INT. J. CONTROL, 1993, VOL. 57, NO. 3, 627–741

A functional approach to LQG balancing

P. A. FUHRMANN†‡ and R. OBER§

The paper has as its theme a circle of problems related to LQG balancing, with a special emphasis on the related problems of model reduction and robust stabilization. The aim is to present a unified point of view to both previously described problems. The unification is achieved by focusing on the study of three functions and the relationships between them and the various operators that are associated with these functions. With an arbitrary transfer function G one can associate, canonically, two objects—the conjugate inner function $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$, which is based on the NRCF, and the R^*, which is associated with the LQG controller of the function G. The approach that is taken is functional emphasizing operators. A balanced realization of a stable g arises as a matrix representation of the shift realization, with respect to a basis made out of suitably normalized Hankel singular vectors. A similar result holds for LQG balanced realizations. Here, the underlying Hankel operator we study is H_{R^*}, where $R = U^*M + V^*N$ and U, V solve the H^∞-Bezout equation $\bar{M}V - \bar{N}U = I$. This Hankel operator has the same singular vectors, though different singular values and Schmidt pairs, as $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$. The basis of singular vectors of H_{R^*} determines canonically a basis for the polynomial model shift realization of g corresponding to which the matrix representation is LQG balanced. One of the central results is that $\begin{bmatrix} \bar{N}_n^* \\ \bar{M}_n^* \end{bmatrix}$, the optimal Hankel norm approximant of $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$, is up to a scaling factor also conjugate inner. Denoting by R_n^* the symbol associated with the LQG controller of g_n we show that $-R_n^*$ is the strictly proper part of the best $n-1$ order Hankel norm approximant of R^*. We will also obtain state-space representations for R_n^* and g_n in terms of the parameters in the LQG balanced state space representation of g. Similar results hold for the case of Nehari complements. These are applied to robust control. As a result of this study the problems of model reduction and robust stabilization can be viewed as dual problems.

List of symbols

Ω_K, Q_{Ω_K}	see Theorem 3.1
Ω_I, Q_{Ω_I}	see Theorem 3.2
J_1, J_2	see Theorem 3.1
Im, Ker	Image, respectively kernel, of an operator
H_+^2, H_-^2	Hardy space of square integrable functions in the complex right, respectively left, half plane
H_+^∞, H_-^∞	Hardy space of bounded functions in the complex right, respectively left, half plane

Received 15 September 1991. Revised 10 March 1992. Communicated by Professor V. Kučera.

† Department of Mathematics, Ben-Gurion University of the Negev, Beer Sheva, Israel.
‡ Earl Katz Family Chair in Algebraic System Theory.
§ Center for Engineering Mathematics, Programs in Mathematical Sciences, The University of Texas at Dallas, Richardson, TX 75083–06688, U.S.A.

$H^{\infty}_{[k]}$	functions in L^{∞} with k poles in the open right half plane
NRCF, NLCF	normalized right, respectively left, coprime factorization, see Definition 3.1
R, $\bar{R}$	see Lemma 4.1 and Lemma 4.2
$\bar{U}_L$, $\bar{V}_L$, U_L, V_L	coprime factors of LQG controller, see Lemma 4.1 and Lemma 4.2
S_K, S_I	see Lemma 5.1
ϕ_K, ϕ_I	see Lemma 5.1 and Corollary 5.1
$A\|\mathcal{A}$	restriction of the operator A to the subspace $\mathcal{A}$
Z_K, Y_K; Z_I, Y_I	see Theorem 5.1; Theorem 5.2
X^p, X_p	rational, respectively polynomial, model; see § 2
$p\|q$	polynomial p divides polynomial q
π_-, π_D	see § 2
$\mathcal{A}^{\perp}$	orthogonal complement of $\mathcal{A}$
H_G, $\hat{H}_G$	Hankel operator, respectively involuted Hankel operator with symbol G, see § 2
$\\|A\\|$	norm of the operator A
A^T	transpose of the matrix A
$\tilde{A}$	complex conjugate transpose of the matrix A
P_+, P_-	orthogonal projection onto H^2_+, respectively H^2_-
$\mathcal{A} \oplus \mathcal{B}$	orthogonal sum of $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \ominus \mathcal{B}$	orthogonal complement of $\mathcal{B}$ in $\mathcal{A}$
I	identity operator
$p \wedge q$	greatest common divisor of p and q
$r, t, p_i, \pi_i, \varepsilon_i, \lambda_i$	see Proposition 8.1
g, e, d	$g = e/d$, with e and d coprime
$\hat{p}_1^{(i)}, \hat{p}_2^{(i)}, \pi_1^{(i)}, \pi_2^{(i)}$	see Equation 92 and Equation 93
$\sigma_1, \ldots, \sigma_n$	see Theorem 6.2
$\mu_1, \ldots, \mu_n$	see Theorem 9.1
$P_{\mathcal{A}}$	orthogonal projection onto the space $\mathcal{A}$
$q_{k,i}$	ith coefficient of the polynomial q_k
δ_{ij}	$\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ for $i \neq j$
α_i	Theorem 9.1
β_i	Theorem 10.1
π_i	Proposition 8.1
ζ_i	Theorem 9.1
α_{ij}	Proposition 8.1
ω_{ij}	Proposition 12.1
$\begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix}$	Lemma 9.3
$\begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \end{bmatrix}$	Lemma 10.1
$\begin{bmatrix} \zeta_1^{(i)} \\ \zeta_2^{(i)} \end{bmatrix}$	Proposition 9.1

$\begin{bmatrix} \omega_1^{(i)} \\ \omega_2^{(i)} \end{bmatrix}$	Equation 184
$\alpha_{i,n-2}$	Equation 204
$\zeta_{i,n-2}$	Equation 240
$\alpha_{ij,n-2}$	Equation 203
$\alpha_{2,n-2}^{(i)}$	Equation 239
$\omega_{ij,n-3}$	Equation 253
$\hat{p}_{1,n-1}$	Equation 205
$\hat{p}_{2,n-1}$	Equation 206
$\pi_{1,n-1}^{(i)}$	Equation 207
$\pi_{2,n-1}^{(i)}$	Equation 208
ν_i	Equation 255
ρ_i	Proposition 12.4
τ_i	Theorem 13.5
η_i	Theorem 13.5
$\left[\begin{array}{c\|c} a_{ij} & b_j \\ \hline c_i & D \end{array}\right]$	matrix elements of the system matrices of the system (A, B, C, D)

1. Introduction

This paper has as its theme a circle of problems related to LQG balancing, with a special emphasis on the related problems of model reduction and robust stabilization. The paper studies these issues from the point of view of normalized coprime factorizations and the associated Hankel operators. It gives a detailed analysis of these operators. The main emphasis, and the main contribution, is in the attempt to clarify the relationships between these objects. We provide an approach that we believe will be central to a wide variety of problems in the general area of model reduction and robust control.

From the technical point of view, this paper is a continuation of Fuhrmann (1991), where polynomial methods for the analysis of Hankel norm approximation problems and those of the related Lyapunov balancing were developed. In that paper a duality theory for Nehari complementation and optimal Hankel norm approximation was established. It turns out in this paper that this duality is the foundation on which a more elaborate duality theory can be developed.

From the conceptual point of view the work in this paper has several roots. As so much else in the general H^∞-control area, it owes a lot to the pioneering work of Adamjan *et al.* (1968 a, b, 1971, 1978) on Hankel norm approximation. The notion of Lyapunov balancing was introduced by Moore (1981). The main principle of balanced realizations is that the solution of a matrix equation in the system's matrices is balanced with respect to the solution of a dual equation. In the case of LQG balanced realizations these equations are the control and filter Riccati equation, as they appear in LQG control theory. In the case of Lyapunov balanced realizations the two equations are the control and observer Lyapunov equation.

From the point of view of model reduction, the idea was to eliminate some dynamical elements which are relatively inessential. Of course, the interest in model reduction is external, i.e. based on input–output considerations. However, external and internal considerations in system theory are very strongly coupled. Thus, it turns out to be convenient to internalize the process of characterizing the non-essential modes. Therefore we look for a coordinate system which exhibits, in terms of certain weights, the contribution of various states to the input–output behaviour (see also Pernebo and Silverman 1982).

This was the approach initiated by Moore. The class of systems considered was the class of stable transfer functions and the weights turned out to be the Hankel singular values. Initially the process was mostly heuristic and the question of finding error bounds for the procedure of Lyapunov balanced truncation was left open. The gap was closed in Glover (1984) and Enns (1984), where the truncation error was bounded by twice the sum of the singular values of the truncated modes. That this bound is tight for the class of relaxation systems was shown in Ober (1987).

The success of the model reduction technique based on balanced coordinates suggested that the approach could be generalized to wider classes of systems. Indeed, Jonckheere and Silverman (1983) introduced LQG balancing which uses algebraic control and filtering Riccati equations rather than the controllability and observability Lyapunov equations which were used in the stable case.

Now, it is very well known that the solution of algebraic Riccati equations is essentially equivalent to spectral factorization problems. This equivalence brings LQG balancing into close proximity to the approach to system theory that is based on coprime factorizations over the ring of stable transfer functions. This approach has long been advocated by Vidyasagar (1985).

The problem of dynamic stabilization of finite-dimensional plants leads to a certain polynomial equation for the denominator of a polynomial coprime factorization. In the framework of H^∞ this polynomial equation can be transformed to a Bezout equation. The role of the H^∞ Bezout equation in the analysis of control problems is absolutely central and its importance cannot be overemphasized. Now, given a rational plant, coprime factorizations over the ring of stable transfer functions are anything but unique. However, in the class of all coprime factorizations there is one pair which is canonical, up to a unitary factor, and these are the normalized coprime factorizations (NCF). These are obtained via spectral factorizations and they establish the connection with the Riccati equation.

Normalized coprime factorizations as representations of transfer functions of linear dynamical systems have gained importance in many areas of control and systems theory. In a paper by Ober and McFarlane (1989) it was shown how they can be used to study LQG balanced realizations. In this paper a close connection was established between the Lyapunov balanced realization of the normalized coprime factors and the LQG balanced realization of the transfer function itself. One of the disadvantages of the state-space approach taken in that paper is that it does not provide good insight into the principles underlying the connection between these two representations. In this paper, we therefore study the problem from an input–output point of view. More precisely we study the problem by using operator theoretic methods and the theory of polynomial models.

In this context two Hankel operators play a distinctive role. One is the Hankel operator based on the normalized coprime factors, and the other one is a symbol related to the LQG controller. These two, or rather two pairs of, Hankel operators are very closely related. In fact they turn out to share the same singular vectors, although with different singular values. This observation allows us to study the Hankel norm approximation problem for one of the Hankel operators in terms of the other. This leads to interesting results concerning vectorial Hankel operators.

At this point the connection to LQG balancing almost suggests itself. In Fuhrmann (1976) a basis free realization method, referred to as the shift realization, based on polynomially coprime factorizations, was developed. In a series of papers, Fuhrmann (1976–1991), showed how various canonical forms can be obtained from the shift realization by choosing a suitable basis and computing the matrix representation of the shift realization with respect to that basis. In fact the last, and very relevant, result in this direction was obtained by Fuhrmann (1991), where it was shown that the Lyapunov balanced realization of an (anti)stable transfer function was obtainable by choosing a basis made of, suitably normalized, Hankel singular vectors. This idea we use here to advantage. However, a modification has to be made. Since we are dealing with not necessarily stable plants, the Hankel operator itself is not the right tool. Rather, we go through the indirect process of obtaining a normalized coprime factorization and, from it, the LQG symbol which is stable. For the associated Hankel operator we obtain a basis of Hankel singular vectors. There is a natural lifting of this basis to the state-space for the shift realization of the original plant. Again, with a suitable normalization, this leads to LQG balancing. In fact the LQG singular values are exactly the Lyapunov singular values of the Hankel operator with the LQG symbol.

Now we are in the position to use the results of Fuhrmann (1991). In particular we can study the Hankel singular values and singular vectors of the optimal Hankel norm approximant to the Hankel operators with the NCF and LQG symbols. Here we get results, generalizing those of Glover (1984) and Fuhrmann (1991), to the situation at hand. The same is true for the Hankel operators based on Nehari complements. In the case of vectorial symbols the previously mentioned results have to be modified somewhat. Either the singular values are weighted by a constant factor, or alternatively, one has to renormalize the Nehari complement or the optimal Hankel norm approximant.

One of the advantages of Lyapunov balancing was that, given a Lyapunov balanced realization of a stable plant, the derivation of a Lyapunov balanced realization of the Nehari complement became immediate through a trivial procedure (Theorem 3.4 in Fuhrmann 1991). Because of the way we interpret LQC balanced realizations, this group of results can be lifted to the context of LQG balancing. Also, the LQG balanced truncation is, in terms of the original plant, equivalent to the approximation in a space with weights arising from both the plant and controller.

However, more connections are illuminated, and duality is one. The relation between optimally robust control, with respect to the coprime factor uncertainty, is seen to be dual to optimal Hankel norm approximation of the coprime factors, which is related to model reduction. This duality can be lifted to the level of the original plant. Now this duality is not completely new, and we digress a bit on this point.

Since the advent of the new H^∞-control theory, as advocated initially by

Zames (1981), interest in algebraic system theory has been declining. Our position is that it is seldom the case that one theory completely supersedes a preceding one. In fact, the insight provided by the algebraic theory can be very helpful in providing motivation and intuition to the new area. Now, one thing that became apparent in the area of algebraic system theory is the duality between problems of stable partial realizations and stabilization by (dynamic) output feedback.

An early instance of this duality can be found in Fuhrmann (1985). It is implicit in some of Antoulas' work (1985) and made explicit, in the scalar case, in Rantzer (1989). There are at least two ways in which one can view partial realizations. The first is the usual, i.e. realization theory based on partial data. However, we can also view partial realization theory as an algebraic model reduction technique. Starting from a given plant, or transfer function, we construct an approximating one, where the degree of approximation is the number of Markov parameters that are equal. With that in mind one sees that the insight into the duality, within the H^∞ context, between model reduction and stabilization, becomes expected and natural.

The approach taken in this paper is mostly a highly explicit and computational one. Since our interest at this stage is to bring out the conceptually new aspects, we have, to a large extent, restricted outselves to scalar plants in the generic situation where all LQG singular values are distinct. This approach, a continuation of Fuhrmann (1991), concentrates on the level of polynomial equations. Most of the results can be easily interpreted in state-space terms, using the theory of polynomial models. Also, the rational models associated with stable or antistable transfer functions, and the LQG symbols are of this type, provide a convenient link between algebra and analysis. In fact, while the derivation of the LQG balanced realization is certainly non-trivial, it has the advantage of bypassing the need to construct continuous time realizations for H^∞ transfer functions, a process that can be problematic.

A word about connections to a geometric approach to H^∞-control theory, is in order. In this connection we would like to mention the striking formula for the stability margin of the optimally robust controller obtained by McFarlane and Glover (1990). Recently there have been a series of attempts at interpreting problems of robust control in geometric terms. By geometric terms we mean principally the geometry of graph spaces of both plant and controller. For some work in this direction we point out Hammer (1985), Verma (1988), Foias *et al.* (1990), Ober and Sefton (1990). Since the results of Fuhrmann (1991) on the geometry of singular vectors of the plant and its best Hankel norm approximant extend, with minor modifications, to the situation studied in the present paper, it is expected that these could be interpreted in terms of graph spaces and their orthogonal complements.

The paper is structured as follows. Section 2 contains the barest outline needed from polynomial model theory, including the derivation of the shift realization.

In § 3 we introduce and investigate normalized coprime factorizations and the Hankel operators associated with them. In particular, we characterize the kernels and images of the various Hankel operators.

The section after introduces the LQG controller and the associated LQG symbol. We obtain an interesting representation of the LQG controller and interpret it geometrically in terms of Hankel operator ranges. Also, we compute a representation of the kernel of the Hankel operator with a normalized coprime factor symbol.

In § 5 it is shown how all the operators that were introduced up to this point are related to one another. In particular we established explicit connections between the singular values and singular vectors of the different operators.

We pass on, in § 7, to re-derive several known state-space formulae. These include state-space representations for the normalized coprime factors. These formulas were originally derived by Meyer and Franklin (1987) and were proved directly, although the intuition remained somewhat obscure. We use polynomial model methods to derive the formulae and we believe that this derivation sheds more light on the problem. The method used seems to be well suited for related state-space derivations. These formulae lead also to state-space representations for the LQG controller and symbol. These can also be found in Glover and McFarlane (1988).

The role of the LQG symbol and its associated Hankel operator in the theory is extremely important and not yet fully understood. It certainly is located at the crossroads of several different research directions; in particular, in connection with geometric analysis.

The LQG symbol provides also a parametrization of arbitrary plants via stable ones. This map from plants to their associated LQG symbol is bijective. A better understanding of this map might shed some light on global topological properties of the space of all transfer functions. Certainly it might be relevant in comparing the cell decompositions of this space and that of the space of all stable transfer functions. In this connection we mention Fuhrmann and Krishnaprasad (1986) and Helmke *et al.* (1988), Ober (1989).

In § 8 we recover the main results on singular-value singular-vector analysis of (anti)stable transfer functions as derived in Fuhrmann (1991). Coupled with the results of § 5 this leads to the singular value analysis of the NCF Hankel operator. Furthermore we analyse the singular values and vectors of the best Hankel norm approximant corresponding to the least-singular value in § 9. By this we generalize results of Glover (1984) and Fuhrmann (1991). The analogous results for the singular value analysis of the Nehari complement are described in § 10.

In § 11, the identification is considered of two instances of balanced realization with matrix representations of the shift realization. We begin by showing that the matrix representation of the shift realization of the normalized coprime factors, taken with respect to a basis of suitably normalized Hankel singular vectors is Lyapunov balanced and coincides with the canonical form obtained previously by Ober and McFarlane (1989). Next we identify the LQG balanced realization with the matrix representations of the shift realization, the representation being with respect to a basis derived from the singular vectors of the LQG Hankel operator.

The LQG approximation problem is studied in § 12. This is done by introducing a special norm, amounting to looking at a weighted space, in the rational model state-space associated with the plant. We compute the balanced realization of the LQG approximant. The passage from the LQG balanced realization of the plant to that of the approximant is equivalent to the analogous case in Lyapunov balancing, see Theorem 8.2 of Fuhrmann (1991).

Finally, in § 13 we discuss briefly the optimally robust stabilization problem. From McFarlane and Glover (1990) we know that the optimally robust controller is related to the Nehari complement of the NCF symbol. This is re-proved differently in the context we have chosen. We proceed to do the Hankel singular value analysis on the renormalized Nehari complement. This is the dual situation to that handled in § 11. In particular, this leads to the direct derivation of a balanced realization of the optimally robust controller from that of the plant.

It will be quite clear to anybody who seriously studies this paper that, in spite of its length, it hardly scratches the surface of the whole body of research in this area of system theory. Much more remains to be done, and those who will participate in the effort will certainly find the experience rewarding.

We would also like to take the opportunity to thank James Sefton for interesting comments and discussions. We refer to his thesis (Sefton 1991) for an alternative point of view and different proofs to a number of the results in this paper.

After submission of the paper we received a reprint (Georgiou and Smith 1991) in which the authors also prove Lemma 9.1.1 and part of Theorem 9.3.1.

For ease of use we include the following reference table to the various balanced realizations handled in this paper.

Symbol	Type	Reference
g	LQG	Theorem 11.2
g_1	LQG	Theorem 13.5
g_n	LQG	Theorem 12.2
R^*	Lyapunov	Corollary 11.1
R_n^*	Lyapunov	Proposition 12.4
R_1^*	Lyapunov	Theorem 13.5
$\begin{bmatrix} -e^*/t^* \\ d^*/t^*) \end{bmatrix}$	Lyapunov	Theorem 11.1

2. Realization theory and preliminaries

We give a very short review of the basics of the theory of polynomial models. For more information on this we refer to Fuhrmann (1981) and the references therein. The theory of polynomial models builds on the abstract module theoretic approach to linear algebra and system theory.

From the fact that, with F an arbitrary field, the ring of polynomials $F[z]$ is a principal ideal domain, it follows that any submodule M of $F^m[z]$ has a representation of the form $M = DF^m[z]$ for some polynomial matrix D. Moreover, the quotient module $F^m[z]/M$ is finite dimensional as a linear space if and only if D is non-singular.

For a non-singular $m \times m$ polynomial matrix D we define the map π_D by

$$\pi_D f = D\pi_- D^{-1} f, \quad \text{for} \quad f \in F^m[z]$$

Then, π_D is a projection in $F^m[z]$ and $\operatorname{Ker} \pi_D = DF^m[z]$. We introduce now an $F[z]$-module structure in $X_D = \operatorname{Im} \pi_D$ by letting

$$p \cdot f = \pi_D(pf)$$

for all p in $F[z]$ and all f in $F^m[z]$. With the previously defined module structure X_D is isomorphic to $F^m[z]/DF^m[z]$.

In X_D we will focus on a special map S_D which corresponds to the action of the identity polynomial z, i.e.

$$S_D f = \pi_D z f \quad \text{for} \quad f \in X_D$$

Thus, the module structure in X_D is identical to the module structure induced by S_D through $p \cdot f = p(S_D)f$. With this definition, the study of S_D is identical to the study of the module structure of X_D. In particular, the invariant subspaces of S_D are just the submodules of X_D. X_D with this module structure is called a *polynomial model*.

In an analogous way the characterization of finite dimensional S_--invariant subspaces of $z^{-1}F^m[[z^{-1}]]$ can be approached. As in the previous case the parametrization proceeds via non-singular polynomial matrices. We define a projection π^D in $z^{-1}F^m[[z^{-1}]]$ by

$$\pi^D h = \pi_- D^{-1} \pi_+ D h \quad \text{for} \quad h \in z^{-1}F^m[[z^{-1}]]$$

and let $X^D = \operatorname{Im} \pi^D$. Then X^D is a submodule of $z^{-1}F^m[[z^{-1}]]$ with the module structure induced by

$$S^D h = S_- h \quad h \in X^D$$

X^D, with this module structure, is called a *rational model*. Actually, it is the rational models that provide the best link between the finite and infinite-dimensional theories, see Fuhrmann (1991). The emphasis on this link is also one of the main tools in the present paper.

The two models X_D and X^D associated with the polynomial D are isomorphic, the isomorphism is given by the map ρ_D: $X^D \to X_D$ defined by $\rho_D h = Dh$ for $h \in X^D$, i.e. we have $\rho_D S^D = S_D \rho_D$.

As in the case of submodules of $F^m[z]$ the key fact is that a subspace M of $z^{-1}F^m[[z^{-1}]]$ is finite dimensional and S_--invariant if and only if $M = X^D$ for some non-singular polynomial matrix D.

The polynomial and rational models provide extremely useful tools for understanding realization theory. As usual, given a proper rational matrix G we will say a system (A, B, C, D) is a *realization of* G if

$$G = D + C(zI - A)^{-1}B$$

We will use the notation $G = [A, B, C, D]$. We will be interested in realizations associated with rational functions having the following representations

$$G = VT^{-1}U + W \tag{1}$$

with T, U, V and W polynomial matrices. Following Rosenbrock (1970) we associate with such a representation the *polynomial system matrix* P

$$P = \begin{bmatrix} T & U \\ -V & W \end{bmatrix} \tag{2}$$

Our approach to the analysis of these systems is to associate with each representation of the form (1), a state space realization in the following way.

We choose X_T as the state space and define the triple (A, B, C), with $A: X_T \to X_T$, $B: \mathbf{R}^m \to X_T$, and $C: X_T \to \mathbf{R}^p$ by, with $\xi \in \mathbf{R}^m$ and $f \in X_T$,

$$\left.\begin{aligned} A &= S_T \\ B\xi &= \pi_T U \xi \\ Cf &= (VT^{-1}f)_{-1} \\ D &= G(\infty) \end{aligned}\right\} \tag{3}$$

We can choose also X^T as the state space and define the triple (A, B, C), with $A: X^T \to X^T$, $B: \mathbf{R}^m \to X^T$, and $C: X^T \to \mathbf{R}^p$ by, with $\xi \in \mathbf{R}^m$ and $f \in X^T$,

$$\left.\begin{aligned} A &= S^T \\ B\xi &= \pi_- T^{-1} U \xi \\ Cf &= (Vf)_{-1} \\ D &= G(\infty) \end{aligned}\right\} \tag{4}$$

We will refer to both realizations as the *associated shift realizations* to the polynomial matrix P, or just the *shift realizations*.

Theorem 2.1: *The systems given by (3) and (4) are realizations of $G = VT^{-1}U + W$. These realizations are reachable if and only if T and U are left coprime and observable if and only if T and V are right coprime.*

The following result, as well as its dual (due to Hautus and Heymann 1978) is extremely useful.

Theorem 2.2: *Let (A, C) be an observable pair, $G(z) = C(zI - A)^{-1}$ be the corresponding state to output transfer function and let*

$$G = T^{-1}U$$

be a left coprime matrix fraction representation. Then, given any polynomial matrix N, the rational function $T^{-1}N$ is strictly proper if and only if there exists a constant matrix K for which $N(z) = U(z)K$. This is equivalent to the columns of U being a basis for X_T.

Theorem 2.3: *Let $G = ND^{-1}$ be a coprime factorization and let (A, B, C) be a minimal realization of G. Let $G' = MD^{-1}$. Then G' has a realization (A, B, C_0) for some C_0.*

Let $G_1 = [A_1, B_1, C_1, D_1]$ and $G_2 = [A_2, B_2, C_2, D_2]$ be two transfer functions realized in the state spaces X_1 and X_2 respectively. If the number of inputs of the second system equals the number of outputs of the first we can feed those outputs to the second system. This gives rise to the *series coupling* and the corresponding transfer function is

$$G_2G_1 = \left[\begin{bmatrix} A_1 & 0 \\ B_2C_1 & A_2 \end{bmatrix}, \begin{bmatrix} B_1 \\ B_2D_1 \end{bmatrix}, (D_2C_1 \;\; C_2),\; D_2D_1\right]$$

We will use also the notation $G_2G_1 = [A_2, B_2, C_2, D_2] \times [A_1, B_1, C_1, D_1]$.

In order to make the paper more accessible we now review a number of results which we will be using frequently in the following developments.

We will call a $p \times m$ matrix $Q \in H_+^\infty$ *inner* if $Q(it)^*Q(it) = I$ a.e. In particular this implies that Q is tall, i.e. that $p \leqslant m$. An inner function will be called a *full inner function* if it is square, i.e. $Q(it)$ is unitary a.e.

The importance of inner functions is the fact that they parametrize invariant subspaces of H^2_+. This is stated next.

Theorem 2.4 (Beurling): *Let $\mathcal{A} \subseteq H^2_+$ be a closed subspace. Then the following are equivalent.*

(1) *$\mathcal{A}$ is a shift invariant subspace, i.e.*

$$\phi\mathcal{A} \subseteq \mathcal{A}$$

for all $\phi \in H^\infty_+$.

(2) *$\mathcal{A}$ can be written as*

$$\mathcal{A} = QH^2_+$$

for some inner function $Q \in H^\infty_+$. The function Q is unique up to a right constant unitary factor U_0.

An invariant subspace $M \subset H^2_+$ will be called a full invariant subspace if it has a Beurling-type representation with Q full inner.

Note that the orthogonal complement of a full shift invariant subspace is characterized as follows,

$$\{QH^2_+\}^\perp = \{Qg;\ \text{with } g \in H^2_- \text{ such that } Qg \in H^2_+\}$$

By Kronecker's theorem we have that the dimension of $\{QH^2_+\}^\perp$ is equal to the McMillan degree of Q.

The projection onto a shift invariant subspace is given by

$$P_{QH^2_+} = QP_+Q^*$$

where P_+ is the orthogonal projection of L^2 onto H^2_+. The projection in H^2_+ onto the orthogonal complement of a full shift invariant subspace is given by

$$P_{\{QH^2_+\}^\perp} = QP_-Q^*$$

and in L^2 it is given by

$$P_{\{QH^2_+\}^\perp} = P_+ - QP_+Q^*$$

One of the central objects of this paper are Hankel operators. If $G \in L^\infty$ we denote by H_G: $H^2_+ \to H^2_-$ the *Hankel operator* with symbol G, i.e.

$$f \mapsto P_-Gf$$

By $\hat{H}_G$ we denote the *involuted Hankel operator*, i.e. the operator $\hat{H}_G$: $H^2_- \to H^2_+$,

$$g \mapsto P_+Gg$$

Schmidt pairs have been a cornerstone of the AAK theory and will also play a prominent role in the sequel. They are formally introduced next.

Let A: $H_1 \to H_2$ be a bounded operator between two Hilbert spaces. $\mu > 0$ will be called a *singular value* and $f \in H_1$ a *singular vector*, if

$$A^*Af = \mu^2 f$$

A pair of vectors $\{f, g\}$, with $f \in H_1$ and $g \in H_2$, will be called a *μ-Schmidt pair* if they satisfy the pair of equations

$$Af = \mu g$$
$$A^*g = \mu f$$

Clearly, if f is a μ-singular vector then, with $g = 1/\mu Af$ the pair $\{f, g\}$ is a μ-Schmidt pair.

3. Normalized coprime factors

At the basis of this paper lies a class of rational factorizations of transfer functions, the so-called normalized coprime factorizations. They were apparently first studied by Vidyasagar in the context of defining a topology describing robustness of a control system (see e.g. Vidyasagar 1985). They have also attracted a considerable amount of attention in robust control (see e.g. McFarlane and Glover 1990), but have also been used in other contexts such as the parametrization problems of linear systems (Ober and McFarlane 1989). In this section we discuss the existence of normalized coprime factorizations and investigate Hankel operators whose symbols are normalized coprime factors.

Definition 3.1: Let G be a strictly proper rational transfer function.

(1) A representation $G = NM^{-1}$ with N, M stable proper rational transfer functions such that M^{-1} is proper and N, M are right coprime, i.e. there exist $\bar{U}$, $\bar{V} \in H_+^\infty$, such that $\bar{V}M - \bar{U}N = I$, is called a *normalized right coprime factorization* (NRCF) of G if it additionally satisfies the relation

$$N^*N + M^*M = I \tag{5}$$

(2) A representation $G = \bar{M}^{-1}\bar{N}$ with $\bar{M}$, $\bar{N}$ stable proper rational transfer functions such that $\bar{M}^{-1}$ is proper and $\bar{N}$, $\bar{M}$ are left coprime, i.e. there exist U, $V \in H_+^\infty$ such that $\bar{M}V - \bar{N}U = I$, is called a *normalized left coprime factorization* (NLCF) of G if it additionally satisfies the relation

$$\overline{NN^*} + \overline{MM^*} = I \tag{6}$$

□

We now quote the existence and uniqueness result for normalized coprime factors (see e.g. Vidyasagar 1985). We indicate how this result can be derived using polynomial methods.

Lemma 3.1: *Let G be a strictly proper rational transfer function. There exists a unique NRCF $G = NM^{-1}$, such that $M(\infty) = I$ and there exists a unique NLCF $G = \bar{M}^{-1}\bar{N}$ such that $\bar{M}(\infty) = I$. All NRCFs of G are given by NU, MU with U a constant unitary matrix. All NLCFs of G are given by $\overline{UN}$, $\overline{UM}$ with $\bar{U}$ a constant unitary matrix.*

Proof: Let $G = ED^{-1}$ be a right coprime polynomial factorization. Put $N = ET^{-1}$ and $M = DT^{-1}$ for some polynomial matrix T. Then (5) reduces to

$$(T^*)^{-1}E^*ET^{-1} + (T^*)^{-1}D^*DT^{-1} = I \tag{7}$$

or

$$E^*E + D^*D = T^*T \tag{8}$$

Thus, we can find a stable T by polynomial spectral factorization, see Coppel (1972). The remaining statements follow from the uniqueness properties of this polynomial spectral factorization. The result for left coprime factorizations follows analogously. □

Corollary 3.1: *Let G be a strictly proper rational transfer function. If $G = ED^{-1}$ is a polynomial right coprime factorization of G, and T a stable spectral factor of $E^*E + D^*D$, i.e.*

$$E^*E + D^*D = T^*T$$

then D and T are right coprime as polynomial matrices.

Proof: Assume D and T are not right coprime. Then there exist a complex number α and a non-zero vector ξ such that $D(\alpha)\xi = T(\alpha)\xi = 0$. The spectral factorization equation implies therefore that $E(\alpha)^*E(\alpha)\xi = 0$. Taking the inner product with ξ we get $\|E(\alpha)\xi\|^2 = 0$, or $E(\alpha)\xi = 0$. But $D(\alpha)\xi = 0$, $E(\alpha)\xi = 0$ together contradict the assumed right coprimeness of D and E. □

By the normalized right (left) coprime factorization we will mean the factorization for which $M(\infty) = I(\bar{M}(\infty) = I)$.

It follows from the construction of the normalized coprime factorization that the McMillan degrees of the transfer functions $[\bar{N}\ \ \bar{M}]$ and $\begin{bmatrix} N \\ M \end{bmatrix}$ are equal to the McMillan degree of G.

In the following lemma we point out how the normalized coprime factorization is related to a normalized factorization which is not necessarily coprime.

Lemma 3.2:

(1) *If $\bar{K}_2^{-1}\bar{K}_1 = \hat{\bar{K}}_2^{-1}\hat{\bar{K}}_1$ with $\bar{K}_1$, $\bar{K}_2 \in H_+^\infty$ a NLCF of $\bar{K}_2^{-1}\bar{K}_1$ and $\hat{\bar{K}}_1$, $\hat{\bar{K}}_2 \in H_+^\infty$ such that*

$$\hat{\bar{K}}_1\hat{\bar{K}}_1^* + \hat{\bar{K}}_2\hat{\bar{K}}_2^* = I$$

then $[\hat{\bar{K}}_1\ \ \hat{\bar{K}}_2] = \bar{Q}[\bar{K}_1\ \ \bar{K}_2]$ for some inner function $\bar{Q} \in H^\infty$.

(2) *If $K_1K_2^{-1} = \hat{K}_1\hat{K}_2^{-1}$ with K_1, $K_2 \in H_+^\infty$ a NRCF of $K_1K_2^{-1}$ and $\hat{K}_1$, $\hat{K}_2 \in H_+^\infty$ such that*

$$\hat{K}_1^*\hat{K}_1 + \hat{K}_2^*\hat{K}_2 = I$$

then

$$\begin{bmatrix} \hat{K}_1 \\ \hat{K}_2 \end{bmatrix} = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} Q$$

for some inner function $Q \in H^\infty$.

Proof:

(1) Since $\bar{K}_1$, $\bar{K}_2$ come from a NLCF of $\bar{K}_2^{-1}\bar{K}_1$, there exist X, $Y \in H_+^\infty$ such that

$$\bar{K}_1X + \bar{K}_2Y = I$$

This implies $\bar{K}_2^{-1}\bar{K}_1X + Y = \bar{K}_2^{-1}$, and hence also

$$\hat{\bar{K}}_2\bar{K}_2^{-1}\bar{K}_1X + \hat{\bar{K}}_2Y = \hat{\bar{K}}_2\bar{K}_2^{-1} \tag{9}$$

Now, the equality $\bar{K}_2^{-1}\bar{K}_1 = \hat{\hat{K}}_2^{-1}\hat{\hat{K}}_1$ implies $\hat{\hat{K}}_2\bar{K}_2^{-1}\bar{K}_1 = \hat{\hat{K}}_1$. So, from (9) we get

$$\hat{\hat{K}}_2\bar{K}_2^{-1} = \hat{\hat{K}}_1 X + \hat{\hat{K}}_2 Y = \bar{Q} \in H_+^\infty$$

or

$$\hat{\hat{K}}_2 = \overline{QK_2}$$

Now $\bar{K}_2^{-1}\bar{K}_1 = \hat{\hat{K}}_2^{-1}\hat{\hat{K}}_1 = \bar{K}_2^{-1}\bar{Q}^{-1}\hat{\hat{K}}_1$, and so also

$$\hat{\hat{K}}_1 = \overline{QK_1}$$

Finally

$$I = \hat{\hat{K}}_1\hat{\hat{K}}_1^* + \hat{\hat{K}}_2\hat{\hat{K}}_2^* = \bar{Q}(\bar{K}_1\bar{K}_1^* + \bar{K}_2\bar{K}_2^*)\bar{Q}^* = \overline{QQ^*}$$

Thus, necessarily $\bar{Q}$ is a full inner function, and hence also $\bar{Q}^*\bar{Q} = I$.

(2) The proof follows from part (1) by duality considerations. □

We now come to analyse Hankel operators whose symbols are normalized coprime factors. We first study the role the coprimeness condition plays in this context.

Lemma 3.3: *Assume* $[-\bar{N} \;\; \bar{M}]\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} = I$ *and* Ker $H\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} = SH_+^2$.

Then

(1) $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$ *has a factorization*

$$\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} = \begin{bmatrix} K_1^* \\ K_2^* \end{bmatrix}\bar{Q}^* \tag{10}$$

with $\bar{Q}$ inner in H_+^∞ and K_1 and K_2 are left coprime.

