribed in the unit cricle and having one vertex at 1 (see Proposition 7 in [4]). It is clear that if z lies in $\mathcal{P}_m$ and is located very close to a vertex of $\mathcal{P}_m$ then the dimension m given in [4] can be lower than our dimension N_1 above. This seems to be a small price we have to pay for the universal applicability of our approach. $\square$

We will now turn to the case of a transfer function with *negative real multiple poles.* We will see that here we can even claim minimality of N_1 in certain cases (see Remark 4 below).

First, we essentially cite the following lemma from [11], Corollary 3.1.13:

Lemma 2: Let $r \in \mathbb{R}$ be such that $|r| < 1$, and let $f \neq 0$ be any real number. Then the Jordan block type matrix of order k

$$M \equiv M(f) \equiv M(r,f,k) := \begin{pmatrix} r & f & 0 & 0 & \ldots 0 \\ 0 & r & f & 0 & \ldots 0 \\ \ddots & & & & \\ 0 & 0 & 0 & 0 & \ldots f \\ 0 & 0 & 0 & 0 & \ldots r \end{pmatrix}$$

is similar to the canonical Jordan block $M(r,1,k) \equiv J(r,k)$. More exactly, defining $d := diag(1, f, f^2, \ldots, f^{k-1})$, we have $dM(r,f,k)d^{-1} = M(r,1,k)$. $\square$

Theorem 2: Let h be a negative number such that $|h| < 1$. If the transfer function $t(v)$ with real coefficients has exclusively the pole h of order k, then there are positive numbers p and w such that $|h| < w < 1$, and the function $t_1(v) := t(v) + \frac{p}{v-w}$ has a nonnegative realization of dimension $N_1 \leq 2k$.

Proof. Consider a real Jordan minimal realization (c_1, A_1, b_1) of $t(v)$. Then A_1 has exactly the single elementary divisor $(v-h)^k$, and A_1 is similar to any matrix $M(h, f, k)$. Let f and w be positive numbers such that $|h| + f < w < 1$. Then there is a real Jordan type minimal realization (c, A, b) of $t(v)$ such that $A = M = M(h, f, k)$ for this value of f. If D and $N \equiv N(f, k)$ denote the diagonal and nilpotent parts of the matrix A, then we have $||A|| = ||D + N|| = |h| + f$.

Moreover, $|h| + f < w < 1$ implies that $||Aw^{-1}|| = \frac{|h|+f}{w} < 1$. Let c_0, b_0 be positive numbers, let I_1 denote the 1×1 identity matrix, and consider the triple

$$[\begin{pmatrix} c_0 & c \end{pmatrix}, wI_1 \oplus M, \begin{pmatrix} b_0 \\ b \end{pmatrix}]. \tag{7}$$

Consider the following vectors $u_j \in \mathbb{R}^{k+1}$ $(j = 1, 2, \ldots, 2k)$, (where ^{T} denotes transpose):
$u_1 := (1, 1, 0, 0, \ldots, 0)^T,\ u_2 := (1, -1, 0, 0, \ldots, 0)^T,\ u_3 := (1, 0, 1, 0, \ldots, 0)^T,\ u_4 := (1, 0, -1, 0, \ldots, 0)^T \ldots$
$\ldots,\ u_{2k} := (1, 0, 0, 0, \ldots, -1)^T$.

Consider now the polyhedral cone $\mathcal{K}_u$ generated by these vectors: $\mathcal{K}_u := cone[u_1, u_2, \ldots, u_{2k}]$. Clearly, any vector $s := (s_0, s_1, \ldots, s_k)^T$ is in $\mathcal{K}_u$ if and only if $\sum_{j=1}^{k} |s_j| \leq s_0$. Hence the matrix $\hat{A} := wI_1 \oplus M$ leaves the polyhedral cone $\mathcal{K}_u$ invariant.

