Remarks on orthogonal polynomials and balanced realizations

J. Hoffmann
Fachbereich Mathematik
Universität Kaiserslautern
Kaiserslautern, Germany

and

P. A. Fuhrmann*†
Department of Mathematics
Ben-Gurion University of the Negev
Beer Sheva, Israel

April 10, 1995

Abstract

Given a proper antistable rational transfer function g, a balanced realization of g is constructed as a matrix representation of the abstract shift realization introduced in Fuhrmann [1976]. The required basis is constructed as a union of sets of polynomials orthogonal with respect to weights given by the square of the absolute values of minimal degree Schmidt vectors of the corresponding Hankel operators. This extends results of Fuhrmann [1991], obtained in the generic case.

*Earl Katz Family Chair in Algebraic System Theory
†Partially supported by the Israeli Academy of Sciences under Grant No. 249/90 and by GIF under Grant No. I 184

1 Introduction

In Moore [1981] the concept of balanced realizations has been introduced as a method of model reduction. Since then an enormous amount of work has been done on balanced realizations and their applications to model reduction and robust control. Kung [1980], Pernebo and Silverman [1982], Glover [1986], McFarlane and Glover [1989] are some papers in this connection. Of course the list is far from exhaustive.

Balancing was introduced first by Moore in the context of stable systems, and has been extended by Jonckheere and Silverman [1983] to arbitrary systems with a pair of Riccati equations replacing the Lyapunov equations in Moore's definition. Fuhrmann and Ober [1993] contains a comprehensive account of various aspects of LQG balancing.

While balanced realizations are usually introduced on the state space level, it is clear, especially from the various balanced canonical forms studied in Ober [1987,1989], that they exhibit certain system invariants. Thus it would be of interest to explore the links between these invariants and the external, i.e. input/output, properties of the system.

In the stable case, that is the case of Lyapunov balancing, it has long been known, see Glover [1984], that the Lyapunov singular values are identical to the singular values of the induced Hankel operator.

However it was not till Fuhrmann [1991] that, at least in the generic case of distinct singular values, the balanced canonical form of Ober was obtained as a matrix representation of the shift realization, introduced in Fuhrmann [1976], with respect to a basis made of suitably normalized Schmidt vectors. Even in the other extreme case, that of all singular values being identical, no such complete identification was made. Rather an approach using continued fractions was taken there. Of course continued fractions relate also to families of orthogonal polynomials, see Akhiezer [1965], Gragg [1972], Szegö [1959], Wall [1948], but the explicit connection, as far as balancing is concerned, was left unexplored.

The present paper closes this gap and produces a construction of a natural orthogonal basis for the state space of the shift realization, such that the corresponding matrix representation is the balanced canonical form. The method we use focuses on the set of all minimal (numerator) degree singular vectors corresponding to the set of all singular values of the Hankel operator. These vectors are uniquely determined, up to a nonzero multiplicative constant. In terms of these vectors we have a simple description of the set of all singular vectors. The degrees of freedom are determined by the degree deficiencies of these singular vectors. By applying a Gram-Schmidt procedure separately in each spectral subspace we get the required basis.

The paper is structured as follows. We begin with a very short review of the shift realization and by recalling the basic results from Fuhrmann [1991]. In section 3 we analyse the case of all singular values coinciding, i.e. of transfer functions of, up to additive constants, antistable inner functions. Finally, in section 4 we state and prove the general result. In the process we also correct an omission in Fuhrmann [1991] by computing also the diagonal elements of the generator matrix of a balanced realization. The present paper is of a technical nature. Yet we believe that it provides some additional insight into the nature of balanced canonical forms. The techniques used in this paper can be applied to derive balanced canonical forms for other classes of functions and this will be the subject of a forthcoming paper.

This work was done while the second author was visiting the Mathematics Department of the University of Kaiserslautern. He would like to express his thanks to D. Prätzel-Wolters for the opportunity and to the DFG for its support.

2 Preliminaries

Polynomial and, even more so, rational models provide the main tool in this paper. We proceed to give the basic definitions. Necessarily the exposition is brief and it is suggested that the interested reader consult some other papers as Fuhrmann [1976,1977,1981,1983,1984,1991] and Helmke and Fuhrmann [1989].

Throughout the paper we will restrict ourselves to the real number field. By $\mathbb{R}[z]$ we denote the ring of polynomials over $\mathbb{R}$, $\mathbb{R}((z^{-1}))$ the set of truncated Laurent series in z^{-1}, i.e. the set of all formal series of the form $\sum_{j=-\infty}^{n_f} f_j z^j$, $n_f \in \mathbb{Z}$. $\mathbb{R}((z^{-1}))$ is a vector space over $\mathbb{R}$ as well as a field. It contains the field $\mathbb{R}(z)$ of rational functions as a subfield. By $\mathbb{R}[[z^{-1}]]$ and $z^{-1}\mathbb{R}[[z^{-1}]]$ we denote the set of all formal power series in z^{-1} and the set of those power series with vanishing constant term respectively. Let π_+ and π_- be the projections of $\mathbb{R}((z^{-1}))$ onto $\mathbb{R}[z]$ and $z^{-1}\mathbb{R}[[z^{-1}]]$ respectively. Since $\mathbb{R}((z^{-1})) = \mathbb{R}[z] \oplus z^{-1}\mathbb{R}[[z^{-1}]]$ they are complementary projections. The space $z^{-1}\mathbb{R}[[z^{-1}]]$ carries a module structure over the ring $\mathbb{R}[z]$, with the module action given by

$$z \cdot h = S_- h = \pi_- z h. \tag{1}$$

Given a monic polynomial q of degree n, we define the associated *rational model* to be the space

$$X^q = Im \pi^q, \tag{2}$$

where π^q is the projection in $z^{-1}\mathbb{R}[[z^{-1}]]$ defined by

$$\pi^q h = \pi_- q^{-1} \pi_+ q h \qquad \text{for} \qquad h \in z^{-1}\mathbb{R}[[z^{-1}]]. \tag{3}$$

X^q is a submodule of $z^{-1}\mathbb{R}[[z^{-1}]]$, its elements being all strictly proper rational functions with q as denominator. The module structure is given by

$$S^q h = S_- h \qquad \text{for} \quad h \in X^q. \tag{4}$$

The great usefullness of these functional models in system theory stems from the fact that, in these terms, realization theory becomes a triviality. Moreover, realization theory provides a link between techniques based on functional and operator methods on the one hand and state space methods on the other. Thus, given a proper rational function $\phi = \dfrac{n}{d}$ the *associated realization* is constructed as follows.

The state space for the realization is chosen to be X^d and (A,B,C,D) are defined through

$$\left\{\begin{array}{lll} A = S^d & & \\ B\xi = \dfrac{n}{d}\xi & \text{for} & \xi \in \mathbb{R}, \\ Cf = (f)_{-1} = (zf)(\infty) & \text{for} & f \in X^d, \\ D = \phi(\infty). & & \end{array}\right. \tag{5}$$

The realization of ϕ is minimal, by the coprimeness of n and d. It is this realization we use as a basis for obtaining a balanced realization.

In Fuhrmann [1991,1993], a detailed analysis of Hankel operators with rational, scalar, antistable symbol, was carried out. We refer to these papers for a more complete introduction of all the spaces. Here we restrict ourselves to the basics. H^2_+ and H^2_- are the Hardy spaces of the right and left half planes respectively. Both spaces are considered as subspaces of the L^2 space of the imaginary axis. H^∞_- is the space of bounded analytic functions in the left half plane.

We assume $\phi \in H^\infty_-$ is rational and $\phi = \dfrac{n}{d}$ is a representation of ϕ as a quotient of coprime polynomials, naturally with d antistable. The Hankel operator $H_\phi : H^2_+ \longrightarrow H^2_-$ is defined by

$$H_\phi f = P_- \phi f \qquad for\ \ f \in H^2_+,$$

where P_- is the orthogonal projection of L^2 onto H^2_-. It has been shown in Fuhrmann [1991] that $KerH_\phi = \dfrac{d}{d^*}H^2_+$, $\{KerH_\phi\}^\perp = \{\dfrac{d}{d^*}H^2_+\}^\perp = X^{d^*}$ and $ImH_\phi = H^2_- \ominus \dfrac{d^*}{d}H^2_- = X^d$. The space X^d is the space of all strictly proper rational functions with d as their denominator.

Thus for the study of the Hankel operator we can restrict the Hankel operator to a map from X^{d^*} to X^d. The advantage is that, by cutting out the kernel, the restriction is a finite dimensional linear transformation.

It has been shown in the quoted papers that the Schmidt pairs with numerator polynomials of minimal degree, corresponding to the Hankel singular value μ, are of the form $\{\dfrac{p}{d^*}, \epsilon\dfrac{p^*}{d}\}$, $\epsilon \in \{\pm 1\}$; here $p \in \mathbb{R}[z]$ is such that the equation

$$\frac{n}{d}\frac{p}{d^*} = \epsilon\mu\frac{p^*}{d} + \frac{\pi}{d^*}$$

or, equivalently, that the fundamental polynomial equation, with $\lambda = \epsilon\mu$,

$$np = \lambda d^* p^* + d\pi \tag{6}$$

is solvable. The minimum degree Schmidt vectors, with different singular values, have been shown to be of particular importance, in the generic case, and were used to construct a basis for X^d and X^{d^*} respectively. In fact bases with a suitable normalization led to balanced realizations.