(2) *If $\bar{M}$ and $\bar{N}$ are left coprime then $\bar{Q}$ is a constant matrix and*

$$\left\| H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} \right\| < 1$$

(3) $\{\bar{Q}H_+^2\}^\perp$ *is the subspace of all f such that*

$$\left\| H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} f \right\| = \|f\|$$

(4) *If $\bar{Q}$ is rational, then the McMillan degree of $\bar{Q}$ is equal to the multiplicity of the singular value of $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ of magnitude* 1.

(5)
$$S = \overline{QQ_0} \tag{11}$$

Proof:

(1) This is an immediate consequence of Lemma 3.2.

(2) That $\bar{Q}$ is a constant unitary function follows from Lemma 3.2. To show

that the norm of the Hankel operator is less than one we use the fact that for $f \in H^2_+$,

$$P_+\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f + P_-\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f = \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f$$

which implies that

$$(H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}})^* H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} + (T_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}})^* T_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} = I$$

where the Toeplitz operator $T_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}: H^2_+ \to H^2_+$ is defined as,

$$T_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} f := P_+\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f$$

for $f \in H^2_+$. Hence, the norm of the Hankel operator is less than one if and only if the kernel of the Toeplitz operator only contains the zero function. Assume that $f \in \operatorname{Ker}(T_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}] \end{bmatrix}})$, then for all $\begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \in H^2_+$, we have,

$$0 = \left\langle \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}, (T_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}) f \right\rangle = \left\langle \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}, \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f \right\rangle$$

In particular for $X, Y \in H^\infty_-$ such that $\bar{M}Y - \bar{N}X = I$ and $h \in H^2_+$ we have that

$$0 = \left\langle \begin{bmatrix} X \\ Y \end{bmatrix} h, \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f \right\rangle = \langle (\bar{M}Y - \bar{N}X)h, \; f \rangle = \langle h, \; f \rangle$$

which implies that $f = 0$, i.e. the kernel of the Toeplitz operator only contains the zero function. This shows the claim.

(3) If $f \in \{\bar{Q}H^2\}^\perp$ then $f = \bar{Q}h^*$, for some $h \in H^2_+$ and hence

$$\|H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} f\| = \left\| P_- \begin{bmatrix} -\bar{K}_1^* \\ \bar{K}_2^* \end{bmatrix} h^* \right\| = \left\| \begin{bmatrix} -\bar{K}_1^* \\ \bar{K}_2^* \end{bmatrix} h^* \right\| = \|h^*\| = \|f\|$$

If $f \in \{\bar{Q}H^2_+\}$, i.e. $f = \bar{Q}h$ for some $h \in H^2_+$, then,

$$\|H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} f\| = \|H_{\begin{bmatrix} -\bar{K}_1^* \\ \bar{K}_2^* \end{bmatrix}} h\| < 1$$

by the coprimeness of K_1 and K_2. Combining the two cases we obtain the result.

(4) This is an immediate consequence of Part (3) and the fact that the dimension of $\{\bar{Q}H^2_+\}^\perp$ equals the McMillan degree of $\bar{Q}$.

(5) Since SH^2_+ is the kernel of $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ it follows that $SH^{2\perp}_+\{\bar{Q}H^2_+\}^\perp$ therefore we have that $SH^2_+ \subseteq \bar{Q}H^2_+$. But this shows that there exists a square inner function $\bar{Q}_0$ such that $S = \overline{QQ_0}$. □

Part (2) of the lemma appears in McFarlane and Glover (1990) with a different proof. A state space proof of Part (4) is given in Yeh and Wei (1990).

We now state without proof the analogous result for the Hankel operator $H_{[M^*\ N^*]}$.

Lemma 3.4: *Assume* $[M^*\ N^*]\begin{bmatrix} M \\ N \end{bmatrix} = I$ *and* $\operatorname{Im} H_{[M^*\ N^*]} = \{S^*H^2_-\}^\perp$, *for some square inner function S. Then*

(1) $[M^*\ N^*]$ *has a factorization*

$$[M^*\ N^*] = Q^*[K_1^*\ K_2^*]$$

with Q inner in H^∞_+ *and* K_1 *and* K_2 *are right coprime.*

(2) *If M and N are right coprime then Q is a constant matrix and*

$$\|H_{[M^*\ N^*]}\| < 1$$

(3) *If Q is rational then the McMillan degree of Q is equal to the multiplicity of the singular value of* $H_{[M^*\ N^*]}$ *of magnitude* 1.

We continue by computing the kernel of the Hankel operator $H_{[M^*\ N^*]}$.

Theorem 3.1: *Let* $G = NM^{-1}$ *be the NRCF of G and* $G = \bar{M}^{-1}\bar{N}$ *its NLCF. Then*

$$\operatorname{Ker} H_{[M^*\ N^*]} = \Omega_K H^2_+ \tag{12}$$

with an inner function of the form

$$\Omega_K = \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} \tag{13}$$

Here

$$\begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} \tag{14}$$

and Q_{Ω_K} *is the minimal inner function for which the last matrix is in* H^∞_+ *and* $Q_{\Omega_K}(\infty) = I$.

Proof: Define the two subspaces

$$W_1 = \left\{ \begin{bmatrix} M \\ N \end{bmatrix} f | f \in H^2_+ \right\} \tag{15}$$

and

$$W_2 = \left\{ \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} : f_1, f_2 \in H^2_+,\ M^*f_1 + N^*f_2 = 0 \right\} \tag{16}$$

Clearly we have W_1, $W_2 \subseteq \operatorname{Ker} H_{(M^*\ N^*)}$. Both W_1 and W_2 are invariant subspaces, so they have, by Beurling's theorem, representations in terms of inner functions. Obviously $W_1 = \begin{bmatrix} M \\ N \end{bmatrix} H_+^2$ with $\begin{bmatrix} M \\ N \end{bmatrix}$ inner by construction. Similarly, W_2 has a representation $W_2 = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} H_+^2$ with $\begin{bmatrix} J_1 \\ J_2 \end{bmatrix}$ inner.

Now consider $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$. There exists a minimal inner function Q_{Ω_K} such that $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} \in H_+^\infty$. Obviously

$$(M^* \;\; N^*)\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} = 0 \tag{17}$$

Therefore $\begin{bmatrix} -\bar{N} \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} H_+^2 \subset \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} H_+^2$. Since $(-\bar{N}^* Q_{\Omega_K}, \bar{M}^* Q_{\Omega_K})$ is the NLCF of $-G^*$ there exists an inner function Q_1 such that $\begin{bmatrix} -\bar{N}^* \\ \bar{M} \end{bmatrix} Q_{\Omega_K} = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} Q_1$. Now $[J_1^* \;\; J_2^*]\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} = Q_2^* Q_{\Omega_K} = Q_1$. This implies $Q_{\Omega_K} = Q_2 Q_1$ and hence $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_2 \in H_+^\infty$. By the minimality of Q_{Ω_K}, Q_1 thus necessarily reduces to a constant unitary matrix. So $\begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K}$ and $W_2 = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} H_+^2$. The subspaces W_1 and W_2 are orthogonal. Indeed, for arbitrary $f, g \in H_+^2$

$$\left(\begin{bmatrix} J_1 \\ J_2 \end{bmatrix} f, \begin{bmatrix} M \\ N \end{bmatrix} g\right) = ((M^* J_1 + N^* J_2) f, g) = ((-M^* \bar{N}^* + N^* \bar{M}^*) Q_{\Omega_K} f, g) = 0$$

as $\bar{M} N = \bar{N} M$.

Now, by orthogonality

$$W_1 \oplus W_2 = \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} H_+^2 = \Omega_K H_+^2$$

Thus, $\operatorname{Ker} H_{(M^* \; N^*)} \supset \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} H_+^2$. We will show that equality holds.

So, let $\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \in \operatorname{Ker} H_{(M^* \; N^*)}$. If $M^* f_1 + N^* f_2 = 0$ then $\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \in W_2 = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} H_+^2 \subset \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} H_+^2$. So, finally, consider $W = \operatorname{Ker} H_{(M^* \; N^*)} \ominus W_2$. This is also an invariant subspace, and invoking Beurling's theorem once more, it is of the form $W = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} H_+^2$ for some inner function. Now, obviously

$$W_1 = \begin{bmatrix} M \\ N \end{bmatrix} H_+^2 \subset W$$

This implies

$$\begin{bmatrix} M \\ N \end{bmatrix} = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} P$$

for some inner function P. If P were a non-trivial inner function we would get, by Lemma 3.3, that $\|H_{\begin{bmatrix} M \\ N \end{bmatrix}}\| = 1$ which is impossible for a NRCF. □

Note that J_1, J_2 as constructed in the theorem is a NRCF of the function $-G^*$.

The following lemma shows that the Hankel operator with symbol $[M^* \ N^*]$ acts as a multiplication operator on the orthogonal complement of its kernel.

Lemma 3.5: *For the Hankel operator $H_{[M^* \ N^*]}$ the following holds.*

$$H_{(M^* \ N^*)}|_{\{\mathrm{Ker}\, H_{(M^* \ N^*)}\}^\perp} = [M^* \ \ N^*]|_{\{\mathrm{Ker}\, H_{(M^* \ N^*)}\}^\perp} \tag{18}$$

Proof: As

$$\mathrm{Ker}\, H_{(M^* \ N^*)} = \Omega_K H_+^2 \supseteq \begin{bmatrix} M \\ N \end{bmatrix} H_+^2$$

it follows that, if

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \in \{\mathrm{Ker}\, H_{(M^* \ N^*)}\}^\perp$$

then, for all $g \in H_+^2$, we have $\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \perp \begin{bmatrix} M \\ N \end{bmatrix} g$ or $M^* f_1 + N^* f_2 \in H_-^2$, and hence

$$P_-(M^* f_1 + N^* f_2) = M^* f_1 + N^* f_2$$ □

A result analogous to the previous theorem can also be obtained for the kernel of the operator $\hat{H}_{[-\bar{N} \ \bar{M}]}$ as stated next. The proof is omitted.

Theorem 3.2: *Let $G = \bar{M}^{-1}\bar{N}$ be the NLCF of G. Then*

$$\mathrm{Ker}\, \hat{H}_{[-\bar{N} \ \bar{M}]} = \Omega_I^* H_-^2 \tag{19}$$

where

$$\Omega_I^* = \begin{bmatrix} -\bar{N}^* & \bar{K}_1 \\ \bar{M}^* & \bar{K}_2 \end{bmatrix} \tag{20}$$

is such that $\Omega_I \in H_+^\infty$ is inner and

$$\begin{bmatrix} \bar{K}_1 \\ \bar{K}_2 \end{bmatrix} = \begin{bmatrix} M \\ N \end{bmatrix} Q_{\Omega_I}^* \tag{21}$$

Here $Q_{\Omega_I}^*$ is the minimal inner function for which the last matrix is in H_-^∞ and $Q_{\Omega_I}(\infty) = I$.

Corollary 3.2: $\mathrm{Im}\, H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} = \{\Omega_I^* H_-^2\}^\perp$

Proof: The operator $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ is the adjoint operator of $\hat{H}_{[-\bar{N} \ \bar{M}]}$ and hence the claim follows. □

4. The LQG controller

Let $G = NM^{-1} = \bar{M}^{-1}\bar{N}$ be the normalized right and left coprime factorizations of the transfer function G.

Definition 4.1: A doubly coprime factorization consists of H_+^∞ functions such that

$$\begin{bmatrix} \bar{V} & -\bar{U} \\ -\bar{N} & \bar{M} \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{22}$$

□

It is well known that doubly coprime factorizations of G exist (see e.g. Vidyasagar 1985). Also, given any doubly coprime factorization as above then

$$K = UV^{-1} = \bar{V}^{-1}\bar{U}$$

is a stabilizing controller for G. This is the case since $K = UV^{-1}$ (internally) stabilizes $G = NM^{-1}$ if and only if $\bar{M}V - \bar{N}U$ is invertible in H_+^∞. The *Kucera–Youla parametrization*, see Kucera (1979), Youla *et al.* (1976), states that all stabilizing controllers of G are parametrized by the solutions to the Bezout equation $\bar{V}M - \bar{U}N = I$, i.e. K stabilizes G if and only if there exists a coprime factorization $K = \bar{V}^{-1}\bar{U}$ of the controller K such that $\bar{V}M - \bar{U}N = I$. The set of all stabilizing controllers can therefore be written as

$$K = (U + MQ)(V + NQ)^{-1} = (\bar{V} + \overline{QN})^{-1}(\bar{U} + Q\bar{M}) \tag{23}$$

for $Q \in H_+^\infty$, such that $V + NQ \neq 0$ and for $\bar{Q} \in H_+^\infty$ such that $\bar{V} + \overline{QN} \neq 0$.

We will now introduce a Hankel operator which is closely related to the LQG controller which is being studied in this section. The following lemma summarizes some basic results concerning this controller.

Lemma 4.1: *Consider the Bezout equation $\bar{M}V - \bar{N}U = I$.*

(1) *There exists a unique solution $[U_L, V_L]$ of the Bezout equation such that*

$$R^* = M^*U_L + N^*V_L \in H_-^\infty$$

and R^* is strictly proper.

(2) *Let $U, V \in H_+^\infty$ be an arbitrary solution to the Bezout equation $\bar{M}V - \bar{N}U = I$. Then R^* is the strictly proper antistable part of $M^*U + N^*V$.*

(3) *The Hankel operator*

$$H_{M^*U+N^*V}$$

is independent of the solution $[U, V]$ of the Bezout equation.

Proof:

(1) Let $[\tilde{U}, \tilde{V}]$ be a solution to the Bezout equation $\bar{M}V - \bar{N}U = I$. Then all solutions $[U, V]$ to this equation can be uniquely parametrized by

$$[U, V] = [\tilde{U} + M\tilde{Q}, \tilde{V} + N\tilde{Q}], \quad \tilde{Q} \in H_+^\infty$$

Since $M^*M + N^*N = I$, we therefore have that,

$$M^*U + N^*V = M^*\tilde{U} + \tilde{N}^*\tilde{V} + \tilde{Q}$$

Hence, there exists a unique Q_0 such that $M^*\tilde{U} + N^*\tilde{V} + Q_0 \in H_-^\infty$ and is strictly proper. Since the parametrization of the solutions $[U, V]$ is unique, this shows that

$$U_L = \tilde{U} + MQ_0, \quad V_L = \tilde{V} + NQ_0$$

is a solution such that $R^* = M^*U_L + N^*V_L$ has the required properties.

(2) and (3) Note that an arbitrary solution $[U, V]$ to the Bezout equation can be written as

$$[U, V] = [U_L + MQ, V_L + NQ]$$

Therefore, the strictly proper antistable part of

$$M^*U + N^*V = M^*U_L + N^*V_L + Q = R^* + Q$$

is given by R^*. Also

$$H_{M^*U+N^*V} = H_{M^*U_L+N^*V_L+Q} = H_{M^*U_L+N^*V_L}$$ □

The stabilizing controller $K = U_L V_L^{-1}$ is called the *LQG controller* of the plant G. It follows from the state space realization for K which is derived in § 7 that this controller is, in fact, the controller that solves the certain LQG problem that is discussed by Jonckheere and Silverman (1983).

In the following Proposition we will establish a useful consequence of the Kucera–Youla parametrization which we will use heavily in the following.

Proposition 4.1: *With the notation of the previous Lemma we have*

(1) $$\begin{bmatrix} U_L \\ V_L \end{bmatrix} = \begin{bmatrix} M \\ N \end{bmatrix} R^* + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} \tag{24}$$

(2) *The pair of H_+^∞ functions $[U, V]$ solves the Bezout equation $\bar{M}V - \bar{N}U = I$ if and only if there exists a $Q \in H_+^\infty$ such that*

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U_L \\ V_L \end{bmatrix} + \begin{bmatrix} M \\ N \end{bmatrix} Q = \begin{bmatrix} M \\ N \end{bmatrix}(R^* + Q) + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

Proof:

(1) Set

$$U := \begin{bmatrix} M^* & N^* \\ \bar{N} & -\bar{M} \end{bmatrix}$$

It is easily verified that $UU^* = I$. But this implies that $U^*U = I$ and therefore

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} = \begin{bmatrix} M & \bar{N}^* \\ N & -\bar{M}^* \end{bmatrix} \begin{bmatrix} M^* & N^* \\ \bar{N} & -\bar{M} \end{bmatrix} \begin{bmatrix} U_L \\ V_L \end{bmatrix}$$
$$= \begin{bmatrix} M & \bar{N}^* \\ N & -\bar{M}^* \end{bmatrix} \begin{bmatrix} R^* \\ -I \end{bmatrix}$$
$$= \begin{bmatrix} M \\ N \end{bmatrix} R^* + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

(2) This follows from the Kucera–Youla parametrization and Part (1). □

Corollary 4.1: *We have*

$$\inf_{\{\bar{M}V-\bar{N}U=I\}} \left\| \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} - \begin{bmatrix} U \\ V \end{bmatrix} \right\| = \|H_{R^*}\| \tag{25}$$

Proof: Using the Kucera–Youla parametrization, any solution of the Bezout equation $\bar{M}V - NU = I$ has the representation

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} M \\ N \end{bmatrix}(R^* + Q) + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

for some $Q \in H_+^\infty$. Thus

$$\inf_{\{\bar{M}V-\bar{N}U=I\}} \left\| \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} - \begin{bmatrix} U \\ V \end{bmatrix} \right\|$$

$$= \inf_{Q\in H_+^\infty} \left\| \begin{bmatrix} M \\ N \end{bmatrix}(R^* + Q) \right\| = \inf_{Q\in H_+^\infty} \|R^* + Q\| = \|H_{R^*}\| \tag{26}$$

□

We have defined the LQG controller and the function R^* by considering solutions to the Bezout equation $\bar{M}V - \bar{N}U = I$. We can go through similar derivations by considering the 'dual' Bezout equation $\bar{V}M - \bar{U}N = I$. The results are considered in the following lemma.

Lemma 4.2: *Consider the Bezout equation $\bar{V}M - \bar{U}N = I$.*

(1) *There exists a unique solution $[\bar{U}_L, \bar{V}_L]$ such that*

$$\bar{R}^* = \bar{U}_L\bar{M}^* + \bar{V}_L\bar{N}^* \in H_-^\infty$$

(2) $$[\bar{U}_L \;\; \bar{V}_L] = \bar{R}^*[\bar{M} \;\; \bar{N}] + [-N^* \;\; M^*]$$

It is important for our later developments that R^* and $\bar{R}^*$ are, in fact, identical as proved by the following proposition. See also Georgiou and Smith (1990 a) for part (1) of the proposition.

Proposition 4.2: *We have*

(1) $R^* = \bar{R}^*$

(2) $K = U_L V_L^{-1} = \bar{V}_L^{-1}\bar{U}_L$

(3) *The coprime factors U_L, V_L and $\bar{U}_L$, $\bar{V}_L$ satisfy the doubly coprime factorization*

$$\begin{bmatrix} \bar{V}_L & -\bar{U}_L \\ -\bar{N} & \bar{M} \end{bmatrix} \begin{bmatrix} M & U_L \\ N & V_L \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

Proof: We know that

$$\begin{bmatrix} M^* & N^* \\ \bar{N} & -\bar{M} \end{bmatrix} \begin{bmatrix} M & \bar{N}^* \\ N & -\bar{M} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

and hence that

$$\begin{bmatrix} M & \bar{N}^* \\ N & -\bar{M}^* \end{bmatrix} \begin{bmatrix} M^* & \bar{N}^* \\ \bar{N} & -\bar{M} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

Therefore

$$\begin{bmatrix} I & \bar{V}_L U_L - \bar{U}_L V_L \\ 0 & I \end{bmatrix} = \begin{bmatrix} \bar{V}_L & -\bar{U}_L \\ -\bar{N} & M \end{bmatrix} \begin{bmatrix} M & U_L \\ N & V_L \end{bmatrix}$$

$$= \begin{bmatrix} \bar{V}_L & -\bar{U}_L \\ -\bar{N} & \bar{M} \end{bmatrix} \begin{bmatrix} M & \bar{N}^* \\ N & -\bar{M}^* \end{bmatrix} \begin{bmatrix} M^* & N^* \\ \bar{N} & -\bar{M} \end{bmatrix} \begin{bmatrix} M & \bar{U}_L \\ N & \bar{V}_L \end{bmatrix}$$

$$= \begin{bmatrix} I & \bar{R}^* \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & R^* \\ 0 & -I \end{bmatrix}$$

$$= \begin{bmatrix} I & R^* - \bar{R}^* \\ 0 & I \end{bmatrix}$$

which implies that

$$R^* - \bar{R}^* = \bar{V}_L U_L - \bar{U}_L V_L$$

Since the left-hand side is in H_-^∞ and strictly proper and the right-hand side is in H_+^∞, we have that both are zero, i.e.

$$0 = R^* - \bar{R}^* = \bar{V}_L U_L - \bar{U}_L V_L$$

from which we obtain Parts (1), (2) and (3). □

We will interpret now Equation (24). To this end we prove the following theorem. An algebraic analogue can be found in Fuhrmann (1984).

Theorem 4.1:

(1) *Let $G \in H_+^\infty$ and $G_1 \in H_+^\infty$ be rational transfer functions. Then*

$$\operatorname{Im} \hat{H}_{G_1} \subseteq \operatorname{Im} \hat{H}_G \tag{27}$$

if and only if

$$G_1 = GK^* + L^* \tag{28}$$

for some $K, L \in H_+^\infty$

(2) *Let $G \in H_-^\infty$ and $G_1 \in H_-^\infty$ be rational transfer functions. Then*

$$\operatorname{Ker} H_{G_1} \supseteq \operatorname{Ker} H_G \tag{29}$$

if and only if

$$G_1 = KG + L \tag{30}$$

for some $K, L \in H_+^\infty$

Proof:

(1) If (28) holds, then for $f \in H_-^2$,

$$\hat{H}_{G_1} f = P_+ G_1 f = P_+(GK^* + L^*)f = H_G(K^* f) + P_- L^* f$$
$$= \hat{H}_G(Kf).$$

and hence $\operatorname{Im} \hat{H}_{G_1} \subseteq \operatorname{Im} \hat{H}_{G_2}$. Conversely, assume the inclusion of (27). Let $G = QH^*$ and $G_1 = Q_1 H_1^*$ be factorizations with Q and Q_1 square inner functions in H_+^∞ and H, $H_1 \in H_+^\infty$, such that (Q, H) and (Q_1, H_1)

are right coprime pairs. Thus, $\operatorname{Im} \hat{H}_G = \{QH_+^2\}^\perp$ and $\operatorname{Im} \hat{H}_{G_1} = \{Q_1 H_+^2\}^\perp$, with Q and Q_1 inner. The inclusion implies the factorization $Q = Q_1 Q_2$, where Q_2 is a square inner function in H^∞. Since Q and H are right coprime, there exist L and K in H_+^∞ such that

$$H_1 Q_2 = LQ + KH$$

Thus

$$\begin{aligned} G_1 &= Q_1 H_1^* = QQ_2^* H_1^* = Q(Q^* L^* + H^* K^*) \\ &= (QH^*)K^* + L^* = GK^* + L^* \end{aligned} \tag{31}$$

(2) If (30) holds, then for $f \in \operatorname{Ker} H_G$, we have

$$\|H_{G_1} f\| = \sup_{\substack{h \in H^2 \\ \|h\|=1}} \langle h, H_{G_1} f\rangle = \sup_{\substack{h \in H^2 \\ \|h\|=1}} \langle K^* h, Gf\rangle = \sup_{\substack{h \in H^2 \\ \|h\|=1}} \langle K^* h, H_G f\rangle$$
$$= 0$$

and therefore $\operatorname{Ker} H_G \subseteq \operatorname{Ker} H_{G_1}$.

Now, assume that $\operatorname{Ker} H_{G_1} \supseteq \operatorname{Ker} H_{G_2}$. Let $G = HQ^*$ and $G_1 = H_1 Q_1^*$ be factorizations such that Q, Q_1 are square inner functions in H_+^∞ and H, $H_1 \in H_+^\infty$ are such that (H, Q) and (H_1, Q_1) are right coprime pairs. Thus, $\operatorname{Ker} H_G = \{QH_+^2\}$ and $\operatorname{Ker} H_{G_1} = \{Q_1 H_+^2\}$. The inclusion implies that $Q = Q_1 Q_2$ for some square inner function $Q_2 \in H_+^\infty$. Since Q and H are right coprime there exist L, $K \in H_+^\infty$ such that

$$H_1 Q_2 = LQ + KH$$

Thus

$$\begin{aligned} G_1 &= H_1 Q_1^* = H_1 Q_2 Q^* = (LQ + KH)Q^* = L + KHQ^* \\ &= L + KG \end{aligned}$$

□

In the following corollary this theorem is exploited to compare the images and kernels of the Hankel operators which are associated with the coprime factors of the transfer function G and its LQG controller.

Corollary 4.2: *Let* $\begin{bmatrix} U_L \\ V_L \end{bmatrix}$ *be the coprime factors associated with LQG controller in Lemma 4.1. Then*

(1) $\operatorname{Ker} H_{[U_L^* \; V_L^*]} \supseteq \operatorname{Ker} H_{[M^* \; N^*]}$

(2) $\operatorname{Im} \hat{H}_{\begin{bmatrix} M \\ N \end{bmatrix}} \supseteq \operatorname{Im} \hat{H}_{\begin{bmatrix} U_L \\ V_L \end{bmatrix}}$

Proof: The proof follows from Proposition 4.1 and Theorem 4.1. □

5. Relations between Hankel operators

In the previous sections, Hankel operators were introduced which corresponded to coprime factors of the transfer function G and the LQG controller. In this section we are going to study how these Hankel operators are related. We begin by proving that Diagram 1 in Fig. 1 commutes.

We need the following lemma.

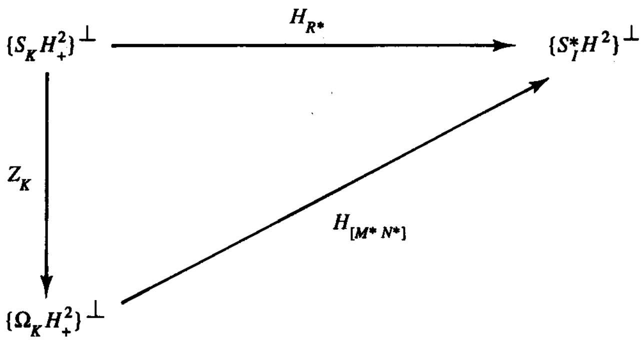

Figure 1. Diagram 1.

Lemma 5.1:

(1) *R^* has a factorization $R^* = \Phi_K S_K^*$, where $S_K \in H_+^\infty$ is a square inner function with $S_K(\infty) = I$ and $\Phi_K \in H_+^\infty$, such that Φ_K and S_K are right coprime. With this we have*

$$\operatorname{Ker} H_{R^*} = S_K H_+^2$$

(2) *R^* has a factorization $R^* = S_I^* \Phi_I$, where $S_I \in H_+^\infty$ is a square inner function with $S_I(\infty) = I$ and $\Phi_I \in H_+^\infty$, such that Φ_I and S_I are left coprime. With this we have*

$$\operatorname{Im} H_{R^*} = \{S_I^* H_-^2\}^\perp$$

(3) *The following identity holds,*

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} S_K = \begin{bmatrix} M \\ N \end{bmatrix} \Phi_K + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} S_K \tag{32}$$

and

$$\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} S_K \in H_+^\infty$$

(4) *S_K is the minimum degree inner function in H_+^∞ such that $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} S_K \in H_+^\infty$* (i.e. $S_K = Q_{\Omega_K}$) *and hence*

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} S_K = \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} \begin{bmatrix} \Phi_K \\ I \end{bmatrix}$$

(5) *The functions $\begin{bmatrix} U_L \\ V_L \end{bmatrix}$ and $\begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix}$ are left coprime. In particular if $\bar{V}_L^{-1} \bar{U}_L$ is a left coprime factorization of $K = U_L V_L^{-1}$ such that $\bar{V}_L M - \bar{U}_L N = I$ we have*

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} (-\bar{N} \;\; \bar{M}) + \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} \begin{bmatrix} \bar{V}_L & \bar{U}_L \\ 0 & 0 \end{bmatrix} = I$$

(6) *S_K and $\begin{bmatrix} \Phi_K \\ \hat{Q} \end{bmatrix}$ are right coprime.*

Proof: Parts (1) and (2) are standard results.

(3) From Proposition 4.1 we know that

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} = \begin{bmatrix} M \\ N \end{bmatrix} R^* + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

This, together with the factorization of R^* in Part (1) implies the claimed identity. Since $\begin{bmatrix} U_L \\ V_L \end{bmatrix} S_K \in H_+^\infty$ and $\begin{bmatrix} M \\ N \end{bmatrix} \Phi_K \in H_+^\infty$ we also have that

$$\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} S_K \in H_+^\infty$$

(4) Since $-G^* = J_1 J_2^{-1}$ is the NRCF of $-G^*$ and $-G^* = (-\bar{N}^* S_K)(\bar{M}^* S_K)^{-1}$ is a factorization with factors in H_+^∞ such that

$$(-\bar{N}^* S_K)^*(-\bar{N}^* S_K) + (\bar{M}^* S_K)^*(\bar{M}^* S_K) = I$$

we have by Lemma 3.2 that

$$\begin{bmatrix} J_1 \\ J_2 \end{bmatrix} Q_0 = \begin{bmatrix} -\bar{N}^* S_K \\ \bar{M}^* S_K \end{bmatrix}$$

for some inner function Q_0. But, therefore, $S_K = Q_{\Omega_K} Q_0$ and hence

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} Q_{\Omega_K} Q_0 = \begin{bmatrix} M \\ N \end{bmatrix} \Phi_K + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} Q_0$$

or equivalently,

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} Q_{\Omega_K} = \begin{bmatrix} M \\ N \end{bmatrix} \Phi_K Q_0^{-1} + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K}$$

Since $\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} Q_{\Omega_K} \in H_+^\infty$ this implies that

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} := \begin{bmatrix} M \\ N \end{bmatrix} \Phi_K Q_0^{-1} \in H_+^\infty$$

Thus

$$\hat{\Phi}_K := [\bar{V}_L \;\; \bar{U}_L] \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} := [\bar{V}_L \;\; \bar{U}_L] \begin{bmatrix} M \\ N \end{bmatrix} \Phi_K Q_0^{-1} = \Phi_K Q_0^{-1}$$

and therefore

$$\Phi_K = \hat{\Phi}_K Q_0$$

By the right coprimeness of $\Phi_K = \hat{\Phi}_K Q_0$ and $S_K = Q_{\Omega_K} Q_0$ this implies that Q_0 is a unitary constant matrix. Since $Q_{\Omega_K}(\infty) = I$ and $S_K(\infty) = I$ we have that $Q_0 = I$ and hence the result.

(5) Note that

$$\begin{bmatrix} \bar{V}_L & -\bar{U}_L \\ -\bar{N} & \bar{M} \end{bmatrix} \begin{bmatrix} M & U_L \\ N & V_L \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

and therefore

$$\begin{bmatrix} M & U_L \\ N & V_L \end{bmatrix} \begin{bmatrix} \bar{V}_L & -\bar{U}_L \\ -\bar{N} & \bar{M} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

The statement in Part (5) can be verified using this identity.

(6) By assumption, S_K and Φ_K are right coprime. Therefore, there exist X, $Y \in H_+^\infty$ such that

$$XS_K + Y\Phi_K = I$$

Hence we have that

$$XS_K + [Y \quad 0] \begin{bmatrix} \Phi_K \\ \hat{Q} \end{bmatrix} = I$$

which shows the claim. □

In § 3 we have seen that $\Omega_K H_+^2$, where $\Omega_K = \begin{bmatrix} M & J_1 \\ N & J_1 \end{bmatrix}$, is the kernel of the Hankel operator $H_{[M^* \ N^*]}$. As the first step to proving that Diagram 1 in Fig. 1 commutes we construct a bijection between the orthogonal complement of the kernels of H_{R^*} and of $H_{[M^* \ N^*]}$, i.e. between $\{S_K H_+^2\}^\perp$ and $\{\Omega_K H_+^2\}^\perp$.

Theorem 5.1: *The map* Z_K: $\{S_K H_+^2\}^\perp \to \{\Omega_K H_+^2\}^\perp$ *defined by*

$$Z_K f = P_{\{\Omega_K H_+^2\}^\perp} \begin{bmatrix} U_L \\ V_L \end{bmatrix} f \tag{33}$$

is invertible. Its inverse is given by Y_K: $\{\Omega_K H_+^2\}^\perp \to \{S_K H_+^2\}^\perp$

$$Y_K \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = P_{\{S_K H_+^2\}^\perp} [-\bar{N} \ \ \bar{M}] \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \tag{34}$$

Proof: From the previous Lemma we know that

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} S_K = \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} \begin{bmatrix} \Phi_K \\ I \end{bmatrix}$$

This and the corresponding coprimeness results in the same Lemma, imply using Theorem 14.8, p. 203, and Theorem 14.11, p. 206, in Fuhrmann (1981) that Z_K is left and right invertible with inverse Y_K. □

Note that the adjoints Z_K^*: $\{\Omega_K H_+^2\}^\perp \to \{S_K H_+^2\}^\perp$ and Y_K^*: $\{S_K H_+^2\}^\perp \to \{\Omega_K H_+^2\}^\perp$ are given by

$$Z_K^* \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = P_{\{S_K H_+^2\}^\perp} P_+ [U_L^* \ \ V_L^*] \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$$

and

$$Y_K^* f = P_{\{\Omega_K H_+^2\}^\perp} P_+ \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f$$

The connection between the Hankel operators H_{R^*}, $H_{[M^* \ N^*]}$ and Z_K is established in the following proposition.

Proposition 5.1: $H_{R^*}|_{\{S_K H_+^2\}^\perp} = H_{[M^* \ N^*]} Z_K$

Proof: We compute

$$H_{[M^* \;\; N^*]}Zf = P_-[M^* \;\; N^*]P_{\{\Omega_K H_+^2\}^\perp}\begin{bmatrix} U_L \\ V_L \end{bmatrix} f$$

$$= P_-[M^* \;\; N^*]\begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} P_- \begin{bmatrix} M^* & N^* \\ J_1^* & J_2^* \end{bmatrix}\begin{bmatrix} U_L \\ V_L \end{bmatrix} f$$

$$= P_-[I \;\; 0]P_-\begin{bmatrix} M^*U_L + N^*V_L \\ J_1^*U_L + J_2^*V_L \end{bmatrix} f = H_{R^*}f \qquad \square$$

As a corollary we can now identify the image of $H_{[M^* \;\; N^*]}$.

Corollary 5.1: Im $H_{[M^* \;\; N^*]} = \{S_I^* H_-^2\}^\perp$.

Proof: The results follows from the proposition and the fact that $\Omega_K H_+^2$ is the kernel of $H_{[M^* \;\; N^*]}$ since,

$$\text{Im}\, H_{[M^* \;\; N^*]} = \text{Im}\, H_{[M^* \;\; N^*]}|_{\{\Omega_K H_+^2\}^\perp} = \text{Im}\, \text{H}_{[M^* \;\; N^*]}Z = \text{Im}\, H_{R^*} = \{S_I^* H_-^2\}^\perp$$

$\square$

Combining the previous results in this section we have proved that Diagram 1 in Fig. 1 indeed commutes. We have also shown that the spaces which appear in the diagram are images, respectively orthogonal complements of kernels, of the Hankel operators. It therefore follows from Diagram 1 that we have established the connection between the operator $H_{[M^* \;\; N^*]}$ and H_{R^*}.

Next we are going to consider Diagram 2 in Fig. 2 which can be seen as being dual to Diagram 1. In Proposition 4.2 we saw that $\bar{R} = R$. Since $(H_{\bar{R}^*})^* = \hat{H}_{\bar{R}} = \hat{H}_R$ we therefore have that $\hat{H}_{\bar{R}}$ maps $\{S_I^* H_-^2\}^\perp$ into $\{S_K H_+^2\}^\perp$.

We now state without proof of the next theorem which shows that Diagram 2 commutes. The proof is analogous to the proof of Theorem 5.1.