We shall show that we can choose the positive numbers c_0, b_0 so that $\begin{pmatrix} b_0 \\ b \end{pmatrix} \in \mathcal{K}_u, \qquad \begin{pmatrix} c_0 & c \end{pmatrix} \mathcal{K}_u \geq 0$ will hold. Indeed, if we take any $b_0 \geq ||b||_1 \equiv \sum_{j=1}^{k} |b_j|$, where b_j denote the components of the vector b, then we have $\begin{pmatrix} b_0 \\ b \end{pmatrix} \in \mathcal{K}_u$. On the other hand, we want to satisfy for every $j = 1, 2, \ldots, 2k$ the inequality $\begin{pmatrix} c_0 & c \end{pmatrix} u_j \geq 0$. Denoting the projection onto the subspace of the last k coordinates (parallel to that of the first one) by P, we obtain the sufficient conditions $c_0 + cPu_j \geq 0$. These are a finite number of conditions on c_0, which are satisfied if we choose $c_0 \geq -cPu_j \qquad (j = 1, 2, \ldots, 2k)$.

The number s of the extreme rays of the polyhedral cone $\mathcal{K}_u$ is clearly $s = 2k$. There exists a real $(k+1) \times s$ matrix S (having as columns the extreme rays of $\mathcal{K}_u$) satisfying $\mathcal{K}_u = S\mathbb{R}^s_+$. The proved properties of the cone $\mathcal{K}_u$ imply the existence of a nonnegative $s \times s$ matrix A_+ such that $\hat{A}S = SA_+$, of a nonnegative $s \times 1$ vector b_+ such that $\begin{pmatrix} b_0 \\ b \end{pmatrix} = Sb_+$, and that the $1 \times s$ vector defined by $c_+ := \begin{pmatrix} c_0 & c \end{pmatrix} S$ is nonnegative. Finally, define $p := c_0 b_0 > 0$. It is easy to check that (7) is a realization of $t_1(v)$, and $c_+(vI - A_+)^{-1} b_+ = t(v) + \frac{p}{v-w} = t_1(v)$. $\square$

Remark 4. As opposed to Theorem 1, here we can claim *minimality* of the dimension $N_1 = 2k$ *in certain cases.* Namely, assume that the location and the order of the negative pole of $t(v)$ satisfy the condition $|h|k \geq k - 1$. Then, any positive realization (c, A, b) of the function $t_1(v)$ must be of dimension at least $N_1 \geq 2k$, due to the following argument (in which we combine ideas from [1] and [9]). The dominant pole of $t_1(z)$ is w, therefore we can assume without loss of generality that all eigenvalues λ of A satisfy $|\lambda| \leq w$ (see Theorem 3.2 in [1]). Also, the trace of A is clearly nonnegative. Combining these facts we get $0 \leq TrA \leq hk + (N_1 - k)w$, which implies $N_1 \geq 2k$.

On the one hand, this argument shows that, in general, the appearence of the pole order k as a factor in the dimension N_1 should not be unexpected (cf. the result of Theorem 1 and Remark 2 above). Moreover, it also shows that an attempt to improve the result of Theorem 2 *must* take into account the value hk in some way. □

Now we turn to the case transfer functions with *nonnegative poles of higher order*. The result here was already obtained in Corollary 1 of [12] with a different approach. However, in order to give a *unified* treatment of all cases in Theorem 4 below, we now show how our present approach works in this case (yielding a result equivalent to that of [12]).

Theorem 3: Let q be a nonnegative number such that $0 \leq q < 1$. If the transfer function t with real coefficients has exclusively the pole q of order k, then there are positive numbers p and w such that $q < w < 1$, and the function $t_1(v) := t(v) + \frac{p}{v-w}$ has a nonnegative realization of order at most $k+1$.

Proof. Consider a real Jordan minimal realization $(\bar{c}, \bar{A}, \bar{b})$ of $t(v)$. Then $\bar{A}$ has exactly the single elementary divisor $(v-q)^k$, and $\bar{A}$ is similar to any matrix $M(q,f,k)$. Let f and w be positive numbers such that $q + f < w < 1$. Then there is a real Jordan type minimal realization (c, A, b) of $t(v)$ such that $A = M = M(q,f,k)$ for this value of f, and the components of the vectors are given by $c = \begin{pmatrix} c_1 & c_2 & \dots & c_k \end{pmatrix}$, $b^T = \begin{pmatrix} b_1 & b_2 & \dots & b_k \end{pmatrix}$, where ^{T} denotes transpose. Considering the partial fraction decomposition of the transfer function $t(v) = c(vI - M)^{-1}b$, we see that the coefficient e_j of $\frac{1}{(v-q)^j}$ is $e_j = (c_1 b_j + c_2 b_{j+1} + \dots + c_{k-j+1} b_k) f^{j-1} \qquad (j = 1, 2, \dots, k)$.