3 All-pass transfer functions

In this section we restrict ourselves to the special case of antistable transfer functions all of whose Hankel singular values coincide. By a result of Glover [1984], see also Fuhrmann [1991], the functions are, up to an additive constant, conjugate inner functions, or inner functions in H^∞_-. This special case has been studied already in Ober [1987] where the connections to continued fractions are indicated. However no attempt has been made there to identify the canonical form in functional terms. Similarly, in Fuhrmann [1991] there was no attempt to identify the basis that leads, via the shift realization, to the balanced canonical form for this class of functions.

The importance of this special case, providing in a sense the building blocks for the general case, has been already recognized by Ober. Now with a continued fraction expansion we can associate a sequence of orthogonal polynomials. This is a classical subject, see Akhiezer [1965].

The theorem that follows explains the connection between the balanced canonical form for conjugate inner functions and a sequence of polynomials orthogonalized relative to a weight function related to the minimum (numerator) degree singular vector of the corresponding Hankel operator. With respect to this particular basis, suitably normalized, the matrix representation of the shift realization is just the balanced canonical form obtained by Ober.

Theorem 3.1 *Let $\phi = \frac{n}{d} \in H^\infty_-$, d monic, $n \wedge d = 1$ and $\deg d = n$. Let us assume that all the singular values of the Hankel operator H_ϕ coincide, that is $\sigma_1 = \cdots = \sigma_n = \sigma > 0$. Let t_i^* be the polynomials obtained from $\{1, z, \ldots, z^{n-1}\}$ via the Gram-Schmidt orthonormalization procedure with respect to the weight function $\frac{1}{|d|^2}$. Then a set of constants $\{g_1, \ldots, g_n\} \subseteq \mathbb{R}$ can be chosen such that for $q_i^* := g_i t_i^*, i \in \underline{n}$ the normalization*

$$\|\frac{q_n^*}{d}\|_2^2 = \sigma, \tag{7}$$

holds and that the matrix representation of the shift realization of ϕ with respect to the basis $\{\frac{q_i^*}{d}, i = 1, \ldots, n\}$ of X^d has the following form:

$$A = \begin{pmatrix} 0 & -\alpha_1 & 0 & \cdots & \cdots & 0 \\ \alpha_1 & 0 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -\alpha_{n-2} & 0 \\ \vdots & & \ddots & \ddots & 0 & -\alpha_{n-1} \\ 0 & \cdots & \cdots & 0 & \alpha_{n-1} & \theta_n \end{pmatrix}, \qquad \alpha_i > 0, i = 1, \ldots, n-1, \tag{8}$$

$$B = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ b_n \end{pmatrix}, \qquad b_n = -\epsilon q_{n,n-1},\ \epsilon = \pm 1, \tag{9}$$

$$C = \begin{pmatrix} 0, & \dots & \dots & ,0 & ,sb_n \end{pmatrix}, \qquad s = (-1)^n \epsilon, \tag{10}$$

$$D = \phi(\infty) \tag{11}$$

and

$$\theta_n = \frac{b_n^2}{2\sigma}. \tag{12}$$

Proof: We start with the construction of the specific basis $\{\frac{q_i^*}{d}, i = 1, \dots, n\}$ of X^d which guarrantees the special form of the matrix A.
Take the basis $\{\frac{z^{i-1}}{d}, i = 1, \dots, n\}$ of X^d and apply the Gram-Schmidt orthonormalization procedure; this yields a new basis $\{\frac{t_i^*}{d}, i = 1, \dots, n\}$ of X^d, where

$$\frac{t_i^*}{d} = \|\frac{\tilde{t}_i^*}{d}\|_2^{-1} \cdot \frac{\tilde{t}_i^*}{d}, \qquad i = 1, \dots n \tag{13}$$

and

$$\frac{\tilde{t}_1^*}{d} = \frac{1}{d}, \tag{14}$$

$$\frac{\tilde{t}_i^*}{d} = \frac{z^{i-1}}{d} - \sum_{j=1}^{i-1} (\frac{z^{i-1}}{d}, \frac{\tilde{t}_j^*}{d}) \frac{\tilde{t}_j^*}{d}, \qquad i = 2, \dots, n. \tag{15}$$

Observe that

$$deg\, t_i^* = i - 1, \qquad i = 1, \dots, n \tag{16}$$

and

$$\frac{t_i^*}{d} \perp span\{\frac{z^j}{d}, j \in \{0, \dots, i-2\}\}, \qquad i = 2, \dots n. \tag{17}$$

Furthermore, equations (13) - (15) show that all the leading coefficients of the polynomials t_i^* are positive. We remark that the basis $\{\frac{t_i^*}{d}, i = 1, \dots, n\}$ is, up to multiplication of the elements of the basis by constants, equal to the desired basis.
Now we proof a recursion formula for the polynomials t_i^*. Obviously there holds

$$\frac{t_2^*}{d} = \frac{\gamma_1 z t_1^*}{d} - \theta_1 \frac{t_1^*}{d}, \qquad \gamma_1, \theta_1 \in \mathbb{R}. \tag{18}$$

Multiplication by d yields

$$t_2^* = \gamma_1 z t_1^* - \theta_1 t_1^* = (\gamma_1 z - \theta_1) t_1^*, \qquad \gamma_1, \theta_1 \in \mathbb{R}. \tag{19}$$

Moreover, for $i = 2, \ldots, n-1$ there holds

$$t_{i+1}^* = (\gamma_i z - \theta_i) t_i^* - \beta_{i-1} t_{i-1}^*, \qquad \gamma_i, \theta_i, \beta_{i-1} \in \mathbb{R}. \tag{20}$$

To prove this, let for $i \in \{2, \ldots, n-1\}$

$$\frac{z t_i^*}{d} = \sum_{j=1}^{i+1} \alpha_j^i \frac{t_j^*}{d}. \tag{21}$$

Now for $k \in \{1, \ldots, i-2\}$, because of the orthonormality, there holds

$$\begin{aligned} \alpha_k^i &= (\frac{z t_i^*}{d}, \frac{t_k^*}{d}) = (\frac{t_i^*}{d}, \frac{z^* t_k^*}{d}) \\ &= -(\frac{t_i^*}{d}, \frac{z t_k^*}{d}) = 0, \end{aligned} \tag{22}$$

since $deg(z t_k^*) = (k-1) + 1 = k < i - 1$, (by (16)) and (17).
Multiplication of equation (21) by d and defining

$$\beta_{i-1} := \frac{\alpha_{i-1}^i}{\alpha_{i+1}^i}, \quad \theta_i := \frac{\alpha_i^i}{\alpha_{i+1}^i} \text{ and } \gamma_i := \frac{1}{\alpha_{i+1}^i}, \qquad i \in \{2, \ldots, n-1\}$$

yields (20).
Since all the leading coefficients of the polynomials t_i^* are positive, it immediately follows that

$$\gamma_i > 0, \qquad i = 1, \ldots, n-1. \tag{23}$$

Moreover, there holds

$$\theta_i = 0, \qquad i = 1, \ldots, n-1. \tag{24}$$

Divide both sides of (19) respectively (20) by d and take the inner product with $\frac{t_i^*}{d}$; because of the orthonormality this results in

$$\theta_i = (\gamma_i \frac{z t_i^*}{d}, \frac{t_i^*}{d}), \qquad i = 1, \ldots, n-1. \tag{25}$$

Hence

$$\begin{aligned} \theta_i &= \gamma_i (\frac{z t_i^*}{d}, \frac{t_i^*}{d}) \\ &= \gamma_i \frac{1}{2\pi} \int_{-i\infty}^{i\infty} z \frac{t_i^* t_i}{d^* d} dz \\ &= \gamma_i \frac{1}{2\pi} \int_{-i\infty}^{i\infty} z \mid \frac{t_i}{d} \mid^2 dz. \end{aligned}$$

Now $\mid \frac{t_i}{d} \mid^2$ is symmetric with respect to the origin, whereas z changes sign. Thus

$$\frac{1}{2\pi} \int_{-i\infty}^{i\infty} z \mid \frac{t_i}{d} \mid^2 dz = 0$$

and (24) follows.
This implies that the polynomial t_i^* does only contain even/odd powers of z for i odd/even, which is easily proved using (20), (24) by an induction argument; observe that t_1^* is constant and that, by (19) and (24),

$$t_2^* = \gamma_1 z t_1^*, \qquad \gamma_1 \in \mathbb{R}.$$

Finally,

$$\beta_{i-1} < 0, \qquad i = 2, \ldots, n-1. \tag{26}$$

Proceeding as for the proof of (24), but now applying inner multiplication by $\frac{t_{i-1}^*}{d}$ yields

$$\begin{aligned} \beta_{i-1} &= \gamma_i(\frac{zt_i^*}{d}, \frac{t_{i-1}^*}{d}) = \gamma_i(\frac{t_i^*}{d}, \frac{z^* t_{i-1}^*}{d}) && (27) \\ &= -\gamma_i(\frac{t_i^*}{d}, \frac{zt_{i-1}^*}{d}). && (28) \end{aligned}$$