Theorem 5.2:

(1) *The map* Z_I: $\{S_I^* H_-^2\}^\perp \to \{\Omega_I^* H_-^2\}^\perp$ *defined by*

$$Z_I f = P_{\{\Omega_I^* H_-^2\}^\perp}\begin{bmatrix} -\bar{V}_L^* \\ \bar{U}_L^* \end{bmatrix} f$$

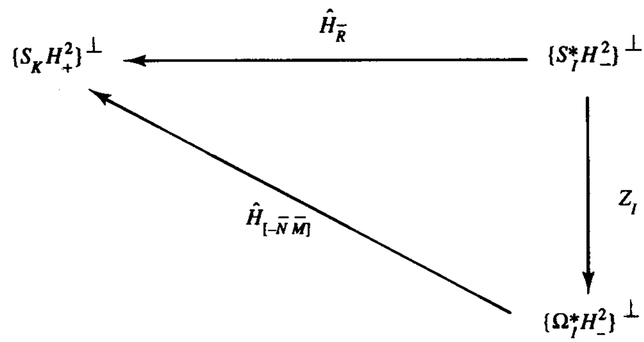

Figure 2. Diagram 2.

is invertible. Its inverse is given by $Y_I\colon \{\Omega_I^* H_-^2\}^\perp \to \{S_I^* H_-^2\}^\perp$

$$Y_I\begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = P_{\{S_I^* H_-^2\}^\perp}[M^* \;\; N^*]\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$$

(2) $\hat{H}_{\bar{R}}\big|_{\{S_I^* H_-^2\}^\perp} = \hat{H}_{[-\bar{N} \;\; \bar{M}]} Z_I$

(3) $\operatorname{Im} \hat{H}_{[-\bar{N} \;\; \bar{M}]} = \{S_K H_+^2\}^\perp$

The adjoints $Z_I^*\colon \{\Omega_I^* H_-^2\}^\perp \to \{S_I^* H_-^2\}^\perp$ and $Y_I^*\colon \{S_I^* H_+^2\}^\perp \to \{\Omega_I^* H_-^2\}^\perp$ are given by

$$Z_I^*\begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = P_{\{S_I^* H_-^2\}^\perp} P_-[-\bar{V}_\mathrm{L} \;\; \bar{U}_\mathrm{L}]\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$$

and

$$Y_I^* f = P_{\{\Omega_I^* H_-^2\}^\perp} P_- \begin{bmatrix} M \\ N \end{bmatrix} f$$

Taking adjoints in Diagram 2 we obtain Diagram 3 in Fig. 3. Combining Diagram 1 with Diagram 3 we obtain the commuting Diagram 4 in Fig. 4. The role played by the operator T which is introduced in Diagram 4 is explained in the following proposition.

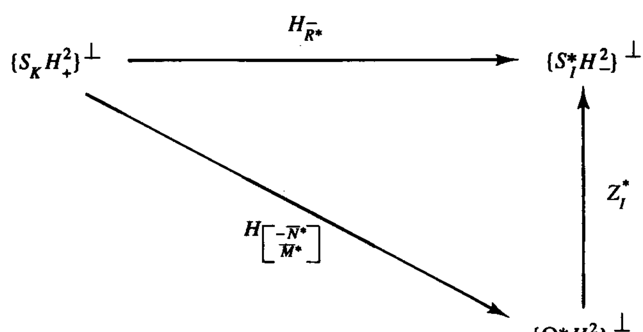

Figure 3. Diagram 3.

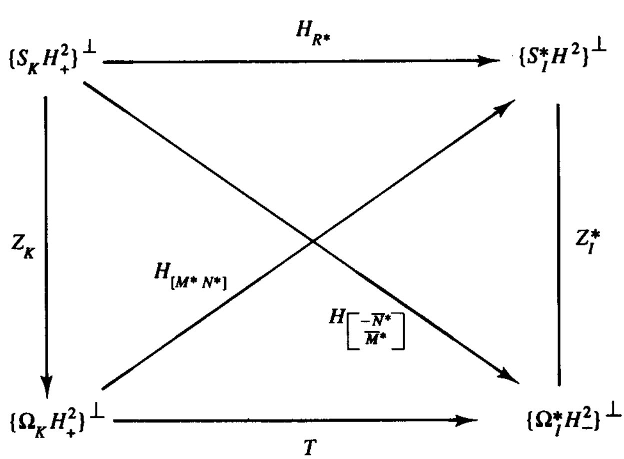

Figure 4. Diagram 4.

Proposition 5.2: *The operator*

$$T: \{\Omega_K H_+^2\}^\perp \to \{\Omega_I^* H_-^2\}^\perp$$

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \mapsto P_-\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}[-\bar{N} \ \ \bar{M}]\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

is such that Diagram 4 commutes.

Proof: We can calculate T to be:

$$T = H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} Y_K = P_-\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} P_{\{S_I H_+^2\}^\perp}[-\bar{N} \ \ \bar{M}]$$

$$= P_-\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}[-\bar{N} \ \ \bar{M}]$$

where the last identity follows since $S_I H_+^2 = \operatorname{Ker} P_-\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$. □

As a corollary to the results in this section we have the following statement about the dimensions of the spaces with which we are dealing.

Corollary 5.2: *If any of the spaces $\{S_K H_+^2\}^\perp$, $\{S_I^* H_-^2\}^\perp$, $\{\Omega_K H_+^2\}^\perp$, $\{\Omega_I^* H_-^2\}^\perp$ is finite dimensional, so are all the other and*

$$\dim(\{S_K H_+^2\}^\perp) = \dim(\{S_I^* H_-^2\}^\perp) = \dim(\{\Omega_K H_+^2\}^\perp) = \dim(\{\Omega_I^* H_-^2\}^\perp).$$

If the McMillan degree of G is n, then n is the dimension of these spaces.

6. Singular values and singular vectors

In the previous section we have introduced a number of operators. In this section we are going to analyse their singular values and singular vectors. For simplicity of presentation we will assume that we are dealing with rational functions.

Denote by $(f_{\Omega,i}, g_{S,i})$, $1 \leq i \leq n$, the Schmidt pairs of the Hankel operator $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$, i.e.

$$H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} g_{S,i} = \sigma_i f_{\Omega,i}$$

for $1 \leq i \leq n$. Similarly, we denote by $(f_{S,i}, g_{\Omega,i})$, $1 \leq i \leq n$, the Schmidt pairs of the operator $H_{[M^* \ N^*]}$ with singular values $\rho_1 \geq \rho_2 \geq \rho_3 \geq \ldots \geq \rho_n > 0$, i.e.

$$H_{[M^* \ N^*]} g_{\Omega,i} = \rho_i f_{S,i}$$

for $1 \leq i \leq n$. Note that $g_{S,i} \in \{S_K H_+^2\}^\perp$, $g_{\Omega,i} \in \{\Omega_K H_+^2\}^\perp$, $f_{S,i} \in \{S_I^* H_-^2\}^\perp$, $f_{\Omega,i} \in \{\Omega_I^* H_-^2\}^\perp$ for $1 \leq i \leq n$.

Proposition 6.1: *The maps Z_K and Y_K, defined by (33) and (34) respectively, satisfy*

(1) $Y_K Y_K^* + H^*_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}\Big|_{\{S_K H_+^2\}^\perp} = I\big|_{\{S_K H_+^2\}^\perp}$

(2) $Z_K^* Z_K = I|_{\{S_K H_+^2\}^\perp} + H_{R^*}^* H_{R^*}|_{\{S_K H_+^2\}^\perp}$

(3) $I|_{\{S_\Omega H_+^2\}^\perp} + (Z_K H^*_{\begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix}})(H_{\begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix}} Z_K^*) = Z_K Z_K^*$

(4) $I|_{\{\Omega_K H_+^2\}^\perp} = Y_K^* Y_K + (Y_K^* H_{R^*}^*)(H_{R^*} Y_K)$

Proof:

(1) For $f \in \{S_K H_+^2\}^\perp$ we have

$$Y_K Y_K^* f = P_{\{S_K H_+^2\}^\perp}[-\bar N \quad \bar M] P_{\{\Omega_K H_+^2\}^\perp} P_+ \begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix} f$$

$$= P_{\{S_K H_+^2\}^\perp}[-\bar N \ \ \bar M]\left[I - \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} P_+ \begin{bmatrix} M^* & N^* \\ J_1^* & J_2^* \end{bmatrix}\right] P_+ \begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix} f$$

$$= P_{\{S_K H_+^2\}^\perp}[-\bar N \ \ \bar M](I - P_-) \begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix} f$$

$$- P_{\{S_K H_+^2\}}{}^\perp[-\overline{N} \ \ \overline{M}] \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} P_+ \begin{bmatrix} M^* & N^* \\ J_1^* & J_2^* \end{bmatrix} P_+ \begin{bmatrix} -\overline{N}^* \\ \overline{M}^* \end{bmatrix} f$$

$$= P_{\{S_K H_+^2\}^\perp} f - P_{\{S_K H_+^2\}^\perp}[-\bar N \ \ \bar M] P_- \begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix} f$$

$$- P_{\{S_K H_+^2\}^\perp}[0 \ \ S_K] P_+ \begin{bmatrix} M^* & N^* \\ J_1^* & J_2^* \end{bmatrix} P_+ \begin{bmatrix} -\overline{N}^* \\ \overline{M}^* \end{bmatrix} f$$

$$= P_{\{S_K H_+^2\}^\perp} f - P_{\{S_K H_+^2\}^\perp}(P_+ + P_-)[-\overline{N} \ \ \overline{M}] P_- \begin{bmatrix} -\overline{N}^* \\ \overline{M}^* \end{bmatrix} f$$

$$= I_{\{S_K H_+^2\}^\perp} f - H^*_{\begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix}} H_{\begin{bmatrix} -\bar N^* \\ \bar M^* \end{bmatrix}} f$$

(2) For $f \in \{S_K H_+^2\}^\perp$ we have

$$Z_K^* Z_K f = P_{\{S_K H_+^2\}^\perp} P_+ [U_L^* \ \ V_L^*] P_{\{\Omega_K H_+^2\}^\perp} \begin{bmatrix} U_L \\ V_L \end{bmatrix} f$$

$$= P_{\{S_K H_+^2\}^\perp} P_+ [U_L^* \ \ V_L^*] \begin{bmatrix} M & -\bar N^* S_K \\ N & \bar M^* S_K \end{bmatrix} P_- \begin{bmatrix} M^* & N^* \\ -S_K^* \bar N & S_K^* \bar M \end{bmatrix} \begin{bmatrix} U_L \\ V_L \end{bmatrix} f$$

$$= P_{\{S_K H_+^2\}^\perp} P_+ [R \ \ S_K] P_- \begin{bmatrix} R^* \\ S_K^* \end{bmatrix} f$$

$$= \hat H_R \hat H_R^* f + I|_{\{S_K H_+^2\}^\perp} f$$

(3) and (4) follow from (1) and (2) by recalling that $Y_K = Z_K^{-1}$. □

In an analogous way to the previous proposition we can obtain results connecting the operators $\hat H_R$, $\hat H_{\begin{bmatrix} M \\ N \end{bmatrix}}$ and Z_I, Y_I.

Proposition 6.2: *The maps Z_I and Y_I, satisfy*

(1) $Y_I Y_I^* + \hat{H}^*_{\begin{bmatrix} M \\ N \end{bmatrix}} \hat{H}_{\begin{bmatrix} M \\ N \end{bmatrix}}|_{\{S_I^* H_-^2\}^\perp} = I|_{\{S_I^* H_-^2\}^\perp}$

(2) $Z_I^* Z_I = I|_{\{S_I^* H_-^2\}^\perp} + \hat{H}_{\bar{R}}^* \hat{H}_{\bar{R}}|_{\{S_I^* H_-^2\}^\perp}$

(3) $I|_{\{S_b^* H_-^2\}^\perp} + (Z_I \hat{H}^*_{\begin{bmatrix} M \\ N \end{bmatrix}})(\hat{H}_{\begin{bmatrix} M \\ N \end{bmatrix}} Z_I = Z_I Z_I^*$

(4) $I|_{\{\Omega_K H_-^2\}^\perp} = Y_I^* Y_I + (Y_I^* \hat{H}_{\bar{R}}^*)(\hat{H}_{\bar{R}} Y_I)$

We will need the following theorem.

Theorem 6.1: *Let T, S be such that*

$$T^*T + S^*S = I$$

or equivalently

$$\|Tx\|^2 + \|Sx\|^2 = \|x\|^2$$

Let the singular values of S be $\sigma_1 \geqslant \ldots \geqslant \sigma_n$ and the singular values of T be $\tau_1 \leqslant \ldots \leqslant \tau_n$. Then

$$\tau_j^2 + \sigma_j^2 = 1 \tag{35}$$

Proof: Let M_j be an arbitrary j-dimensional subspace. Then, using the Min–Max characterization of singular values, we have

$$\min_{x \perp M_{j-1}} \frac{\|Tx\|^2}{\|x\|^2} + \max_{x \perp M_{j-1}} \frac{\|Sx\|^2}{\|x\|^2} = 1$$

and hence

$$\max_{M_{j-1}} \min_{x \perp M_{j-1}} \frac{\|Tx\|^2}{\|x\|^2} + \min_{M_{j-1}} \max_{x \perp M_{j-1}} \frac{\|Sx\|^2}{\|x\|^2} = 1$$

i.e. $\tau_j^2 + \sigma_j^2 = 1$. □

The following theorem summarizes how the singular values and singular vectors of the different operators are related.

Theorem 6.2: *There exist Schmidt pairs $(f_{S,i}, g_{\Omega,i})$ of the Hankel operator $H_{[M^*\ N^*]}$ with singular values $\rho_1 \geqslant \rho_2 \geqslant \ldots \geqslant \rho_n > 0$ and Schmidt pairs $(f_{\Omega,i}, g_{S,i})$ of the Hankel operator $H_{\begin{bmatrix} -\tilde{N}^* \\ \tilde{M}^* \end{bmatrix}}$ with singular values $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_n > 0$ such that*

(1) *The singular values of $H_{\begin{bmatrix} -\tilde{N}^* \\ \tilde{M}^* \end{bmatrix}}$ and $H_{[M^*\ N^*]}$ coincide, i.e. $\sigma_i = \rho_i$, $1 \leqslant i \leqslant n$.*

(2) *The Schmidt pairs of Z_K are $(g_{\Omega,i}, g_{S,i})_{1 \leqslant i \leqslant n}$ with singular values $(1 - \sigma_i^2)^{1/2}$ $1 \leqslant i \leqslant n$.*

(3) *The Schmidt pairs of Z_I are $(f_{\Omega,i}, f_{S,i})_{1 \leqslant i \leqslant n}$ with singular values $(1 - \sigma_i^2)^{1/2}$, $1 \leqslant i \leqslant n$.*

(4) *The Schmidt pairs of H_{R^*} are $(f_{S,i}, g_{S,i})_{1 \leqslant i \leqslant n}$ with singular values $\dfrac{\sigma_i}{(1 - \sigma_i^2)^{1/2}}$, $1 \leqslant i \leqslant n$.*

(5) *The Schmidt pairs of T are $(f_{\Omega,i}, g_{\Omega,i})_{1 \leqslant i \leqslant n}$ with singular values $\dfrac{\sigma_i}{(1 - \sigma_i^2)^{1/2}}$, $1 \leqslant i \leqslant n$.*

Proof: For $1 \leq i \leq n$ set $h_i := \frac{1}{(1-\sigma_i^2)^{1/2}} Y_K^* g_{S,i}$. Since

$$Y_K Y_K^* + H^*_{\left[\begin{smallmatrix}-\bar{N}^*\\ \bar{M}^*\end{smallmatrix}\right]} H_{\left[\begin{smallmatrix}-\bar{N}^*\\ \bar{M}^*\end{smallmatrix}\right]}\Big|_{\{S_K H_+^2\}^\perp} = I|_{\{S_K H_+^2\}^\perp}$$

we have that

$$Y_K h_i = \frac{1}{(1-\sigma_i^2)^{1/2}} Y_K Y_K^* g_{S,i} = \frac{1}{(1-\sigma_i^2)^{1/2}} \left(I - H^*_{\left[\begin{smallmatrix}-\bar{N}^*\\ \bar{M}^*\end{smallmatrix}\right]} H_{\left[\begin{smallmatrix}-\bar{N}^*\\ \bar{M}^*\end{smallmatrix}\right]}\right) g_{S,i}$$
$$= (1-\sigma_i^2)^{1/2} g_{S,i}$$

Furthermore, for $1 \leq l,\ k \leq n$,

$$\langle h_l, h_k\rangle = \frac{1}{(1-\sigma_l^2)^{1/2}(1-\sigma_k^2)^{1/2}} \langle Y_K Y_K^* g_{S,l}, g_{S,k}\rangle$$
$$= \frac{1-\sigma_l^2}{(1-\sigma_l^2)^{1/2}(1-\sigma_k^2)^{1/2}} \delta_{k,l}$$

Hence, $(h_i, g_{S,i})$ are the Schmidt pairs of Y_K with singular values $(1-\sigma_i^2)^{1/2}$, $1 \leq i \leq n$. As $Z_K = Y_K^{-1}$ we also know that $(g_{S,i}, h_i)$ are Schmidt vectors of Z_K with singular values $1/(1-\sigma_i^2)^{1/2}$, $1 \leq i \leq n$.

Now, consider the equation

$$Z_K^* Z_K = I|_{\{S_K H_+^2\}^\perp} + H_{R^*}^* H_{R^*}|_{\{S_K H_+^2\}^\perp}$$

from which we obtain for $1 \leq i \leq n$ that

$$\frac{1}{1-\sigma_i^2} g_{S,i} = Z_K^* Z_K g_{S,i} = g_{S,i} + H_{R^*}^* H_{R^*} g_{S,i}$$

and hence

$$H_{R^*}^* H_{R^*} g_{S,i} = \frac{\sigma_i^2}{1-\sigma_i^2} g_{S,i}$$

Let $k_i := \frac{(1-\sigma_i^2)^{1/2}}{\sigma_i} H_{R^*} g_{S,i}$. Then $(k_i, g_{S,i})$ are the Schmidt pairs of H_{R^*} with singular values $\frac{\sigma_i}{(1-\sigma_i^2)}$, $1 \leq i \leq n$.

In a dual fashion we can infer that Y_I has singular vectors $(f_{S,i}, \bar{h}_i)$ with singular values $(1-\rho_i^2)^{1/2}$, where $\bar{h}_i := \frac{1}{(1-\rho_i^2)^{1/2}} Y_I^* f_{S,i}$, $1 \leq i \leq n$, and $\hat{H}_{\bar{R}} = H_{R^*}^*$ has singular vectors $(\bar{k}_i, f_{S,i})$ with singular values $\frac{\rho_i}{(1-\rho_i^2)}$, where $\bar{k}_i := \frac{(1-\rho_i^2)^{1/2}}{\rho_i} H_{R^*}^* f_{S,i}$, $1 \leq i \leq n$.

Comparing the two singular values, singular vector decompositions of H_R, we can conclude that for all $1 \leq i \leq n$

$$\frac{\sigma_i}{(1-\sigma_i^2)^{1/2}} = \frac{\rho_i}{(1-\rho_i^2)^{1/2}}$$

and therefore that $\rho_i = \sigma_i$

$$k_i = f_{S,i} \quad \bar{k}_i = g_{S,i}$$

Strictly speaking to obtain these identities one set of singular vectors might have to be redefined, due to the fact that singular vectors are not unique in case of repeated singular values and since, in any other case, they are non-unique up to a unitary factor.

We have, therefore, shown that $(f_{S,i}, g_{S,i})$ are Schmidt pairs of H_{R^*} with singular values $\sigma_i/(1-\sigma_i^2)^{1/2}$, $1 \leq i \leq n$. Having proved Parts (1) and (4) we now complete the proof of Parts (2) and (3). Recall that $H_{[M^* \ N^*]} = H_{R^*} Z_K^{-1}$. This decomposition shows that $(f_{S,i}, h_i)$ are the Schmidt vectors of $H_{[M^* \ N^*]}$. But, by assumption, $(f_{S,i}, g_{\Omega,i})$ are the Schmidt pairs of this operator. This implies that $h_i = g_{\Omega,i}$, $1 \leq i \leq n$, which completes the proof of Part (2). We can show analogously that $\bar{h}_i = f_{\Omega,i}$, $1 \leq i \leq n$, and hence Part (3).

To prove Part (5) consider

$$Tg_{\Omega,i} = Z_I^{-*} H_{[M^* \ N^*]} g_{\Omega,i} = \sigma_i Z_I^{-*} f_{S,i} = \frac{\sigma_i}{(1-\sigma_i^2)^{1/2}} f_{\Omega,i}$$

$1 \leq i \leq n$, which proves the claim. □

Note that McFarlane and Glover (1990) had proved that $\|H_{R^*}\| = \sigma_1/(1-\sigma_1^2)^{1/2}$.

The previous theorem is of great importance inasmuch as it allows the Hankel singular value analysis of a normalized coprime factorization to be reduced to the scalar case. This interplay between a vector and a scalar case will be used extensively in the rest of the paper.

Theorem 6.3: *Let* $W: H_+^2 \to H_+^2 \ominus \begin{bmatrix} N \\ M \end{bmatrix} H_+^2$ *be defined by*

$$Wf = P_{\{[{N \atop M}]H_+^2\}^\perp} \begin{bmatrix} U_L \\ V_L \end{bmatrix} f \tag{36}$$

Then the following diagram is commutative.

$$\begin{array}{ccc} H_+^2 & \xrightarrow{\quad W \quad} & H_+^2 \ominus \begin{bmatrix} M \\ N \end{bmatrix} H_+^2 \\ P_{\{S_K H^2\}^\perp} \downarrow & & \downarrow P_{\{\Omega_K H_+^2\}^\perp} \\ \{S_K H_+^2\}^\perp & \xrightarrow{\quad Z_K \quad} & \{\Omega_K H_+^2\}^\perp \end{array} \tag{37}$$

and

$$\|Z_K\| = \|W\| = \inf_{Q \in H_+^\infty} \left\| \begin{bmatrix} U_L \\ V_L \end{bmatrix} + \begin{bmatrix} M \\ N \end{bmatrix} Q \right\|_\infty \tag{38}$$

Proof: Clearly, to show the commutativity of the diagram it suffices to show that, with the notation of Theorem 3.1, for $WS_K H_+^2 \subset \Omega_K H_+^2$ which follows from (32), and that for $f \in \{S_K H_+^2\}^\perp$ we have $\left[{U_L \atop V_L}\right] f \perp \left[{J_1 \atop J_2}\right] H_+^2$. So let us assume $f \in \{S_K H_+^2\}^\perp$ and $g \in H_+^2$. We know that

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} f = \begin{bmatrix} M \\ N \end{bmatrix} R^* f + \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f$$

So

$$\left(\begin{bmatrix} U_L \\ V_L \end{bmatrix} f, \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} g\right) = \left(\left(\begin{bmatrix} M \\ N \end{bmatrix} R^* + \begin{bmatrix} -\bar{N}^* \\ \bar{M} \end{bmatrix}\right) f, \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} g\right)$$
$$= \left(\begin{bmatrix} M \\ N \end{bmatrix} R^* f, \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} g\right) + \left(\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} f, \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} g\right)$$
$$= (R^* f, (M^* J_1 + N^* J_2) g) + (f, (-\bar{N}^* J_1 + -\bar{M}^* J_2) g) = 0$$

Here we used the fact that

$$(M^* J_1 + N^* J_2) = (-M^* \bar{N}^* S + N^* \bar{M}^* S_K) = (-M^* \bar{N}^* + N^* \bar{M}^*) S_K \tag{39}$$

and

$$-\bar{N} J_1 + \bar{M} J_2 = (\overline{NN^*} + \overline{MM^*}) S_K = S_K \tag{40}$$

and $f \perp S_K H_+^2$ by assumption.

By the commutant lifting theorem, see Sz.-Nagy and Foias (1970), there exists a map $\bar{W}$ such that $\|\bar{W}\| = \|Z_K\|$ and the diagram

$$\begin{array}{ccc} H_+^2 & \xrightarrow{\bar{W}} & H_+^2 \\ P_{\{S_K H^2\}^\perp} \downarrow & & \downarrow P_{\{\Omega_K H_+^2\}^\perp} \\ \{S_K H_+^2\}^\perp & \xrightarrow{Z_K} & \{\Omega_K H_+^2\}^\perp \end{array} \tag{41}$$

commutes.

Obviously

$$\|\bar{W}\| = \inf \left\| \begin{bmatrix} U_L \\ V_L \end{bmatrix} + \begin{bmatrix} M & J_1 \\ N & J_2 \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \right\| \leq \inf \left\| \begin{bmatrix} U_L \\ V_L \end{bmatrix} + \begin{bmatrix} M \\ N \end{bmatrix} Q \right\| = \|W\|$$

So

$$\|W\| = \|\bar{W}\| = \|Z_K\| \leq \|W\|$$

and equality follows. □

7. State-space realizations

In this section we are going to derive state space realizations for the transfer functions which were introduced in the previous sections.

We begin with the derivation of state space formulas for normalized right and left coprime factors. These formulae were first obtained by Meyer and Franklin (1987). Our proof is however different and is based on a derivation of the realization from the spectral factorization underlying the normalized coprime factorizations. The main tool is Theorem 2.3.

Lemma 7.1:

(1) *Let G be a strictly proper rational transfer function, and let NM^{-1} be its normalized right coprime factorization. Let G have a minimal state-space realization (A, B, C). Then a state-space realization for $\begin{bmatrix} M \\ N \end{bmatrix}$ is given by*

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A - BB^*X & B \\ \hline -B^*X & I \\ C & 0 \end{array}\right] \tag{42}$$

where X is the positive definite solution to the Control Algebraic Riccati Equation (CARE),

$$A^*X + XA + C^*C - XBB^*X = 0$$

(2) *Let G be a strictly proper rational transfer function, and let $\bar{M}^{-1}\bar{N}$ be its normalized left coprime factorization. Let G have a minimal state-space realization (A, B, C). Then a state-space realization for $(-\bar{N}\ \ \bar{M})$ is given by*

$$(-\bar{N}\ \ \bar{M}) = \left[\begin{array}{c|cc} A - ZC^*C & B & ZC^* \\ \hline -C & 0 & I \end{array}\right] \tag{43}$$

where Z is the positive definite solution to the Filter Algebraic Riccati Equation (FARE),

$$AZ + ZA^* + BB^* - XC^*CX = 0$$

Proof:

(1) Let $\begin{bmatrix} M \\ N \end{bmatrix} = \begin{bmatrix} DT^{-1} \\ ET^{-1} \end{bmatrix}$ be the normalized coprime factorization of $G = ED^{-1}$ as constructed in Lemma 3.1 and consider

$$\begin{aligned} \Phi(s) &= I + G^*(s)G(s) = I + (D^*)^{-1}E^*ED^{-1} \\ &= (D^*)^{-1}(D^*D + E^*E)D^{-1} = (D^*)^{-1}T^*TD^{-1} \end{aligned} \tag{44}$$

Now, from $\Phi(s) = I + G^*(s)G(s)$ we have a series realization for Φ of the form

$$\left(\begin{bmatrix} A & 0 \\ C^*C & -A^* \end{bmatrix}, \begin{bmatrix} B \\ 0 \end{bmatrix}, (0\ \ -B^*), I\right) \tag{45}$$

On the other hand, $\Phi(s) = ((D^*)^{-1}T^*)(TD^{-1})$, and since, by Theorem 2.3, the transfer function TD^{-1} has a realization of the form (A, B, C_0, I) and $(D^*)^{-1}T^*$ a realization of the form $(-A^*, C_0^*, -B^*, I)$, by cascading the two, Φ must have also a realization of the form

$$\left(\begin{bmatrix} A & 0 \\ C_0^*C_0 & -A^* \end{bmatrix}, \begin{bmatrix} B \\ C_0^* \end{bmatrix}, (C_0\ \ -B^*), I\right) \tag{46}$$

By the state-space isomorphism theorem these two realizations are isomorphic and the isomorphism is unique. Assume X is such that

$$\begin{bmatrix} I & 0 \\ X & J \end{bmatrix}\begin{bmatrix} A & 0 \\ C^*C & -A^* \end{bmatrix} = \begin{bmatrix} A & 0 \\ C_0^*C_0 & -A^* \end{bmatrix}\begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \tag{47}$$

This is equivalent to the pair of matrix equations

$$\begin{aligned} XA + C^*C &= C_0^*C_0 - A^*X \\ XB &= C_0 \end{aligned} \tag{48}$$

In turn this is equivalent to the Control Algebraic Riccati Equation (CARE):

$$XA + A^*X + C^*C - XBB^*X = 0 \tag{49}$$

In particular, any solution of the CARE leads to a pair N, M, however, only the stabilizing solution leads to a stable pair.

Now we saw that $TD^{-1} = (A, B, B^*X, I)$ and so, by inverting the realization in a standard way,

$$M = DT^{-1} = (A - BB^*X, B, -B^*X, I) \tag{50}$$

Assuming, as we did, that G is strictly proper we have, by another recourse to Theorem 2.3, that

$$N = ET^{-1} = (A - BB^*X, B, C_1, 0) \tag{51}$$

Now

$$\begin{aligned} G &= (ET^{-1})(TD^{-1}) \\ &= C_1(sI - A + BB^*X)^{-1}B(I + B^*X(sI - A)^{-1}B) \\ &= C_1(sI - A + BB^*X)^{-1}(I + BB^*X(sI - A)^{-1})B \\ &= C_1(sI - A + BB^*X)^{-1}(sI - A + BB^*X)(sI - A)^{-1}B \\ &= C_1(sI - A)^{-1}B \end{aligned} \tag{52}$$

But $G = C(sI - A)^{-1}B$ so we get $C_1 = C$ and

$$ET^{-1} = (A - BB^*X, B, C, 0) \tag{53}$$

Taking (50) and (53) together is equivalent to (42).

(2) The proof is similar and omitted. □

We now come to the derivation of state-space realizations for the coprime factors of the LQG controller K. In the proof it can be seen how the range inclusion result of Corollary 4.2 leads to the desired state-space realization.

Lemma 7.2: *Let G be a strictly proper rational transfer function, and let NM^{-1} be its normalized right coprime factorization. Let G have a minimal state space realization (A, B, C). Then state-space realizations for the coprime factors of the LQG controller are given by*

$$\begin{bmatrix} U_{\mathrm{L}} \\ V_{\mathrm{L}} \end{bmatrix} = \left[\begin{array}{c|c} A - BB^*X & ZC^* \\ \hline -B^*X & 0 \\ C & I \end{array}\right] \tag{54}$$

and

$$(\bar{V}_{\mathrm{L}} \;\; -\bar{U}_{\mathrm{L}}) = \left[\begin{array}{c|cc} A - ZC^*C & B & ZC^* \\ \hline B^*X & I & 0 \end{array}\right] \tag{55}$$

Proof: Since $\operatorname{Im} \hat{H}_{\begin{bmatrix} M \\ N \end{bmatrix}} = \{\Omega_K H_+^2\}^{\perp}$ we have the following left coprime representations

$$\begin{bmatrix} M \\ N \end{bmatrix} = \Omega_K H^* \tag{56}$$

for some $H \in H_+^\infty$. But

$$\operatorname{Im} \hat{H}_{\begin{bmatrix} U_L \\ V_L \end{bmatrix}} \subset \operatorname{Im} \hat{H}_{\begin{bmatrix} M \\ N \end{bmatrix}}$$

so we have the following, not necessarily coprime, representation

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} = \Omega_K F^* \tag{57}$$

for some $F \in H_+^\infty$. Factorizations (56) and (57) imply that if

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A - BB^*X & B \\ \hline -B^*X & I \\ C & 0 \end{array}\right] \tag{58}$$

then

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} = \left[\begin{array}{c|c} A - BB^*X & L \\ \hline -B^*X & 0 \\ C & I \end{array}\right] \tag{59}$$

for some linear map L. Recalling the realization (43) for $(-\bar{N} \;\; \bar{M})$ and the fact that $\bar{M} V_L - \bar{N} U_L = I$, we compute

$$\begin{aligned} I &= (I - C(sI - A + ZC^*C)^{-1} ZC^*)(I + C(sI - A + BB^*X)^{-1} L) \\ &\quad + C(sI - A + ZC^*C)^{-1} BB^*X(sI - A + BB^*X)^{-1} L \\ &= I - C(sI - A + ZC^*C)^{-1} ZC^* + C(sI - A + BB^*X)^{-1} L \\ &\quad - C(sI - A + ZC^*C)^{-1} ZC^*C(sI - A + BB^*X)^{-1} L \\ &\quad + C(sI - A + ZC^*C)^{-1} BB^*X(sI - A + BB^*X)^{-1} L \end{aligned}$$

Hence

$$\begin{aligned} & C(sI - A + ZC^*C)^{-1} ZC^* \\ &\quad = C\{I - (sI - A + ZC^*C)^{-1} ZC^*C + (sI - A + ZC^*C)^{-1} BB^*X\} \\ &\qquad \times (sI - A + BB^*X)^{-1} L) \\ &\quad = C(sI - A + ZC^*C)^{-1}\{sI - A + ZC^*C - ZC^*C + BB^*X\} \\ &\qquad \times (sI - A + BB^*X)^{-1} L \\ &\quad = C(sI - A + ZC^*C)^{-1}(sI - A + BB^*X)(sI - A + BB^*X)^{-1} L \\ &\quad = C(sI - A + ZC^*C)^{-1} L \end{aligned}$$

since the pair $(C, A - ZC^*C)$ is observable, the equality

$$L = ZC^* \tag{60}$$

follows. Thus, (54) is proved.

Putting (42) and (54) together we obtain

$$\begin{bmatrix} M & U_L \\ N & V_L \end{bmatrix} = \left[\begin{array}{c|cc} A - BB^*X & B & ZC^* \\ \hline -B^*X & I & 0 \\ C & 0 & I \end{array}\right] \tag{61}$$

Now, starting from

$$\begin{bmatrix} \overline{V}_L & -\overline{U}_L \\ -\overline{N} & \overline{M} \end{bmatrix} \begin{bmatrix} M & U_L \\ N & V_L \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{62}$$

and using the state-space representation for inverse transfer functions, we get

$$\begin{bmatrix} \overline{V}_L & -\overline{U}_L \\ -\overline{N} & \overline{M} \end{bmatrix} = \left[\begin{array}{c|cc} A - ZC^*C & B & ZC^* \\ \hline B^*X & I & 0 \\ -C & 0 & I \end{array}\right] \tag{63}$$

□

From this result we immediately obtain a state-space realization for the controller K.

Lemma 7.3: *Let G be a strictly proper rational transfer function, and let NM^{-1} be its normalized right coprime factorization. Let G have a minimal state space realization (A, B, C). Then a state-space realization of the LQG controller $K = U_L V_L^{-1}$ is given by*

$$K = U_L V_L^{-1} = \left[\begin{array}{c|c} A - BB^*X - ZC^*C & ZC^* \\ \hline -B^*X & 0 \end{array}\right] \tag{64}$$

Proof: Starting from (54) we have

$$V_L^{-1} = I - C(sI - A + BB^*X + ZC^*C)^{-1} ZC^* \tag{65}$$

and hence

$$\begin{aligned} U_L V_L^{-1} &= -B^*X(sI - A + BB^*X)^{-1} ZC^* \\ &\quad \times [I - C(sI - A + BB^*X + ZC^*C)^{-1} ZC^*] \\ &= -B^*X(sI - A + BB^*X)^{-1} \\ &\quad \times [I - ZC^*C(sI - A + BB^*X + ZC^*C)^{-1}] ZC^* \\ &= -B^*X(sI - A + BB^*X)^{-1} [sI - A + BB^*X + ZC^*C - ZC^*C] \\ &\quad \times (sI - A + BB^*X + ZC^*C)^{-1} ZC^* \\ &= -B^*X(sI - A + BB^*X)^{-1} \\ &\quad \times (sI - A + BB^*X)(sI - A + BB^*X + ZC^*C)^{-1} ZC^* \\ &= -B^*X(sI - A + BB^*X + ZC^*C)^{-1} ZC^* \end{aligned}$$

□

We would like to comment on the state-space representation (64). Clearly the poles of $U_L V_L^{-1}$ are equal to the zeros of V_L. Now

$$V_L = \left[\begin{array}{c|c} A - BB^*X & ZC^* \\ \hline C & I \end{array}\right] \tag{66}$$

So the zeros of V_L are the zeros of the polynomial system matrix, see Fuhrmann and Hautus (1980).

$$\left[\begin{array}{c|c} sI - A + BB^*X & ZC^* \\ \hline -C & I \end{array}\right] \tag{67}$$

By elementary transformations these are the zeros of $sI - A + BB^*X + ZC^*C$.

We now come to derive state space realizations for R. These realizations were derived for the first time by Glover and McFarlane (1988) using different methods. In particular, we avoid the use of the Bucy relationships by invoking Proposition 4.2.

Theorem 7.1: *Let R be defined as in Lemma 4.1 i.e. $R = U_L^*M + V_L^*N$. Then R has the state-space realizations*

$$R = \left[\begin{array}{c|c} A - BB^*X & B \\ \hline C(I + ZX) & 0 \end{array}\right] \tag{68}$$

and

$$R = \left[\begin{array}{c|c} A - ZC^*C & (I + ZX)B \\ \hline C & 0 \end{array}\right] \tag{69}$$

Here X, Y are the unique positive definite solutions of CARE and FARE.