These equations show that we can obtain the same coefficients e_j, hence *the same transfer function* $t(v)$, *if we redefine* $b_1 := 0,\ b_2 := 0,\ \dots,\ b_{k-1} := 0,\ b_k := 1$. Indeed, we can then evaluate the (uniquely redefined) components of the vector c from the equations $e_j = c_{k-j+1} f^{j-1} \qquad (j = 1, 2, \dots, k)$. In what follows we shall make essential use of the form of these redefined vectors, and we shall *denote this redefined realization again* by (c, A, b).

If D and $N \equiv N(f,k)$ denote the diagonal and nilpotent parts of the matrix A, then we have $||A|| = ||(D+N)|| = q + f$. Moreover, $q + f < w < 1$ implies that $||Aw^{-1}|| = \frac{q+f}{w} < 1$. Let c_0, b_0 be positive numbers, let I_1 denote the 1×1 identity matrix, and consider the triple

$$[\begin{pmatrix} c_0 & c \end{pmatrix}, wI_1 \oplus M, \begin{pmatrix} b_0 \\ b \end{pmatrix}]. \tag{8}$$

Consider the following vectors $u_j \in \mathbb{R}^{k+1}$ $(j = 0, 1, 2, \dots, k)$, (where ^{T} denotes transpose): $u_0 := (1,0,0,0,\dots,0)^T,\ u_1 := (1,1,0,0,\dots,0)^T,\ u_2 := (1,0,1,0,\dots,0)^T,\ u_3 := (1,0,0,1,\dots,0)^T, \dots, u_k := (1,0,0,0,\dots,1)^T$, and consider the polyhedral cone $\mathcal{K}_u$ generated by these vectors: $\mathcal{K}_u := cone[u_0, u_1, u_2, \dots, u_k]$.

Clearly, any vector $s := (s_0, s_1, \dots, s_k)^T$ is in $\mathcal{K}_u$ if and only if for $j = 1, 2, \dots, k$ we have $0 \leq s_j \leq \sum_{i=1}^{k} s_i \leq s_0$. Hence the *nonnegative* matrix $\hat{A} := wI_1 \oplus M$ leaves the polyhedral cone $\mathcal{K}_u$ invariant.

We shall show that we can choose the positive numbers c_0, b_0 so that $\begin{pmatrix} b_0 \\ b \end{pmatrix} \in \mathcal{K}_u, \qquad \begin{pmatrix} c_0 & c \end{pmatrix} \mathcal{K}_u \geq 0$ will hold. Indeed, if we take any $b_0 \geq ||b||_1 = 1$ then, as a consequence of the redefined form of b, we have $\begin{pmatrix} b_0 \\ b \end{pmatrix} \in \mathcal{K}_u$. On the other hand, we want to satisfy for every $j = 0, 1, 2, \ldots, k$ the inequality $\begin{pmatrix} c_0 & c \end{pmatrix} u_j \geq 0$. Denoting the projection onto the subspace of the last k coordinates (parallel to that of the first one) by P, we obtain $c_0 + cPu_j \geq 0$. These are a finite number of conditions on c_0, which are satisfied if we choose $c_0 \geq -cPu_j \qquad (j = 0, 1, 2, \ldots, k)$.

The number s of the extreme rays of the polyhedral cone $\mathcal{K}_u$ is clearly $s = k + 1$. There exists a real $(k+1) \times s$ matrix S (having as columns the extreme rays of $\mathcal{K}_u$) satisfying $\mathcal{K}_u = S\mathbb{R}^s_+$. The proved properties of the cone $\mathcal{K}_u$ imply the existence of a nonnegative $s \times s$ matrix A_+ such that $\hat{A}S = SA_+$, of a nonnegative $s \times 1$ vector b_+ such that $\begin{pmatrix} b_0 \\ b \end{pmatrix} = Sb_+$, and that the $1 \times s$ vector defined by $c_+ := \begin{pmatrix} c_0 & c \end{pmatrix} S$ is nonnegative. Finally, define $p := c_0 b_0 > 0$. It is easy to check that (8) is a realization of $t_1(v)$, and $c_+(vI - A_+)^{-1} b_+ = t(v) + \frac{p}{v-w}$. $\square$

With the help of Theorems 1, 2 and 3 we can now give a *unified and universal* treatment of the positive decomposition problem for all asymptotically stable transfer functions $t(v)$.