Furthermore, by (20) and (27)

$$\begin{aligned} \beta_{i-1} &= -\gamma_i(\frac{\gamma_{i-1} z t_{i-1}^* - \beta_{i-2} t_{i-2}^*}{d}, \frac{zt_{i-1}^*}{d}) \\ &= -\gamma_i\langle \gamma_{i-1}(\frac{zt_{i-1}^*}{d}, \frac{zt_{i-1}^*}{d}) - \beta_{i-2}(\frac{t_{i-2}^*}{d}, \frac{zt_{i-1}^*}{d})\rangle \\ &= -\gamma_i\langle \gamma_{i-1}\|\frac{zt_{i-1}^*}{d}\|_2^2 - \beta_{i-2}(\frac{t_{i-2}^*}{d}, \frac{zt_{i-1}^*}{d})\rangle \\ &= -\gamma_i\langle \gamma_{i-1}\|\frac{zt_{i-1}^*}{d}\|_2^2 - \beta_{i-2}\frac{\beta_{i-2}}{\gamma_{i-1}}\rangle \\ &= -\gamma_i\langle \gamma_{i-1}\|\frac{zt_{i-1}^*}{d}\|_2^2 - \frac{\beta_{i-2}^2}{\gamma_{i-1}}\rangle. \end{aligned}$$

But

$$\begin{aligned} \beta_{i-2}^2 &= \mid \beta_{i-2} \mid^2 = \gamma_{i-1}^2(\tfrac{zt_{i-1}^*}{d}, \tfrac{t_{i-2}^*}{d})^2 \\ &\leq \gamma_{i-1}^2\|\tfrac{zt_{i-1}^*}{d}\|_2^2 \cdot \|\tfrac{t_{i-2}^*}{d}\|_2^2 \\ &= \gamma_{i-1}^2\|\tfrac{zt_{i-1}^*}{d}\|_2^2 \end{aligned} \tag{29}$$

by the Cauchy - Schwartz inequality, and hence

$$\beta_{i-1} \leq 0, \qquad i = 2, \ldots, n-1.$$

Equality in the Cauchy - Schwartz inequality does hold iff the two factors in the inner product are linearly dependent which obviously is not the case here. Hence (26) is true.

Observe that by dividing equations (19) and (20) by d one obtains a representation of

$$S^d\,\frac{t_i^*}{d} = \frac{zt_i^*}{d}, \qquad i = 1, \ldots n-1$$

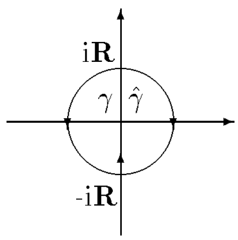
with respect to the basis $\{\frac{t_i^*}{d}, i = 1, \ldots, n\}$.
What we will do next is to find a representation of $S^d\,\frac{t_n^*}{d}$ with respect to this basis.
The technique we will use is to replace inner products in H_+^2, which are given by integrals on the imaginary axis, by limits of contour integrals. These contour integrals are more amenable to computation using partial fraction decompositions.
Because of the monicity of d there holds

$$S^d\,\frac{t_n^*}{d} = \frac{zt_n^* - (-1)^{n-1} t_{n,n-1}\, d}{d}. \tag{30}$$

Moreover, for $k \in \{1, \ldots, n-2\}$ we have

$$\begin{aligned}
(S^d\,\frac{t_n^*}{d}, \frac{t_k^*}{d}) &= (\frac{zt_n^* - (-1)^{n-1} t_{n,n-1}\, d}{d}, \frac{t_k^*}{d}) \\
&= \frac{1}{2\pi} \int_{-i\infty}^{i\infty} \frac{zt_n^* - (-1)^{n-1} t_{n,n-1}\, d}{d} \cdot \frac{t_k}{d^*} dz \\
&= \lim_{R \mapsto \infty} \langle\, \frac{1}{2\pi i} \int_{\hat{\gamma}} \frac{zt_n^* t_k}{dd^*} dz - \frac{(-1)^{n-1} t_{n,n-1}}{2\pi i} \int_{\hat{\gamma}} \frac{t_k}{d^*} dz \,\rangle,
\end{aligned}$$

where γ and $\hat{\gamma}$ denote the semicircular contours defined below.



Note that γ is positively oriented, whereas $\hat{\gamma}$ is negatively oriented. Moreover, the degree deficiency of numerator relative to denominator in the integrand $\frac{zt_n^* - (-1)^{n-1} t_{n,n-1}\, d}{d} \cdot \frac{t_k}{d^*}$ is at least two. This permits the switch to contour integrals.
Since d^* is stable, i.e. d^* has only poles in LHP, there holds

$$\lim_{R \mapsto \infty} \frac{1}{2\pi i} \int_{\hat{\gamma}} \frac{t_k}{d^*} dz = 0.$$

So we obtain

$$(S^d\,\frac{t_n^*}{d},\frac{t_k^*}{d}) = \lim_{R\mapsto\infty}\frac{1}{2\pi i}\int_{\hat{\gamma}}\frac{zt_n^*t_k}{dd^*}dz \tag{31}$$

$$= \frac{1}{2\pi}\int_{-i\infty}^{i\infty}\frac{t_n^*zt_k}{dd^*}dz \tag{32}$$

$$= (\frac{t_n^*}{d},\frac{z^*t_k^*}{d}) = -(\frac{t_n^*}{d},\frac{zt_k^*}{d}). \tag{33}$$

But for $k\in\{1,\ldots,n-2\}$ equation (16) gives

$$deg(zt_k^*) = (deg\,t_k^*)+1 \leq (n-3)+1 = n-2,$$

and by (17) this implies

$$(S^d\,\frac{t_n^*}{d},\frac{t_k^*}{d}) = 0, \qquad k=1,\ldots,n-2.$$

Hence we have

$$S^d\,\frac{t_n^*}{d} = \theta_n\frac{t_n^*}{d}+\beta_{n-1}\frac{t_{n-1}^*}{d}, \qquad \theta_n,\beta_{n-1}\in\mathbb{R}. \tag{34}$$

Now we verify that

$$\beta_{n-1} < 0. \tag{35}$$

From (34) one obtains, using the orthonormality of the $\frac{t_i^*}{d^*}$,

$$\begin{aligned}
\beta_{n-1} &= (S^d\,\frac{t_n^*}{d},\frac{t_{n-1}^*}{d})\\
&= (\frac{zt_n^*-(-1)^{n-1}t_{n,n-1}d}{d},\frac{t_{n-1}^*}{d})\\
&= \frac{1}{2\pi}\int_{-i\infty}^{i\infty}\frac{zt_n^*-(-1)^{n-1}t_{n,n-1}d}{d}\cdot\frac{t_{n-1}}{d^*}dz\\
&= \frac{1}{2\pi}\int_{-i\infty}^{i\infty}\frac{zt_n^*t_{n-1}}{dd^*}dz-\frac{(-1)^{n-1}t_{n,n-1}}{2\pi}\int_{-i\infty}^{i\infty}\frac{t_{n-1}}{d^*}dz\\
&= \lim_{R\mapsto\infty}\langle\frac{1}{2\pi i}\int_{\hat{\gamma}}\frac{zt_n^*t_{n-1}}{dd^*}dz-\frac{(-1)^{n-1}t_{n,n-1}}{2\pi i}\int_{\hat{\gamma}}\frac{t_{n-1}}{d^*}dz\rangle.
\end{aligned}$$

The last integral is zero because of the stability of d^*. So

$$\begin{aligned}
\beta_{n-1} &= \lim_{R\mapsto\infty}\frac{1}{2\pi i}\int_{\hat{\gamma}}\frac{zt_n^*t_{n-1}}{dd^*}dz = \frac{1}{2\pi}\int_{-i\infty}^{i\infty}\frac{t_n^*zt_{n-1}}{dd^*}dz\\
&= (\frac{t_n^*}{d},\frac{z^*t_{n-1}^*}{d}) = -(\frac{t_n^*}{d},\frac{zt_{n-1}^*}{d}).
\end{aligned}$$

Furthermore, by (20) and (27)

$$\beta_{n-1} = -(\frac{t_n^*}{d},\frac{zt_{n-1}^*}{d})$$

$$\begin{aligned}
&= -(\frac{\gamma_{n-1} z t_{n-1}^* - \beta_{n-2} t_{n-2}^*}{d}, \frac{z t_{n-1}^*}{d}) \\
&= -\gamma_{n-1}(\frac{z t_{n-1}^*}{d} \frac{z t_{n-1}^*}{d}) + \beta_{n-2}(\frac{t_{n-2}^*}{d}, \frac{z t_{n-1}^*}{d}) \\
&= -\gamma_{n-1} \| \frac{z t_{n-1}^*}{d} \|_2^2 + \frac{\beta_{n-2}^2}{\gamma_{n-1}}. \qquad (36)
\end{aligned}$$

Analogously, as in (29), one sees

$$\beta_{n-2}^2 < \gamma_{n-1}^2 \| \frac{z t_{n-1}^*}{d} \|_2^2, \qquad (37)$$

which proves (35) by substituting it in (36).