Proof: From (54) we get

$$(U_L^* \quad V_L^*) = \left[\begin{array}{c|cc} -A^* + BB^*X & -XB & C^* \\ \hline -CZ & 0 & I \end{array}\right] \tag{70}$$

and using (42) we compute

$$\begin{aligned} R(s) &= U_L(s)^*M(s) + V_L(s)^*N(s) \\ &= CZ(sI + A^* - XBB^*)^{-1}XB(I - B^*X(sI - A + BB^*X)^{-1}B) \\ &\quad + (I - CZ(sI + A^* - XBB^*)^{-1}C^*)C(sI - A + BB^*X)^{-1}B \\ &= CZ(sI + A^* - XBB^*)^{1}XB \\ &\quad - CZ(sI + A^* - XBB^*)^{-1}XBB^*X(sI - A + BB^*X)^{-1}B \\ &\quad + C(sI - A + BB^*X)^{-1}B \\ &\quad - CZ(sI + A^* - XBB^*)^{-1}C^*C(sI - A + BB^*X)^{-1}B \end{aligned}$$

Hence

$$\begin{aligned} R &- C(sI - A + BB^*X)^{-1}B \\ &= CZ(sI + A^* - XBB^*)^{-1}\{X - XBB^*X(sI - A + BB^*X)^{-1} \\ &\quad - C^*C(sI - A + BB^*X)^{-1}\}B \\ &= CZ(sI + A^* - XBB^*)^{-1} \\ &\quad \times \{X(sI - A + BB^*X) - XBB^*X - C^*C\}(sI - A + BB^*X)^{-1}B \\ &= CZ(sI + A^* - XBB^*)^{-1}\{sX - XA - C^*C\}(sI - A + BB^*X)^{-1}B \\ &= CZ(sI + A^* - XBB^*)^{-1}\{sX + A^*X - XBB^*X\} \\ &\quad \times (sI - A + BB^*X)^{-1}B \\ &= CZ(sI + A^* - XBB^*)^{-1}(sI + A^* - XBB^*)X(sI - A + BB^*X)^{-1}B \\ &= CZX(sI - A + BB^*X)^{-1}B \end{aligned}$$

or

$$R = C(I + ZX)(sI - A + BB^*X)^{-1}B \tag{71}$$

i.e. (68) follows.

In an analogous way we show that $\bar{R} = \overline{MU_{\mathrm{L}}^*} + \overline{NV_{\mathrm{L}}^*}$ has a state space realization given by

$$\bar{R}(s) = C(sI - A + ZC^*C)^{-1}(I + ZX)B$$

Since $R(s) = \bar{R}(s)$ we have therefore derived the second state-space realization for R. □

We have seen how to obtain state-space realizations of R. The realization approach can be bypassed and we can derive procedures for the computation of R working directly with polynomial data. We only consider the scalar case.

Thus, let $G = e/d$, and let $G = N/M$, with $N = e/t$, $M = d/t$ the normalized coprime factors of G. By Corollary 4.2 the coprime factors of the LQG controller have the form $U_{\mathrm{L}} = b/t$, $V_{\mathrm{L}} = a/t$, with $\deg a = \deg t$ and $\deg b < \deg t$. The Bezout equation is now $MV_{\mathrm{L}} - NU_{\mathrm{L}} = 1$ or

$$\frac{d}{t}\frac{a}{t} - \frac{e}{t}\frac{b}{t} = 1 \tag{72}$$

which is equivalent to the polynomial equation

$$da - eb = t^2 \tag{73}$$

Note that both d and t are monic polynomials. Let us write $s = t - d$. This implies $t^2 = d^2 + 2ds + s^2$. Thus, it is enough to solve the following three equations,

$$da - eb = \begin{cases} d^2 \\ 2ds \\ s^2 \end{cases} \tag{74}$$

The equation $da' - eb' = s^2$ has, by the fact that $e \wedge d = 1$, a unique solution a', b' with $\deg a' < \deg e$ and $\deg b' < \deg d$. The equation $da'' - eb'' = 2ds$ is solved by $a'' = 2s$, $b'' = 0$ whereas the equation $da''' - eb''' = d^2$ is solved by $a''' = d$, $b''' = 0$. So a solution of (73) is given by

$$a = 2t - d + a', \quad b = b'$$

The LQG controller is therefore $K = U_L V_L^{-1} = ba^{-1} = b'/2t - d + a'$

We compute a simple example using both the state-space and the polynomial methods. Let $G(s) = 1/(s+1)$. Then a realization is given by

$$G = \left[\begin{array}{c|c} -1 & 1 \\ \hline 1 & 0 \end{array}\right] \tag{75}$$

Both Riccati equations, i.e. the CARE and FARE, reduce in this case to $x^2 + 2x - 1 = 0$. So $x = -1 \pm \sqrt{2}$. Since we look for $x = z > 0$ we must have $x = z = \sqrt{2} - 1$. Now

$$\begin{bmatrix} U_L \\ V_L \end{bmatrix} = \left[\begin{array}{c|c} A - BB^*X & ZC^* \\ \hline -B^*X & 0 \\ C & I \end{array}\right] = \left[\begin{array}{c|c} -\sqrt{2} & \sqrt{2} - 1 \\ \hline 1 - \sqrt{2} & 0 \\ 1 & 1 \end{array}\right] \tag{76}$$

So

$$U_L = \frac{2\sqrt{2} - 3}{s + \sqrt{2}}$$

$$V_L = \frac{s + 2\sqrt{2} - 1}{s + \sqrt{2}} \tag{77}$$

and therefore the LQG controller is given by $K = \dfrac{2\sqrt{2} - 3}{s + 2\sqrt{2} - 1}$. The function R is given by

$$R = \left[\begin{array}{c|c} A - BB^*X & B \\ \hline C(I + ZX) & 0 \end{array}\right] = \left[\begin{array}{c|c} -\sqrt{2} & 1 \\ \hline 2(2 - \sqrt{2}) & 0 \end{array}\right] \tag{78}$$

and so $R(s) = 2(2 - \sqrt{2})/s + \sqrt{2}$. Of course this also implies that

$$R^*(s) = \frac{-2(2 - \sqrt{2})}{s - \sqrt{2}}$$

We now repeat the computation polynomially. From the spectral factorization $dd^* + ee^* = tt^*$ we get $t = s + \sqrt{2}$. Solving $da - eb = t^2$ by the method outlined before we get $a = s + 2\sqrt{2} - 1$, $b = 2\sqrt{2} - 3$. This leads to (77).

$$\begin{aligned} R^* &= M^*U_L + N^*V_L = \frac{d^*}{t^*} U_L + \frac{e^*}{t^*} V_L \\ &= \frac{1 - s}{\sqrt{2} - s} \, \frac{2\sqrt{2} - 3}{\sqrt{2} + s} + \frac{1}{\sqrt{2} - s} \, \frac{s + 2\sqrt{2} - 1}{\sqrt{2} + s} \\ &= \frac{s(4 - 2\sqrt{2}) - (4 - 4\sqrt{2})}{(\sqrt{2} - s)(\sqrt{2} + s)} \\ &= \frac{4 - 2\sqrt{2}}{\sqrt{2} - s} \end{aligned} \tag{79}$$

or

$$R^* = \frac{-2(2 - \sqrt{2})}{s - \sqrt{2}} \tag{80}$$

8. A detailed analysis of singular vectors

In this section we will present a detailed analysis of the singular values and singular vectors of Hankel operators associated with a scalar, normalized coprime factorization. We will, however, only restrict ourselves to the case where the transfer function G is scalar.

The keys to this analysis are the results of Theorem 6.2 that relate the singular values and singular vectors of this Hankel operator to those of a related Hankel operator with a scalar, antistable symbol, specifically to the Hankel operator associated with the symbol R^*.

To the analysis of this Hankel operator we can bring to bear all the results of Fuhrmann (1991). These are summarized in Proposition 8.1. The results of this section will be used in the following sections to study a Hankel norm approximation problem and a Nehari extension problem.

The following proposition summarizes a number of results on scalar functions. Since we will apply these to the function R^* we state the results for $R^* = r^*/t^*$.

Proposition 8.1: *Let $r^*/t^* \in H_-^\infty$ be a scalar, strictly proper, transfer function, with r and t coprime polynomials and t is monic of degree n. Assume that $\mu_1 \geq \ldots \geq \mu_n > 0$ are the singular value of H_{r^*/t^*}.*

(1) *There exist uniquely determined signs ε_i and polynomials p_i, $1 \leq i \leq n$, such that*

(*a*) $\left\{\frac{p_i}{t}, \varepsilon_i \frac{p_i^*}{t^*}\right\}_{i=1}^n$ *are* (non-normalized) *Schmidt vectors of H_{r^*/t^*}.*

(*b*) *if μ is a singular value of H_{r^*/t^*}, then there exists an index k such that the Schmidt pair $\left(\frac{p_k}{t}, \varepsilon_k \frac{p_k^*}{t^*}\right)$ has the following property. Amongst all the numerators q^1 of the Schmidt vector q^1/t of the Schmidt pairs $\left(\frac{q^1}{t}, \frac{q^2}{t^*}\right)$ of H_{r^*/t^*} with singular value σ, the numerator p_k of p_k/t has the smallest degree.*

(2) *There exist polynomials π_i, $1 \leq i \leq n$, such that with $\lambda_i = \varepsilon_i \mu_i$, $1 \leq i \leq n$,*

$$\frac{r^*}{t^*}\frac{p_i}{t} = \lambda_i \frac{p_i^*}{t^*} + \frac{\pi_i}{t} \tag{81}$$

$$r^* p_i = \lambda_i t p_i^* + t^* \pi_i \tag{82}$$

or

$$\frac{r^*}{t^*} = \lambda_i \frac{t p_i^*}{t^* p_i} + \frac{\pi_i}{p_i} \tag{83}$$

(3) *There exist polynomials α_{ij} of degree less than or equal to $n-2$ with the properties*

$$\alpha_{ij} = -\alpha_{ij},\ \alpha_{ii} = 0$$

for $1 \leqslant i,\ j \leqslant n$, *such that*

(*a*)
$$\lambda_i p_i^* p_j - \lambda_j p_j^* p_i = t\alpha_{ij} \tag{84}$$

(*b*) *if* i, j *are such that* $\mu_i \neq \mu_j$, *then* α_{ij} *is non-zero.*

(*c*) *if* i, j *are such that* $\mu_i \neq \mu_j$, *then*

$$p_i p_j^* = \frac{1}{\lambda_i^2 - \lambda_j^2}\{\lambda_j t^* \alpha_{ij} + \lambda_i t \alpha_{ij}^*\} \tag{85}$$

for $1 \leqslant i,\ j \leqslant n$.

Proof: The proof for, (1) and (2) is essentially Theorem 3.3 in Fuhrmann (1991).

(3) By eliminating the left terms of the two equations

$$\left.\begin{aligned} r^* p_i &= \lambda_i t p_i^* + t^* \pi_i \\ r^* p_j &= \lambda_j t p_j^* + t^* \pi_j \end{aligned}\right\} \tag{86}$$

it follows that

$$0 = t\{\lambda_i p_i^* p_j - \lambda_j p_j^* p_i\} + t^*\{\pi_i p_j - \pi_j p_i\}$$

Since t and t^* are coprime, there exist polynomials α_{ij} such that

$$\lambda_i p_i^* p_j - \lambda_j p_j^* p_i = t\alpha_{ij}$$

$1 \leqslant i,\ j \leqslant n$. Since the degree of p_i and the degree of p_j is less than or equal to $n-1$, and the degree of t is n, the degrees of α_{ij} are less than or equal to $n-2$. We can also see from this expression that $\alpha_{ij} = -\alpha_{ij}$ and $\alpha_{ii} = 0$, $1 \leqslant i$, $j \leqslant n$. Now assume that i, j are such that $\sigma_i \neq \sigma_j$. Solving the equation

$$\lambda_i p_i^* p_j - \lambda_j p_j^* p_i = t\alpha_{ij}$$

and its complex conjugate

$$-\lambda_i p_i^* p_j + \lambda_i p_i p_j^* = t\alpha_{ij}^*$$

for $p_i p_j^*$, we obtain that

$$p_i p_j^* = \frac{1}{\lambda_i^2 - \lambda_j^2}\{\lambda_j t^* \alpha_{ij} + \lambda_i t \alpha_{ij}^*\}$$

Since $p_i p_j^*$ is non-zero this expression also implies that α_{ij} has to be non-zero. □

Our aim is to study the Hankel operator $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ given a NLCF $g = \bar{M}^{-1}\bar{N}$ of the transfer function $g = \dfrac{e}{d}$, where e and d are coprime polynomials, with d monic of degree n. Note that since g is assumed to be a scalar transfer function and that the NLCF and the NRCF coincide. As before, our study will be based on the relationship between $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ and H_{R^*} as established in § 5. We recall the basic set-up and at the same time introduce the notation which we will be using in the following. To this end, let us take a polynomial spectral factorization

$$ee^* + dd^* = tt^* \tag{87}$$

or

$$\left(\frac{e}{t}\right)\left(\frac{e}{t}\right)^* + \left(\frac{d}{t}\right)\left(\frac{d}{t}\right)^* = 1 \tag{88}$$

Here t is stable and normalized to be monic. In this case

$$\begin{bmatrix} M \\ N \end{bmatrix} = \begin{bmatrix} d/t \\ e/t \end{bmatrix} \tag{89}$$

and

$$(-\bar{N} \;\; \bar{M}) = \left(-\frac{e}{t} \;\; \frac{d}{t}\right) \tag{90}$$

For the associated function R^* we have $R^* =: r^*/t^*$. Note that the polynomials r and t are coprime, since the McMillan degrees of R and g are the same (see Theorem 6.2). Then $\operatorname{Ker} H_{R^*} = \dfrac{t^*}{t} H_+^2$ and so

$$\{\operatorname{Ker} H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}\}^\perp = \{\operatorname{Ker} H_{R^*}\}^\perp = \{S_K H_+^2\}^\perp = X^t \tag{91}$$

with $S_K = t^*/t$. In the previous proposition p_i/t was defined to be a (non-normalized) Schmidt vector of H_{R^*}. By Theorem 6.2 p_i/t is therefore also a (non-normalized) Schmidt vector of $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$. By the results in § 5 the functions in $\operatorname{Im} H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ are of the form $\begin{bmatrix} a/t^* \\ b/t^* \end{bmatrix}$ for some polynomials a and b. Therefore the ith Schmidt pair of $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ with singular value σ_i can be written as $\left(\begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}, \frac{p_i}{t}\right)$ for some polynomials $\hat{p}_1^{(i)}$, $\hat{p}_2^{(i)}$ whose degrees are less then the degree of t.

By Lemma 3.5 the operator $\hat{H}_{[-\bar{N} \;\; \bar{M}]}$ restricted to $\{\operatorname{Ker} \hat{H}_{(-\bar{N} \;\; \bar{M})}\}^\perp$ acts by multiplication. Therefore the *singular value/singular vector equations* for $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$ can be written as

$$\left.\begin{aligned} H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} \frac{p_i}{t} &= P_- \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} \frac{p_i}{t} = \sigma_i \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} \\ H^*_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} &= (-\bar{N} \;\; \bar{M}) \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} = \sigma_i \frac{p_i}{t} \end{aligned}\right\} \tag{92}$$

By partial fraction decomposition there exist polynomials $\pi_1^{(i)}$, $\pi_2^{(i)}$ of degree at most $n-1$, such that

$$\left.\begin{aligned} \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} \frac{p_i}{t} &= \sigma_i \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} + \begin{bmatrix} \pi_1^{(i)}/t \\ \pi_2^{(i)}/t \end{bmatrix} \\ \begin{bmatrix} \frac{-e}{t} & \frac{d}{t} \end{bmatrix} \begin{bmatrix} p_1^{(i)}/t^* \\ p_2^{(i)}/t^* \end{bmatrix} &= \sigma_i \frac{p_i}{t} \end{aligned}\right\} \tag{93}$$

or, rewritten polynomially,

$$\left.\begin{aligned} \begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i &= \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix} \\ -e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} &= \sigma_i t^* p_i \end{aligned}\right\} \tag{94}$$

Another way to rewrite (93) is

$$\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} = \sigma_i \frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i \end{bmatrix} + \begin{bmatrix} \pi_1^{(i)}/p_i \\ \pi_2^{(i)}/p_i \end{bmatrix} \tag{95}$$

We now proceed to analyse in more detail the Schmidt pairs of the Hankel operator $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$. These results will be proved in a series of Lemmas.

In the following lemma we prove that the ratio of the vectors of a Schmidt pair is an all-pass function. This result is one of the corner stones for the derivations in the subsequent sections on Hankel norm approximation and Nehari extension.

Lemma 8.1: *Let* $\left\{ \frac{p}{t}, \begin{bmatrix} \hat{p}_1/t^* \\ \hat{p}_2/t^* \end{bmatrix} \right\}$ *be a Schmidt pair associated with the singular value* σ. *Then* $\begin{bmatrix} \hat{p}_1/p \\ \hat{p}_2/p \end{bmatrix}$ *is all-pass.*

Proof: Taking the adjoint of the first of the singular vector equations

$$\begin{aligned} \begin{bmatrix} -e^* \\ d^* \end{bmatrix} p &= \sigma t \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \end{bmatrix} + t^* \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} \\ -e\hat{p}_1 + d\hat{p}_2 &= \sigma t^* p \end{aligned} \tag{96}$$

we get

$$\begin{aligned} p^*(-e \;\; d) &= \sigma t^*(\hat{p}_1^* \;\; \hat{p}_2^*) + t(\pi_1^* \;\; \pi_2^*) \\ -e\hat{p}_1 + d\hat{p}_2 &= \sigma t^* p \end{aligned} \tag{97}$$

Multiplying the first equation on the right by $\begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \end{bmatrix}$ the second by p^* and subtracting we get

$$0 = \sigma t^*\{pp^* - (\hat{p}_1\hat{p}_1^* + \hat{p}_2\hat{p}_2^*)\} - t(\pi_1^*\hat{p}_1 + \pi_2^*\hat{p}_2) \tag{98}$$

Therefore, as t and t^* are coprime polynomials, we have $t | pp^* - (\hat{p}_1\hat{p}_1^* + \hat{p}_2\hat{p}_2^*)$ and by symmetry also $t^* | pp^* - (\hat{p}_1\hat{p}_1^* + \hat{p}_2\hat{p}_2^*)$. By a degree argument we therefore have that $pp^* - (\hat{p}_1\hat{p}_1^* + \hat{p}_2\hat{p}_2^*) = 0$ or

$$p_i p_i^* = (\hat{p}_1^{(i)}(\hat{p}_1^{(i)})^* + \hat{p}_2^{(i)}(\hat{p}_2^{(i)})^*) \tag{99}$$

This is equivalent to $\begin{bmatrix} \hat{p}_1^*/p \\ \hat{p}_2^*/p \end{bmatrix}$ being all-pass. □

In the following Lemma all Schmidt pairs are described that correspond to a particular singular value. We will use Lemma 3.3 in Fuhrmann (1991).

Lemma 8.2:

(1) *Let σ_i be a singular value of the Hankel operator $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. Then there exists an index k such that the Schmidt pair $\left\{ \frac{p_k}{t}, \begin{bmatrix} \hat{p}_1^{(k)}/t^* \\ \hat{p}_2^{(k)}/t^* \end{bmatrix} \right\}$ has the following property: amongst all the numerators of all Schmidt vectors $\frac{q}{t}$ of the Schmidt pairs $\left\{ \frac{q}{t}, \begin{bmatrix} \hat{q}_1/t^* \\ \hat{q}_2/t^* \end{bmatrix} \right\}$ corresponding to the singular value σ_i, the numerator of p_k/t, i.e. p_k has the smallest degree.*

(2) *Let $\left\{ \frac{p_i}{t}, \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} \right\}$ be a minimal degree σ_i-Schmidt pair of $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. All Schmidt pairs corresponding to the singular value σ_i are of the form $\left\{ \frac{q_i}{t}, \begin{bmatrix} \hat{q}_1^{(i)}/t^* \\ \hat{q}_2^{(i)}/t^* \end{bmatrix} \right\}$ with*

$$\left. \begin{aligned} q_i &= p_i a \\ \begin{bmatrix} \hat{q}_1^{(i)} \\ \hat{q}_2^{(i)} \end{bmatrix} &= \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} \end{aligned} \right\} \tag{100}$$

with a any polynomial such that $\deg a < \deg t - \deg p_i$.

In particular the multiplicity of σ_i as a singular value is equal to $n - \deg p_i$.

Proof:

(1) This follows from Proposition 8.1 and the fact that H_{r^*/t^*} and $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$ share the same Schmidt vectors in H_+^2.

(2) By Lemma 3.3 in Fuhrmann (1991) we have $q_i = ap_i$ for some polynomial a of degree $\leq n - \deg p_i$. Consider now the two pairs of singular

vector and singular value equations:

$$\begin{aligned} \begin{bmatrix} -e^* \\ d^* \end{bmatrix} p &= \sigma t \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \end{bmatrix} + t^* \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} \\ [-e \ \ d] \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \end{bmatrix} &= \sigma t^* p \end{aligned} \tag{101}$$

and

$$\begin{aligned} \begin{bmatrix} -e^* \\ d^* \end{bmatrix} q &= \sigma t \begin{bmatrix} \hat{q}_1 \\ \hat{q}_2 \end{bmatrix} + t^* \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \\ [-e \ \ d] \begin{bmatrix} \hat{q}_1 \\ \hat{q}_2 \end{bmatrix} &= \sigma t^* q \end{aligned} \tag{102}$$

(For ease of notation we have dropped the subscript i.) Since $q = ap$ we have also

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} ap = \sigma t \begin{bmatrix} a\hat{p}_1 \\ a\hat{p}_2 \end{bmatrix} + t^* \begin{bmatrix} a\pi_1 \\ a\pi_2 \end{bmatrix} \tag{103}$$

On the other hand

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} ap = \begin{bmatrix} -e^* \\ d^* \end{bmatrix} q + \sigma t \begin{bmatrix} \hat{q}_1 \\ \hat{q}_2 \end{bmatrix} + t^* \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \tag{104}$$

Subtracting, we get

$$0 = \sigma t \begin{bmatrix} \hat{q}_1 - a\hat{p}_1 \\ \hat{q}_2 - a\hat{p}_2 \end{bmatrix} + t^* \begin{bmatrix} \rho_1 - a\pi_1 \\ \rho_2 - a\pi_2 \end{bmatrix} \tag{105}$$

Since t and t^* are coprime and $\deg(\hat{q}_i - a\hat{p}_i) < n$ it follows that $\hat{q}_i = a\hat{p}_i$ and $\rho_i = a\pi_i$. □

Schmidt pairs corresponding to the same singular value are non-unique; however, the ratio of the σ_i-Schmidt pair vectors is invariant. This result, due originally to Adamjan *et al.* (1971) in the scalar case, is given next in our context.

Lemma 8.3: *Let* $\left\{ \frac{p_i}{t}, \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} \right\}$ *and* $\left\{ \frac{q_i}{t}, \begin{bmatrix} \hat{q}_1^{(i)}/t^* \\ \hat{q}_2^{(i)}/t^* \end{bmatrix} \right\}$ *be two Schmidt pairs of the Hankel operator* $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$, *corresponding to the same singular value* σ_i. *Then*

$$\frac{p_i}{q_i} = \frac{\hat{p}_1^{(i)}}{\hat{q}_1^{(i)}} = \frac{\hat{p}_2^{(i)}}{\hat{q}_2^{(i)}} \tag{106}$$

i.e. the ratio $\begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i) \end{bmatrix}$ *is independent of the Schmidt pair.*

Proof: The proof follows from the previous Lemma. □

We are going to show that the functions we have constructed indeed form sets of orthogonal vectors.

Lemma 8.4: *Let*

$\left\{\frac{p_i}{t}, \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}\right\}$ *be the Schmidt pairs of* $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. *Then* $\left\{\frac{p_i}{t}\right\}$ *is an orthogonal set in* $\{\text{Ker } H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}\}^{\perp}$ *and* $\left\{\begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}\right\}$, *is an orthogonal set in* $\text{Im } H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$.

Proof: Of course, this follows from the fact that eigenvectors of a self adjoint operator that correspond to different eigenvalues are necessarily orthogonal. We give, in addition, a polynomial proof. Since $\{p_i/t\}$ are Schmidt vectors of H_{R^*} we know that these vectors form an orthogonal set by Remark 6.1 in Fuhrmann (1991). For $\left\{\begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}\right\}$ we proceed as follows. From the singular value/singular vector equations

$$\left.\begin{aligned} \begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i &= \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix} \\ -e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} &= \sigma_i t^* p_i \end{aligned}\right\} \tag{107}$$

we obtain, multiplying the first equation by $((\hat{p}_1^{(j)})^* \; (\hat{p}_2^{(j)})^*)$

$$\sigma_j t p_j^* p_i = \sigma_i t((\hat{p}_1^{(j)})^* \hat{p}_1^{(i)} + (\hat{p}_2^{(j)})^* \hat{p}_2^{(i)}) + t^*((\hat{p}_1^{(j)})^* \pi_1^{(i)} + (\hat{p}_2^{(j)})^* \pi_2^{(i)}) \tag{108}$$

Since $t \wedge t^* = 1$ t divides $(\hat{p}_1^{(j)})^* \pi_1^{(i)} + (\hat{p}_2^{(j)})^* \pi_2^{(i)}$. So there exist polynomials a_{ij} such that

$$(\hat{p}_1^{(j)})^* \pi_1^{(i)} + (\hat{p}_2^{(j)})^* \pi_2^{(i)} = a_{ij} t$$

We divide the equality (108) by $t^2 t^*$ to get

$$\sigma_j \frac{p_j^* p_i}{t^* t} = \sigma_i \frac{(\hat{p}_1^{(j)})^* \hat{p}_1^{(i)} + (\hat{p}_2^{(j)})^* \hat{p}_2^{(i)}}{t^* t} + \frac{a_{ij}}{t}$$

Integrating this equality over a semicircular contour, as shown in Fig. 5 and taking the limit as $R \to \infty$, noting that $\int_\gamma a_{ij}/t \, dz = 0$ by the stability of t, leads

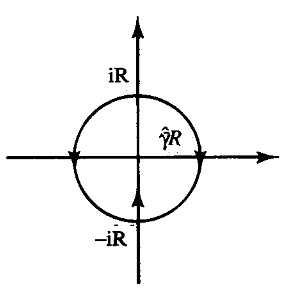

Figure 5.

to

$$0 = \sigma_j \int_{-\infty}^{\infty} \frac{p_j^* p_i}{t^* t} \mathrm{d}\tau = \sigma_i \int_{-\infty}^{\infty} \frac{(\hat{p}_1^{(j)})^* \hat{p}_1^{(i)} + (\hat{p}_2^{(j)})^* \hat{p}_2^{(i)}}{t^* t} \mathrm{d}\tau \qquad \square$$

We now derive a number of relationships which will be central to our later developments.

Theorem 8.1: *Let $g = e/d$ and let $\bar{N} = e/t$, $\bar{M} = d/t$ be the normalized coprime factors of g. Let ε_i, $1 \leq i \leq n$, be the signs and $\left\{\frac{p_i}{t}, \ \varepsilon_i \frac{p_i^*}{t^*}\right\}_{i=1}$ the (non-normalized) singular vectors associated with H_{R^*} as in Proposition 8.1.*
Then,

(1) *For $i = 1, \ldots, n$, we have*

$$d\pi_2^{(i)} - e\pi_1^{(i)} = (1 - \sigma_i^2) t p_i \tag{109}$$

(2) *The following relation holds*

$$(\pi_1^{(i)})^* \hat{p}_1^{(i)} + (\pi_2^{(i)})^* \hat{p}_2^{(i)} = 0 \tag{110}$$

(3) *The functions*

$$\frac{1}{(1 - \sigma_i^2)^{1/2}} \begin{bmatrix} \pi_1^{(i)}/p_i \\ \pi_2^{(i)}/p_i \end{bmatrix}$$

are all-pass functions.

(4) *For $i = 1, \ldots, n$, we have*

$$d^* \pi_1^{(i)} + e^* \pi_2^{(i)} = \varepsilon_i \sigma_i (1 - \sigma_i^2 t p_i^*)^{1/2} \tag{111}$$

Proof:

(1) Multiplying the first singular vector equation

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i = \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$$

on the left by $(-e \ \ d)$ we obtain, using $ee^* + dd^* = tt^*$, that

$$tt^* p_i = \sigma_i t (d\hat{p}_2^{(i)} - e\hat{p}_1^{(i)}) + t^* (d\pi_2^{(i)} - e\pi_1^{(i)})$$

Using the second singular vector equation

$$-e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} = \sigma_i t^* p_i$$

we obtain the result.

(2) In the proof of Lemma 8.1 we showed that

$$0 = \sigma t^* \{pp^* - (\hat{p}_1 \hat{p}_1^* + \hat{p}_2 \hat{p}_2^*)\} - t(\pi_1^* \hat{p}_1 + \pi_2^* \hat{p}_2)$$

and that $pp^* - (\hat{p}_1 \hat{p}_1^* + \hat{p}_2 \hat{p}_2^*) = 0$. This implies the result.

(3) Multiply the first singular vector equation,

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i = \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$$

on the left by $((\pi_1^{(i)})^* \ (\pi_2^{(i)})^*)$ and using

$$\pi_1^{(i)*}\hat{p}_1^{(i)} + \pi_2^{(i)*}\hat{p}_2^{(i)} = 0$$

we have

$$\begin{aligned}(-e^*\pi_1^{(i)*} + d^*\pi_2^{(i)*})p_i &= \sigma_i t(\pi_1^{(i)*}\hat{p}_1^{(i)} + \pi_2^{(i)*}\hat{p}_2^{(i)}) + t^*(\pi_1^{(i)*}\pi_1^{(i)} + \pi_2^{(i)*}\pi_2^{(i)}) \\ &= t^*(\pi_1^{(i)*}\pi_1^{(i)} + \pi_2^{(i)*}\pi_2^{(i)})\end{aligned}$$

Now, by (1) we have that $-e\pi_1^{(i)} + d\pi_2^{(i)} = (1 - \sigma_i^2)tp_i$ and so

$$(1 - \sigma_i^2)p_i p_i^* = \pi_1^{(i)*}\pi_1^{(i)} + \pi_2^{(i)*}\pi_2^{(i)} \tag{112}$$

i.e. $\dfrac{1}{(1 - \sigma_i^2)^{1/2}}\begin{bmatrix}\pi_1^{(i)}/p_i \\ \pi_2^{(i)}/p_i\end{bmatrix}$ is isometric or all-pass.

(4) From Part (1) we have, with $\bar{M} = d/t$ and $\bar{N} = e/t$, the H_+^∞ Bezout equation

$$\bar{M}V - \bar{N}U = -\frac{e}{t}\frac{\pi_1^{(1)}}{t(1 - \sigma_1^2)p_1} + \frac{d}{t}\frac{\pi_2^{(1)}}{(1 - \sigma_1^2)p_1} = 1 \tag{113}$$

The general solution of this equation is given by

$$\begin{bmatrix}U \\ V\end{bmatrix} = \frac{1}{1 - \sigma_1^2}\begin{bmatrix}\pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1\end{bmatrix} + \begin{bmatrix}d/t \\ e/t\end{bmatrix}h \tag{114}$$

with $h \in H_+^\infty$. Therefore

$$M^*U + N^*V = \frac{1}{1 - \sigma_1^2}\left(\frac{d^*\pi_1^{(1)} + e^*\pi_2^{(1)}}{t^*p_1}\right) + h \tag{115}$$

To get the LQG controller we choose $h \in H_+^\infty$ so that

$$\frac{r^*}{t^*} = \frac{1}{1 - \sigma_1^2}\left\{\frac{d^*\pi_1^{(1)} + e^*\pi_2^{(1)}}{t^*p_1}\right\} + h \tag{116}$$

We will show now that $h = \pi_1/p_1$ and $d^*\pi_1^{(1)} + e^*\pi_2^{(1)} = \lambda_1(1 - \sigma_1^2)tp_1^*$.

To this end we multiply the equation

$$\begin{bmatrix}-e^* \\ d^*\end{bmatrix}p_1 = \sigma_i t\begin{bmatrix}\hat{p}_1^{(1)} \\ \hat{p}_2^{(1)}\end{bmatrix} + t^*\begin{bmatrix}\pi_1^{(1)} \\ \pi_2^{(1)}\end{bmatrix} \tag{117}$$

by $(d^* \ e^*)$ to obtain

$$0 = \sigma_1 t(d^*\hat{p}_1^{(1)} + e^*\hat{p}_2^{(1)}) + t^*(d^*\pi_1^{(1)} + e^*\pi_2^{(1)}) \tag{118}$$

Thus, there exists a polynomial l_1 such that $(d^*\pi_1^{(1)} + e^*\pi_2^{(1)}) = l_1 t$.

If we substitute this into equation (116) we obtain

$$\frac{r^*}{t^*}\frac{p_1}{t} = \frac{1}{1 - \sigma_1^2}\frac{l_1}{t^*} + h\frac{p_1}{t}$$

Applying the projection P_- to this equality and recalling that $\left\{\dfrac{p_1}{t}, \dfrac{\varepsilon_1 p_1^*}{t2}\right\}$ is a Schmidt pair of H_{R^*} with singular value μ_1 we get

$$H_{\frac{r^*}{t^*}}\frac{p_1}{t} = \frac{1}{1 - \sigma_1^2}\frac{l_1}{t^*} = \lambda_1\frac{p_1^*}{t^*}$$

This implies the equality

$$l_1 = \lambda_1(1 - \sigma_1^2)p_1^*$$

This, substituted in (116), yields

$$\frac{r^*}{t^*} = \lambda_1 \frac{tp_1^*}{t^* p_1} + h$$

Comparing this with the equation

$$\frac{r^*}{t^*} = \lambda_1 \frac{tp_1^*}{t^* p_1} + \frac{\pi_1}{p_1}$$

leads to $h = \frac{\pi_1}{p_1}$.