Theorem 4: Assume that the transfer function $t(v)$ with real coefficients has exactly the nonnegative real poles q_j of orders g_j for $j = 1, \ldots, j_1$, further the negative poles n_j of orders h_j for $j = 1, \ldots, j_2$, and finally the nonreal pole pairs z_j and $\overline{z_j}$ of orders k_j for $j = 1, \ldots, j_3$, , all in the open unit disk. Denote the absolute values of the poles by r_j $(j = 1, \ldots, j_1 + j_2 + j_3)$, choose a positive number w and a small positive number f such that $r_j + f < w < 1$ $(j = 1, \ldots, j_1 + j_2 + j_3)$.

Then there is a positive number p such that the transfer function $t_1(v) := t(v) + \frac{p}{v-w}$ has a nonnegative realization of order at most $1 + \sum_{j=1}^{j_1} g_j + 2\sum_{j=1}^{j_2} h_j + 4\sum_{j=1}^{j_3} k_j(Q_j + 1)$. Here the nonnegative integers $Q_j := Q(r_{j_1+j_2+j}, f, w)$ $(j = 1, \ldots, j_3)$ can be determined as in formula (5).

Proof. The preceding methods apply separately to each group of the poles (nonnegative, negative poles, nonreal conjugate pole pairs).

Consider a *real minimal Jordan type realization* (c, A, b) of $t(v)$. Recalling the preceding methods, we can assume that A is the direct sum of real Jordan type blocks (with "nilpotent parameter" f everywhere), and that for nonnegative poles the corresponding *parts* of b have the form $(0, 0, \ldots, 0, 1)^T$, as needed in the proof of Theorem 3 above. Introduce the notation $G := \sum_{j=1}^{j_1} g_j, \quad H := \sum_{j=1}^{j_2} h_j, \quad K := \sum_{j=1}^{j_3} k_j$. In the real vector space $\mathbb{R}^{1+G+H+2K}$ consider the vectors

$u_0 := (1, 0, 0, \ldots, 0)^T$, $u_1 := (1, 1, 0, 0, \ldots, 0)^T$, $\ldots$, $u_G := (1, 0, \ldots, 0, 1, 0, \ldots, 0)^T$,

where the second components 1 stand on the places $1, 2, \ldots, G$. Further, consider the vectors $(1, 0, \ldots, 1, 0, \ldots, 0)^T$, $(1, 0, \ldots, -1, 0, \ldots, 0)^T$, where the second nonzero components stand on the places $G + 1, \ldots, G + H + 2K$, and apply the corresponding notation $u_{G+1}, \ldots, u_{G+2H+4K}$.

Next we complete this system, according to Theorem 1, by the vectors obtained by applying (possibly repeatedly) the Jordan type matrix $wI_1 \oplus A$ to the vectors u_r $(r = 1+G+2H+1, \ldots, 1+G+2H+4K)$. By Proposition 1, the number of the needed "new" vectors in this completed system is not greater than $4\sum_{j=1}^{j_3} k_j(Q_j+1)$. The methods of Theorems 1, 2, 3 show that the cone $\mathcal{C}$ generated by this completed system of vectors is a system invariant cone for the triple

$$[\begin{pmatrix} c_0 & c \end{pmatrix}, \quad wI_1 \oplus A, \quad \begin{pmatrix} b_0 \\ b \end{pmatrix}],$$

if the positive numbers c_0, b_0 are sufficiently large. Define then $p := c_0 b_0$. It is clear that the above triple is a minimal Jordan type realization of the transfer function $t_1(v)$. Consequently, $t_1(v)$ has a positive realization of order not greater than $1 + G + 2H + 4\sum_{j=1}^{j_3} k_j(Q_j+1)$, as stated. $\square$