Taking equation (34), substituting expression (30) for $S^d \dfrac{t_n^*}{d}$ and multiplying by d, we obtain

$$z t_n^* - (-1)^{n-1} t_{n,n-1} d = \theta_n t_n^* + \beta_{n-1} t_{n-1}^*. \qquad (38)$$

Now the comparison of the leading coefficients in (38) yields

$$(-1)^{n-2} t_{n,n-2} - (-1)^{n-1} t_{n,n-1} d_{n-1} = \theta_n (-1)^{n-1} t_{n,n-1} \qquad (39)$$

or

$$\theta_n = -\frac{t_{n,n-2}}{t_{n,n-1}} - d_{n-1}. \qquad (40)$$

But t_n^* contains only even/odd powers of z depending on whether n is odd/even. Since by (16)

$$deg\, t_n^* = n - 1,$$

we have $t_{n,n-2} = 0$ and hence

$$\theta_n = -d_{n-1}. \qquad (41)$$

Observe that d^* is stable, i.e. for

$$d^*(z) = (-1)^n z^n + d_{n-1}^* z^{n-1} + \cdots + d_0^*$$

there holds

$$d_i^* \neq 0, \quad sgn\, d_i^* = (-1)^n, \qquad i = 0, \ldots, n-1.$$

But

$$d_i^* = (-1)^i d_i, \qquad i = 0, \ldots, n-1.$$

Hence the signs of the coefficients of d interlace. Since d is monic, we finally get $d_{n-1} < 0$ and hence

$$\theta_n > 0. \qquad (42)$$

Looking at equations (19), (20) and (34) we have derived the following matrix representation for the map

$$S^d : X^d \mapsto X^d$$

with respect to the basis $\tilde{\mathcal{B}} := \{\frac{t_i^*}{d}, i = 1, \dots n\}$:

$$[S^d]_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}} = \begin{pmatrix} 0 & \frac{\beta_1}{\gamma_2} & 0 & \cdots & \cdots & 0 & 0 \\ \frac{1}{\gamma_1} & 0 & \frac{\beta_2}{\gamma_3} & \ddots & & \vdots & \vdots \\ 0 & \frac{1}{\gamma_2} & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \frac{\beta_{n-2}}{\gamma_{n-1}} & 0 \\ \vdots & & & \ddots & \ddots & 0 & \beta_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{\gamma_{n-1}} & \theta_n \end{pmatrix} \tag{43}$$

Now we can finally determine the basis $\mathcal{B} := \{\frac{q_i^*}{d}, i = 1, \dots, n\}$ from the statement of the theorem. This is done by

$$q_i^* := g_i t_i^*, \qquad g_i \in \mathbb{R}, i = 1, \dots, n, \tag{44}$$

where the g_i are constructed in the following such that A is of the form (8) and $\|\frac{q_n^*}{d}\|_2^2 = \sigma$.

From the orthogonality of the basis $\mathcal{B}$ we get

$$a_{ij} = \frac{(S^d \frac{q_j^*}{d}, \frac{q_i^*}{d})}{(\frac{q_i^*}{d}, \frac{q_i^*}{d})} = \frac{g_j}{g_i}(S^d \frac{t_j^*}{d}, \frac{t_i^*}{d}),$$

and hence A is of the following form:

$$A = diag^{-1}(g_1, \dots, g_n)[S^d]_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}} diag(g_1, \dots, g_n)$$

$$= \begin{pmatrix} 0 & \frac{\beta_1 g_2}{\gamma_2 g_1} & 0 & \cdots & \cdots & 0 & 0 \\ \frac{g_1}{\gamma_1 g_2} & 0 & \frac{\beta_2 g_3}{\gamma_3 g_2} & \ddots & & \vdots & \vdots \\ 0 & \frac{g_2}{\gamma_2 g_3} & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \frac{\beta_{n-2} g_{n-1}}{\gamma_{n-1} g_{n-2}} & 0 \\ \vdots & & & \ddots & \ddots & 0 & \frac{\beta_{n-1} g_n}{g_{n-1}} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{g_{n-1}}{\gamma_{n-1} g_n} & \theta_n \end{pmatrix}. \tag{45}$$

By (26) and (35) there holds

$$\beta_i < 0, \qquad i = 1, \dots, n-1.$$

Write

$$\beta_i := -\delta_i^2, \quad \delta_i > 0, \qquad i = 1, \dots n-1.$$

Then we have to solve the following set of equations:

$$\begin{aligned} \|\frac{q_n}{d}\|_2^2 &= \sigma \\ -\frac{g_i}{\gamma_i g_{i+1}} &= \frac{-\delta_i^2 g_{i+1}}{\gamma_{i+1} g_i}, \qquad i = 1, \dots, n-1 \end{aligned} \tag{46}$$

or

$$\begin{aligned} \|\frac{q_n}{d}\|_2^2 &= \sigma \\ \frac{g_{i+1}^2}{g_i^2} &= \frac{\gamma_{i+1}}{\delta_i^2 \gamma_i}, \qquad i = 1, \dots, n-1 \end{aligned} \tag{47}$$

(with $\gamma_n = 1$). Observe that equations (47) are solvable, since by (23)

$$\gamma_i > 0, \qquad i = 1, \dots, n-1.$$

Now

$$\|\frac{t_n^*}{d}\|_2^2 = 1,$$

since the basis $\tilde{\mathcal{B}}$ was constructed by the Gram-Schmidt orthonormalization procedure; in view of the first equation of (47) this implies

$$g_n = \pm\sqrt{\sigma}; \tag{48}$$

the sign can be chosen arbitrarily. Then one solves the last $n-1$ equations in (47) for $i = n-1, \dots, 1$:

$$g_i = sgn(g_n)\sqrt{\frac{\gamma_i \delta_i^2}{\gamma_{i+1}} g_{i+1}^2}\,; \tag{49}$$

here the sign is now uniquely determined by the choice in (48) and the additional requirement

$$\alpha_i := \frac{g_i}{\gamma_i g_{i+1}} = -\frac{\beta_i g_{i+1}}{\gamma_{i+1} g_i} > 0, \tag{50}$$

which was formulated in the statement of the theorem.

It remains to calculate the maps B, C and D; for this purpose we take a closer look at ϕ. Since the multiplicity of the singular value σ is n, by Glover [1984] or Corollary 3.2,3. in Fuhrmann [1991], there exists a constant k such that

$$\frac{n}{d} + k = \lambda \frac{d^*}{d}, \tag{51}$$

where $\lambda = \epsilon\sigma$ and $\epsilon \in \{\pm 1\}$ is determined from (6), i.e.

$$nq_1 = \epsilon\sigma d^* q_1^* + d\pi. \tag{52}$$

Rewrite (51) as

$$\frac{n}{d} = \lambda \frac{d^*}{d} - k. \tag{53}$$

Now define

$$\frac{\tilde{n}}{d} := \frac{n}{d} - \phi(\infty). \tag{54}$$

Then $\frac{\tilde{n}}{d}$ is strictly proper, i.e.

$$\frac{\tilde{n}}{d} \in X^d.$$

Furthermore, by (53)

$$\frac{\tilde{n}}{d} = \lambda \frac{d^*}{d} - k - \phi(\infty) =: \lambda \frac{d^*}{d} - \bar{k}. \tag{55}$$

Moreover, comparison of the leading coefficients of the numerator polynomials in (55) yields

$$0 = \lambda(-1)^n - \bar{k}, \tag{56}$$

as d is monic and $deg\, \tilde{n} = n - 1$. Hence

$$\bar{k} = (-1)^n \lambda\, ,$$

and (55) gives

$$\frac{\tilde{n}}{d} = \lambda\left(\frac{d^*}{d} - (-1)^n\right). \tag{57}$$

Now, by setting

$$D := \phi(\infty),$$

it suffices to consider the shift realization of $\frac{\tilde{n}}{d}$. We begin by computing the representation of the input map

$$B = (b_1, \ldots, b_n)^T : \mathbb{R} \mapsto X^d$$

of the realization. We can write

$$\frac{\tilde{n}}{d} = \sum_{j=1}^{n} b_j \frac{q_j^*}{d}. \tag{58}$$

By the orthogonality of the basis we have

$$b_i = \frac{\left(\frac{\tilde{n}}{d}, \frac{q_i^*}{d}\right)}{\left(\frac{q_i^*}{d}, \frac{q_i^*}{d}\right)}, \qquad i = 1, \ldots, n. \tag{59}$$

Observe that from (52) we can conclude that

$$\tilde{n} q_i = \epsilon\sigma d^* q_i + d\tilde{\pi}_i, \qquad i = 1, \ldots, n. \tag{60}$$