We proceed now to prove equality (111) for all i. To this end we set

$$\begin{bmatrix} U_k \\ V_k \end{bmatrix} = \frac{1}{1 - \sigma_k^2} \begin{bmatrix} \pi_1^{(k)}/p_k \\ \pi_2^{(k)}/p_k \end{bmatrix} \tag{119}$$

Then $\begin{bmatrix} U_k \\ V_k) \end{bmatrix} \in H^\infty_{[k-1]}$, i.e. its unstable part has at most McMillan degree $k-1$, $k = 1, \ldots, n$ and $\bar{M}V_k - \bar{N}U_k = 1$. Of course $\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} \in H^\infty_+$. The general solution of the equation

$$\bar{M}X_k - \bar{N}Y_k = 1 \tag{120}$$

with $\begin{bmatrix} Y_k \\ X_k \end{bmatrix} \in H^\infty_{[k-1]}$, is given by

$$\begin{bmatrix} Y_k \\ X_k \end{bmatrix} = \begin{bmatrix} U_1 \\ V_1 \end{bmatrix} - \begin{bmatrix} M \\ N \end{bmatrix} q$$

with $q \in H^\infty_{[k-1]}$. We look now for the solution $\begin{bmatrix} Y_k \\ X_k \end{bmatrix}$ of the Bezout equation (120) with minimal L^∞ norm. This is a two-block problem that is easily reduced to a one-block problem, using a standard trick. We use the fact that $\begin{bmatrix} M^* & N^* \\ -\bar{N} & \bar{M} \end{bmatrix}$ is an all-pass function.

$$\inf_{q \in H^\infty_{[k-1]}} \left\| \begin{bmatrix} Y_k \\ X_k \end{bmatrix} \right\|_\infty = \inf_{q \in H^\infty_{[k-1]}} \left\| \begin{bmatrix} M^* & N^* \\ -\bar{N} & \bar{M} \end{bmatrix} \left[\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} - \begin{bmatrix} M \\ N \end{bmatrix} q \right] \right\|_\infty$$

$$= \inf_{q \in H^\infty_{[k-1]}} \left\| \begin{bmatrix} M^*U_1 + N^*V_1 - q \\ I \end{bmatrix} \right\|_\infty$$

$$= \{1 + \inf_{q \in H^\infty_{[k-1]}} \|M^*U_1 + N^*V_1 - q\|_\infty^2\}^{1/2} = \{1 + \mu_k^2\}^{1/2}$$

Now, by our proof for the case $i = 1$ we have

$$M^*U_1 + N^*V_1 = \frac{1}{1 - \sigma_1^2} \left(\frac{d^*\pi_1^{(1)} + e^*\pi_2^{(1)}}{t^* p_1} \right) = \lambda_1 \frac{tp_1^*}{t^* p_1}$$

However, since

$$\frac{r^*}{t^*} = \lambda_1 \frac{tp_1^*}{t^*p_1} + \frac{\pi_1}{p_1} = \lambda_k \frac{tp_k^*}{t^*p_k} + \frac{\pi_k}{p_k}$$

we have

$$M^*U_1 + N^*V_1 - q = \lambda_1 \frac{tp_1^*}{t^*p_1} - q = \lambda_k \frac{tp_k^*}{t^*p_k} + \frac{\pi_k}{p_k} - \frac{\pi_1}{p_1} - q \tag{121}$$

and

$$\inf_{q\in H^\infty_{[k-1]}} \|M^*U_1 + N^*V_1 - q\|_\infty = \inf_{q\in H^\infty_{[k-1]}} \left\|\frac{r^*}{t^*} - \frac{\pi_1}{p_1} - q\right\|_\infty$$
$$= \inf_{q'\in H^\infty_{[k-1]}} \left\|\frac{r^*}{t^*} - q'\right\|_\infty = \mu_k$$

Comparing this with (121), we have $q = \frac{\pi_k}{p_k} - \frac{\pi_1}{p_1}$ for the optimizing q. Therefore, the minimizing solution of (120) is given by

$$\begin{bmatrix} Y_k \\ X_k \end{bmatrix} = \begin{bmatrix} U_1 \\ V_1 \end{bmatrix} - \begin{bmatrix} M \\ N \end{bmatrix} q$$
$$= \frac{1}{1-\sigma_1^2}\begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix} - \begin{bmatrix} d/t \\ e/t \end{bmatrix}\left(\frac{\pi_k}{p_k} - \frac{\pi_1}{p_1}\right)$$

On the other hand we claim that

$$\begin{bmatrix} Y_k \\ X_k \end{bmatrix} = \frac{1}{1-\sigma_k^2}\begin{bmatrix} \pi_1^{(k)}/p_k \\ \pi_2^{(k)}/p_k \end{bmatrix}$$

Indeed, by AAK theory, p_k has $k-1$ unstable zeros. So

$$\begin{bmatrix} \pi_1^{(k)}/p_k \\ \pi_2^{(k)}/p_k \end{bmatrix} \in H^\infty_{[k-1]}.$$

By (109) it solves the Bezout equation $\bar{M}X_k - \bar{N}Y_k = 1$. Finally, by Part (3),

$$\left\|\frac{1}{1-\sigma_k^2}\begin{bmatrix} \pi_1^{(k)}/p_k \\ \pi_2^{(k)}/p_k \end{bmatrix}\right\|_\infty = \frac{1}{(1-\sigma_k^2)^{1/2}} = (1+\mu_k^2)^{1/2}$$

Thus, we have

$$\begin{bmatrix} Y_k \\ X_k \end{bmatrix} = \frac{1}{1-\sigma_k^2}\begin{bmatrix} \pi_1^{(k)}/p_k \\ \pi_2^{(k)}/p_k \end{bmatrix} = \frac{1}{1-\sigma_1^2}\begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix} - \begin{bmatrix} d/t \\ e/t \end{bmatrix}\left(\frac{\pi_k}{p_k} - \frac{\pi_1}{p_1}\right)$$

In turn, this implies

$$M^*Y_k + N^*X_k = \frac{1}{1-\sigma_k^2}\left(\frac{d^*\pi_1^{(k)} + e^*\pi_2^{(k)}}{t^*p_k}\right)$$
$$= \frac{1}{1-\sigma_1^2}\left(\frac{d^*\pi_1^{(1)} + e^*\pi_2^{(1)}}{t^*p_1}\right) - \left(\frac{\pi_k}{p_k} - \frac{\pi_1}{p_1}\right)$$
$$= \lambda_1\frac{tp_1^*}{t^*p_1} + \frac{\pi_1}{p_1} - \frac{\pi_k}{p_k} = \lambda_k \frac{tp_k^*}{t^*p_k}$$

From this the required equality

$$d^*\pi_1^{(k)} + e^*\pi_2^{(k)} = \lambda_k(1 - \sigma_k^2)tp_k^* = \varepsilon_k\sigma_k(1 - \sigma_k^2)^{1/2}tp_k^*$$

clearly follows. □

As a corollary we can obtain a closed form representation for the Schmidt vectors of $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$.

Corollary 8.1: *Let* $\left\{\frac{p_i}{t}, \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}\right\}$ *be a minimal degree* σ_i*-Schmidt pair of* $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. *Then the following relations hold true:*

$$\left.\begin{aligned} \hat{p}_1^{(i)} &= \frac{1}{t}\{-\sigma_i e^* p_i - \varepsilon_i(1 - \sigma_i^2)^{1/2} dp_i^*\} \\ \hat{p}_2^{(i)} &= \frac{1}{t}\{\sigma_i d^* p_i - \varepsilon_i(1 - \sigma_i^2)^{1/2} ep_i^*\} \end{aligned}\right\} \tag{122}$$

Proof: Multiplying the first singular vector/singular value equation

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i = \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$$

on the left by $(d^* \;\; e^*)$ we get

$$0 = \sigma_i t(d^*\hat{p}_1^{(1)} + e^*\hat{p}_2^{(1)}) + t^*(d^*\pi_1^{(1)} + e^*\pi_2^{(1)})$$

Using that $d^*\pi_1^{(i)} + e^*\pi_2^{(i)} = \lambda_i(1 - \sigma_i^2)tp_i^*$, we obtain that

$$\sigma_i t(d^*\hat{p}_1^{(1)} + e^*\hat{p}_2^{(1)}) = -\lambda_i(1 - \sigma_i^2)^{1/2} t^* p_i^*$$

or

$$d^*\hat{p}_1^{(1)} + e^*\hat{p}_2^{(1)} = -\varepsilon_i(1 - \sigma_i^2)^{1/2} t^* p_i^* \tag{123}$$

This equation and the second singular vector equation

$$-e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} = \sigma_i t^* p_i$$

we can write as

$$\begin{bmatrix} -e & d \\ d^* & e^* \end{bmatrix} \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} = \begin{bmatrix} \sigma_i t^* p_i \\ -\varepsilon_i(1 - \sigma_i^2)^{1/2} t^* p_i^* \end{bmatrix} \tag{124}$$

Since

$$\begin{bmatrix} -e & d \\ d^* & e^* \end{bmatrix} \begin{bmatrix} -e^* & d \\ d^* & e \end{bmatrix} = tt^* \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and hence also

$$\begin{bmatrix} -e^* & d \\ d^* & e \end{bmatrix} \begin{bmatrix} -e & d \\ d^* & e^* \end{bmatrix} = tt^* \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

it follows from (124) that

$$tt^*\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix} = \begin{bmatrix}-e^* & d\\ d^* & e\end{bmatrix}\begin{bmatrix}\sigma_i t^* p_i\\ -\varepsilon_i(1-\sigma_i^2)^{1/2} t^* p_i^*\end{bmatrix} \tag{125}$$

This proves (122). □

The following proposition provides a generalization of the scalar results in Proposition 8.1 to the case of normalized coprime factorizations.

Proposition 8.2: *Let* $\sigma_1 \geqslant \ldots \geqslant \sigma_n$ *be the singular values of the Hankel operator* $H_{\begin{bmatrix}-\bar{N}^*\\ \bar{M}^*\end{bmatrix}}$ *and let* $p_i, \begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}, \begin{bmatrix}\pi_1^{(i)}\\ \pi_2^{(i)}\end{bmatrix}$ *be defined by the s.v. equations (94). Then there exist polynomials* $\alpha_1^{(i,j)}, \alpha_2^{(i,j)}$, $1 \leqslant i, j \leqslant n$, *of degree less then or equal to* $n-2$ *such that*

(1) $$\sigma_i\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}p_j - \sigma_j\begin{bmatrix}\hat{p}_1^{(j)}\\ \hat{p}_2^{(j)}\end{bmatrix}p_i = \sigma_i t^*\begin{bmatrix}\alpha_1^{(i,j)}\\ \alpha_2^{(i,j)}\end{bmatrix} \tag{126}$$

(2) $$\begin{bmatrix}\pi_1^{(j)}\\ \pi_2^{(j)}\end{bmatrix}p_i - \begin{bmatrix}\pi_1^{(i)}\\ \pi_2^{(i)}\end{bmatrix}p_j = \sigma_i t\begin{bmatrix}\alpha_1^{(i,j)}\\ \alpha_2^{(i,j)}\end{bmatrix} \tag{127}$$

(3) *if i, j are such that* $\sigma_i \neq \sigma_j$, *then* $\begin{bmatrix}\alpha_1^{(i,j)}\\ \alpha_2^{(i,j)}\end{bmatrix}$ *is non-zero*

Proof:

(1) From the singular value equations

$$\begin{bmatrix}-e^*\\ d^*\end{bmatrix}p_i = \sigma_i t\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix} + t^*\begin{bmatrix}\pi_1^{(i)}\\ \pi_2^{(i)}\end{bmatrix} \tag{128}$$

and

$$\begin{bmatrix}-e^*\\ d^*\end{bmatrix}p_j = \sigma_n t\begin{bmatrix}\hat{p}_1^{(j)}\\ \hat{p}_2^{(j)}\end{bmatrix} + t^*\begin{bmatrix}\pi_1^{(j)}\\ \pi_2^{(j)}\end{bmatrix} \tag{129}$$

we get, by eliminating the left-side terms, that

$$0 = t\left\{\sigma_i\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}p_j - \sigma_j\begin{bmatrix}\hat{p}_1^{(j)}\\ \hat{p}_2^{(j)}\end{bmatrix}p_i\right\} + t^*\left\{\begin{bmatrix}\pi_1^{(i)}\\ \pi_2^{(i)}\end{bmatrix}p_j - \begin{bmatrix}\pi_1^{(j)}\\ \pi_2^{(j)}\end{bmatrix}p_i\right\} \tag{130}$$

Since t and t^* are coprime, there exist $\alpha_1^{(i,j)}, \alpha_2^{(i,j)}$ such that (126) holds.

(2) Substituting (126) into the above equation yields

$$0 = \sigma_i tt^*\begin{bmatrix}\alpha_1^{(i,j)}\\ \alpha_2^{(i,j)}\end{bmatrix} + t^*\left\{\begin{bmatrix}\pi_1^{(i)}\\ \pi_2^{(i)}\end{bmatrix}p_j - \begin{bmatrix}\pi_1^{(j)}\\ \pi_2^{(j)}\end{bmatrix}p_i\right\}$$

which is equivalent to (127).

(3) Let i, j be such that $\sigma_i \neq \sigma_j$ and assume that $\alpha_1^{(i,j)} = \alpha_2^{(i,j)} = 0$. Since

$$\sigma_i\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}p_j - \sigma_j\begin{bmatrix}\hat{p}_1^{(j)}\\ \hat{p}_2^{(j)}\end{bmatrix}p_i = \sigma_i t^*\begin{bmatrix}\alpha_1^{(i,j)}\\ \alpha_2^{(i,j)}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

we therefore have that

$$\begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i \end{bmatrix} = \frac{\sigma_j}{\sigma_i} \begin{bmatrix} \hat{p}_1^{(j)}/p_j \\ \hat{p}_2^{(j)}/p_j \end{bmatrix}$$

But $\begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i \end{bmatrix}$ and $\begin{bmatrix} \hat{p}_1^{(j)}/p_j \\ \hat{p}_2^{(j)}/p_j \end{bmatrix}$ are all-pass which leads to a contradiction since $\dfrac{\sigma_j}{\sigma_i} \neq 1$. □

9. Hankel norm approximation

We now come to apply some results of the previous section to the case of Hankel norm approximation. As in the case of scalar functions the study of Hankel norm approximation is closely linked to the study of the singular vectors.

The following theorem summarizes the results on the Hankel norm approximation of scalar functions that are necessary for our later development. Again, we will state all the results immediately in terms of the transfer function $R^* = r^*/t^*$.

For the sake of simplicity of presentation we will assume that the smallest singular value has multiplicity one.

Theorem 9.1: *Let $r^*/t^* \in H_-^\infty$ be a scalar, strictly proper, transfer function, with r and t coprime polynomials and t is monic of degree n. Assume that $\mu_1 \geqslant \mu_2 \ldots \geqslant \mu_{n-1} > \mu_n > 0$ are the singular values of H_{r^*/t^*}. Then*

(1) *There exist non-zero polynomials α_i, $i = 1, \ldots, n-1$ of degree less than or equal to $n-2$, and signs ε_i, $i = 1, \ldots, n-1$, such that*

$$\lambda_i p_n^* p_i - \lambda_n p_i^* p_n = \lambda_i t \alpha_i \tag{131}$$

where $\lambda_i = \varepsilon_i \mu_i$.

(2) *The polynomial p_n has degree $n-1$ and has its zeros in the right-half plane, i.e. $\pi_n/p_n \in H_-^\infty$, and H_{π_n/p_n} has rank $n-1$.*

(3) $$\mu_n = \inf\left\{\left\| H_{\frac{r^*}{t^*}} - H_q \right\|;\ H_q \text{ has rank at most } n-1\right\} = \left\| H_{\frac{r^*}{t^*}} - H_{\frac{\pi_n}{p_n}} \right\|$$

(4) $$\mu_n = \inf\left\{\left\| \frac{r^*}{t^*} - q \right\|_\infty;\ q \in L^\infty \text{ and has at most } n-1 \text{ poles in the open left-half plane}\right\} = \left\| \frac{r^*}{t^*} - \frac{\pi_n}{p_n} \right\|_\infty$$

(5) $H_{\frac{\pi_n}{p_n}}$ *has the singular values* μ_i, $i = 1, \ldots, n-1$. *The Schmidt pairs of* $H_{\frac{\pi_n}{p_n}}$ *are given by*

$$\left\{\frac{\alpha_i}{p_n^*}, \varepsilon_i \frac{\alpha_i^*}{p_n}\right\}, \; i = 1, \ldots, n-1$$

(6) *There exist polynomials* ζ_i, $1 \leq i \leq n$, *such that*

$$\frac{\pi_n}{p_n}\frac{\alpha_i}{p_n^*} = \lambda_i \frac{\alpha_i^*}{p_n} + \frac{\zeta_i}{p_n^*}, \tag{132}$$

with $\lambda_i = \varepsilon_i \mu_i$, $1 \leq i \leq n-1$.

(7) *We have*

$$\frac{\alpha_i}{p_n^*} = P_{\left\{\frac{p_n}{p_n^*}H_+^2\right\}^{\perp}} \frac{p_i}{t} = P_{X^{p_n^*}} \frac{p_i}{t} \tag{133}$$

i.e. the singular vectors of H_{π_n/p_n} *are projections of the singular vectors of* H_{r^*/t^*} *onto* $X^{p_n^*}$ *the orthogonal complement of*

$$\operatorname{Ker} H_{\frac{\pi_n}{p_n}} = \frac{p_n}{p_n^*} H_+^2$$

For $i = 1, \ldots, n-1$,

$$\left\|\frac{\alpha_i}{p_n^*}\right\|^2 = \left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\left\|\frac{p_i}{t}\right\|^2 \tag{134}$$

Proof:

(1) This, with an obvious change of notation, is a special case of (84).

(2), (3), (4) and (5). This is the content of Theorem 5.1 in Fuhrmann (1991).

(6) This follows from Part (5).

(7) Rewrite (131) in the form

$$\frac{p_i}{t} = \frac{\alpha_i}{p_n^*} + \frac{\lambda_n}{\lambda_i}\frac{p_n}{p_n^*}\frac{p_i^*}{t} \tag{135}$$

This is the orthogonal decomposition of p_i/t relative to the direct sum

$$H_+^2 = X^{p_n^*} \oplus \frac{p_n}{p_n^*} H_+^2$$

So (133) follows.

(8) From (135) we get, by orthogonality,

$$\left\|\frac{\alpha_i}{p_n^*}\right\|^2 + \frac{\mu_n^2}{\mu_i^2}\left\|\frac{p_i^*}{t}\right\|^2 = \left\|\frac{p_i}{t}\right\|^2$$

or, as

$$\left\|\frac{p_i^*}{t}\right\| = \left\|\frac{p_i}{t}\right\|$$

$$\left\|\frac{\alpha_i}{p_n^*}\right\|^2 = \left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\left\|\frac{p_i}{t}\right\|^2 \tag{136}$$

□

As in the case of scalar functions, we can characterize the Schmidt vectors of the Hankel norm approximant as projections of the Schmidt vectors of the original Hankel operator onto the image of the orthogonal complement of the kernel of the Hankel norm approximant.

The function discussed in the following lemma will be shown to be related to the solution to a Hankel norm approximation problem in the subsequent Theorem.

Lemma 9.1:

(1) $$\frac{1}{(1-\sigma_n^2)^{1/2}}\begin{bmatrix}\pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n\end{bmatrix}$$

is inner in H_-^∞ *and has McMillan at most* $n-1$.

(2) $$\frac{1}{(1-\sigma_n^2)}[\pi_1^{(n)*}/p_n^* \;\; \pi_2^{(n)*}/p_n^*]$$

is a NLCF of $\left(\dfrac{\pi_1^{(n)*}}{\pi_2^{(n)*}}\right)$

Proof:

(1) This follows from Theorem 8.1 and the fact that p_n has its zeros in the open right-half plane.

(2) That the factors are coprime follows from Theorem 8.1. Indeed, assume to the contrary that τ is a non-trivial common factor of $\pi_1^{(n)}$ and $\pi_2^{(n)}$. From (109) and (111) it follows that $\tau | tp_n$ and $\tau | tp_n^*$. This implies that τ and t have a common factor which is necessarily stable. Now from

$$\begin{bmatrix}-e^* \\ d^*\end{bmatrix} p_n = \sigma_n t\begin{bmatrix}\hat{p}_1^{(n)} \\ \hat{p}_2^{(n)}\end{bmatrix} + t^*\begin{bmatrix}\pi_1^{(n)} \\ \pi_2^{(n)}\end{bmatrix}$$

it follows that $\tau \left| \begin{bmatrix}-e^* \\ d^*\end{bmatrix} p_n\right.$. Since τ is stable it is necessarily a common factor of e^* and d^*, which contradicts the assumption that e and d are coprime.

Incidentally, the coprimeness also follows from Lemma 3.3 and Theorem 9.3; this will be proved, independently, later.

The normalization is proved in Part 1. □

Next we prove that $H_{\begin{bmatrix}\pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n\end{bmatrix}}$ is the best Hankel norm approximant of $H_{\begin{bmatrix}-\tilde{N}^* \\ \tilde{M}^*\end{bmatrix}}$ of rank less than or equal to $n-1$.

Theorem 9.2: *Assume that* $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$ *has singular values* $\sigma_1 \geq \dots \geq \sigma_{n-1} > \sigma_n > 0$. *We have*

(1) *The function* $\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$ *has McMillan degree* $n-1$.

(2) $$\sigma_n = \inf\left\{ \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right\|_\infty ; \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \in L^\infty \right.$$

$$\left. \text{and the antistable part of } (q_1^{\mathrm{T}}\ q_2^{\mathrm{T}})^{\mathrm{T}} \text{ has McMillan degree at most } n-1 \right\}$$

$$= \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \right\|_\infty$$

(3) $$\sigma_n = \inf\left\{ \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - H_{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} \right\| ; \begin{bmatrix} q_1 \\ q_2 \in L^\infty \end{bmatrix} \right.$$

$$\left. \text{and the rank of } H_{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} \text{ is at most } n-1 \right\}$$

$$= \left\| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} - H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \right\|$$

Proof:

(2) We know that

$$\sigma_n \leq \left\| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} - H_{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} \right\| \leq \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right\| \leq \Big\|_\infty$$

if the anti-stable part of $(q_1^{\mathrm{T}}\ q_2^{\mathrm{T}})^{\mathrm{T}}$ has at most McMillan degree $n-1$. The first inequality follows since the error incurred by approximating an operator A of rank n by an operator B of rank $n-1$ is at least $\sigma_n(A)$, i.e. the nth singular value of A.

Specializing (95) to the case of the last singular value we have,

$$\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} = \sigma_n \frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix}$$

As $\frac{t}{t^*}\begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix}$ is all-pass and using the fact that $\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$ is in H_-^∞ with degree at most $n-1$ we have proved the result since

$$\sigma_n \leq \inf\left\{ \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right\|_\infty ; \text{the anti-stable part of} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right.$$

$$\left. \text{has McMillan degree at most } n-1 \right\}$$

$$\leq \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \right\|_\infty$$

$$= \left\| \sigma_n \frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \right\|_\infty$$

$$= \sigma_n$$

which proves the required equality.

(3) This follows since

$$\sigma_n \leq \inf \{ \| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} - H_{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} \|; \text{ the rank of } H_{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} \text{ is at most } n-1 \}$$

$$\leq \| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} - H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \| \leq \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \right\|_\infty$$

$$= \sigma_n$$

which implies the required equalities, since

$$\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$$

is anti-stable and has McMillan degree at most $n-1$, and therefore

$$H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$$

has rank at most $n-1$.

(1) Assume that

$$H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$$

has rank less than or equal to $n-2$, then

$$\sigma_{n-1} = \inf \{ \| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} - A \|; A \text{ an operator with rank}(A) \leq n-2 \}$$

$$\leq \inf \{ \| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} - H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \|$$

$$= \sigma_n < \sigma_{n-1}$$

which is a contradiction. Since the Hankel operator has at most rank $n-1$ this proves the claim. □

We now come to the main theorem of this section in which the Hankel norm approximant is analysed in some detail. Before stating the theorem we need to prove two lemmas. We will make use of the following simple result concerning computation of singular values.

Lemma 9.2: *Let H_1, H_2 be two Hilbert spaces and let $T: H_1 \to H_2$ be a bounded operator. Let $\{\phi_i\}$ and $\{\psi_i\}$ be complete orthogonal sets in H_1 and H_2 respectively. Assume*

$$T\phi_i = \psi_i$$

Then the singular values of T are given by

$$\rho_i = \frac{\|\psi_i\|}{\|\phi_i\|} \tag{137}$$

and

$$T^*\psi_i = \rho_i^2\phi_i \tag{138}$$

Proof: Assume $i \neq j$. Then

$$0 = (\psi_i, \psi_j) = (T\phi_i, \psi_j) = (\phi_i, T^*\psi_j) \tag{139}$$

By expanding $T^*\psi_j = \sum \gamma_{ij}\phi_i$, we see that the previous equality implies that $\gamma_{ij} = 0$ for $i \neq j$. So $T^*\psi_i = \gamma_{ii}\phi_i$. Now

$$(T\phi_i, \psi_i)(\psi_i, \psi_i) = (\phi_i, T^*\psi_i) = \overline{\gamma_{ii}}(\phi_i, \phi_i)$$

or

$$\rho_i^2 = \overline{\gamma_{ii}} = \gamma_{ii} = \frac{\|\psi_i\|^2}{\|\phi_i\|^2} > 0 \tag{140}$$

Thus, we get the two equations

$$\begin{aligned} T\phi_i &= \psi_i \\ T^*\psi_i &= \frac{\|\psi_i\|^2}{\|\phi_i\|^2}\phi_i \end{aligned} \tag{141}$$

These are equivalent to the singular value equations for T and hence the singular values are

$$\rho_i = \frac{\|\psi_i\|}{\|\phi_i\|}$$

□

In the following lemma the numerators of the Schmidt vectors of the $(n-1)$-order Hankel norm approximant are constructed.

Lemma 9.3: *Let $\sigma_1 \geq \cdots \geq \sigma_{n-1} > \sigma_n > 0$ be the singular values of the Hankel operator $H\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$ and let p_i, $\begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix}$, $\begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$ be defined by the singular value/singular vector equations (94). Then*

(1) *there exist polynomials $\alpha_1^{(i)}$, $\alpha_2^{(i)}$ $i = 1, \ldots, n-1$, with $\begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix}$ non-zero, of degree less than or equal to $n-2$, such that*

$$\sigma_i\begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} p_n - \sigma_n\begin{bmatrix} \hat{p}_1^{(n)} \\ \hat{p}_2^{(n)} \end{bmatrix} p_i = \sigma_i t^*\begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} \tag{142}$$

$$(2) \qquad \begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} p_i - \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix} p_n = \sigma_i t \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} \tag{143}$$

$$(3) \qquad \left\| \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \right\|^2 = \left(1 - \frac{\sigma_n^2}{\sigma_i^2}\right) \left\| \frac{p_i}{t^*} \right\|^2 \tag{144}$$

$$(4) \qquad \frac{\left\| \begin{pmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{pmatrix} \right\|}{\left\| \dfrac{\alpha_i}{p_n^*} \right\|} = (1 - \sigma_n^2)^{1/2} \tag{145}$$

Proof:

(1), (2) The proof follows from Proposition 8.2 by setting $\alpha_k^{(i)} = \alpha_k^{(i,n)}$ for all $1 \leqslant i \leqslant n$.

(3) The equation

$$\sigma_i \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} p_n - \sigma_n \begin{bmatrix} \hat{p}_1^{(n)} \\ \hat{p}_2^{(n)} \end{bmatrix} p_i = \sigma_i t^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} \tag{146}$$

can be rewritten as

$$\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} + \frac{\sigma_n}{\sigma_i} \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \frac{p_i}{t^*} = \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} = \begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i \end{bmatrix} \begin{bmatrix} p_i \\ t^* \end{bmatrix} \tag{147}$$

We claim that $\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}$ and $\begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \dfrac{p_i}{t^*}$ are orthogonal in H_-^2. Now, using the fact that $\begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix}$ is inner in H_-^∞,

$$\left(\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}, \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \frac{p_i}{t^*} \right) = \left(\frac{(\hat{p}_1^{(n)})^* \alpha_1^{(i)} + (\hat{p}_2^{(n)})^* \alpha_2^{(i)}}{p_n p_n^*}, \frac{p_i}{t^*} \right)$$

For this inner product to vanish it suffices to show that $p_n | (\hat{p}_1^{(n)})^* \alpha_1^{(i)} + (\hat{p}_2^{(n)})^* \alpha_2^{(i)}$. To this end we multiply (146) by $((\hat{p}_1^{(n)})^* (\hat{p}_2^{(n)})^*)$. This leads to

$$\sigma_i ((\hat{p}_1^{(n)})^* \hat{p}_1^{(i)} + (\hat{p}_2^{(n)})^* \hat{p}_2^{(i)}) p_n = \sigma_n p_n p_n^* p_i + \sigma_i t^* ((\hat{p}_1^{(n)})^* \alpha_1^{(i)} + (\hat{p}_2^{(n)})^* \alpha_2^{(i)}) \tag{148}$$

Now we show that p_n and t^* are coprime. Indeed, from the singular value equation

$$r^* p_n = \lambda_n t p_n^* + t^* \pi_n$$

it is clear that if p_n and t^* have a non-trivial, necessarily anti-stable, common factor $p_n \wedge t^*$ then it has to divide also tp_n^*. However, tp_n^* is stable so this is impossible. Thus, from (148) the required division relation follows.

We proceed, using orthogonality, to get

$$\left\| \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \right\|^2 + \frac{\sigma_n^2}{\sigma_i^2} \left\| \frac{p_i}{t^*} \right\|^2 = \left\| \frac{p_i}{t^*} \right\|^2$$

or

$$\left\| \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \right\|^2 = \left(1 - \frac{\sigma_n^2}{\sigma_i^2}\right) \left\| \frac{p_i}{t^*} \right\|^2 \tag{149}$$

(4) By Theorem 6.2, we have $\mu_i^2 = \sigma_i^2/1 - \sigma_i^2$ so

$$\left(1 - \frac{\mu_n^2}{\mu_i^2}\right) = 1 - \frac{\sigma_n^2}{1 - \sigma_n^2} \frac{1 - \sigma_i^2}{\sigma_i^2} = \frac{(1 - \sigma_n^2)\sigma_i^2 - \sigma_n^2(1 - \sigma_i^2)}{\sigma_i^2(1 - \sigma_n^2)}$$

$$= \frac{\sigma_i^2 - \sigma_n^2}{\sigma_i^2(1 - \sigma_n^2)} = \frac{1}{1 - \sigma_n^2}\left(1 - \frac{\sigma_n^2}{\sigma_i^2}\right)$$

Hence

$$\frac{\left\| \begin{pmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{pmatrix} \right\|^2}{\left\| \dfrac{\alpha_i}{p_n^*} \right\|^2} = \frac{\left(1 - \dfrac{\sigma_n^2}{\sigma_i^2}\right)}{\left(1 - \dfrac{\mu_n^2}{\mu_i^2}\right)} = 1 - \sigma_n^2 \tag{150}$$

which implies (145). □

In the following theorem we characterize the singular values and singular vectors of the Hankel norm approximant. Contrary to Lemma 9.1 in Glover (1984) or Theorem 5.1 in Fuhrmann (1991) the Hankel singular values of $H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$ are not $\sigma_1 \geqslant \cdots \geqslant \sigma_{n-1}$ but they have to be slightly modified. An independent proof of this fact can be found in Sefton (1991). We prefer to state the result with the modification occurring in the symbol.

Theorem 9.3: *Assume $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_{n-1} > \sigma_n > 0$ to be the singular values of the Hankel operator $H_{\begin{bmatrix} -N^* \\ M^* \end{bmatrix}}$. Let*

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_n = \sigma_n t \begin{bmatrix} \hat{p}_1^{(n)} \\ \hat{p}_2^{(n)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix}$$

$$-e\hat{p}_1^{(n)} + d\hat{p}_2^{(n)} = \sigma_n t^* p_n$$

be the singular value/singular vector equations corresponding to the nth-singular value. Then

(1) *The Hankel operator $\dfrac{1}{(1 - \sigma_n^2)^{1/2}} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$ has singular values $\sigma_1 \geqslant \ldots \geqslant \sigma_{n-1}$, and the σ_i-Schmidt pairs are given by*

$$\left\{ \frac{\alpha_i}{p_n^*}, \frac{1}{(1 - \sigma_n^2)^{1/2}} \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \right\}$$

where the α_i are defined by Theorem 9.1.7 and $\alpha_1^{(i)}$, $\alpha_2^{(i)}$ by Lemma 9.3.

(2) *We have*

$$\frac{\alpha_i}{p_n^*} = P_{X^{p_n^*}} \frac{p_i}{t} \tag{151}$$

i.e. the singular vectors of $\frac{1}{(1-\sigma_n^2)^{.0242}} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$ *are the projections of the singular vectors of*

$$H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}} \text{ on } \left\{ \operatorname{Ker} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \right\}$$

(3) *We have*

$$\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} = P_{\{\operatorname{Ker} \hat{H}(\pi_1^{(n)*}/p_n^* \pi_2^{(n)*}/p_n^*)\}^\perp} \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} \tag{152}$$

i.e. the $\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}$ *are the projections of the singular vectors of* $H_{(-\bar{N}\ \bar{M})}$ *onto*

$$\{\operatorname{Ker} \hat{H}_{(\pi_1^{(n)*}/p_n^* \pi_2^{(n)*}/p_n^*)}\}^\perp$$

Proof:

(1) Let $\left\{\frac{p_i}{t} : i = 1, \ldots, n\right\}$ be the joint singular vectors of H_{r^*/t^*}, and $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. By AAK theory p_n is anti-stable. First we will show that for $1 \leqslant i \leqslant n-1$,

$$H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{\alpha_i}{p_n^*} = \sigma \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \tag{153}$$

To this end we note that the equation

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_n = \sigma_n t \begin{bmatrix} \hat{p}_1^{(n)} \\ \hat{p}_2^{(n)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix}$$

can be rewritten as

$$\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} = \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \sigma_n \frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \tag{154}$$

Since p_n is anti-stable, we have $\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \in H_-^\infty$.

Now $\operatorname{Ker} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} = \frac{p_n}{p_n^*} H_+^2$. From (135), i.e. from

$$\frac{p_i}{t} = \frac{\alpha_i}{p_n^*} + \frac{\lambda_n}{\lambda_i} \frac{p_n}{p_n^*} \frac{p_i^*}{t} \tag{155}$$

it follows that

$$\frac{p_i}{t} - \frac{\alpha_i}{p_n^*} \in \operatorname{Ker} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \tag{156}$$

Hence

$$H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{\alpha_i}{p_n^*} = H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{p_i}{t} \tag{157}$$

We compute, using (154), (157) and (142),

$$\begin{aligned}
H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{\alpha_i}{p_n^*} &= H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{p_i}{t} = P_-\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \frac{p_i}{t} \\
&= P_-\left\{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \sigma_n \frac{t}{t^*}\begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix}\right\}\frac{p_i}{t} \\
&= P_-\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} \frac{p_i}{t} - \sigma_n \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \frac{p_i}{t^*} \\
&= P_-\left\{\sigma_i \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} - \sigma_n \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \frac{p_i}{t^*}\right\} \\
&= P_- \frac{1}{p_n t^*}\left\{\sigma_i \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} p_n - \sigma_n \begin{bmatrix} \hat{p}_1^{(n)} \\ \hat{p}_2^{(n)} \end{bmatrix} p_i\right\} \\
&= P_-\left\{\frac{1}{p_n t^*}\sigma_i t^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix}\right\} = \sigma_i \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}
\end{aligned} \tag{158}$$

i.e.

$$H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{\alpha_i}{p_n^*} = \sigma_i \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \tag{159}$$

This can be rewritten as

$$\frac{1}{(1-\sigma_n^2)^{1/2}} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} \frac{\alpha_i}{p_n^*} = \sigma_i \frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \tag{160}$$

By Equation (145),

$$\left\|\frac{\alpha_i}{p_n^*}\right\| = \frac{1}{(1-\sigma_n^2)^{1/2}} \left\|\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}\right\|$$

Now, from (147) and the orthogonality of the p_i/t^* and the

$$\begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}$$

proved in Lemma 8.4, it follows that the

$$\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}$$

are also orthogonal. We can therefore apply Lemma 9.2 to deduce that

$$\left\{ \frac{\alpha_i}{p_n^*}, \frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \right\}$$

are the σ_i-Schmidt pairs of

$$\frac{1}{(1-\sigma_n^2)^{1/2}} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$$

(2) This is just a restatement of Theorem 9.1.7 coupled with the fact that the $\{\alpha_i/p_n^*\}$ are the joint singular vectors of H_{π_n/p_n} and $H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$ and

$$\operatorname{Ker} H_{\pi_n/p_n} = \operatorname{Ker} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} = \frac{p_n}{p_n^*} H_+^2$$

(3) It is clear that

$$\operatorname{Ker} \hat{H}_{\left(\frac{\pi_1^{(n)*}}{p_n^*} \; \frac{\pi_2^{(n)*}}{p_n^*}\right)} \supseteq \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} H_-^2$$

for, let $h \in H_-^2$, then we have

$$P_+ \left(\frac{\pi_1^{(n)*}}{p_n^*} \; \frac{\pi_2^{(n)*}}{p_n^*} \right) \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} h = 0$$

as $(\pi_1^{(i)})^* \hat{p}_1^{(i)} + (\pi_2^{(i)})^* \hat{p}_2^{(i)} = 0$, by (110). On the other hand, by (159),

$$\begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \in \operatorname{Im} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}} = \left\{ \operatorname{Ker} \hat{H}_{\left(\frac{\pi_1^{(n)*}}{p_n^*} \; \frac{\pi_2^{(n)*}}{p_n^*}\right)} \right\}^{\perp}$$

Now rewrite (142), i.e.

$$\sigma_i \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} p_n - \sigma_n \begin{bmatrix} \hat{p}_1^{(n)} \\ \hat{p}_2^{(n)} \end{bmatrix} p_i = \sigma_i t^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} \tag{161}$$

as

$$\begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} = \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} + \frac{\sigma_n}{\sigma_i} \begin{bmatrix} \hat{p}_1^{(n)}/p_n \\ \hat{p}_2^{(n)}/p_n \end{bmatrix} \frac{p_i}{t^*} \tag{162}$$

This is an orthogonal decomposition, and so (152) follows. □

Corollary 9.1: *The Schmidt vectors in H_+^2 of the $n-1$ degree Hankel norm approximant of $H_{\frac{r^*}{t^*}}$, i.e. of $H_{\frac{\pi_n}{p_n}}$, coincide with the Schmidt vectors in H_+^2*

of $H_{\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}}$, *i.e. with those of the* $n-1$ *degree Hankel norm approximant* $H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$.

In the following proposition further results are collected on the Schmidt vectors of the rank $n-1$ Hankel norm approximant.

Proposition 9.1:

(1) *There exist polynomials* $\zeta_1^{(i)}$, $\zeta_2^{(i)}$ *of degree* $\leqslant n-2$ *such that*

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} \alpha_i = \sigma_i p_n^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} + p_n \begin{bmatrix} \zeta_1^{(i)} \\ \zeta_2^{(i)} \end{bmatrix} \tag{163}$$

(2) *We have*

$$(\pi_1^{(n)})^* \alpha_1^{(i)} + (\pi_2^{(n)})^* \alpha_2^{(i)} = \sigma_i(1-\sigma_n^2) p_n \alpha_i \tag{164}$$

(3)
$$-e\alpha_1^{(i)} + d\alpha_2^{(i)} = \sigma_i \left(1 - \frac{\sigma_n^2}{\sigma_i^2}\right) p_i p_n \tag{165}$$

Proof:

(1) From the equation

$$P_- \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \frac{\alpha_i}{p_n^*} = \sigma_i \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} \tag{166}$$

it follows, by partial fraction decomposition, that there exist $\zeta_1^{(i)}$, $\zeta_2^{(i)}$ such that

$$\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix} \frac{\alpha_i}{p_n^*} = \sigma_i \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} + \begin{bmatrix} \zeta_1^{(i)}/p_n^* \\ \zeta_2^{(i)}/p_n^* \end{bmatrix} \tag{167}$$

This is equivalent to (163).