IV. Examples

1. The transfer function of a low-pass digital Chebyshev filter of order 3 is given in [5], p. 184 by

$$t(z) := \frac{0.1253986950 + 0.3331328522\, z^2 + 0.1984152016\, z}{z^3 - 0.6905561900\, z^2 + 0.8018906100\, z - 0.3892083200}$$

with partial fraction decomposition

$$\frac{-0.01050864690 + 0.1411896961\, i}{z - 0.07522998673 + 0.8455579204\, i} - \frac{0.01050864690 + 0.1411896961\, i}{z - 0.07522998673 - 0.8455579204\, i} + \frac{0.3541501460}{z - 0.5400962165}$$

(We remark here that since the poles are simple one could also apply the different approach of [4]; see Example 2 below for a case when multiple poles are present). The maximal column norm of the pertaining matrix A of order 3 is 0.9207879071. We can choose $w := 0.93$, and $Q := 0$. Indeed, $m > 0$ implies $||(A/w)^m|| = 0.9900945238^m < 1$. We have then

$$\hat{A} = \begin{pmatrix} 0.93 & 0 & 0 & 0 \\ 0 & 0.07522998673 & 0.8455579204 & 0 \\ 0 & -0.8455579204 & 0.07522998673 & 0 \\ 0 & 0 & 0 & 0.5400962165 \end{pmatrix}$$

and we define

$$S := \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

By solving the corresponding linear program with nonnegative constraints, we obtain the matrix

$$A_+ = \begin{pmatrix} 0.07522998673 & 0.0 & 0.8455579204 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.07522998673 & 0.0 & 0.8455579204 & 0.0 & 0.0 \\ 0.0 & 0.8455579204 & 0.07522998673 & 0.0 & 0.0 & 0.0 \\ 0.8455579204 & 0.0 & 0.0 & 0.07522998673 & 0.0 & 0.0 \\ 0.009212092870 & 0.009212092870 & 0.00921209287 & 0.009212092870 & 0.9300000000 & 0.3899037835 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.5400962165 \end{pmatrix}$$

Let $c := (1,1,1,1)$, $b^T := (b_0, -0.151698343, 0.130681049, 0.3541501461)$. The triple $(c, \hat{A}, b)$ is a (minimal) realization of the "augmented" transfer function

$$t_1(z) := \frac{b_0}{z - 0.93} + t(z).$$

Choosing $b_0 > 0$ sufficiently large, say $b_0 := 5$, the linear program $Sx = b$ has the nonnegative solution

$$b_+ = (0,\ 0.1516983430,\ 0.1306810490,\ 0,\ 4.363470462,\ 0.3541501461)^T$$

Defining $c_+ := cS = (2,0,2,0,1,2)$, we see that (c_+, A_+, b_+) is a nonnegative realization of the function $t_1(z)$. Hence the desired decomposition of the transfer function of the Chebyshev filter is $t(z) = t_1(z) - \frac{b_0}{z-0.93}$.

2. We also sketch a theoretical example where higher order poles are present. Assume that the transfer function

$$t(v) = \frac{0.1}{1000v^3 + 2700v^2 + 2430v + 729} - \frac{100v - 100}{50v^2 - 70v + 29} - \frac{50v^2 - 70v + 20}{2500v^4 - 7000v^3 + 7800v^2 - 4060v + 841}$$

is given. Then $t(v)$ has the pole -0.9 of order 3 and the nonreal (conjugate) poles $x + yi := 0.7 + 0.3i$ and $0.7 - 0.3i$, both of order 2. The transfer function $t(v)$ is determined by the minimal Jordan realization (c, A, b), where $c := (1,0,0,1,-1,1,-1)$, $b^T := (0,0,1,1,1,-1,1)$, and

$$A = \begin{pmatrix} \frac{-9}{10} & \frac{1}{100} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{-9}{10} & \frac{1}{100} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-9}{10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{7}{10} & \frac{3}{10} & \frac{1}{100} & 0 \\ 0 & 0 & 0 & \frac{-3}{10} & \frac{7}{10} & 0 & \frac{1}{100} \\ 0 & 0 & 0 & 0 & 0 & \frac{7}{10} & \frac{3}{10} \\ 0 & 0 & 0 & 0 & 0 & \frac{-3}{10} & \frac{7}{10} \end{pmatrix}$$

We calculate that $r_1 := \sqrt{x^2 + y^2}$ is approximately 0.762, and we choose the values $f := 0.01$ and $w := 0.99$. Hence we can check that $Q = Q(r_1, f, w) = 1$.