Dividing this by dd^* and integrating over the contour $\hat{\gamma}$ one can calculate the numerator in expression (59); this yields

$$b_i = 0, \qquad i = 1, \ldots, n-1 \tag{61}$$

and in view of the normalization condition (7)

$$b_n = -\frac{1}{\|\frac{q_n^*}{d}\|_2^2} \cdot \epsilon\sigma q_{n,n-1} = -\epsilon q_{n,n-1} \tag{62}$$

The output map

$$C = (c_1, \ldots, c_n) : X^d \mapsto \mathbb{R}$$

is easily computed as

$$c_i = C(\frac{q_i^*}{d}) = (\frac{q_i^*}{d})_{-1} = \begin{cases} 0 & , i = 1, \ldots, n-1 \\ (-1)^{n-1} q_{n,n-1} & , i = n \end{cases}. \tag{63}$$

Setting $s := (-1)^n \epsilon$ we have verified relations (9) and (10).
Finally we have to check that equation (12) is true. Take equation (58), substitute expression (57) and multiply by d to obtain

$$\lambda(d^* - (-1)^n d) = \sum_{j=1}^{n} b_j q_j^* = b_n q_n^* \tag{64}$$

in view of (61). Comparison of the leading coefficients results in

$$\lambda((-1)^{n-1} d_{n-1} - (-1)^n d_{n-1}) = b_n (-1)^{n-1} q_{n,n-1}. \tag{65}$$

Furthermore, using (62) one obtains

$$\lambda(d_{n-1} + d_{n-1}) = -\epsilon q_{n,n-1}^2 \tag{66}$$

and

$$d_{n-1} = -\frac{\epsilon}{2\lambda} q_{n,n-1}^2 = -\frac{q_{n,n-1}^2}{2\sigma}, \tag{67}$$

which by (41) results in

$$\theta_n = \frac{q_{n,n-1}^2}{2\sigma}; \tag{68}$$

this is relation (12). ■

4 General result

Now we turn to the study of the general case, i.e. the case of several singular values whose multiplicity may be greater than 1. Again we identify the balanced canonical form obtained by Ober with a matrix representation of the shift realization. The matrix representation is with respect to a basis constructed from local families of orthogonal polynomials. Specifically, for each singular value σ of multiplicity ν we consider the minimum degree solution of the fundamental polynomial equation (6). For the corresponding Schmidt vector, whose numerator polynomial has degree $n-\nu$, we take the square of its absolute value as a weight function and compute a set of ν polynomials orthogonal with respect to this weight. With these polynomials we identify a corresponding set of Schmidt vectors. The union, over all singular values, of these sets of vectors provides an orthogonal basis. With an appropriate normalization we obtain the required basis.

Theorem 4.1 *Let $\phi = \frac{n}{d} \in H^\infty_-$, d monic, $n \wedge d = 1$ and $\deg d = n$. Let $\sigma_1 > \sigma_2 > \cdots > \sigma_k > 0$ be the singular values of the associated Hankel operator H_ϕ, where σ_j is of multiplicity n_j, $j = 1, \dots, k$, $\sum_{i=1}^{k} n_i = n$, and let $p_j^{(1)}$, $j = 1, \dots, k$ denote the minimal degree solutions, corresponding to σ_j, of the fundamental polynomial equation (6) such that $(p_j^{(1)})^*$ is monic. Finally, let $(t_j^{(\ell)})^*$, $\ell = 1, \dots n_j$ be the polynomials obtained from $\{1, z, \dots, z^{n_j-1}\}$ via the Gram-Schmidt orthonormalization procedure with respect to the weight function $\frac{|(p_j^{(1)})^*|^2}{|d|^2}$, for $j = 1, \dots, k$. Then a set of constants $\{g_\ell^j,\ \ell = 1, \dots n_j,\ j = 1, \dots, k\} \subseteq \mathbb{R}$ can be chosen such that for $(q_j^{(\ell)})^* := g_\ell^j (t_j^{(\ell)})^*$, $\ell = 1, \dots n_j$, $j = 1, \dots, k$ the normalization*

$$\left\|\frac{(q_j^{(n_j)})^*}{d}\right\|_2^2 = \sigma_j, \qquad j = 1, \cdots, k \tag{69}$$

holds and that the matrix representation of the shift realization of ϕ with respect to the basis $\{\frac{(q_j^{(\ell)})^*}{d},\ \ell = 1, \dots n_j,\ j = 1, \dots, k\}$ of X^d has the following form:

$$A = (A_{ij})_{i,j=1,\dots,k}, \qquad A_{ij} \in \mathbb{R}^{n_i + n_j},\ i, j = 1, \dots, k,$$

$$A_{ii} = \begin{pmatrix} 0 & -\alpha_1^i & 0 & \cdots & \cdots & 0 \\ \alpha_1^i & 0 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -\alpha_{n_i-2}^i & 0 \\ \vdots & & \ddots & \ddots & 0 & -\alpha_{n_i-1}^i \\ 0 & \cdots & \cdots & 0 & \alpha_{n_i-1}^i & a_{ii} \end{pmatrix}, \qquad i = 1, \dots, k, \tag{70}$$

$$\alpha_j^i > 0, j = 1, \dots, n_i - 1,\ i = 1, \dots, k,$$

$$A_{ij} = \begin{pmatrix} 0 & \cdots & \cdots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ \vdots & & & \vdots & 0 \\ 0 & \cdots & \cdots & 0 & a_{ij} \end{pmatrix}, \qquad i,j = 1,\ldots,k,\ i \neq j, \tag{71}$$

$$B = (\underbrace{0,\ldots,0,b_1}_{n_1}, \underbrace{0,\ldots,0,b_2}_{n_2},\ldots,\underbrace{0,\ldots,0,b_k}_{n_k})^T, \tag{72}$$

$$b_i = (-1)^{n-n_i-1}\epsilon_i q^{(n_i)}_{i,n-1},\ \epsilon_i = \pm 1,\ i = 1,\ldots,k,$$

$$C = (\underbrace{0,\ldots,0,c_1}_{n_1}, \underbrace{0,\ldots,0,c_2}_{n_2},\ldots,\underbrace{0,\ldots,0,c_k}_{n_k}), \tag{73}$$

$$c_i = (-1)^{n-1} q^{(n_i)}_{i,n-1} = (-1)^{n_i}\epsilon_i b_i,\ i = 1,\ldots,k,$$

$$D = \phi(\infty) \tag{74}$$

and

$$a_{ij} = \left(\frac{b_i b_j}{\epsilon_i(-1)^{n_i-1}\epsilon_j(-1)^{n_j-1}\sigma_i + \sigma_j} \right), \qquad i,j = 1,\ldots,k \tag{75}$$

Proof: Setting

$$D := \phi(\infty)$$

we may assume without loss of generality that $\frac{n}{d} \in H^\infty_-$ is strictly proper, i.e.

$$\frac{n}{d} \in X^d.$$

We start with the construction of the basis $\{\frac{(q_j^{(\ell)})^*}{d},\ \ell = 1,\ldots n_j,\ j = 1,\ldots,k\}$ of X^d. Let

$$p_j^{(1)}, \qquad j = 1,\ldots,k$$

denote the minimal degree solutions of the equations

$$np_j^{(1)} = \lambda_j d^*(p_j^{(1)})^* + d\pi_j, \qquad j = 1,\ldots,k \tag{76}$$

such that

$$(p_j^{(1)})^* \text{ is monic }, \qquad j = 1,\ldots,k \tag{77}$$

and

$$\lambda_j \in \mathbb{R},\ \lambda_j = \epsilon_j\sigma_j,\ \epsilon_j = \pm 1.$$

Then, by Corollary 3.2,1. in Fuhrmann [1991], we have

$$deg\, p_j^{(1)} = n - n_j, \qquad j = 1,\ldots,k. \tag{78}$$

Furthermore, again by Corollary 3.2,1. from the above reference

$$\mathcal{B}_j := \{\frac{(p_j^{(1)})^*}{d}, \frac{z(p_j^{(1)})^*}{d}, \ldots, \frac{z^{n_j-1}(p_j^{(1)})^*}{d}\}, \qquad j = 1, \ldots, k \tag{79}$$

is a basis for the space spanned by the singular vectors of H_ϕ corresponding to the singular value σ_j; denote

$$X_j := span\mathcal{B}_j, \qquad j = 1, \ldots, k. \tag{80}$$

Now apply to each of the bases $\mathcal{B}_j$ the Gram-Schmidt orthonormalization procedure to obtain bases

$$\tilde{\mathcal{B}}_j := \{\frac{(t_j^{(1)})^*}{d}, \frac{(t_j^{(2)})^*}{d}, \ldots, \frac{(t_j^{(n_j)})^*}{d}\}, \qquad j = 1, \ldots, k \tag{81}$$

of X_j with

$$deg\, t_j^{(\ell)} = n - n_j + \ell - 1, \qquad \ell = 1, \ldots, n_j,\ j = 1, \ldots, k. \tag{82}$$