(2) As the function $\dfrac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$ is inner in H_-^∞, it follows, by Lemma 3.5, that the Hankel operator

$$\frac{1}{(1-\sigma_n^2)^{1/2}} \hat{H}_{\left(\frac{\pi_1^{(n)*}}{p_n^*} \; \frac{\pi_2^{(n)*}}{p_n^*}\right)}$$

acts on

$$\left\{ \operatorname{Ker} \hat{H}_{\left(\frac{\pi_1^{(n)*}}{p_n^*} \; \frac{\pi_2^{(n)*}}{p_n^*}\right)} \right\}^\perp$$

by multiplication. Therefore

$$\frac{1}{(1-\sigma_n^2)^{1/2}} \left(\frac{\pi_1^{(n)*}}{p_n^*} \; \frac{\pi_2^{(n)*}}{p_n^*} \right) \frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix} = \sigma_i \frac{\alpha_i}{p_n^*}$$

and so (164) follows.

(3) We start from (142), multiply by $(-e \;\; d)$ and use the second equation in (94), to obtain

$$\sigma_i t^*(-e\alpha_1^{(i)} + d\alpha_2^{(i)}) = \sigma_i^2 t^* p_i p_n - \sigma_n^2 t^* p_i p_n \tag{168}$$

This is equivalent to the statement. □

10. Nehari extension

We now come to the analysis of the Nehari extension problem. We will again first quote a summary of scalar results before proceeding to derive results concerning the Nehari extension of normalized coprime factors.

Theorem 10.1: *Let $r^*/t^* \in H_-^\infty$ be a scalar, strictly proper, transfer function, with r and t coprime polynomials and t monic of degree n. Assume $\mu_1 > \mu_2 \geq \cdots \geq \mu_n > 0$ are the singular value of H_{r^*/t^*}.*

(1) *There exist non-zero polynomials β_i, $i = 2, \ldots, n$, of degree $\leq n-2$ and signs ε_i, $i = 2, \ldots, n$, such that*

$$\lambda_1 p_1^* p_i - \lambda_i p_i^* p_1 = \lambda_1 t^* \beta_i \tag{169}$$

where $\lambda_i = \varepsilon_i \mu_i$, $1 \leq i \leq n$.

(2) *The polynomial p_1 has degree $n-1$ and has its zeros in the open left-half plane, i.e.*

$$\frac{\pi_1^*}{p_1^*} \in H_-^\infty$$

and $H_{\pi_1^*/p_1^*}$ has rank $n-1$.

(3) *The singular values of the Hankel operator $H_{\pi_1^*/p_1^*}$ are $\mu_2 \geq \cdots \geq \mu_n$ and the corresponding Schmidt pairs are* $\left\{\frac{\beta_i}{p_1}, \varepsilon_i \frac{\beta_i^*}{p_1^*}\right\}$

(4) $$\mu_1 = \inf\left\{\left\|\frac{r^*}{t^*} - q\right\|_\infty, q \in H_+^\infty\right\} = \left\|\frac{r^*}{t^*} - \frac{\pi_1}{p_1}\right\|_\infty$$

(5) *We have*

$$\frac{\beta_i}{p_1^*} = P_{X^{p_1^*}} \frac{p_i}{t^*} = P_{\left\{\frac{p_i}{p_1^*} H_-^2\right\}^\perp} \frac{p_i}{t^*} \tag{170}$$

i.e. the singular vectors of $H_{\pi_1^*/p_1^*}$ are projections of the singular vectors of H_{r^*/t^*} onto the orthogonal complement of $\operatorname{Im} H_{\pi_1^*/p_1^*} = \frac{p_1}{p_1^*} H_-^2$.

(6) $$\left\|\frac{\beta_i}{p_1^*}\right\|^2 = \left(1 - \frac{\mu_i^2}{\mu_1^2}\right)\left\|\frac{p_i}{t^*}\right\|^2 \tag{171}$$

(7) *With α_i defined by (131) we have*

$$\alpha_1 = \beta_n^* \tag{172}$$

Proof:

(1) This is, with an obvious change of notation, a special case of Proposition 8.1.

(2), (3) and (4) This is the content of Theorem 7.2 in Fuhrmann (1991).

(5) Rewrite (169) in the form

$$\frac{p_i}{t^*} = \frac{\beta_i}{p_1^*} + \frac{\lambda_i}{\lambda_1}\frac{p_1}{p_1^*}\frac{p_i^*}{t^*} \tag{173}$$

This is the orthogonal decomposition of p_i/t^* relative to the direct sum

$$H_-^2 = X^{p_1^*} \oplus \frac{p_1}{p_1^*}H_-^2$$

So (170) follows.

(6) From (173) it follows, using orthogonality, that $\mu_i = |\lambda_i|$

$$\left\|\frac{p_i}{t^*}\right\|^2 = \left\|\frac{\beta_i}{p_1^*}\right\|^2 + \frac{\mu_i^2}{\mu_1^2}\left\|\frac{p_i^*}{t^*}\right\|^2$$

Now, $\left\|\frac{p_i}{t^*}\right\|^2 = \left\|\frac{p_i^*}{t^*}\right\|^2$, and so (171) follows.

(7) The proof follows by comparing (131) and (169). □

Before we come to discuss the solution to the Nehari extension problem we need to state the following proposition.

Proposition 10.1: *Let $g = e/d$ and let $\bar{N} = e/t$, $\bar{M} = d/t$ be the normalized coprime factors of g. Then,*

(1) *Let*

$$U = \frac{1}{(1-\sigma_1^2)}\frac{\pi_1^{(1)}}{p_1}, \; V = \frac{1}{(1-\sigma_1^2)}\frac{\pi_2^{(1)}}{p_1}$$

Then U, V solve the Bezout equation $\bar{M}V - \bar{N}U = 1$, i.e.

$$\frac{d}{t}\frac{\pi_2^{(1)}}{(1-\sigma_1^2)p_1} - \frac{e}{t}\frac{\pi_1^{(1)}}{(1-\sigma_1^2)p_1} = 1$$

(2) *We have*

$$d^*\pi_1^{(1)} + e^*\pi_2^{(1)} = \lambda_1(1-\sigma_1^2)tp_1^* \tag{175}$$

(3) $\frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix}\pi_1^{(n)}/p_1 \\ \pi_2^{(n)}/p_1\end{bmatrix}$ *is inner in H_+^∞*

Proof: The proof follows from Theorem 8.1 by noting that both t and p_1 are stable polynomials. □

We proceed to adapt the proof of Nehari's theorem, given in Fuhrmann (1991), to the case where the symbol is derived from a NCF of $g = e/d$.

Theorem 10.2 (Nehari): *Assume that* $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$ *has singular values* $\sigma_1 > \sigma_2 \geqslant \cdots \geqslant \sigma_n$. *Then*

$$\sigma_1 = \left\| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} \right\| = \inf \left\{ \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right\|_\infty, \ q_j \in H_+^\infty \right\}$$

$$= \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(n)}/p_1 \\ \pi_2^{(n)}/p_1 \end{bmatrix} \right\|_\infty$$

Proof: Since, for $\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \in H_+^\infty$, we have $H_{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}} = 0$ we have

$$\sigma_1 = \left\| H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}} \right\| + \left\| H_{\left[\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}\right]} \right\|$$

$$= \leqslant \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right\|_\infty$$

So

$$\sigma_1 \leqslant \inf \left\{ \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \right\|_\infty, \ q_j \in H_+^\infty \right\}$$

For $i = 1$ the singular value/singular vector Equation (94) can be rewritten as

$$\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(n)}/p_1 \\ \pi_2^{(n)}/p_1 \end{bmatrix} = \sigma_1 \frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(1)}/p_1 \\ \hat{p}_2^{(1)}/p_1 \end{bmatrix} \tag{176}$$

and using the fact that $\frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(1)}/p_1 \\ \hat{p}_2^{(1)}/p_1 \end{bmatrix}$ is all-pass we get

$$\sigma_1 = \left\| \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix} - \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix} \right\|_\infty$$

$$= \left\| \sigma_1 \frac{t}{t^*} \begin{bmatrix} \hat{p}_1^{(1)}/p_1 \\ \hat{p}_2^{(1)}/p_1 \end{bmatrix} \right\|_\infty = \sigma_1$$

Noting that $\begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix} \in H_+^\infty$ by the previous proposition this proves the result. □

We now come to the study of a Hankel operator that is associated with the Nehari extension. For this study we need the following lemma.

Lemma 10.1: *Let* $\sigma_1 \geqslant \cdots \geqslant \sigma_n > 0$ *be the singular values of the Hankel operator* $H_{(-\bar{N}^* \ \bar{M}^*)}$ *and let* p_i, $\begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix}$, $\begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$ *be defined by the singular value/*

singular vector equations (94). Then

(1) *there exist non-zero polynomials* $\begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \end{bmatrix}$, $i = 2, \ldots, n$, *such that*

$$\sigma_1 \begin{bmatrix} \hat{p}_1^{(1)} \\ \hat{p}_2^{(1)} \end{bmatrix} p_i - \sigma_i \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} p_1 = \sigma_1 t^* \begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \end{bmatrix} \tag{177}$$

(2)
$$\begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix} p_1 - \begin{bmatrix} \pi_1^{(1)} \\ \pi_2^{(1)} \end{bmatrix} p_i = \sigma_i t \begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \end{bmatrix} \tag{178}$$

(3)
$$\left\| \begin{bmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{bmatrix} \right\|^2 = \left(1 - \frac{\sigma_i^2}{\sigma_1^2}\right) \left\| \frac{p_i}{t^*} \right\|^2 \tag{179}$$

(4)
$$\frac{\left\| \begin{pmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{pmatrix} \right\|}{\left\| \dfrac{\beta_i}{p_1} \right\|} = (1 - \sigma_i^2)^{1/2} \sigma_i$$

(5) *The following equality holds*

$$\begin{bmatrix} \alpha_1^{(1)} \\ \alpha_2^{(1)} \end{bmatrix} = \begin{bmatrix} \beta_1^{(n)} \\ \beta_2^{(n)} \end{bmatrix} \tag{180}$$

Proof:

(1) and (2) follow directly from Proposition 8.2 with an obvious change of notation.

(3) Equation (177) can be rewritten as

$$\begin{aligned} \begin{bmatrix} \hat{p}_1^{(1)}/p_1 \\ \hat{p}_2^{(1)}/p_1 \end{bmatrix} \frac{p_i}{t^*} &= \begin{bmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{bmatrix} + \frac{\sigma_i}{\sigma_1} \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix} \\ &= \begin{bmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{bmatrix} + \frac{\sigma_i}{\sigma_1} \begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i \end{bmatrix} \frac{p_i}{t^*} \end{aligned} \tag{181}$$

Since

$$\begin{bmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{bmatrix} \in H_+^2 \quad \text{and} \quad \begin{bmatrix} \hat{p}_1^{(1)}/t^* \\ \hat{p}_2^{(1)}/t^* \end{bmatrix} \in H_-^2$$

they are orthogonal. Recalling that

$$\begin{bmatrix} \hat{p}_1^{(1)}/p_1 \\ \hat{p}_2^{(1)}/p_1 \end{bmatrix}$$

is all-pass, yields

$$\left\| \begin{bmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{bmatrix} \right\|^2 + \frac{\sigma_i^2}{\sigma_1^2} \left\| \begin{bmatrix} \hat{p}_1^{(i)}/p_i \\ \hat{p}_2^{(i)}/p_i \end{bmatrix} \frac{p_i}{t^*} \right\|^2 = \left\| \begin{bmatrix} \hat{p}_1^{(1)}/p_1 \\ \hat{p}_2^{(1)}/p_1 \end{bmatrix} \frac{p_i}{t^*} \right\|^2$$

Thus, (179) follows.

(4) From (179) and (171) we get

$$\frac{\left\|\begin{pmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{pmatrix}\right\|^2}{\left\|\dfrac{\beta_i}{p_1}\right\|^2} = \frac{1-\dfrac{\sigma_i^2}{\sigma_1^2}}{1-\dfrac{\mu_i^2}{\mu_1^2}}$$

$$= \frac{\left(1-\dfrac{\sigma_i^2}{\sigma_1^2}\right)}{\left(1-\dfrac{\sigma_i^2}{1-\sigma_i^2}\,\dfrac{1-\sigma_1^2}{\sigma_1^2}\right)} = \frac{(\sigma_1^2-\sigma_i^2)(1-\sigma_i^2)}{\sigma_1^2(1-\sigma_i^2)-\sigma_i^2(1-\sigma_1^2)} \tag{182}$$

$$= 1-\sigma_i^2$$

Note that this ratio is i-dependent, contrary to the case in (145).

(5) Choosing $i = 1$ in (142) and $i = n$ in (177) and comparing proves (180). □

In the following theorem we examine the Nehari extension in some detail.

Theorem 10.3: *Let $\sigma_1 > \sigma_2 \geq \sigma_3 \ldots \geq \sigma_n > 0$ be the singular values of the Hankel operator $H_{\begin{bmatrix}-e^*/t^*\\ d^*/t^*\end{bmatrix}}$. Let*

$$\begin{bmatrix}-e^*p_i\\ d^*p_i\end{bmatrix} = \sigma_i t\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix} + t^*\begin{bmatrix}\pi_1^{(i)}\\ \pi_2^{(i)}\end{bmatrix}$$

$$-e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} = \sigma_i t^* p_i$$

be the singular vector/singular value equations. Then

(1) *the involuted Hankel operator $\dfrac{1}{(1-\sigma_1^2)^{1/2}}\hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}$ has singular values $\sigma_2 \geq \cdots \geq \sigma_n$ and its σ_i-Schmidt pairs are*

$$\left\{\frac{\beta_i^*}{p_1^*}, \frac{\varepsilon_1\varepsilon_i}{(1-\sigma_i^2)^{1/2}}\begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}\right\}$$

$i = 2, \ldots, n$ with the β_i defined in Theorem 10.1 and $\beta_1^{(i)}$, $\beta_2^{(i)}$ by Lemma 10.1.

(2) *The operator $\hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}$ has rank $n-1$.*

(3) $$\operatorname{Ker}\hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}} = \frac{p_1}{p_1^*}H_-^2 \tag{183}$$

(4) *There exist polynomials $\omega_1^{(i)}$, $\omega_2^{(i)}$ of degree $\leq n-2$ such that*

$$\frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix}\pi_1^{(1)}\\ \pi_2^{(1)}\end{bmatrix}\beta_i^* = \sigma_i p_1^* \frac{\varepsilon_1\varepsilon_i}{(1-\sigma_i^2)^{1/2}}\begin{bmatrix}\beta_1^{(i)}\\ \beta_2^{(i)}\end{bmatrix} + p_1\begin{bmatrix}\omega_1^{(i)}\\ \omega_2^{(i)}\end{bmatrix} \tag{184}$$

(5) *We have*

$$(\pi_1^{(1)})^*\beta_1^{(i)} + (\pi_2^{(1)})^*\beta_2^{(i)} = \sigma_i\varepsilon_1\varepsilon_i(1-\sigma_1^2)^{1/2}(1-\sigma_i^2)^{1/2}\beta_i^* p_1 \tag{185}$$

(6) *We have*

$$(\pi_1^{(1)})^*\omega_1^{(i)} + (\pi_2^{(1)})^*\omega_2^{(i)} = (1-\sigma_1^2)^{1/2}(1-\sigma_i^2)\beta_i^* p_1^* \tag{186}$$

Proof:

(1) We will show that

$$\frac{1}{(1-\sigma_1^2)^{1/2}}\hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}\frac{\beta_i^*}{p_1^*} = \sigma_i\frac{\varepsilon_1\varepsilon_i}{(1-\sigma_i^2)^{1/2}}\begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}$$

From (143), it follows that

$$\begin{bmatrix}\pi_1^{(n)}/p_n\\ \pi_2^{(n)}/p_n\end{bmatrix} - \begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix} = \sigma_1\frac{t}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} \tag{187}$$

and hence

$$\hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}} = -\sigma_1\hat{H}_{\frac{t}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}} \tag{188}$$

Now Equation (169) can be rewritten as

$$t\beta_i^* = p_1p_i^* - \frac{\lambda_i}{\lambda_1}p_ip_1^* \tag{189}$$

and

$$\frac{t\beta_i^*}{p_1p_n} = \frac{p_i^*}{p_n} - \frac{\lambda_i}{\lambda_1}\frac{p_ip_1^*}{p_1p_n} \tag{190}$$

which implies the following

$$\frac{t\beta_i^*}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}\frac{1}{p_1^*} = \frac{p_i^*}{p_np_1^*}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} - \frac{\lambda_i}{\lambda_1}\frac{p_i}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} \tag{191}$$

Since $\frac{p_i^*}{p_np_1^*}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} \in H_-^2$ it follows that

$$P_+\frac{t}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}\frac{\beta_i^*}{p_1^*} = -\frac{\lambda_i}{\lambda_1}P_+\frac{p_i}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} \tag{192}$$

So we have to obtain a partial fraction decomposition of

$$\frac{p_i}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}$$

Going back to equations

$$\left.\begin{aligned} t^*\begin{bmatrix}\beta_1^{(i)}\\ \beta_2^{(i)}\end{bmatrix} &= \begin{bmatrix}\hat{p}_1^{(1)}\\ \hat{p}_2^{(1)}\end{bmatrix}p_i - \frac{\sigma_i}{\sigma_1}\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}p_1 \\ t^*\begin{bmatrix}\beta_1^{(n)}\\ \beta_2^{(n)}\end{bmatrix} &= \begin{bmatrix}\hat{p}_1^{(1)}\\ \hat{p}_2^{(1)}\end{bmatrix}p_n - \frac{\sigma_n}{\sigma_1}\begin{bmatrix}\hat{p}_1^{(n)}\\ \hat{p}_2^{(n)}\end{bmatrix}p_1 \end{aligned}\right\} \tag{193}$$

it follows, by eliminating the middle terms, that

$$\begin{aligned} t^*\left\{\begin{bmatrix}\beta_1^{(n)}\\ \beta_2^{(n)}\end{bmatrix}p_i - \begin{bmatrix}\beta_1^{(i)}\\ \beta_2^{(i)}\end{bmatrix}p_n\right\} &= \frac{1}{\sigma_1}\left\{\sigma_i\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}p_1p_n - \sigma_n\begin{bmatrix}\hat{p}_1^{(n)}\\ \hat{p}_2^{(n)}\end{bmatrix}p_1p_i\right\} \\ &= \frac{p_1}{\sigma_1}\left\{\sigma_i\begin{bmatrix}\hat{p}_1^{(i)}\\ \hat{p}_2^{(i)}\end{bmatrix}p_n - \sigma_n\begin{bmatrix}\hat{p}_1^{(n)}\\ \hat{p}_2^{(n)}\end{bmatrix}p_i\right\} \\ &= \frac{p_1}{\sigma_1}\left\{\sigma_i t^*\begin{bmatrix}\alpha_1^{(i)}\\ \alpha_2^{(i)}\end{bmatrix}\right\} \end{aligned} \tag{194}$$

or

$$\begin{bmatrix}\beta_1^{(n)}\\ \beta_2^{(n)}\end{bmatrix}p_i - \begin{bmatrix}\beta_1^{(i)}\\ \beta_2^{(i)}\end{bmatrix}p_n = \frac{\sigma_i}{\sigma_1}p_1\begin{bmatrix}\alpha_1^{(i)}\\ \alpha_2^{(i)}\end{bmatrix} \tag{195}$$

Now, by Lemma 10.1.5, $\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} = \begin{bmatrix}\beta_1^{(n)}\\ \beta_2^{(n)}\end{bmatrix}$ and so

$$\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}p_i - \begin{bmatrix}\beta_1^{(i)}\\ \beta_2^{(i)}\end{bmatrix}p_n = \frac{\sigma_i}{\sigma_1}p_1\begin{bmatrix}\alpha_1^{(i)}\\ \alpha_2^{(i)}\end{bmatrix} \tag{196}$$

Now, by Lemma 10.1.5, $\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} = \begin{bmatrix}\beta_1^{(n)}\\ \beta_2^{(n)}\end{bmatrix}$ and so

$$\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}p_i - \begin{bmatrix}\beta_1^{(i)}\\ \beta_2^{(i)}\end{bmatrix}p_n = \frac{\sigma_i}{\sigma_1}p_1\begin{bmatrix}\alpha_1^{(i)}\\ \alpha_2^{(i)}\end{bmatrix} \tag{197}$$

Dividing through by p_1p_n we get

$$\frac{p_i}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} = \begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix} + \frac{\sigma_i}{\sigma_1}\begin{bmatrix}\alpha_1^{(i)}/p_n\\ \alpha_2^{(i)}/p_n\end{bmatrix} \tag{198}$$

Hence

$$P_+\frac{p_i}{p_1p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix} = \begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix} \tag{199}$$

Summing up, we have

$$\hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}} \frac{\beta_i^*}{p_1^*} = -\sigma_1 \hat{H}_{\frac{t}{p_1 p_n}\begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}} \frac{\beta_i^*}{p_1^*}$$

$$= \sigma_1 \frac{\lambda_i}{\lambda_1} P_+ \frac{p_i}{p_1 p_n} \begin{bmatrix}\alpha_1^{(1)}\\ \alpha_2^{(1)}\end{bmatrix}$$

$$= \sigma_1 \frac{\lambda_i}{\lambda_1} \begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}$$

We rewrite this in the form

$$\frac{1}{(1-\sigma_1^2)^{1/2}} \hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}} \frac{\beta_i^*}{p_1^*} = \sigma_i \varepsilon_1 \varepsilon_i \frac{1}{(1-\sigma_i^2)^{1/2}} \begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}$$

Now (177) can be rewritten as

$$\begin{bmatrix}\hat{p}_1^{(1)}/p_1\\ \hat{p}_2^{(1)}/p_1\end{bmatrix} \frac{p_i}{t^*} = \begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix} + \frac{\sigma_i}{\sigma_1} \begin{bmatrix}\hat{p}_1^{(i)}/t^*\\ \hat{p}_2^{(i)}/t^*\end{bmatrix} \tag{200}$$

This is an orthogonal decomposition as $\begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix} \in H_+^2$ and $\begin{bmatrix}\hat{p}_1^{(i)}/t^*\\ \hat{p}_2^{(i)}/t^*\end{bmatrix} \in H_-^2$.

From here the orthogonality of the set $\left\{\begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}\right\}$ follows. Since the β_i^*/p_1^* are also orthogonal, we can apply Lemma 9.2 to conclude that the singular values of $\frac{1}{(1-\sigma_1^2)^{1/2}} \hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}$ are $\sigma_2 \geqslant \cdots \geqslant \sigma_n > 0$ and the Schmidt pairs are $\left\{\frac{\beta_i^*}{p_1^*}, \frac{\varepsilon_1\varepsilon_i}{(1-\sigma_i^2)^{1/2}} \begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}\right\}$.

(2) It suffices to show that $\dim \operatorname{Im} \hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}} = n-1$. This follows from Part (1), as $\operatorname{Im} \hat{H}_{\frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}$ is spanned by the $\left\{\begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}\right\}$.

(3) Indeed, for every $h \in H_-^2$, we have

$$P_+ \begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix} \frac{p_1}{p_1^*} h = P_+ \begin{bmatrix}\pi_1^{(1)}/p_1^*\\ \pi_2^{(1)}/p_1^*\end{bmatrix} h = 0 \tag{201}$$

and hence

$$\frac{p_1}{p_1^*} H_-^2 \subseteq \operatorname{Ker} \hat{H}_{\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}$$

To show equality it suffices to note that $\{p_1/p_1^*\}^\perp$ has dimension $n-1$, i.e. the same dimension as the Hankel operator.

(4) These are the singular vector/singular value equations for the operator

$$\hat{H}_{\frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}}$$

(5) By Lemma 3.5 the Hankel operator

$$\frac{1}{(1-\sigma_1^2)^{1/2}}H_{\left[\frac{(\pi_1^{(1)})^*}{p_1^*}\ \frac{(\pi_2^{(1)})^*}{p_1^*}\right]}$$

acts by multiplication.

Therefore

$$\frac{1}{(1-\sigma_1^2)^{1/2}}\left(\frac{(\pi_1^{(1)})^*}{p_1^*}\ \frac{(\pi_2^{(1)})^*}{p^*}\right)\frac{\varepsilon_1\varepsilon_i}{(1-\sigma_i^2)^{1/2}}\begin{bmatrix}\beta_1^{(i)}/p_1\\ \beta_2^{(i)}/p_1\end{bmatrix}=\sigma_i\frac{\beta_i^*}{p_1^*}$$

and (185) follows.

(6) The proof follows by left multiplying (184) by $((\pi_1^{(1)})^*\ (\pi_2^{(1)})^*)$, using Part (5) and the fact that $\frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix}\pi_1^{(1)}/p_1\\ \pi_2^{(1)}/p_1\end{bmatrix}$ is all-pass. □

We now give a control theoretic interpretation of the previous results.

Corollary 10.1: *The rational function $k=\pi_1^{(1)}/\pi_2^{(1)}$ is a stabilizing controller for g.*

Proof: The proof follows from the Bezout equation (174). Also we can check directly that

$$\frac{g}{1-kg}=\frac{\dfrac{e}{d}}{1-\dfrac{\pi_1^{(1)}}{\pi_2^{(1)}}\dfrac{e}{d}}=\frac{1}{1-\sigma_1^2}\frac{e\pi_2^{(1)}}{tp_1}\in H_+^\infty$$ □

Actually, the controller given in Corollary 10.1 is a special controller. In fact it turns out to be the optimally robust controller, see McFarlane and Glover (1990). We will return to a discussion of the robust control in § 13.

11. LQG balancing

In Ober and McFarlane (1989) it was shown how normalized coprime factorizations can be used to study LQG balanced realizations. A close relationship was established between the Lyapunov balanced realization of a coinner function based on a NLCF of a transfer function G, and an LQG balanced realization of G itself. Part of the motivation for the present paper was to gain a better understanding of this relation.

We base our approach on the results of Fuhrmann (1991), establishing a Lyapunov balanced realization of an asymptotically (anti) stable transfer function g as a matrix representation of the shift realization of g with respect to a basis of, suitably normalized, singular vectors of the Hankel operator H_g. We begin by proving an analogous result for the normalized coprime factors of a, not necessarily stable, transfer function g. It may not come as a total surprise that the Lyapunov balanced realization we derive for the coinner function coincides with the canonical form in Ober and McFarlane (1989) specialized to the case at hand.

Following that, we recall the definition of LQG balancing and, in Theorem 11.2, we prove an analogous result for g itself. The basis we use is of course no longer made up of Hankel singular vectors. However, it is very closely related to the basis used in the realization of the normalized coprime factors.

Definition 11.1: (Moore 1981): A minimal, asymptotically stable system (A, B, C, D) is called *Lyapunov balanced* if there exists a diagonal matrix $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$, such that

$$\left.\begin{aligned} A\Sigma + \Sigma\tilde{A} &= -B\tilde{B} \\ \tilde{A}\Sigma + \Sigma A &= -\tilde{C}C \end{aligned}\right\} \tag{202}$$

The matrix Σ is called the *gramian* of the system (A, B, C, D) and its diagonal entries are called the *Hankel singular values* of the system. □

Before we come to derive a state-space realization of the normalized coprime factors we have to summarize a number of relationships between coefficients of the polynomials that we are interested in. Recall that by $q_{j,i}$ we denote the ith coefficient of the polynomial q_j, i.e. $q_j = \sum_{i=0}^{n} s^i q_{j,i}$.

Proposition 11.1:

(1) *Let r^*/t^* be a scalar, strictly proper, antistable transfer function with t monic. Assume the notation of Proposition 8.1 and assume that $\mu_1 > \ldots > \mu_n > 0$ are the singular value of H_{r^*/t^*}.*

(*a*) *We have, for the α_{ij} defined in (84),*

$$\alpha_{ij,n-2} = -(\lambda_i - \lambda_j)p_{i,n-1}p_{j,n-1} \tag{203}$$

(*b*) *We have for the polynomials defined in* (131),

$$\alpha_{i,n-2} = (-1)^{n-1}\frac{\lambda_i - \lambda_n}{\lambda_i}p_{n,n-1}p_{i,n-1} \tag{204}$$

(2) *Assume the notation of Proposition 8.1 and assume that $\sigma_1 > \sigma_2 > \ldots > \sigma_n > 0$. Let $\left\{p_i, \begin{bmatrix}\hat{p}_1^{(i)} \\ \hat{p}_2^{(i)}\end{bmatrix}\right\}$ be a minimal degree solution pair corresponding to the singular value σ_i of $H_{\begin{bmatrix}-e^*/t^* \\ d^*/t^*\end{bmatrix}}$. Then the following relations hold true,*

$$\hat{p}_{1,n-1}^{(i)} = (-1)^n \varepsilon_i (1 - \sigma_i^2)^{1/2} p_{i,n-1} \tag{205}$$

$$\hat{p}_{2,n-1}^{(i)} = (-1)^n \sigma_i p_{i,n-1} \tag{206}$$

$$\pi_{1,n-1}^{(i)} = -\varepsilon_i \sigma_i (1 - \sigma_i^2)^{1/2} p_{i,n-1} \tag{207}$$

$$\pi_{2,n-1}^{(i)} = (1 - \sigma_i^2) p_{i,n-1} \tag{208}$$

Proof:

(1)(a) We equate the highest degree coefficient in (85) to get

$$
\begin{aligned}
p_{i,n-1}(-1)^{n-1}p_{j,n-1} &= \frac{1}{\lambda_i^2-\lambda_j^2}\{\lambda_j(-1)^n\alpha_{ij,n-2}+\lambda_i(-1)^{n-2}\alpha_{ij,n-2}\}\\
&= (-1)^n\frac{\lambda_i+\lambda_j}{\lambda_i^2-\lambda_j^2}\alpha_{ij,n-2}\\
&= (-1)^n\frac{\alpha_{ij,n-2}}{\lambda_i-\lambda_j}
\end{aligned}
\tag{209}
$$

(1)(b) We equate the highest degree coefficients in

$$\lambda_i p_n^* p_i - \lambda_n p_i^* p_n = \lambda_i t \alpha_i \tag{210}$$

to get (204).

(2) We will use the singular value equations

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i = \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix} \tag{211}$$

and

$$-e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} = \sigma_i t^* p_i \tag{212}$$

Equating the highest degree coefficients in (212) proves (206) as the degree of e is less then the degree of d.

From (211) we get $(-1)^n p_{i,n-1} = \sigma_i \hat{p}_{2,n-1}^{(i)} + (-1)^n \pi_{2,n-1}^{(i)}$. Substituting (206) proves (208). From the first coordinates in (211) we get

$$0 = \sigma_i \hat{p}_{1,n-1}^{(i)} + (-1)^n \pi_{1,n-1}^{(i)}$$

or

$$\pi_{1,n-1}^{(i)} = (-1)^{n-1}\sigma_i \hat{p}_{1,n-1}^{(i)} \tag{213}$$

So it remains to compute $\hat{p}_{1,n-1}^{(i)}$. For this we equate the highest order coefficients in (94), i.e.

$$d^*\pi_1^{(i)} + e^*\pi_2^{(i)} = \varepsilon_i \sigma_i (1-\sigma_i^2)^{1/2} t p_i^*$$

This leads to

$$(-1)^n \pi_{1,n-1}^{(i)} = \varepsilon_i\sigma_i(1-\sigma_i^2)^{1/2}(-1)^{n-1}p_{i,n-1}$$

which proves (207). Substituting this in (213) proves (205). □

It can be shown easily using state-space methods that an asymptotically stable minimal system has a Lyapunov balanced realization (see e.g. Moore 1981). We now come to rederive a canonical form for normalized coprime factors (Ober and McFarlane 1989). We show here that this canonical form is in fact a shift realization whose matrix representation is calculated with respect to the singular vectors in H_+^2 of $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. We restrict ourselves to the case of non-repeated singular values. The case of repeated singular values can be analysed in the same way as was done in Fuhrmann (1991) for the case of scalar functions.

Theorem 11.1 *Let $g = e/d$ be a strictly proper transfer function with e and d coprime polynomials. Let e/t, d/t be the normalized coprime factors of g and let $\sigma_1 > \cdots > \sigma_n > 0$ be the singular values of the Hankel operator $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. Let ε_i, $1 \leqslant i \leqslant n$, be the signs and $\left\{\frac{p_i}{t}, \varepsilon_i \frac{p_i^*}{t^*}\right\}$, $1 \leqslant i \leqslant n$, the Schmidt vectors H_{R^*}, normalized so that*

$$\left\|\frac{p_i}{t}\right\|^2 = \sigma_i \tag{214}$$

$1 \leqslant i \leqslant n$. Then the matrix representation of the shift realization of the function $(-e/t \;\; d/t)$ with respect to this basis is Lyapunov balanced with $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ and is given by (A, B, C, D) with

$$\left.\begin{aligned}
\alpha_{ji} &= -\frac{1}{\sigma_j}\frac{\lambda_j}{\lambda_i + \lambda_j} p_{i,n-1} p_{j,n-1} \\
&= -\frac{\varepsilon_j(1-\sigma_i^2)^{1/2}}{\varepsilon_i\sigma_i(1-\sigma_j^2)^{1/2} + \varepsilon_j\sigma_j(1-\sigma_i^2)^{1/2}} p_{i,n-1} p_{j,n-1} \\
B_i &= (b_i' \;\; b_i'') \\
&= (-\varepsilon_i(1-\sigma_i^2)^{1/2} p_{i,n-1} \;\; -\sigma_i p_{i,n-1}) \\
c_i &= p_{i,n-1} \\
D &= (0 \;\; 1)
\end{aligned}\right\} \tag{215}$$

where a_{ji} is the jith entry of A, B_i the ith row of B, c_i the ith entry of the c-vector, $p_{i,n-1}$ the leading coefficient of p_i and $\lambda_i = \varepsilon_i\sigma_i/(1-\sigma_i^2)^{1/2}$, $1 \leqslant i \leqslant n$.

Proof: The constant term is given by

$$D = \begin{bmatrix} -\frac{e}{t} & \frac{d}{t} \end{bmatrix}(\infty) = (0 \;\; 1)$$

The computation of the output map C is simple. We have

$$c_i = C\frac{p_i}{t} = \left(\frac{p_i}{t}\right)_{-1} = p_{i,n-1}$$

We now compute the matrix representation of the input map B. Since $e/t \in X^t$ and $(d-t)/t \in X^t$ and the vectors p_i/t, $1 \leqslant i \leqslant n$, form a basis in X^t, there are (b_i', b_i''), $1 \leqslant i \leqslant n$, such that we can write

$$\begin{bmatrix} -\frac{e}{t} & \frac{d}{t} \end{bmatrix} - [0 \;\; 1] = \begin{bmatrix} -\frac{e}{t} & \frac{d-t}{t} \end{bmatrix} = \sum_{i=1}^{n} [b_i' \;\; b_i''] \frac{p_i}{t}$$

Note that by the orthogonality of the p_i/t, $1 \leqslant i \leqslant n$, we have that

$$\left(\begin{bmatrix} -\frac{e}{t} & \frac{d}{t} \end{bmatrix} - [0 \;\; 1], \frac{p_i}{t}\right) = [b_i' \;\; b_i'']$$

This implies that

$$-\frac{e}{t} = \sum_{i=1}^{n} b_i' \frac{p_i}{t}$$

and

$$\frac{d-t}{t} = \sum_{i=1}^{n} b_i'' \frac{p_i}{t}$$

We use the orthogonality of the p_i/t and our normalization to get

$$b_i' = -\frac{\left(\frac{e}{t}, \frac{p_i}{t}\right)}{\left(\frac{p_i}{t}, \frac{p_i}{t}\right)} = -\frac{\left(\frac{e}{t}, \frac{p_i}{t}\right)}{\sigma_i}$$

and

$$b_i'' = \frac{\left(\frac{d-t}{t}, \frac{p_i}{t}\right)}{\left(\frac{p_i}{t}, \frac{p_i}{t}\right)} = \frac{\left(\frac{d-t}{t}, \frac{p_i}{t}\right)}{\sigma_i}$$

We proceed to compute the two inner products. For this we resort to contour integration and the residue calculus. Let γ_R and $\hat{\gamma}_R$ be the semicircular contours shown in Fig. 6

Note that γ_R is positively orientated whereas $\hat{\gamma}_R$ is negatively orientated. The singular value/singular vector equation (128) implies

$$-\frac{e}{t}\frac{p_i^*}{t^*} = \sigma_i \frac{(\hat{p}_1^{(i)})^*}{t} + \frac{(\pi_1^{(i)})^*}{t^*} \tag{216}$$

Integrating over γ_R, we have

$$\begin{aligned} -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e}{t}\frac{p_i^*}{t^*} d\tau &= \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\gamma_R} \left\{\sigma_i \frac{(\hat{p}_1^{(i)})^*}{t} + \frac{(\pi_1^{(i)})^*}{t^*}\right\} dz \\ &= \lim_{R\to\infty} \frac{\sigma_i}{2\pi i}\int_{\gamma_R} \frac{(\hat{p}_1^{(i)})^*}{t} dz \\ &= (-1)^{n-1}\sigma_i \hat{p}_{1,n-1}^{(i)} \end{aligned}$$

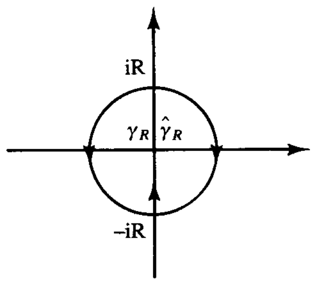

Figure 6.