As described in Theorems 1, 2, 3 above, we will choose $\hat{A} = wI_1 \oplus A$, $\hat{c} := (c_0, 1, 0, 0, 1, -1, 1, -1)$, $\hat{b}^T := (b_0, 0, 0, 1, 1, 1, -1, 1)$, $p = c_0 b_0$, with appropriately large positive values of b_0, c_0, such that the minimal Jordan realization $\hat{c}(vI - \hat{A})^{-1}\hat{b} = t(v) + \frac{p}{v-w} = t_1(v)$ will lead to a sytem invariant cone $\mathcal{C}$, and consequently to a positive realization (c_+, A_+, b_+) of $t_1(v)$.

The matrix S containing as columns the system of vectors described in Theorem 4 above is a 8×22 matrix. The calculation of the jth column $A_+(j)$ of one nonnegative 22×22 matrix A_+ can be done by finding one nonnegative solution $y_j \in \mathbb{R}^{22}$ of the linear equation $\hat{A}S(j) = Sy_j$ $(j = 1, \ldots, 22)$. The method can be, e.g., the application of a suitable linear program.

Then the definitions $c_0 := 2$, $\quad b_0 := 5$ are sufficient to ensure that $t_1(v)$ has a nonnegative realization (c_+, A_+, b_+) of dimension 22, where $c_+ := \hat{c} \cdot S$ and $\hat{b} = S \cdot b_+$.

V. Conclusion

In this paper we considered decompositions $t(z) = t_1(z) - t_2(z)$ of an *asymptotically stable* transfer function $t(z)$ as a difference of two *positive and a.s.* systems $t_1(z)$ and $t_2(z)$. Such decompositions are important due to the positivity of certain networks in applications, such as CRN's. Here we extended earlier results of Benvenuti, Farina and Anderson [4], and provided a unified and universal solution to the positive decomposition problem for *any a.s.* transfer function $t(z)$. An essential feature of the main result is that one resulting positive system is 1-dimensional, while the dimension of the other is reasonably low, which enhances the possibility of a practical application. Furthermore, our approach is easy-to-compute, leading to a general and efficient computer algorithm as explained in Theorem 4. Only in some cases can we claim minimality of the obtained positive realizations, and would be interested to see improvements in this direction.

References

[1] B.D.O. Anderson, M. Deistler, L. Farina, and L. Benvenuti, "Nonnegative realization of a linear system with nonnegative impulse response," *IEEE Trans. Circuits Syst. I*, vol. 43, no.2, pp 134-142, Feb. 1996.

[2] L. Benvenuti and L. Farina, "A tutorial on the positive realization problem", *IEEE Trans. Automat. Control* vol. 49, no. 5, pp. 651–664, 2004.

[3] L. Benvenuti and L. Farina, "The design of fiber-optic filters", *Journal of Lightwave Technology*, vol. 19, pp. 1366-1375, 2001.

[4] L. Benvenuti, L. Farina, and B.D.O. Anderson, "Filtering through a combination of positive filters," *IEEE Trans. Circuits Syst. I*, vol. 46, pp. 1431-1440, Dec. 1999.

[5] N.K. Bose, *Digital filters: Theory and Applications.* New York: North-Holland, 1985.

[6] L. Farina and S. Rinaldi, *Positive linear systems: Theory and applications.* New York: Wiley, 2000.

[7] K.-H. Förster and B. Nagy, "Nonnegative realizations of matrix transfer functions", *Lin. Algebra Appl.*, vol. 311, pp. 107-129, 2000.

[8] A. Gersho and B. Gopinath, "Charge-routing networks", *IEEE Trans. Circuits Syst.*, vol. 26, pp. 81-92, Febr. 1979.

[9] C. Hadjicostis, "Bounds on the size of minimal nonnegative realizations for discrete-time LTI systems," *Syst. Contr. Lett.*, vol. 37, pp 39-43, 1999.

[10] A. Halmschlager and M. Matolcsi, "Minimal positive realizations for a class of transfer functions", *IEEE Trans. Circuits Syst. II*, **52**, vol. 4, 177-180, 2005.

[11] R.A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.

[12] B. Nagy, M. Matolcsi, "Minimal Positive Realizations of Transfer Functions with Nonnegative Multiple Poles," *IEEE Trans. Automat. Control* vol. 50, no. 9, pp. 1447-1450, September, 2005.