Observe that

$$\begin{aligned} \frac{(t_j^{(\ell)})^*}{d} \quad &\perp \quad span\{\frac{p}{d} \in X_j,\ deg\, p < n - n_j + \ell - 1\} \\ &= \quad span\{\frac{(p_j^{(1)})^*}{d} q,\ deg\, q < \ell - 1\} \\ &= \quad span\{\frac{(p_j^{(1)})^*}{d}, \frac{z(p_j^{(1)})^*}{d}, \ldots, \frac{z^{\ell-2}(p_j^{(1)})^*}{d}\}; \end{aligned} \tag{83}$$

this is immediately clear from the construction. One also obtains the recursion formulas

$$(t_j^{(2)})^* = (\gamma_1^j z - \theta_1^j)(t_j^{(1)})^* \tag{84}$$

and

$$(t_j^{(\ell+1)})^* = (\gamma_\ell^j z - \theta_\ell^j)(t_j^{(\ell)})^* - \beta_{\ell-1}^j (t_j^{(\ell-1)})^*, \qquad \ell = 2, \ldots, n_j - 1 \tag{85}$$

for $j = 1, \ldots, k$, which are analogous to (19) and (20). Equation (84) is obvious from the Gram-Schmidt procedure. Since

$$\frac{z(t_j^{(\ell)})^*}{d} \in X_j, \qquad \ell = 2, \ldots, n_j - 1,\ j = 1, \ldots, k \tag{86}$$

one can prove (85) with the help of (83) exactly as (20).
The monicity of the $(p_j^{(1)})^*$, $j = 1, \ldots, k$ and the applied construction yield

$$\gamma_\ell^j > 0, \qquad \ell = 1, \ldots, n_j - 1,\ j = 1, \ldots, k. \tag{87}$$

Moreover, there holds

$$\theta_\ell^j = 0, \qquad \ell = 1, \ldots, n_j - 1,\ j = 1, \ldots, k; \tag{88}$$

this follows in the same way as (24).
Finally,

$$\beta^j_{\ell-1} < 0, \qquad \ell = 2, \dots, n_j - 1,\ j = 1, \dots, k; \tag{89}$$

this is proved analogously as relation (26).
By dividing equations (84) and (85) by d we obtain a representation of

$$S^d \frac{(t_j^{(\ell)})^*}{d} = \frac{z(t_j^{(\ell)})^*}{d}, \qquad \ell = 1, \dots, n_j - 1,\ j = 1, \dots, k$$

with respect to the bases $\tilde{\mathcal{B}}_j$, $j = 1, \dots, k$, respectively. Since

$$S^d \frac{(t_j^{(\ell)})^*}{d} \in X_j, \qquad \ell = 1, \dots, n_j - 1,\ j = 1, \dots, k$$

and

$$X_j \perp X_\ell, \quad \ell \neq j, \tag{90}$$

the representations obtained so far are also representations with respect to the basis

$$\tilde{\mathcal{B}} := \{\frac{(t_1^{(1)})^*}{d}, \dots, \frac{(t_1^{(n_1)})^*}{d}, \frac{(t_2^{(1)})^*}{d}, \dots, \frac{(t_2^{(n_2)})^*}{d}, \dots, \frac{(t_k^{(1)})^*}{d}, \dots, \frac{(t_k^{(n_k)})^*}{d}\} \tag{91}$$

of X^d. It remains to find a representation of

$$S^d \frac{(t_j^{(n_j)})^*}{d} = \frac{z(t_j^{(n_j)})^* - (-1)^{n-1} t_{j,n-1}^{(n_j)} d}{d}, \qquad j = 1, \dots, k \tag{92}$$

with respect to $\tilde{\mathcal{B}}$. Observe that in general

$$S^d \frac{(t_j^{(n_j)})^*}{d} \notin X_j, \qquad j = 1, \dots, k.$$

Let for $j \in \{1, \dots, k\}$

$$S^d \frac{(t_j^{(n_j)})^*}{d} = \sum_{i=1}^{k} \sum_{\ell=1}^{n_i} \alpha^j_{i,\ell} \frac{(t_i^{(\ell)})^*}{d}. \tag{93}$$

Now because of (90) and the orthonormality of the bases $\tilde{\mathcal{B}}_j$ there holds

$$\alpha^j_{i,\ell} = (S^d \frac{(t_j^{(n_j)})^*}{d}, \frac{(t_i^{(\ell)})^*}{d}), \qquad \ell = 1, \dots, n_i,\ i = 1, \dots, k. \tag{94}$$

The calculations done to obtain equation (31) are also valid here:

$$\alpha^j_{i,\ell} = (S^d \frac{(t_j^{(n_j)})^*}{d}, \frac{(t_i^{(\ell)})^*}{d}) = -(\frac{(t_j^{(n_j)})^*}{d}, \frac{z(t_i^{(\ell)})^*}{d}) \tag{95}$$

for $\ell = 1, \dots, n_i - 1$, $i = 1, \dots, k$.
First consider the case $i \neq j$; then

$$\frac{z(t_i^{(\ell)})^*}{d} \in X_i \perp X_j, \qquad \ell = 1, \dots, n_i - 1,$$

and hence

$$\alpha^j_{i,\ell} = 0, \qquad \ell = 1, \ldots, n_i - 1,\ i = 1, \ldots, k, i \neq j. \tag{96}$$

For $i = j$ we obtain by (82), (83) and (95) that

$$\alpha^j_{j,\ell} = 0, \qquad \ell = 1, \ldots, n_j - 2, \tag{97}$$

since

$$deg(z(t_j^{(\ell)})^*) = deg(t_j^{(\ell)})^* + 1 \leq (n - n_j + n_j - 3) + 1 = n - 2$$

for $\ell \in \{1, \ldots, n_j - 2\}$. The argument for the proof of (35) can be used to show that

$$\alpha^j_{j,n_j-1} < 0, \qquad j = 1, \ldots, k. \tag{98}$$

Summarizing, equations (84), (85) and (93) provide the following representation of the mapping

$$S^d : X^d \mapsto X^d$$

with respect to the basis $\tilde{\mathcal{B}}$:

$$[S^d]^{\tilde{\mathcal{B}}}_{\tilde{\mathcal{B}}} = (\tilde{A}_{ij})_{i,j=1,\ldots,k}, \tag{99}$$

where

$$\tilde{A}_{ij} = \begin{pmatrix} 0 & \cdots & \cdots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ \vdots & & & \vdots & 0 \\ 0 & \cdots & \cdots & 0 & \alpha^j_{i,n_i} \end{pmatrix}, \qquad i, j = 1, \ldots, k,\ i \neq j \tag{100}$$

and

$$\tilde{A}_{jj} = \begin{pmatrix} 0 & \frac{\beta^j_1}{\gamma^j_2} & 0 & \cdots & \cdots & 0 & 0 \\ \frac{1}{\gamma^j_1} & 0 & \frac{\beta^j_2}{\gamma^j_3} & \ddots & & \vdots & \vdots \\ 0 & \frac{1}{\gamma^j_2} & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \frac{\beta^j_{n_j-2}}{\gamma^j_{n_j-1}} & 0 \\ \vdots & & & \ddots & \ddots & 0 & \alpha^j_{j,n_j-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{\gamma^j_{n_j-1}} & \alpha^j_{j,n_j} \end{pmatrix}. \tag{101}$$

Finally, we can now determine the basis $\mathcal{B} := \{\frac{(q_j^{(\ell)})^*}{d}, \ell = 1, \ldots n_j,\ j = 1, \ldots, k\}$ from the statement of the theorem by defining

$$(q_j^{(\ell)})^* := g^j_\ell \cdot (t_j^{(\ell)})^*, \qquad g^j_\ell \in \mathbb{R}, \ell = 1, \ldots, n_j,\ j = 1, \ldots, k \tag{102}$$

Observe that for each $j \in \{1, \ldots, k\}$ the matrices $\tilde{A}_{jj}$ are of the form (43); also (87), (89) and (98) show that the signs of the elements of $\tilde{A}_{jj}$ coincide with the ones in (43); hence

$$A_{jj} := diag^{-1}(g_1^j, \ldots, g_{n_j}^j)\tilde{A}_{jj} diag(g_1^j, \ldots, g_{n_j}^j)$$

can be brought to the form (70) by solving the same set of equations as in the all-pass case. Moreover, because of (69) and the orthonormality of the bases $\tilde{\mathcal{B}}_j$, we have

$$g_{n_j}^j = \pm\sqrt{\sigma_j}, \qquad j = 1, \ldots, k. \tag{103}$$

Hence we have proved the following:

$$A = T^{-1} \, [S^d]_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}} T = (A_{ij})_{i,j=1,\ldots,k}, \tag{104}$$

where

$$T := diag(g_1^1, \ldots, g_{n_1}^1, g_1^2, \ldots, g_{n_2}^2, \ldots, g_1^k, \ldots, g_{n_k}^k)$$

with

$$A_{ii} := \begin{pmatrix} 0 & -\alpha_1^i & 0 & \cdots & \cdots & 0 & 0 \\ \alpha_1^i & 0 & \ddots & \ddots & & \vdots & \vdots \\ 0 & \ddots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & \ddots & -\alpha_{n_i-2}^i & 0 \\ \vdots & & & \ddots & \ddots & 0 & -\alpha_{n_i-1}^i \\ 0 & \cdots & \cdots & \cdots & 0 & -\alpha_{n_i-1}^i & a_{ii} \end{pmatrix}, \qquad i = 1, \ldots, k \tag{105}$$

and

$$A_{ij} := \begin{pmatrix} 0 & \cdots & \cdots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ \vdots & & & \vdots & 0 \\ 0 & \cdots & \cdots & 0 & a_{ij} \end{pmatrix}, \qquad i, j = 1, \ldots, k, \; i \neq j; \tag{106}$$

here the

$$\alpha_j^i, \qquad j = 1, \ldots, n_i - 1, \; i = 1, \ldots, k$$

are defined completely analogous as in (50); furthermore, observe that

$$a_{ij} = \frac{g_{n_j}^j}{g_{n_i}^i} \alpha_{i,n_i}^j, \qquad i, j = 1, \ldots, k. \tag{107}$$