Using equality (205), i.e. $\hat{p}_{1,n-1}^{(i} = (-1)^n \varepsilon_i (1-\sigma_i^2)^{1/2} p_{i,n-1}$, we get

$$b_i' = -\frac{\left(\frac{e}{t}, \frac{p_i}{t}\right)}{\left(\frac{p_i}{t}, \frac{p_i}{t}\right)} = \frac{(-1)^{n-1}(-1)^n \varepsilon_i \sigma_i (1-\sigma_i^2)^{1/2} p_{i,n-1}}{\sigma_i}$$

$$= -\varepsilon_i (1-\sigma_i^2)^{1/2} p_{i,n-1}$$

Similarly,

$$b_i'' = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d-t}{t} \frac{p_i^*}{t^*} \mathrm{d}\tau$$

As the degree deficiency of the numerator with respect to the denominator is at least two, this integral can be, using a partial fraction decomposition, simply computed by contour integration. We start with the following observation.

$$b_i'' = \frac{1}{\sigma_i} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d-t}{t} \frac{p_i^*}{t^*} \mathrm{d}\tau = \lim_{R\to\infty} \frac{1}{\sigma_i} \frac{1}{2\pi i} \int_{\gamma_R} \frac{d}{t} \frac{p_i^*}{t^*} - \frac{p_i^*}{t^*} \mathrm{d}z$$

$$= \lim_{R\to\infty} \frac{1}{\sigma_i} \frac{1}{2\pi i} \int_{\gamma_R} \frac{d}{t} \frac{p_i^*}{t^*} \mathrm{d}z - \lim_{R\to\infty} \frac{1}{\sigma_i} \frac{1}{2\pi i} \int_{\gamma_R} \frac{p_i^*}{t^*} \mathrm{d}z$$

$$= \lim_{R\to\infty} \frac{1}{\sigma_i} \frac{1}{2\pi i} \int_{\gamma_R} \frac{d}{t} \frac{p_i^*}{t^*} \mathrm{d}z$$

From the singular value/singular vector equations we obtain that

$$\frac{d^*}{t^*} \frac{p_i}{t} = \sigma_i \frac{\hat{p}_2^{(i)*}}{t} + \frac{\pi_2^{(i)*}}{t^*}$$

and therefore

$$b_i'' = \lim_{R\to\infty} \frac{1}{\sigma_i} \frac{1}{2\pi i} \int_{\gamma_R} \sigma_i \frac{\hat{p}_2^{(i)*}}{t} + \frac{\pi_2^{(i)*}}{t^*} \mathrm{d}z$$

$$= \lim_{R\to\infty} \frac{1}{\sigma_i} \frac{1}{2\pi i} \int_{\gamma_R} \sigma_i \frac{\hat{p}_2^{(i)*}}{t} \mathrm{d}z$$

$$= (-1)^{n-1} \hat{p}_{2,n-1}^{(i)}$$

Using (206) we get

$$b_i'' = (-1)^{n-1}(-1)^n \sigma_i p_{i,n-1} = -\sigma_i p_{i,n-1}$$

So the ith row of B is given by (215).

Finally, we compute the generator matrix A. With $\xi_i = p_{i,n-1}$ we put

$$S^t \frac{p_i}{t} = \sum_{j=1}^{n} a_{ji} \frac{p_j}{t}$$

Now

$$S^t \frac{p_i}{t} = \pi^t \frac{z p_i}{t} = \frac{z p_i - \xi_i t}{t}$$

Utilizing our inner product as well as the orthogonality of the set $\{p_i/t | i = 1, \ldots, n\}$, we have

$$a_{ji} = \frac{\left(S^t\frac{p_i}{t}, \frac{p_j}{t}\right)}{\left(\frac{p_j}{t}, \frac{p_j}{t}\right)} = \frac{\left(\frac{zp_i - \xi_i t}{t}, \frac{p_j}{t}\right)}{\sigma_i} \tag{217}$$

We proceed to compute the numerator inner product. Since $\deg(zp_i - \xi_i t)p_j^* \leq 2n - 2$ a standard estimate on contour integrals yields

$$\begin{aligned}\left(\frac{zp_i - \xi_i d}{t}, \frac{p_j}{t}\right) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{zp_i - \xi_i t}{t}\,\frac{p_j^*}{t^*}\,\mathrm{d}\tau \\ &= \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\gamma_R} \frac{zp_i - \xi_i t}{t}\,\frac{p_j^*}{t^*}\,\mathrm{d}z \\ &= \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\gamma_R} \frac{zp_ip_j^*}{tt^*}\,\mathrm{d}z - \lim_{R\to\infty} \frac{\xi_i}{2\pi_i}\int_{\gamma_R} \frac{p_j^*}{t^*}\,\mathrm{d}z \\ &= \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\gamma_R} \frac{zp_ip_j^*}{tt^*}\,\mathrm{d}z\end{aligned}$$

Now, from (85) we get

$$\frac{zp_ip_j^*}{tt^*} = \frac{1}{\lambda_i^2 - \lambda_j^2}\left\{\lambda_j\frac{z\alpha_{ij}}{t} + \lambda_i\frac{z\alpha_{ij}^*}{t^*}\right\} \tag{218}$$

So, using (203), we get

$$\begin{aligned}\frac{1}{2\pi i}\int_{\gamma_R} \frac{zp_ip_j^*}{tt^*}\,\mathrm{d}z &= \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2}\alpha_{ij,n-2} \\ &= \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2}(-1)(\lambda_i - \lambda_j)p_{i,n-1}p_{j,n-1} \\ &= -\frac{\lambda_j}{\lambda_i + \lambda_j}p_{i,n-1}p_{j,n-1}\end{aligned}$$

and hence

$$a_{ji} = -\frac{1}{\sigma_j}\,\frac{\lambda_j}{\lambda_i + \lambda_j}p_{i,n-1}p_{j,n-1} \tag{219}$$

Recalling that $\lambda_i = \varepsilon_i\mu_i$ and $\mu_i = \dfrac{\sigma_i}{(1 - \sigma_i^2)^{1/2}}$, we can rewrite (219) as

$$\begin{aligned}a_{ji} &= -\frac{1}{\sigma_j}\,\frac{\lambda_j}{\lambda_i + \lambda_j}\,p_{i,n-1}p_{j,n-1} \\ &= -\frac{1}{\sigma_j}\,\frac{\dfrac{\varepsilon_j\sigma_j}{(1 - \sigma_j^2)^{1/2}}}{\dfrac{\varepsilon_j\sigma_j}{(1 - \sigma_i^2)^{1/2}} + \dfrac{\varepsilon_j\sigma_j}{(1 - \sigma_j^2)^{1/2}}}\,p_{i,n-1}p_{j,n-1} \\ &= -\frac{\varepsilon_j(1 - \sigma_i^2)^{1/2}}{\varepsilon_i\sigma_i(1 - \sigma_j^2)^{1/2} + \varepsilon_j\sigma_j(1 - \sigma_i^2)^{1/2}}\,p_{i,n-1}p_{j,n-1}\end{aligned} \tag{220}$$

We proceed to show that the realization so obtained is Lyapunov balanced. We begin by computing the i, j-element of $\tilde{A}\Sigma + \Sigma A + \tilde{C}C$.

$$
\begin{aligned}
(\tilde{A}\Sigma + \Sigma A + \tilde{C}C)_{ij} &= a_{ji}\sigma_j + \sigma_i a_{ij} + c_i c_j \\
&= -\frac{1}{\sigma_j}\frac{\lambda_j}{\lambda_i + \lambda_j} p_{i,n-1}p_{j,n-1}\sigma_j \\
&\quad - \sigma_i \frac{1}{\sigma_i}\frac{\lambda_i}{\lambda_i + \lambda_j} p_{i,n-1}p_{j,n-1} \\
&\quad + p_{i,n-1}p_{j,n-1} \\
&= 0
\end{aligned}
$$

In the same way

$$
\begin{aligned}
(A\Sigma + \Sigma\tilde{A} + B\tilde{B})_{ij} &= a_{ij}\sigma_j + \sigma_i\alpha_{ji} + b'_i b'_j + b''_i b''_j \\
&= -\frac{\varepsilon_i(1-\sigma_j^2)^{1/2}\sigma_j}{\varepsilon_i\sigma_i(1-\sigma_j^2)^{1/2} + \varepsilon_j\sigma_j(1-\sigma_i^2)^{1/2}} p_{i,n-1}p_{j,n-1} \\
&\quad - \frac{\varepsilon_j(1-\sigma_i^2)^{1/2}\sigma_i}{\varepsilon_i\sigma_i(1-\sigma_j^2)^{1/2} + \varepsilon_j\sigma_j(1-\sigma_i^2)^{1/2}} p_{i,n-1}p_{j,n-1} \\
&\quad + \varepsilon_i\varepsilon_j(1-\sigma_i^2)^{1/2}(1-\sigma_j^2)^{1/2} + \sigma_i\sigma_j p_{i,n-1}p_{j,n-1} = 0 \qquad \square
\end{aligned}
$$

Following Jonckheere and Silverman (1983), we define LQG balancing. Since we have to deal with non-strictly proper systems in later parts of this paper, we give the general definition for non-strictly proper systems (see Ober 1989).

Definition 11.2: A minimal system (A, B, C, D) is called *LQG balanced* if there exists a diagonal matrix $\Sigma = \text{diag}(\mu_1, \ldots, \mu_n) > 0$, $\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n > 0$, such that

$$
\left.
\begin{aligned}
(A - BS^{-1}D^TC)^T\Sigma + \Sigma(A - BS^{-1}D^TC) - \Sigma BS^{-1}B^T\Sigma + C^T R^{-1}C = 0 \\
(A - BS^{-1}D^TC)\Sigma + \Sigma(A - BS^{-1}D^TC)^T - \Sigma C^T R^{-1}C\Sigma + BS^{-1}B^T = 0
\end{aligned}
\right\} \tag{221}
$$

where $R = I + DD^T$ and $S = I + D^TD$. The diagonal entries of the matrix Σ are called the *LQG singular values* of the system.

We now come to derive a LQG balanced realization of the transfer fucntion g. Again this realization will be shown to be the matrix representation of the shift realization with respect to abasis that is constructed from the Schmidt vectors of H_{R^*}.

Theorem 11.2: *Let $g = e/d$ and let e/t, d/t be the normalized coprime factors of g. Let $R^* = r^*/t^*$ be the function associated with the LQG controller. Assume the singular values of H_{R^*} are $\mu_1 > \ldots > \mu_n > 0$. Let $\left\{\frac{p_i}{t}, \varepsilon_i\frac{p_i^*}{t^*}\right\}$ be the μ_i-Schmidt pairs of H_{R^*}. Then*

(1) $\left\{\frac{p_i}{d}\right\}_{t=1}^{n}$ *is a basis for X^d.*

(2) *If we normalize the basis so that*

$$\left\|\frac{p_i}{t}\right\|^2 = \sigma_i(1-\sigma_i^2)^{1/2} \tag{222}$$

then the matrix represenation of the shift realization (4) of g with respect to the basis $\left\{\frac{p_i}{d}\right\}_{i=1}^{n}$ *is LQG balanced. Specifically we have*

$$\left.\begin{aligned} a_{ij} &= -\varepsilon_j p_{i,n-1}p_{j,n-1}\left(\frac{1-\lambda_i\lambda_j}{\lambda_i+\lambda_j}\right) \\ b_i &= \varepsilon_i p_{i,n-1} \\ c_i &= p_{i,n-1} = \varepsilon_i b_i \end{aligned}\right\} \tag{223}$$

where $\lambda_i = \varepsilon_i \mu_i$.

(3) *The previous LQG balanced realization is signature symmetric. Specifically, with* $J = \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_n)$ *we have*

$$\begin{aligned} JA &= \tilde{A}J \\ \tilde{C} &= JB \end{aligned} \tag{224}$$

(4) *With respect to the constructed LQG balanced realization we have*

$$\frac{p_i}{d} = C(Zi - A)^{-1}e_i \tag{225}$$

Proof:

(1) Since d and t are of equal degreee the spaces X^d and X^t have the same direction. Moreover the multiplication map by t/d is an invertible map of X^t on X^d. Since the set $\left\{\frac{p_i}{t} | i = 1, \ldots, n\right\}$ is a basis for X^t it follows that $\left\{\frac{p_i}{d} | i = 1, \ldots, n\right\}$ is a basis for X^d.

(2) We will compute the matrix representation of the shift realization of $g = e/d$ in this basis. To this end we turn X^d into a Hilbert space by introducing in it an inner product through

$$\left[\frac{p}{d}, \frac{q}{d}\right] = \left(\frac{p}{t}, \frac{q}{t}\right)_{H_+^2} \tag{226}$$

This definition is equivalent to considering X^d as a subspace of an L^2 space relative to the measure, or weight, given by $\|d/t\|^2$. Note that the vectors p_i/d, $1 \leq i \leq n$, are orthogonal with respect to the inner product $[\cdot,\cdot]$. To compute the input map we put

$$\frac{e}{d} = \sum_{i=1}^{n} b_i \frac{p_i}{d}$$

so

$$b_i = \frac{\left[\frac{e}{d}, \frac{p_i}{d}\right]}{\left[\frac{p_i}{d}, \frac{p_i}{d}\right]} = \frac{\left(\frac{e}{t}, \frac{p_i}{t}\right)}{\left(\frac{p_i}{t}\right)}$$

We use now the normalization $\|p_i/t\|^2 = \sigma_i(1-\sigma_i^2)^{1/2}$ and the computation of $\left(\frac{e}{t}, \frac{p_i}{t}\right)$ carried out in the proof of Theorem 11.1 to get

$$b_i = \frac{\left(\frac{e}{t}, \frac{p_i}{t}\right)}{\left(\frac{p_i}{t}\right)} = -\frac{(-1)^{n-1}(-1)^n \varepsilon_i \sigma_i (1-\sigma_i^2)^{1/2} p_{i,n-1}}{\sigma_i(1-\sigma_i^2)^{1/2}}$$

$$= \varepsilon_i p_{i,n-1}$$

The computation of the c_i is easy as

$$c_i = C\frac{p_i}{d} = \left(\frac{p_i}{d}\right)_{-1} = p_{i,n-1} \tag{227}$$

To compute the generator matrix we set

$$S^d \frac{p_i}{d} = \sum_{j=1}^{n} a_{ji} \frac{p_j}{d}$$

Now

$$S^d \frac{p_i}{d} = \pi^d \frac{zp_i}{d} = \frac{zp_i - \xi_i d}{d}$$

where $\xi_i = p_{i,n-1}$. Utilizing out inner product as well as the orthogonality of the set $\{p_i/t \mid i = 1, \ldots, n\}$, we have

$$a_{ji} = \frac{\left[S^d \frac{p_i}{d}, \frac{p_j}{d}\right]}{\left[\frac{p_j}{d}, \frac{p_j}{d}\right]} = \frac{\left[\frac{zp_i - \xi_i d}{d}, \frac{p_j}{d}\right]}{\left[\frac{p_j}{d}, \frac{p_j}{d}\right]}$$

$$= \frac{\left(\frac{zp_i - \xi_i d}{t}, \frac{p_j}{t}\right)}{\left(\frac{p_j}{t}, \frac{p_j}{t}\right)}$$

$$= \frac{\left(\frac{zp_i - \xi_i d}{t}, \frac{p_j}{t}\right)}{\sigma_j(1-\sigma_j^2)^{1/2}}$$

We proceed now to compute the numerator. Since $\deg(zp_i - \xi_i d)p_j^* \leq 2n-2$

a standard estimate on contour integrals yields

$$\left(\frac{zp_i - \xi_i d}{t}, \frac{p_j}{t}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{zp_i - \xi_i d}{t}\,\frac{p_j^*}{t^*}\,\mathrm{d}\tau$$

$$= \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\gamma R} \frac{zp_i - \xi_i d}{t}\,\frac{p_j^*}{t^*}\,\mathrm{d}z$$

We compute the integral on the contour γ_R, with R large enough. Again we use (85) to get

$$\frac{zp_i p_j^*}{tt^*} = \frac{1}{\lambda_i^2 - \lambda_j^2}\left\{\lambda_j\,\frac{z\alpha_{ij}}{t} + \lambda_i\,\frac{z\alpha_{ij}^*}{t^*}\right\} \tag{229}$$

So, using the stability of t,

$$\frac{1}{2\pi i}\int_{\gamma R} \frac{zp_i}{t}\,\frac{p_j^*}{t^*}\,\mathrm{d}z = \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2}\,\frac{1}{2\pi i}\int_{\gamma R} \frac{z\alpha_{ij}}{t}\,\mathrm{d}z \tag{230}$$

Deforming γ_R to a large circular, positively orientated, contour and expanding the integrand at ∞, we get, using (203) that

$$\begin{aligned}\frac{1}{2\pi i}\int_{\gamma R} \frac{zp_i}{t}\,\frac{p_j^*}{t^*}\,\mathrm{d}z &= \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2}\,\alpha_{ij,n-2}\\ &= \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2}\,(-1)(\lambda_i - \lambda_j)p_{i,n-1}p_{j,n-1}\\ &= -\frac{\lambda_j}{\lambda_i + \lambda_j}\,p_{i,n-1}p_{j,n-1}\end{aligned} \tag{231}$$

To compute the integral $\frac{1}{2\pi i}\int_{\gamma R} \frac{dp_j}{tt^*}\,\mathrm{d}z$ we use the singular value/singular vector equation (128) which implies

$$\frac{dp_j^*}{tt^*} = \sigma_j\,\frac{(\hat{p}_2^{(j)})^*}{t} + \frac{(\pi_2^{(j)})^*}{t^*} \tag{232}$$

Hence

$$\begin{aligned}\frac{1}{2\pi i}\int_{\gamma R} \frac{dp_j^*}{tt^*}\,\mathrm{d}z &= \frac{1}{2\pi i}\int_{\gamma R}\left\{\sigma_j\,\frac{(\hat{p}_2^{(j)})^*}{t} + \frac{(\pi_2^{(j)})^*}{t^*}\right\}\mathrm{d}z\\ &= \frac{\sigma_j}{2\pi i}\int_{\gamma R}\frac{(\hat{p}_2^{(j)})^*}{t}\,\mathrm{d}z = \sigma_j(-1)^{n-1}\hat{p}_{2,n-1}^{(j)}\\ &= (-1)^{n-1}\sigma_j(-1)^n\sigma_j p_{j,n-1}\\ &= -\sigma_j^2 p_{j,n-1}\end{aligned} \tag{233}$$

where we have used (206). Equations (231) and (233) taken together with $\xi_i = p_{i,n-1}$ imply

$$\begin{aligned}\left(\frac{zp_i - \xi_i d}{t}, \frac{p_j}{t}\right) &= -\frac{\lambda_j}{\lambda_i + \lambda_j}\,p_{i,n-1}p_{j,n-1} - p_{i,n-1}(-1)\sigma_j^2 p_{j,n-1}\\ &= \frac{1}{\lambda_i + \lambda_j}\,p_{i,n-1}p_{j,n-1}\{-\lambda_j + \sigma_j^2(\lambda_i + \lambda_j)\}\end{aligned}$$

$$= \frac{1}{\lambda_i + \lambda_j} p_{i,n-1} p_{j,n-1} \{-\lambda_j(1 - \sigma_j^2) + \lambda_i \sigma_j^2\}$$

$$= \frac{(1 - \sigma_j^2)}{\lambda_i + \lambda_j} p_{i,n-1} p_{j,n-1} \left\{-\lambda_j + \lambda_i \frac{\sigma_j^2}{(1 - \sigma_j^2)}\right\}$$

$$= \frac{(1 - \sigma_j^2)}{\lambda_i + \lambda_j} p_{i,n-1} p_{j,n-1} \{-\lambda_j + \lambda_i \lambda_j^2\}$$

$$= - \frac{\lambda_j(1 - \sigma_j^2)}{\lambda_i + \lambda_j} p_{i,n-1} p_{j,n-1} \{1 - \lambda_i \lambda_j\}$$

and so

$$a_{ji} = -\varepsilon_j p_{i,n-1} p_{j,n-1} \frac{1 - \lambda_i \lambda_j}{\lambda_i + \lambda_j} \tag{234}$$

Summing up, we have the matrix representation

$$\left.\begin{aligned} a_{ji} &= -\varepsilon_j p_{i,n-1} p_{j,n-1} \left(\frac{1 - \lambda_i \lambda_j}{\lambda_i + \lambda_j}\right) \\ b_i &= \varepsilon_i p_{i,n-1} \\ c_i &= p_{i,n-1} \end{aligned}\right\} \tag{235}$$

Now it is trivial to check directly that this is a LQG balanced realization, with $\Sigma = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n)$.

(3) From the realization (235) it follows immediately that, with $j = \operatorname{diag}(\varepsilon_1, \ldots \varepsilon_n)$,

$$\begin{aligned} JA &= \tilde{A} J \\ \tilde{C} &= JB \end{aligned} \tag{236}$$

i.e. the realization is signature symmetric.

(4) The proof follows from Theorem 8.1.4 in Fuhrmann (1991). □

It is worthwhile noting how the parameters of the realization of

$$\left(-\frac{e}{t}, \frac{d}{t}\right)$$

are related to the parameters of the realization of $g = e/d$. The parameters λ_i, ε_i, σ_i, $1 \leq i \leq n$, are the same for both realizations. The parameters $p_{i,n-1}$ are different, but related as follows,

$$\hat{p}_{i,n-1} = (1 - \sigma_i^2)^{1/4} \hat{p}_{i,n-1}$$

where $\hat{p}_{i,n-1}$ are the parameters used in the realization of

$$\left(-\frac{e}{t}, \frac{d}{t}\right)$$

and $\hat{p}_{i,n-1}$ are the parameters used in the realization of

$$\left(\frac{e}{d}\right)$$

We now come to derive a Lyapunov balanced realization of R. Before during this we quote the following result from Fuhrmann (1991) in which it is shown that a Lyapunov balanced realization can be seen to be the matrix representation of a shift realization with respect to a basis made up from suitably normalized Schmidt vectors.

Proposition 11.2: *Let $\phi = n/d \in H^\infty$ with d monic. Let $\eta_1 > \eta_2 > \cdots > \eta_n > 0$ be the singular values of H_ϕ with singular vectors $\left\{\frac{q_i}{d}, \varepsilon_i \frac{q_i^*}{d^*}\right\}$, $\varepsilon_i = \pm 1$, $i = 1, 2, \ldots, n$. Assume that the $\{q_i/d\}$ are normalized such that*

$$\left\|\frac{g_i}{d}\right\|^2 = \eta_i$$

$i = 1, 2, \ldots, n$. Then the matrix representation (A, B, C, D) of the shift realization of ϕ^* with respect to the basis $\{q_i/d\}$, $i = 1, \ldots, n$, is given by

$$A = \left(\frac{\varepsilon_j b_i b_j}{\varepsilon_i \eta_i + \varepsilon_j \eta_j}\right)_{1 \leqslant i,j \leqslant n}$$

$$B = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$$

$$C = -(\varepsilon_1 b_1, \varepsilon_2 b_2, \ldots, \varepsilon_n b_n)$$

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

where $b_i = \varepsilon_i g_{i,n-1}$, and $q_{i,n-1}$ is the leading coefficient of q_i. Moreover, $(-A^*, -C^*, B^*, D^*)$ is a Lyapunov balanced realization of ϕ with Lyapunov gramian $\Sigma = \operatorname{diag}(\eta_1, \eta_2, \ldots, \eta_n)$.

We can now apply this result to obtain a Lyapunov balanced realization of R.

Corollary 11.1: *Assume $\mu_1 > \mu_2 > \ldots > \mu_n > 0$, then the matrix representation of the shift realization of R^* with respect to the basis $\{p_i/t^*\}$, which is normalized such that*

$$\left\|\frac{p_i^*}{t^*}\right\|^2 = \mu_i$$

$1 \leqslant i \leqslant n$, is given by

$$\left[\begin{array}{c|c} \left(\dfrac{\varepsilon_j b_i b_j}{\lambda_i + \lambda_j}\right) & b_i \\ \hline -\varepsilon_j b_j & 0 \end{array}\right], \quad i, j = 1, \ldots, n \tag{237}$$

where $\lambda_i = \varepsilon_i \mu_i$ and $b_i = \varepsilon_i p_{i,n-1}$. The parameters $p_{i,n-1}$ are the leading coefficients of the polynomials p_i, $1 \leqslant i \leqslant n$. The realization

$$\left[\begin{array}{c|c} \left(\dfrac{\varepsilon_j b_i b_j}{\lambda_i + \lambda_j}\right) & b_i \\ \hline \varepsilon_j b_j & 0 \end{array}\right], \quad i, j = 1, \ldots, n \tag{238}$$

is a Lyapunov balanced realization of R with gramian $\Sigma = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n)$.

We would like to point out one result that deserves further study. Starting with an arbitrary transfer function g of McMillan degree n, we have associated with it a unique stable transfer function, namely the function R^*. This function was constructed from the normalized coprime factorization of g and a factorization of the LQG controller of g, see § 4. Now, as a result of Theorem 11.2, this map can be inverted. In fact, to get back from r/t to g all we have to do is to take a Lyapunov balanced realizations of r/t and use the LQG realization (223) to reconstruct g.

The net effect of this is that we have constructed a bijective map of the space of all rational functions of McMillan degree n, i.e. Rat(n), onto the space of all stable transfer functions of the same McMillan degree. It would be of interest to study further the topological properties of this map.

12. The approximants

The present paper can be seen as focusing on a study of the three functions

$$G, \quad R^*, \quad \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

and the relationships between them and between the various operators that are associated with these functions. In § 9 we have studied the Hankel norm approximation problem of

$$R^* \quad \text{and} \quad \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

An interesting result of this study was that $(1-\sigma_n^2)^{1/2}$ times the complex conjugate of the Nankel norm approximant of

$$\begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix} = \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}, \quad \text{i.e.} \frac{1}{(1-\sigma_n^2)^{1/2}} (\pi_1^{(n)^*}/p_n^* \; \pi_2^{(n)^*}/p_n^*)$$

is a normalized coprime factorization of the function $g_n^* := \pi_1^{(n)^*}/\pi_2^{(n)^*}$. We can therefore initiate the same study as we have done for g and its associated functions, now for g_n^* and its normalized coprime factorization $[\bar{N}_n \; \bar{M}_n] := 1/(1-\sigma_n^2)^{1/2} \; [\pi_1^{(n)^*}/p_n^* \pi_2^{(n)^*}/p_n^*]$. In particular we can construct the coprime factorization of its LQG controller and derive the associated function R_n^*, as in § 4. The question we are trying to answer in this section is how the triple of functions

$$g_n, \quad R_n^* \quad \text{and} \quad \begin{bmatrix} -\bar{N}_n^* \\ \bar{M}_n^* \end{bmatrix}$$

relates to the original triple

$$g, \quad R^* \quad \text{and} \quad \begin{bmatrix} -\bar{N}^* \\ \bar{M}^* \end{bmatrix}$$

One of the main results of this study will be that R_n^*, is, in fact, the strictly proper part of the $n-1$ order Hankel norm approximant of R^*. We will also obtain state-space representations for R_n^* and g_n in terms of the parameters in the LQG balanced state-space representation of g.

We need the following proposition.

Proposition 12.1: *Let $\alpha_1^{(i)}$, $\alpha_2^{(i)}$ be defined by Lemma 9.3 and $\zeta_1^{(i)}$, $\zeta_2^{(i)}$ by Proposition 9.1. Then the leading coefficients of the polynomials satisfy,*

(1) $$\alpha_{1,n-2}^{(i)} + \lambda_n \alpha_{2,n-2}^{(i)} = \frac{1}{\sigma_i}(1 - \sigma_i^2)(\lambda_i - \lambda_n) p_{i,n-1} p_{n,n-1} \tag{239}$$

(2) $$\zeta_{1,n-2}^{(i)} + \lambda_n \zeta_{2,n-2}^{(i)} = (-1)^n \frac{\lambda_i - \lambda_n}{1 + \lambda_i^2} p_{n,n-1} p_{i,n-1} \tag{240}$$

Proof: (1) From Equation (143), i.e.

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} p_i - \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix} p_n = \sigma_i t \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix}$$

we get

$$(\pi_1^{(n)} + \lambda_n \pi_2^{(n)}) p_i - (\pi_1^{(i)} + \lambda_n \pi_2^{(i)}) p_n = \sigma_i t (\alpha_1^{(i)} + \lambda_n \alpha_2^{(i)})$$

Comparing the highest-order coefficients,

$$(\pi_{1,n-1}^{(n)} + \lambda_n \pi_{2,n-1}^{(n)}) p_i - (\pi_{1,n-1}^{(i)} + \lambda_n \pi_{2,n-1}^{(i)}) p_n = \sigma_i (\alpha_{1,n-1}^{(i)} + \lambda_n \alpha_{2,n-1}^{(i)}) \tag{241}$$

We use now (207) and (207) and note that

$$\begin{aligned} \pi_{1,n-1}^{(i)} + \lambda_n \pi_{2,n-1}^{(i)} &= [-\varepsilon_i \sigma_i (1 - \sigma_i^2)^{1/2} + \lambda_n (1 - \sigma_i^2)] p_{i,n-1} \\ &= -(1 - \sigma_i^2)(\lambda_i - \lambda_n) p_{i,n-1} \end{aligned}$$

and therefore we also have

$$\pi_{1,n-1}^{(n)} + \lambda_n \pi_{2,n-1}^{(n)} = 0$$

Using these two identities we obtain (239) from (241).

(2) From (163), i.e.

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} \alpha_i = \sigma_i p_n^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} + p_n \begin{bmatrix} \zeta_1^{(i)} \\ \zeta_2^{(i)} \end{bmatrix}$$

we obtain

$$(\pi_1^{(n)} + \lambda_n \pi_2^{(n)}) \alpha_i = \sigma_i p_n^* (\alpha_1^{(i)} + \lambda_n \alpha_2^{(i)}) + p_n (\zeta_1^{(i)} + \lambda_n \zeta_2^{(i)})$$

Comparing highest degree coefficients we obtain

$$\begin{aligned} &(\pi_{1,n-1}^{(n)} + \lambda_n \pi_{2,n-1}^{(n)}) \alpha_{i,n-2} \\ &\quad = \sigma_i (-1)^{n-1} p_{n,n-1} (\alpha_{1,n-2}^{(i)} + \lambda_n \alpha_{2,n-2}^{(i)}) + p_{n,n-1} (\zeta_{1,n-2}^{(i)} + \lambda_n \zeta_{2,n-2}^{(i)}) \end{aligned}$$

The left-side term vanishes in view of (207) and (208). So

$$\zeta_{1,n-2}^{(i)} + \lambda_n \zeta_{2,n-2}^{(i)} = (-1)^n \sigma_i (\alpha_{1,n-2}^{(i)} + \lambda_n \alpha_{2,n-2}^{(i)})$$

From part (1) we therefore have

$$\zeta_{1,n-2}^{(i)} + \lambda_n \zeta_{2,n-2}^{(i)} + (-1)^n (1 - \sigma_i^2)(\lambda_i - \lambda_n) p_{n,n-1} p_{i,n-1}$$ □

In the following proposition results are collected on the Schmidt vectors of the Hankel operator $H\pi_n/p_n$.

Proposition 12.2: *Assume the notation of Theorem 9.1.*

(1) *There exist polynomials* ζ_i, $1 \leq i \leq n$, *such that*

$$\frac{\pi_n}{p_n}\frac{\alpha_i}{p_n^*} = \lambda_i \frac{\alpha_i^*}{p_n} + \frac{\zeta_i}{p_n^*}$$

with $\lambda_i = \varepsilon_i \mu_i$, $1 \leq i \leq n-1$.

(2) *There exist polynomials* ω_{ij} *of degree less than or equal to* $n-3$ *with the properties*

$$\omega_{ij} = -\omega_{ij},\ \omega_{ii} = 0$$

for $1 \leq i, j \leq n-1$, *such that*

(*a*) $\lambda_i \alpha_i^* \alpha_j - \lambda_j \alpha_j^* \alpha_i = p_n^* \omega_{ij}$

(*b*) *if* i, j *are such that* $\mu_i \neq \mu_j$, *then* ω_{ij} *is non-zero.*

(*c*) *if* i, j *are such that* $\mu_i \neq \mu_j$, *then*

$$\alpha_i \alpha_j^* = \frac{1}{\lambda_i^2 - \lambda_j^2}\{\lambda_j \omega_{ij} p_n + \lambda_i p_n^* \omega_{ij}^*\} \tag{242}$$

Proof:

(1) and (2). Applying Theorem 9.1 and setting up the singular value/singular vector equations to the case π_n/p_n and the polynomials α_i, $1 \leq i \leq n-1$ we obtain the existence of the polynomials ω_{ij} with the required properties. □

We now come to derive a number of important identities related to the Hankel norm approximant of $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$.

Theorem 12.1: *Let* $g = e/d$ *and let* e/t, d/t *be the normalized coprime factors of* g. *Let* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{n-1} > \sigma_n > 0$ *be the singular values of* $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$ *and let* $\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$ *be the optimal Hankel approximant corresponding to* σ_n. *Let* $\zeta_1^{(i)}$, $\zeta_2^{(i)}$ *be as defined by Proposition 9.1. Set*

$$[\bar{N}_n, \bar{M}_n,] := \begin{bmatrix} N_n \\ M_n \end{bmatrix}^{\mathrm{T}} := \frac{1}{(1-\sigma_n^2)^{1/2}} [\pi_1^{(n)*}/p_n^*\ \pi_2^{(n)*}/p_n^*]$$

and

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} := \frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)} \begin{bmatrix} -\zeta_1^{(i)}/\alpha_i \\ \zeta_2^{(i)}/\alpha_i \end{bmatrix} \in H^\infty_{[i-1]}$$

Then we have

(1) $\dfrac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \alpha_1^{(i)}/\alpha_i \\ \alpha_2^{(i)}/\alpha_i \end{bmatrix}$ *is all pass.*

(2)
$$(\alpha_1^{(i)})^* \zeta_1^{(i)} + (\alpha_2^{(i)})^* \zeta_2^{(i)} = 0 \tag{243}$$

(3) $$(\pi_1^{(n)})^*\zeta_1^{(i)} + (\pi_2^{(n)})^*\zeta_2^{(i)} = (1-\sigma_i^2)(1-\sigma_n^2)p_n^*\alpha_i \tag{244}$$

(4) $$\frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)^{1/2}}\begin{bmatrix}\zeta_1^{(i)}/\alpha_i\\ \zeta_2^{(i)}/\alpha_i\end{bmatrix} \tag{245}$$

is all pass.