The calculation of the c_i, $i = 1, \ldots, k$ is again obvious; hence (73) holds true. For the calculation of the B-matrix we make the following ansatz:

$$\frac{n}{d} = \sum_{i=1}^{k} \sum_{\ell=1}^{n_i} b_\ell^i \frac{(q_i^{(\ell)})^*}{d} \tag{108}$$

which by the orthogonality of the basis yields

$$b_\ell^i = \frac{(\frac{n}{d}, \frac{(q_i^{(\ell)})^*}{d})}{(\frac{(q_i^{(\ell)})^*}{d}, \frac{(q_i^{(\ell)})^*}{d})} \qquad \ell = 1, \ldots, n_i, \; i = 1, \ldots, k. \tag{109}$$

Since $\frac{(q_i^{(1)})^*}{d}$, $i = 1, \ldots, k$ are minimal degree Schmidt vectors, they satisfy the fundamental polynomial equation (6), i.e.

$$nq_i^{(1)} = \epsilon_i \sigma_i d^* (q_i^{(1)})^* + d\pi_i, \qquad i = 1, \ldots, k, \tag{110}$$

where $\epsilon_i \in \{\pm 1\}$ and π_i polynomial for $i \in \{1, \ldots, k\}$. Furthermore, all numerator polynomials of elements of X_i are of the form $q_i^{(1)} p$ with p a polynomial of degree less than n_i (see Fuhrmann [1991], Lemma 3.3). Thus in particular:

$$q_i^{(\ell)} = q_i^{(1)} a_\ell^i, \qquad \ell = 1, \ldots, n_i, \; i = 1, \ldots, k, \tag{111}$$

where the a_ℓ^i are polynomials of degree $\ell - 1$. In view of (84) and (85) it is clear that a_ℓ^i only contains even/odd powers of z for ℓ odd/even. So we obtain from (110) that

$$\begin{aligned} nq_i^{(\ell)} &= \epsilon_i \sigma_i d^* (q_i^{(1)})^* a_\ell^i + d(\pi_i a_\ell^i) \\ &= \epsilon_i \sigma_i (-1)^{\ell-1} d^* (q_i^{(1)})^* (a_\ell^i)^* + d(\pi_i a_\ell^i) \\ &= \epsilon_i \sigma_i (-1)^{\ell-1} d^* (q_i^{(\ell)})^* + d(\pi_i a_\ell^i) \end{aligned} \tag{112}$$

Division by dd^* and contour integration over $\hat{\gamma}$ gives

$$(\frac{n}{d}, \frac{(q_i^{(\ell)})^*}{d}) = \begin{cases} 0 & , \ell = 1, \ldots, n_i - 1 \\ -\epsilon_i \sigma_i (-1)^{n-n_i} q_{i,n-1}^{(n_i)} & , \ell = n_i \end{cases} \tag{113}$$

for $i = 1, \ldots, k$, which in view of normalization (69) results in (72).
Next we calculate a_{ij}, $i \neq j$; observe that from (112) we get

$$0 = d^* \{\epsilon_i \sigma_i (-1)^{n_i - 1} (q_i^{(n_i)})^* q_j^{(n_j)} - \epsilon_j \sigma_j (-1)^{n_j - 1} (q_j^{(n_j)})^* q_i^{(n_i)}\}$$
$$+ d\{\pi_i a_{n_i}^i q_j^{(n_j)} - \pi_j a_{n_j}^j q_i^{(n_i)}\}.$$

Since d and d^* are coprime, there exists a polynomial x_{ij} of degree less than $n - 1$ such that

$$dx_{ij} = \epsilon_i \sigma_i (-1)^{n_i - 1} (q_i^{(n_i)})^* q_j^{(n_j)} - \epsilon_j \sigma_j (-1)^{n_j - 1} (q_j^{(n_j)})^* q_i^{(n_i)} \tag{114}$$

After some algebra one obtains

$$q_i^{(n_i)} (q_j^{(n_j)})^* = \frac{\epsilon_j \sigma_j (-1)^{n_j - 1}}{\sigma_i^2 - \sigma_j^2} dx_{ij} + \frac{\epsilon_i \sigma_i (-1)^{n_i - 1}}{\sigma_i^2 - \sigma_j^2} d^* x_{ij}^* \tag{115}$$

(see also Fuhrmann [1991]). Now from (107), (95), (102) and (103) we get the following representation of a_{ij}:

$$\begin{aligned} a_{ij} &= \frac{g^j_{n_j}}{g^i_{n_i}}\alpha^j_{i,n_i} \\ &= \frac{g^j_{n_j}}{g^i_{n_i}}(S^d\,\frac{(t_j^{(n_j)})^*}{d},\frac{(t_i^{(n_i)})^*}{d}) \\ &= \frac{g^j_{n_j}}{g^i_{n_i}}\frac{1}{g^j_{n_j}g^i_{n_i}}(S^d\,\frac{(q_j^{(n_j)})^*}{d},\frac{(q_i^{(n_i)})^*}{d}) \\ &= \frac{1}{\sigma_i}(S^d\,\frac{(q_j^{(n_j)})^*}{d},\frac{(q_i^{(n_i)})^*}{d}). \end{aligned} \tag{116}$$

Moreover, using

$$S^d\,\frac{(q_j^{(n_j)})^*}{d} = \frac{z(q_j^{(n_j)})^* - (-1)^{n-1}q^{(n_j)}_{j,n-1}d}{d}$$

and integrating over contour $\hat{\gamma}$ yields

$$(S^d\,\frac{(q_j^{(n_j)})^*}{d},\frac{(q_i^{(n_i)})^*}{d}) = \lim_{R\mapsto\infty}\frac{1}{2\pi i}\int_{\hat{\gamma}}\frac{z(q_j^{(n_j)})^* q_i^{(n_i)}}{dd^*}dz \tag{117}$$

which in view of (115) gives

$$(S^d\,\frac{(q_j^{(n_j)})^*}{d},\frac{(q_i^{(n_i)})^*}{d}) = \frac{\epsilon_i\sigma_i(-1)^{n_i-1}}{\sigma_i^2-\sigma_j^2}(-1)^{n-1}x_{ij,n-2} \tag{118}$$

On the other hand, integration over γ and application of (115) results in

$$\begin{aligned} (S^d\,\frac{(q_j^{(n_j)})^*}{d},\frac{(q_i^{(n_i)})^*}{d}) &= \lim_{R\mapsto\infty}\frac{1}{2\pi i}\int_{\gamma}\frac{z(q_j^{(n_j)})^* q_i^{(n_i)}}{dd^*}dz + q^{(n_j)}_{j,n-1}q^{(n_i)}_{i,n-1} \\ &= \frac{\epsilon_j\sigma_j(-1)^{n_j-1}}{\sigma_i^2-\sigma_j^2}(-1)^n x_{ij,n-2} + q^{(n_j)}_{j,n-1}q^{(n_i)}_{i,n-1}. \end{aligned} \tag{119}$$

Equating (118) and (119) shows that

$$x_{ij,n-2} = \frac{\sigma_i^2-\sigma_j^2}{\epsilon_i\sigma_i(-1)^{n_i-1}+\epsilon_j\sigma_j(-1)^{n_j-1}}(-1)^{n-1}q^{(n_j)}_{j,n-1}q^{(n_i)}_{i,n-1} \tag{120}$$

which plugged in in (118) and (116) gives

$$a_{ij} = \frac{\epsilon_i(-1)^{n_i-1}}{\epsilon_i\sigma_i(-1)^{n_i-1}+\epsilon_j\sigma_j(-1)^{n_j-1}}q^{(n_j)}_{j,n-1}q^{(n_i)}_{i,n-1}\,;$$

this is (75) for the case $i \neq j$.
Finally, by (94) and (107) there holds

$$a_{ii} = \alpha^i_{i,n_i} = (S^d\,\frac{(t_i^{(n_i)})^*}{d},\frac{(t_i^{(n_i)})^*}{d}) \tag{121}$$

and in view of (102) and (103)

$$
\begin{aligned}
a_{ii} &= \frac{1}{(g^i_{n_i})^2}(S^d \frac{(q_i^{(n_i)})^*}{d}, \frac{(q_i^{(n_i)})^*}{d}) \\
&= \frac{1}{\sigma_i}(S^d \frac{(q_i^{(n_i)})^*}{d}, \frac{(q_i^{(n_i)})^*}{d}). \qquad (122)
\end{aligned}
$$

Moreover, using the previously defined contours γ and $\hat{\gamma}$ we have

$$
\begin{aligned}
(S^d \frac{(q_i^{(n_i)})^*}{d}, \frac{(q_i^{(n_i)})^*}{d}) &= (\frac{z(q_i^{(n_i)})^* - (-1)^{n-1} q_{i,n-1}^{(n_i)} d}{d}, \frac{(q_i^{(n_i)})^*}{d}) \\
&= \frac{1}{2\pi}\int_{-i\infty}^{i\infty} \frac{z(q_i^{(n_i)})^* - (-1)^{n-1} q_{i,n-1}^{(n_i)} d}{d} \cdot \frac{q_i^{(n_i)}}{d^*} dz \\
&= \lim_{R\mapsto\infty} \frac{1}{2\pi i}\langle \int_{\hat{\gamma}} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz \qquad (123) \\
&\quad -(-1)^{n-1} q_{i,n-1}^{(n_i)} \int_{\hat{\gamma}} \frac{q_i^{(n_i)}}{d^*} dz \rangle \\
&= \lim_{R\mapsto\infty} \frac{1}{2\pi i}\int_{\hat{\gamma}} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz
\end{aligned}
$$

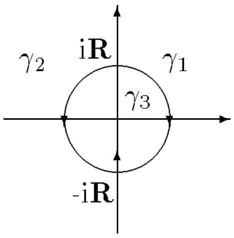
because of the stability of d^*. On the other hand

$$
\begin{aligned}
(S^d \frac{(q_i^{(n_i)})^*}{d}, \frac{(q_i^{(n_i)})^*}{d}) &= \lim_{R\mapsto\infty} \frac{1}{2\pi i}\langle \int_{\gamma} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz \\
&\quad -(-1)^{n-1} q_{i,n-1}^{(n_i)} \int_{\gamma} \frac{q_i^{(n_i)}}{d^*} dz \rangle \qquad (124) \\
&= \lim_{R\mapsto\infty} \frac{1}{2\pi i}\int_{\gamma} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz \\
&\quad -(-1)^{n-1}(-1)^n (q_{i,n-1}^{(n_i)})^2.
\end{aligned}
$$

Hence summing (123) and (124) one obtains

$$
\begin{aligned}
(S^d \frac{(q_i^{(n_i)})^*}{d}, \frac{(q_i^{(n_i)})^*}{d}) &= \frac{1}{2}\lim_{R\mapsto\infty} \frac{1}{2\pi i}\langle \int_{\gamma} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz \\
&\quad + \int_{\hat{\gamma}} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz \rangle + \frac{1}{2}(q_{i,n-1}^{(n_i)})^2. \qquad (125)
\end{aligned}
$$

Now we decompose γ and $\hat{\gamma}$ in the way indicated below.



Observe that $\gamma = \gamma_2 + \gamma_3$ and $\hat{\gamma} = \gamma_1 + \gamma_3$. Then

$$\lim_{R\mapsto\infty} \frac{1}{2\pi i} \int_{\gamma_3} \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*} dz = \lim_{R\mapsto\infty} \frac{1}{2\pi i} \int_{\gamma_3} z \mid \frac{q_i^{(n_i)}}{d} \mid^2 dz = 0. \tag{126}$$

Let

$$\varphi_1 : [0, \pi] \mapsto \mathcal{C}, \quad t \mapsto Re^{i(\frac{\pi}{2}-t)}$$

be a parametrization of γ_1 and

$$\varphi_2 : [0, \pi] \mapsto \mathcal{C}, \quad t \mapsto Re^{i(\frac{\pi}{2}+t)}$$

be a parametrization of γ_2. It is easily calculated that

$$\varphi_1 = -\varphi_2(\pi - t). \tag{127}$$

Furthermore, for

$$g(z) := \frac{z(q_i^{(n_i)})^* q_i^{(n_i)}}{dd^*}$$

there holds

$$g(-z) = -g(z). \tag{128}$$

So finally we can show that

$$\begin{aligned} \int_{\gamma_1} g(z)dz &= \int_0^\pi g(\varphi_1(t))\dot{\varphi}_1(t)dt \\ &= \int_0^\pi g(-\varphi_2(\pi - t))\dot{\varphi}_2(\pi - t)dt \\ &= -\int_0^\pi g(\varphi_2(\pi - t))\dot{\varphi}_2(\pi - t)dt \\ &= \int_\pi^0 g(\varphi_2(\tau))\dot{\varphi}_2(\tau)d\tau \\ &= -\int_0^\pi g(\varphi_2(\tau))\dot{\varphi}_2(\tau)d\tau \\ &= -\int_{\gamma_2} g(z)dz. \end{aligned} \tag{129}$$

Because of (126) and (129) equation (125) is reduced to

$$(S^d \frac{(q_i^{(n_i)})^*}{d}, \frac{(q_i^{(n_i)})^*}{d}) = \frac{1}{2} (q_{i,n-1}^{(n_i)})^2.$$

Together with (122) one obtains

$$a_{ii} = \frac{(q_{i,n-1}^{(n_i)})^2}{2\sigma_i} = \frac{b_i^2}{2\sigma_i}, \qquad i = 1, \ldots, k,$$

which is (75) for the case $i = j$. ■

References

[1965] N.I. Akhiezer, *The classical moment problem*, Haffner Publishing Co., New York.

[1976] P. A. Fuhrmann, "Algebraic system theory: An analyst's point of view", *J. Franklin Inst.*, 301, 521-540.

[1977] P. A. Fuhrmann, "On strict system equivalence and similarity", *Int. J. Contr.*, 25, 5-10.

[1981] P. A. Fuhrmann, *Linear systems and operators in Hilbert space*, McGraw-Hill, New York.

[1983] P. A. Fuhrmann, "On symmetric rational transfer functions", *Lin. Alg. Appl.*, 50, 167-250.

[1984] P. A. Fuhrmann, "On Hamiltonian transfer fuctions", *Lin. Alg. Appl.*, 84, 1-93.

[1991] P. A. Fuhrmann, "A polynomial approach to Hankel norm and balanced approximations", *Lin. Alg. Appl.*, 146, 133-220.

[1993] P. A. Fuhrmann, "An algebraic approach to Hankel norm approximation problems", *L. Markus Festschrift.*

[1993] P. A. Fuhrmann and R. Ober, " A functional approach to LQG balancing", to appear *Int. J. Contr.*.

[1984] K. Glover, "All optimal Hankel-norm approximations and their L^∞-error bounds", *Int. J. Contr.*, 39, 1115-1193.

[1986] K. Glover, "Robust stabilization of linear multivariable systems, relations to approximation", *Int. J. Contr.*, 43, 741-766.

[1974] W.B. Gragg, "Matrix interpretation and applications of the continued fraction algorithm", *Rocky Mount. J. of Math.*, 4, 213-225.

[1976] U. Helmke and P. A. Fuhrmann, "Bezoutians", *Lin. Alg. Appl.*, 122-124, 1039-1097.

[1983] E. Jonckheere and L. Silverman, "A new set of invariants for linear systems - application to reduced order compensator design", *IEEE Trans. on Auto. Contr.*, 28, 953-964.

[1980] S. Kung, "Optimal Hankel-norm reductions: scalar systems", *1980 Proc. Joint Automat. Contr. Conf.*, San Francisco.

[1989] D. McFarlane and K. Glover, "Robust controller design using normalized coprime factor plant descriptions", *Lecture Notes in Control and Information Sciences*, vol. 10, Springer Verlag.

[1981] B. C. Moore, "Principal component analysis in linear systems: controllability, observability and model reduction", *IEEE Trans. on Auto. Contr.*, 26, 17-32.

[1987a] R. Ober, "Balanced realizations: canonical form, parametrization, model reduction", *Int. J. Contr.*, 46, 643-670.

[1987b] R. Ober, "Asymptotically stable allpass transfer functions: canonical form, parametrization and realization", *Proceedings IFAC World Congress*, Munich 1987.

[1989] R. Ober, "Balanced parametrization of classes of linear systems", *SIAM J. Contr. and Optim.*, 29, 1251-1287.

[1982] L. Pernebo and L. Silverman, "Model reduction via balanced state space representations", *IEEE Trans. on Auto. Contr.*, 37, 382-387.

[1959] G. Szegö, *Orthogonal Polynomials*, Amer. Math. Soc., Providence, R.I..

[1948] H.S. Wall, *Analytic Theory of Continued Fractions*, D. Van Nostrand Co., Princeton.