(5) $$-\pi_2^{(n)}\zeta_1^{(i)} + \pi_1^{(n)}\zeta_2^{(i)} = -\lambda_i(1-\sigma_n^2)(1-\sigma_i^2)\alpha_i^*p_n^*$$

(6) $$\bar{M}_nV_i - \bar{N}_nU_i = I \tag{246}$$

(7) $$M_n^*U_i + N_n^*V_i = -\lambda_i\frac{p_n^*\alpha_i^*}{p_n\alpha_i} \tag{247}$$

Proof:

(1) Left multiplying Equation (163), i.e.

$$\begin{bmatrix}\pi_1^{(n)}\\ \pi_2^{(n)}\end{bmatrix}\alpha_i = \sigma_i p_n^*\begin{bmatrix}\alpha_1^{(i)}\\ \alpha_2^{(i)}\end{bmatrix} + p_n\begin{bmatrix}\zeta_1^{(i)}\\ \zeta_2^{(i)}\end{bmatrix}$$

by $((\alpha_1^{(i)})^*\ (\alpha_2^{(i)})^*)$ we have,

$$((\alpha_1^{(i)})^*\pi_1^{(n)} + (\alpha_2^{(i)})^*\pi_2^{(n)})\alpha_i = \alpha_i p_n^*((\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}) + p_n((\alpha_1^{(i)})^*\zeta_1^{(i)} + (\alpha_2^{(i)})^*\zeta_2^{(i)})$$

We use now Equation (164) to get

$$\alpha_i(1-\sigma_n^2)p_n^*\alpha_i^*\alpha_i = \sigma_i p_n^*((\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}) + p_n((\alpha_1^{(i)})^*\zeta_1^{(i)} + (\alpha_2^{(i)})^*\zeta_2^{(i)})$$

or

$$\alpha_i(1-\sigma_n^2)p_n^*\left[\alpha_i^*\alpha_i - \frac{(\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}}{1-\sigma_n^2}\right] = p_n((\alpha_1^{(i)})^*\zeta_1^{(i)} + (\alpha_2^{(i)})^*\zeta_2^{(i)}) \tag{248}$$

So, we get the division relation $p_n \mid \left[\alpha_i^*\alpha_i - \frac{(\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}}{1-\sigma_n^2}\right]$ and by symmetry, also $p_n^* \mid \left[\alpha_i^*\alpha_i - \frac{(\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}}{1-\sigma_n^2}\right]$. From degree considerations we obtain that

$$\left[\alpha_i^*\alpha_i - \frac{(\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}}{1-\sigma_n^2}\right] = 0$$

and hence that

$$(\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)} = (1-\sigma_n^2)\alpha_i^*\alpha_i.$$

(2) Follows from Equation (248) using that

$$\left[\alpha_i^*\alpha_i - \frac{(\alpha_1^{(i)})^*\alpha_1^{(i)} + (\alpha_2^{(i)})^*\alpha_2^{(i)}}{1-\sigma_n^2}\right] = 0$$

by the proof of (1).

(3) Multiplying Equation (163), i.e.

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} \alpha_i = \sigma_i p_n^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} + p_n \begin{bmatrix} \zeta_1^{(i)} \\ \zeta_2^{(i)} \end{bmatrix}$$

by $((\pi_1^{(n)})^* \ (\pi_2^{(n)})^*)$ we have,

$$((\pi_1^{(n)})^*\pi_1^{(n)} + (\pi_2^{(n)})^*\pi_2^{(n)})\alpha_i$$
$$= \sigma_i p_n^*((\pi_1^{(n)})^*\alpha_1^{(i)} + (\pi_2^{(n)})^*\alpha_2^{(i)}) + p_n((\pi_1^{(n)})^*\zeta_1^{(i)} + (\pi_2^{(n)})^*\zeta_2^{(i)})$$

Recalling that $\frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$ is all-pass and (164), i.e.

$$(\pi_1^{(n)})^*\alpha_1^{(i)} + (\pi_2^{(n)})^*\alpha_2^{(i)} = \alpha_i(1-\sigma_n^2)p_n\alpha_i$$

we get

$$(1-\sigma_n^2)p_n p_n^*\alpha_i = \sigma_i p_n^*\sigma_i(1-\sigma_n^2)p_n\alpha_i + p_n((\pi_1^{(n)})^*\zeta_1^{(i)} + (\pi_2^{(n)})^*\zeta_2^{(i)}$$

from which (244) follows.

(4) Left multiplying Equation (163) by $((\zeta_1^{(i)})^* \ (\zeta_2^{(i)})^*)$ we have,

$$((\zeta_1^{(i)})^*\pi_1^{(n)} + (\zeta_2^{(i)})^*\pi_2^{(n)})\alpha_i$$
$$= \sigma_i p_n^*((\zeta_1^{(i)})^*\alpha_1^{(i)} + (\zeta_2^{(i)})^*\alpha_2^{(i)}) + p_n((\zeta_1^{(i)})^*\zeta_1^{(i)} + (\zeta_2^{(i)})^*\zeta_2^{(i)})$$

We use now Equations (244) and (243) to obtain the equality

$$(1-\sigma_i^2)(1-\sigma_n^2)p_n\alpha_i^*\alpha_i = p_n((\zeta_1^{(i)})^*\zeta_1^{(i)} + (\zeta_2^{(i)})^*\zeta_2^{(i)})$$

This is equivalent to the statement.

(6) The result follows directly from (1). We note that both p_n^* and α_1 are stable polynomials.

(5) and (7) We prove this first for the case $i = 1$. Since

$$M_n^* = \frac{1}{(1-\sigma_n^2)^{1/2}} \frac{\pi_2^{(n)}}{p_n}, \quad N_n^* = \frac{1}{(1-\sigma_n^2)^{1/2}} \frac{\pi_1^{(n)}}{p_n}$$

we have

$$M_n^*U_1 + N_n^*V_1 = \frac{1}{(1-\sigma_n^2)(1-\sigma_1^2)} \frac{-\pi_2^{(n)}\zeta_1^{(1)} + \pi_1^{(n)}\zeta_2^{(2)}}{p_n\alpha_1} \tag{249}$$

Now the Hankel operator $H_{M_n^*U_1+N_n^*V_1}$ has the same singular vectors as

$$\frac{1}{(1-\sigma_n^2)^{1/2}} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$$

and these are $\{\alpha_i/p_n^*\}$. Also, the singular values of $\frac{1}{(1-\sigma_n^2)^{1/2}} H_{\begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}}$, are $\sigma_1 > \cdots > \sigma_{n-1}$ and so the singular values of $H_{M_n^*U_1+N_n^*V_1}$ are $\mu_1 > \cdots > \mu_{n-1}$. In particular

$$H_{M_n^* U_1 + N_n^* V_1} \frac{\alpha_1}{p_n^*} = s\mu_1 \frac{\alpha_1^*}{p_n} \tag{250}$$

with $s = \pm 1$. From Equation (163), namely

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} \alpha_i = \alpha_i p_n^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} + p_n \begin{bmatrix} \zeta_1^{(i)} \\ \zeta_2^{(i)} \end{bmatrix}$$

we get, left multiplying by $(-\pi_2^{(n)}\ \pi_1^{(n)})$,

$$0 = \sigma_i p_n^*(-\pi_2^{(n)}\ \alpha_1^{(i)} + \pi_1^{(n)}\alpha_2^{(i)}) + p_n(-\pi_2^{(n)}\zeta_1^{(i)} + \pi_1^{(n)}\pi_2^{(i)}) \tag{251}$$

Therefore we get the division relation $p_n^* | -\pi_2^{(n)}\zeta_1^{(i)} + \pi_1^{(n)}\zeta_2^{(i)}$, and hence the existence of a polynomial l_i such that

$$-\pi_2^{(n)}\zeta_1^{(i)} + \pi_1^{(n)}\zeta_2^{(i)} = l_i p_n^*$$

Using (250) we have for $i = 1$,

$$P_- \frac{1}{(1-\sigma_n^2)(1-\sigma_1^2)} \frac{-\pi_2^{(n)}\zeta_1^{(1)} + \pi_1^{(n)}\zeta_2^{(1)}}{p_n\alpha_1} \frac{\alpha_1}{p_n^*}$$

$$= \frac{1}{(1-\sigma_n^2)(1-\sigma_1^2)} P_- \frac{l_1 p_n^*}{p_n\alpha_1} \frac{\alpha_1}{p_n^*}$$

$$= \frac{1}{(1-\sigma_n^2)(1\ \sigma_1^2)} \frac{l_1}{p_n} = s\mu_1 \frac{\alpha_1^*}{p_n}$$

and this implies

$$l_1 = s\mu_1(1-\sigma_n^2)(1-\sigma_1^2)\alpha_1^*$$

and hence

$$-\ \pi_2^{(n)}\zeta_1^{(1)} + \pi_1^{(n)}\zeta_2^{(1)}\ s\mu_1(1-\sigma_n^2)(1-\sigma_1^2)\alpha_1^* p_n^* \tag{252}$$

We will show that $s = -\varepsilon_1$. We observe that from Equation (251) we get the equality

$$\frac{1}{(1-\sigma_n^2)(1-\sigma_1^2)} \frac{-\pi_2^{(n)}\zeta_1^{(1)} + \pi_1^{(n)}\zeta_2^{(1)}}{p_n\alpha_1} = - \frac{\sigma_1}{(1-\sigma_n^2)(1-\sigma_1^2)} \frac{p_n^*}{p_n} \frac{-\pi_2^{(n)}\alpha_1^{(1)} + \pi_1^{(n)}\alpha_2^{(1)}}{p_n\alpha_1}$$

At ∞, the left term has, by (252), the value $-s\mu_1$. However, this value can be evaluated also from the right-hand side. This, using (204), (207), (208), and (239), leads to

$$-s\mu_1 = -\frac{\sigma_1}{(1-\sigma_n^2)(1-\sigma_1^2)} (-1)^{n-1} \left[\frac{-\pi_{2,n-1}^{(n)}\alpha_{1,n-2}^{(1)} + \pi_{1,n-1}^{(n)}\alpha_{2,n-2}^{(1)}}{p_{n,n-1}\alpha_{1,n-2}}\right]$$

$$= (-1)^n \frac{\sigma_1}{(1-\sigma_n^2)(1-\sigma_1^2)} \left[-(1-\sigma_n^2)\frac{\alpha_{1,n-2}^{(1)}}{\alpha_{1,n-2}} - \varepsilon_n\sigma_n(1-\sigma_n^2)^{1/2}\frac{\alpha_{2,n-2}^{(1)}}{\alpha_{1,n-2}}\right]$$

$$= (-1)^{n-1} \frac{\sigma_1}{(1-\sigma_1^2)} \left[\frac{\alpha_{1,n-2}^{(1)} + \lambda_n\alpha_{2,n-2}^{(1)}}{\alpha_{1,n-2}}\right]$$

$$= \lambda_1$$

So $-s\mu_1 = \lambda_1$ and hence $s = -\varepsilon_1$. This proves the case $i = 1$. We also get,

$$M_n^* U_1 + N_n^* V_1 = -\lambda_1 \frac{p_n^* \alpha_1^*}{p_n \alpha_1}$$

i.e. for $i = 1$, Equation (247) is proved.

We proceed to the proof of the general case. By (12.2),

$$\frac{\pi_n}{p_n} = \lambda_1 \frac{p_n^* \alpha_1^*}{p_n \alpha_1} + \frac{\zeta_1}{\alpha_1} = \lambda_i \frac{p_n^* \alpha_i^*}{p_n \alpha_i} + \frac{\zeta_i}{\alpha_i}$$

So

$$\pi_-(M_n^* U_1 + N_n^* V_1) = -\pi_- \frac{\pi_n}{p_n}$$

We compute now

$$M_n^* U_i + N_n^* V_i = \frac{1}{(1 - \sigma_n^2)(1 - \sigma_i^2)} \frac{-\pi_2^{(n)} \zeta_1^{(i)} + \pi_1^{(n)} \zeta_2^{(i)}}{p_n \alpha_i}$$

Observe that

$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \in H_{[i-1]}^{\infty}$$

and solves the Bezout type equation $\bar{M}_n V_i - \bar{N}_n U_i = 1$. The general solution of this equation, with

$$\begin{bmatrix} Y_i \\ X_i \end{bmatrix} \in H_{[i-1]}^{\infty}$$

is given by

$$\begin{bmatrix} Y_i \\ X_i \end{bmatrix} = \begin{bmatrix} U_1 \\ V_1 \end{bmatrix} - \begin{bmatrix} M_n \\ N_n \end{bmatrix} q, \quad q \in H_{[i-1]}^{\infty}$$

We look now for the minimum norm solution. Clearly, using the fact that

$$\begin{bmatrix} M^* & N^* \\ -\bar{N} & \bar{M} \end{bmatrix}$$

is all-pass,

$$\left\| \begin{bmatrix} X_i \\ Y_i \end{bmatrix} \right\| = \left\| \begin{bmatrix} M_n^* U_1 + N_n^* V_1 - q \\ I \end{bmatrix} \right\|_{\infty} \leq \{1 + \mu_i^2\}^{1/2}$$

Now, using the previously obtained result,

$$M_n^* U_1 + N_n^* V_1 = -\lambda_1 \frac{p_n^* \alpha_1^*}{p_n \alpha_1} = -\left[\lambda_i \frac{p_n^* \alpha_i^*}{p_n \alpha_i} + \left(\frac{\zeta_i}{\alpha_i} - \frac{\zeta_1}{\alpha_1} \right) \right]$$

and so

$$M_n^* U_1 + N_n^* V_1 - q = -\left[\lambda_i \frac{p_n^* \alpha_i^*}{p_n \alpha_i} + \left(\frac{\zeta_i}{\alpha_i} - \frac{\zeta_1}{\alpha_1} \right) + q \right]$$

Therefore, the infimum of $\inf_{q \in H_{[i-1]}^{\infty}} \|M^* U_1 + N^* V_1 - q\|_{\infty}$ is μ_i and attained with

$$q = -\frac{\zeta_i}{\alpha_i} + \frac{\zeta_1}{\alpha_1}$$

This leads to

$$\begin{bmatrix} Y_i \\ X_i \end{bmatrix} = \begin{bmatrix} U_1 \\ V_1 \end{bmatrix} - \begin{bmatrix} M_n \\ N_n \end{bmatrix} q$$

$$= \frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_1^2)} \begin{bmatrix} -\zeta_1^{(1)}/\alpha_1 \\ \zeta_2^{(1)}/\alpha_1 \end{bmatrix} - \frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} (\pi_2^{(n)})^*/p_n^* \\ (\pi_1^{(n)})^*/p_n^* \end{bmatrix} \left(\frac{\zeta_1}{\alpha_1} - \frac{\zeta_i}{\alpha_i} \right)$$

However we will show also that

$$\begin{bmatrix} Y_i \\ X_i \end{bmatrix} = \frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)} \begin{bmatrix} -\zeta_1^{(i)}/\alpha_i \\ \zeta_2^{(i)}/\alpha_i \end{bmatrix}$$

Indeed, $\frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)} \begin{bmatrix} -\zeta_1^{(i)}/\alpha_i \\ \zeta_2^{(i)}/\alpha_i \end{bmatrix}$ is in $H^\infty_{[i-1]}$ and solves the Bezout type equation (246). Finally, by (245),

$$\frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)^{1/2}} \begin{bmatrix} -\zeta_1^{(i)}/\alpha_i \\ \zeta_2^{(i)}/\alpha_i \end{bmatrix}$$

is all-pass. So

$$\left\| \frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)} \begin{bmatrix} -\zeta_1^{(i)}/\alpha_i \\ \zeta_2^{(i)}/\alpha_i \end{bmatrix} \right\|_\infty = \frac{1}{(1-\sigma_i^2)^{1/2}} = (1+\mu_i^2)^{1/2}$$

Therefore, we necessarily have

$$\frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_i^2)} \begin{bmatrix} -\zeta_1^{(i)}/\alpha_i \\ \zeta_2^{(i)}/\alpha_i \end{bmatrix}$$

$$= \frac{1}{(1-\sigma_n^2)^{1/2}(1-\sigma_1^2)} \begin{bmatrix} -\zeta_1^{(1)}/\alpha_1 \\ \zeta_2^{(1)}/\alpha_1 \end{bmatrix} - \frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} (\pi_2^{(n)})^*/p_n^* \\ (\pi_1^{(n)})^*/p_n^* \end{bmatrix} \left(\frac{\zeta_1}{\alpha_1} - \frac{\zeta_i}{\alpha_i} \right)$$

This implies

$$\begin{aligned} M_n^* U_i + N_n^* V_i &= \frac{1}{(1-\sigma_n^2)(1-\sigma_i^2)} \frac{-\pi_2^{(n)}\zeta_1^{(i)} + \pi_1^{(n)}\zeta_2^{(i)}}{p_n\alpha_i} \\ &= \frac{1}{(1-\sigma_n^2)(1-\sigma_1^2)} \frac{-\pi_2^{(n)}\zeta_1^{(1)} + \pi_1^{(n)}\zeta_2^{(1)}}{p_n\alpha_1} - \left(\frac{\zeta_1}{\alpha_{\mathrm{i}}} - \frac{\zeta_i}{\alpha_i} \right) \\ &= -\lambda_1 \frac{p_n^*\alpha_1^*}{p_n\alpha_1} + \left(\frac{\zeta_i}{\alpha_i} - \frac{\zeta_1}{\alpha_1} \right) \\ &= -\lambda_i \frac{p_n^*\alpha_i^*}{p_n\alpha_i} \end{aligned}$$

and this proves (247) for all i.

The following corollary is one of the main results of this section. It shows that the function $-R_n^*$ is the strictly proper part of the $n-1$ degree Hankel norm approximant of R^*.

Corollary 12.1: *The function R_n^*, i.e. the strictly proper anti-stable part of*

$M_n^* U_n + N_n^* V_n$, is given by

$$R_n^* = -\pi_- \frac{\pi_n}{p_n}$$

i.e. the strictly proper part of $-\pi_n/p_n$.

Proof: The result was established in the proof of the previous theorem. □

We also need the following proposition.

Proposition 12.3:

(1) *The leading coefficient $\omega_{ij,n-3}$ of the polynomial ω_{ij}, defined in Proposition 12.2 is given by*

$$\omega_{ij,n-3} = (-1)^n \frac{1}{\lambda_i \lambda_j} (\lambda_i - \lambda_j)(\lambda_i - \lambda_n)(\lambda_j - \lambda_n) p_{i,n-1} p_{j,n-1} p_{n,n-1} \quad (253)$$

(2) *The leading coefficient $\alpha_{2,n-2}^{(i)}$ of the polynomial $\alpha_2^{(i)}$, defined in Lemma 9.3 is given by*

$$\alpha_{2,n-2}^{(i)} = \frac{\varepsilon_i(\lambda_i^2 - \lambda_n^2)}{\lambda_i(1 + \lambda_i^2)^{1/2}(1 + \lambda_n^2)} p_{i,n-1} p_{n,n-1} \quad (254)$$

Proof:

(1) In Proposition 12.2 it was shown that

$$\alpha_i \alpha_j^* = \frac{1}{\lambda_i^2 - \lambda_j^2} \{\lambda_j \omega_{ij} p_n + \lambda_i p_n^* \omega_{ij}^*\}$$

Equating the leading coefficients we obtain that

$$\begin{aligned}\alpha_{i,n-2}(-1)^{n-2}\alpha_{j,n-2} &= \frac{1}{\lambda_i^2 - \lambda_j^2} \{\lambda_j \omega_{ij,n-3} p_{n,n-1} + \lambda_i (-1)^{n-1} p_{n,n-1} (-1)^{n-3} \omega_{ij,n-3}\} \\ &= \frac{1}{\lambda_i - \lambda_j} \omega_{ij,n-3} p_{n,n-1}\end{aligned}$$

and

$$\begin{aligned}(-1)^{n-2}(-1)^{n-1} \frac{\lambda_i - \lambda_n}{\lambda_i} p_{n,n-1} p_{i,n-1} (-1)^{n-1} \frac{\lambda_j - \lambda_n}{\lambda_j} p_{n,n-1} p_{j,n-1} \\ = (-1)^{n-2} \frac{\lambda_i - \lambda_n}{\lambda_i} \frac{\lambda_j - \lambda_n}{\lambda_j} p_{n,n-1}^2 p_{i,n-1} p_{j,n-1} \\ = \frac{1}{\lambda_i - \lambda_j} \omega_{ij,n-3} p_{n,n-1}\end{aligned}$$

This shows that

$$\omega_{ij,n-3} = (-1)^{n-2} \frac{1}{\lambda_i \lambda_j} (\lambda_i - \lambda_j)(\lambda_i - \lambda_n)(\lambda_j - \lambda_n) p_{n,n-1} p_{i,n-1} p_{j,n-1}$$

(2) From Lemma 9.3 we have that

$$\sigma_i t^* \alpha_2^{(i)} = \sigma_i \hat{p}_2^{(i)} p_n - \sigma_n \hat{p}_2^{(n)} p_i$$

Evaluating the leading coefficients of these polynomials we have that

$$\begin{aligned}(-1)^n \sigma_i \alpha_{2,n-2}^{(i)} &= \sigma_i \hat{p}_{2,n-1}^{(i)} p_{n,n-1} - \sigma_n \hat{p}_{2,n-1}^{(n)} p_{i,n-1} \\ &= \sigma_i (-1)^n \sigma_i p_{i,n-1} p_{n,n-1} - \sigma_n (-1)^n \sigma_n p_{n,n-1} p_{i,n-1} \\ &= (-1)^n (\sigma_i^2 - \sigma_n^2) p_{i,n-1} p_{n,n-1}\end{aligned}$$

where we have used Proposition 11.1 and therefore

$$\begin{aligned}\alpha_{2,n-2}^{(i)} &= \frac{\sigma_i^2 - \sigma_n^2}{\sigma_i} p_{i,n-1} p_{n,n-1} \\ &= \frac{\varepsilon_i (\lambda_j^2 - \lambda_n^2)}{\lambda_i (1 + \lambda_i^2)^{1/} (1 + \lambda_n^2)} p_{i,n-1} p_{n,n-1}\end{aligned}$$ □

We are now going to collect further results on R_n^*. In particular, we are going to give a realization of R_n^* whose adjoint realization is a Lyapunov balanced realization of R_n.

Proposition 12.4:

(1) *The Hankel singular values of* $H_{R_n^*}$ *are* $\mu_1 \geq \ldots \geq \mu_{n-1}$ *and the Schmidt pairs are* $\left\{ \frac{\alpha_i^*}{p_n}, -\varepsilon_i \frac{\alpha_i}{p_n^*}, \right\}_{i=1}^{n-1}$.

(2) *Assume* $\mu_1 > \mu_2 > \cdots > \mu_{n-1} > 0$. *The matrix representation of the shift realization with respect to the basis* $\left\{ \frac{v_i \, \alpha_i^*}{p_n} \right\}$,

$$\nu_i = \frac{\mu_i}{(\mu_i^2 - \mu_n^2)^{1/2}} \tag{255}$$

which is normalized as

$$\left\| \nu_i \frac{\alpha_i^*}{p_n} \right\|^2 = \mu_i$$

$1 \leq i \leq n-1$, *is given by*

$$\left[\begin{array}{c|c} \left(\dfrac{\varepsilon_j \rho_i \rho_j b_i b_j}{\lambda_i + \lambda_j} \right) & \rho_i b_i \\ \hline \varepsilon_j \rho_j b_j & 0 \end{array} \right], \quad i, j = 1, \ldots, n-1 \tag{256}$$

where

$$\rho_i = \left(\frac{\lambda_i - \lambda_n}{\lambda_i + \lambda_n} \right)^{1/2} \quad \lambda_i = \varepsilon_i \mu_i$$

and $b_i = \varepsilon_i p_{i,n-1}$. *The parameters* $p_{i,n-1}$ *are the leading coefficients of the polynomials* p_i *which are normalized such that* $\|p_i/t\|^2 = \mu_i$, $1 \leq i \leq n-1$. *The realization*

$$\left[\begin{array}{c|c} \left(\dfrac{-\varepsilon_j \rho_i \rho_j b_i b_j}{\lambda_i + \lambda_j} \right) & \rho_i b_i \\ \hline -\varepsilon_j \rho_j b_j & 0 \end{array} \right], \quad i, j = 1, \ldots, n-1 \tag{257}$$

is a Lyapunov balanced realization of R_n with gramian $\Sigma = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_{n-1})$.

Proof:

(1) With slight modifications this is Theorem 5.1 in Fuhrmann (1991).

(2) We will use here the normalization $\|p_i/t\|^2 = \mu_i$. From the equality

$$\left\|\frac{\alpha_i}{p_n^*}\right\|^2 = \left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\left\|\frac{p_i}{t}\right\|^2$$

we can determine the constant v_i which is such that

$$\mu_i = \left\|v_i^2 \frac{\alpha_i}{p_n^*}\right\|^2 = v_i^2\left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\mu_i$$

to be

$$v_i = \left(\frac{\mu_i^2}{\mu_i^2 - \mu_n^2}\right)^{1/2}$$

In order to obtain a Lyapunov balanced realization of R_n we have to determine the leading coefficients $\hat{\alpha}_{i,n-2}$ of the numerator polynomials of the Schmidt vectors assuming that the denominator polynomials are monic. We have that

$$\begin{aligned}
\hat{\alpha}_{i,n-2} &= v_i \frac{(-1)^{n-2}\alpha_{i,n-2}}{p_{n,n-1}} \\
&= \frac{1}{p_{n,n-1}}\left(\left[\frac{\mu_i^2}{\mu_i^2 - \mu_n^2}\right]^{1/2}(-1)^{n-2}(-1)^{n-1}\frac{\lambda_i - \lambda_n}{\lambda_i}\, p_{n,n-1}p_{i,n-1}\right) \\
&= -\left[\frac{\lambda_i - \lambda_n}{\lambda_i + \lambda_n}\right]^{1/2} p_{i,n-1} \\
&= -\rho_i\varepsilon_i b_i
\end{aligned}$$

Therefore, by Proposition 11.2 the shift realization has a matrix representation given by

$$\begin{aligned}
A &= \left(\frac{\varepsilon_j\rho_i\rho_j b_i b_j}{\lambda_i + \lambda_j}\right)_{1\le i,j\le n-1} \\
B &= (\ldots, \rho_i b_i, \ldots)^{\mathrm{T}} \\
C &= (\ldots, \rho_i b_i, \ldots) \\
D &= 0
\end{aligned}$$

It is straightforward to verify that $(A^*, -C^*, B^*, D^*)$ is a Lyapunov balanced realization of R_n with gramian $\Sigma = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_{n-1})$. □

We also need a lemma in which a connection between the various parameters is established.

Lemma 12.1: *Assume that i is such that $\mu_i > \mu_n$ and let*

$$v_i = \frac{\mu_i}{(\mu_i^2 - \mu_n^2)^{1/2}}$$

and

$$\rho_i = \left(\frac{\lambda_i - \lambda_n}{\lambda_i + \lambda_n}\right)^{1/2} \quad 1 \leq i \leq n-1$$

then

$$v_i \frac{\alpha_{i,n-2}}{\pi_{2,n-1}^{(n)}} = (-1)^{n-1}(1+\lambda_n^2)\rho_i p_{i,n-1}$$

Proof:

$$\begin{aligned} v_i \frac{\alpha_{i,n-2}}{\pi_{2,n-1}^{(n)}} &= \frac{\mu_i}{(\mu_i^2 - \mu_n^2)^{1/2}} (-1)^{n-1} \frac{\lambda_i - \lambda_n}{\lambda_i} p_{n,n-1} p_{i,n-1} \frac{1}{(1-\sigma_n^2) p_{n,n-1}} \\ &= (-1)^{n-1} \frac{1}{1-\sigma_n^2} \left(\frac{\lambda_i - \lambda_n}{\lambda_i + \lambda_n}\right)^{1/2} p_{i,n-1} \\ &= (-1)^{n-1}(1+\lambda_n^2)\rho_i p_{i,n-1} \end{aligned}$$

□

In the following theorem the function $g_n := \pi_1^{(n)}/\pi_2^{(n)}$ is examined. In particular, a matrix representation of the shift realization will be established.

Theorem 12.2:

(1) $\left\{ v_i \frac{\alpha_i}{\pi_2^{(n)}} \right\}_{i=1}^{n-1}$ *is a basis of* $X^{\pi_2^{(n)}}$.

(2) *If* $\mu_1 > \mu_2 > \cdots > \mu_{n-1} > \mu_n > 0$, *then the matrix representation of the shift realization of* $g_n = \pi_1^{(n)}/\pi_2^{(n)}$ *with respect to this basis is given by*

$$\left[\begin{array}{c|c} -\varepsilon_i \rho_i \rho_j b_i b_j \left[\frac{(1-\lambda_i \lambda_j)}{\lambda_i + \lambda_j} - \lambda_n\right] & -\rho_i b_i \\ \hline -(1+\lambda_n^2)\varepsilon_i \rho_i b_i & -\lambda_n \end{array}\right] \tag{258}$$

Here

$$\rho_i = \left(\frac{\lambda_i - \lambda_n}{\lambda_i + \lambda_n}\right)^{1/2} \quad \lambda_i = \varepsilon_i \mu_i$$

and $b_i = \varepsilon_i p_{i,n-1}$. *The parameters* $p_{i,n-1}$ *are the leading coefficients of the polynomials* p_i *which are normalized such that*

$$\left\| \frac{p_i}{t} \right\| = \sigma_i (1-\sigma_i^2)^{1/2} \quad 1 \leq i \leq n-1$$

Proof:

(1) and (2) we compute the matrix representation of the shift realization. Since $\left\{ \frac{\alpha_i^*}{p_n} \right\}_{i=1}^{n-1}$ is a basis of X^{p_n} then $\left\{ v_i \frac{\alpha_i^*}{\pi_2^{(n)}} \right\}_{i=1}^{n-1}$ is a basis for $X^{\pi_2^{(n)}}$,

with constants v_i. We introduce the inner product

$$\left[v_i \frac{\alpha_i^*}{\pi_2^{(n)}}, v_j \frac{\alpha_j^*}{\pi_2^{(n)}}\right] := \left(v_i \frac{\alpha_i^*}{p_n}, v_j \frac{\alpha_j^*}{p_n}\right)$$

In Theorem 11.2 we used the normalization $\|p_i/t\|^2 = \sigma_i(1-\sigma_i^2)^{1/2}$. Here we will use the same normalization, as well as the equality $\left\|\frac{\alpha_i}{p_n^*}\right\|^2 = \left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\left\|\frac{p_i}{t}\right\|^2$. We therefore have to determine the constant v_i such that

$$\left[v_i \frac{\alpha_i^*}{\pi_2^{(n)}}, v_i \frac{\alpha_i^*}{\pi_2^{(n)}}\right] = \sigma_i(1-\sigma_i^2)^{1/2}$$

We have that

$$\left(v_i \frac{\alpha_i^*}{\pi_2^{(n)}}, v_i \frac{\alpha_i^*}{\pi_2^{(n)}}\right] = v_i^2 \left\|\frac{\alpha_i}{p_n^*}\right\|^2 = v_i^2\left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\left\|\frac{p_i}{t}\right\|^2$$

$$= v_i^2\left(1 - \frac{\mu_n^2}{\mu_i^2}\right)\sigma_i\,(1-\sigma_i^2)^{1/2}$$

This implies

$$v_i = \frac{\mu_i}{(\mu_i^2 - \mu_n^2)^{1/2}} \tag{259}$$

and

$$\left\|\frac{\alpha_i}{p_n^*}\right\|^2 = \frac{\varepsilon_i(\lambda_i^2 - \lambda_n^2)}{\lambda_i(1+\lambda_i^2)}$$

Using Proposition 11.1 we compute the constant term of the realization

$$D = g_n(\infty) = \frac{\pi_{1,n-1}^{(n)}}{\pi_{2,n-1}^{(n)}} = -\frac{\varepsilon_n\sigma_n(1-\sigma_n^2)^{1/2}p_{n,n-1}}{(1-\sigma_n^2)p_{n,n-1}} = -\varepsilon_n\mu_n = -\lambda_n$$

This implies that $\dfrac{\pi_1^{(n)} + \lambda_n\pi_2^{(n)}}{\pi_2^{(n)}}$ is strictly proper.

To compute the output map we note that,

$$c_i' = C\left(v_i \frac{\alpha_i^*}{\pi_2^{(n)}}\right)_{-1} = v_i(-1)^{n-2}\frac{\alpha_{i,n-2}}{\pi_{2,n-1}^{(n)}} = -(1+\lambda_n^2)\rho_i p_{i,n-1}$$

or, as $b_i = \varepsilon_i p_{i,n-1}$, we have that

$$c_i' = -(1+\lambda_n^2)\varepsilon_i\rho_i b_i \tag{260}$$

Here $\rho_i = \left(\dfrac{\lambda_i - \lambda_n}{\lambda_i + \lambda_n}\right)^{1/2}$ Note that since $|\lambda_n| < |\lambda_i|$, the term under the square root is positive.

To compute the input map we put

$$\frac{\pi_1^{(n)} + \lambda_n\pi_2^{(n)}}{\pi_2^{(n)}} = \sum_{i=1}^{n-1} b_i' v_i \frac{\alpha_i^*}{\pi_2^{(n)}}$$

So

$$b_i' = \frac{\left[\dfrac{\pi_1^{(n)} + \lambda_n \pi_2^{(n)}}{\pi_2^{(n)}}, \nu_i \dfrac{\alpha_i^*}{\pi_2^{(n)}}\right]}{\nu_i^2 \left[\dfrac{\alpha_i^*}{\pi_2^{(n)}}, \dfrac{\alpha_i^*}{\pi_2^{(n)}}\right]} = \frac{\left(\dfrac{\pi_1^{(n)} + \lambda_n \pi_2^{(n)}}{p_n}, \nu_i \dfrac{\alpha_i^*}{p_n}\right)}{\nu_i^2 \left(\dfrac{\alpha_i^*}{p_n}, \dfrac{\alpha_i^*}{p_n}\right)}$$

$$= \frac{1 + \lambda_i^2}{(\lambda_i^2 - \lambda_n^2)^{1/2}} \left(\frac{\pi_1^{(n)} + \lambda_n \pi_2^{(n)}}{p_n}, \frac{\alpha_i^*}{p_n}\right)$$

$$= \frac{1 + \lambda_i^2}{(\lambda_i^2 - \lambda_n^2)^{1/2}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(\pi_1^{(n)} + \lambda_n \pi_2^{(n)})\alpha_i}{p_n p_n^*} \, d\tau$$

From (163), i.e.

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} \alpha_i = \sigma_i p_n^* \begin{bmatrix} \alpha_1^{(i)} \\ \alpha_2^{(i)} \end{bmatrix} + p_n \begin{bmatrix} \zeta_1^{(i)} \\ \zeta_2^{(i)} \end{bmatrix}$$

we get the equality

$$(\pi_1^{(n)} + \lambda_n \pi_2^{(n)})\alpha_i = \sigma_i p_n^*(\alpha_1^{(i)} + \lambda_n \alpha_2^{(i)}) + p_n(\zeta_1^{(i)} + \lambda_n \zeta_2^{(i)})$$

or

$$\frac{(\pi_1^{(n)} + \lambda_n \pi_2^{(n)})\alpha_i}{p_n p_n^*} = \sigma_i \frac{\alpha_1^{(i)} + \lambda_n \alpha_2^{(i)}}{p_n} + \frac{\zeta_1^{(i)} + \lambda_n \zeta_2^{(i)}}{p_n^*}$$

Integrating over the contour γ_R we get

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(\pi_1^{(n)} + \lambda_n \pi_2^{(n)})\alpha_i}{p_n p_n^*} = \lim_{R\to\infty} \frac{1}{2\pi i} \int_{\gamma_R} \frac{\zeta_1^{(i)} + \lambda_n \zeta_2^{(i)}}{p_n^*} \, dz$$

$$= (-1)^n \frac{\zeta_{1,n-2}^{(i)} + \lambda_n \zeta_{2,n-2}^{(i)}}{p_{n,n-1}} \tag{261}$$

$$= -\frac{\lambda_i - \lambda_n}{1 + \lambda_i^2} p_{i,n-1}$$

Therefore,

$$b_i' = -\frac{1 + \lambda_i^2}{(\lambda_i^2 - \lambda_n^2)^{1/2}} \frac{\lambda_i - \lambda_n}{1 + \lambda_i^2} p_{i,n-1}$$

$$= -\frac{\lambda_i - \lambda_n}{(\lambda_i^2 - \lambda_n^2)^{1/2}} p_{i,n-1} = \varepsilon_i \rho_i p_{i,n-1}$$

$$= -\rho_i b_i$$

To conclude the proof we compute the generator matrix elements. To this end we put

$$S^{\pi_2^{(n)}} \nu_i \frac{\alpha_i^*}{\pi_2^{(n)}} = \sum_{j=1}^{n-1} \alpha_{ji} \nu_j \frac{\alpha_j^*}{\pi_2^{(n)}}$$

Now

$$S^{\pi_2^{(n)}} \nu_i \frac{\alpha_i^*}{\pi_2^{(n)}} = \frac{z\nu_i \alpha_i^* - \eta_i \pi_2^{(n)}}{\pi_2^{(n)}}$$

where η_i is chosen so that

$$\eta_i = \frac{(-1)^{n-2} v_i \alpha_{i,n-2}}{\pi_{2,n-1}^{(n)}} = -(1 + \lambda_n^2)\rho_i p_{i,n-1}$$

where we have used Lemma 12.1. Now

$$\alpha_{ji} = \frac{\left[\dfrac{z v_i \alpha_i^* - \eta_i \pi_2^{(n)}}{\pi_2^{(n)}}, v_j \dfrac{\alpha_j^*}{\pi_2^{(n)}}\right]}{\left[v_j \dfrac{\alpha_j^*}{\pi_2^{(n)}}, v_j \dfrac{\alpha_j^*}{\pi_2^{(n)}}\right]} = \frac{\left(\dfrac{z v_i \alpha_i^* - \eta_i \pi_2^{(n)}}{p_n}, v_i \dfrac{\alpha_j^*}{p_n}\right)}{\left(v_j \dfrac{\alpha_j^*}{p_n}, v_j \dfrac{\alpha_j^*}{p_n}\right)}$$

$$= \frac{1 + \lambda_j^2}{(\lambda_j^2 - \lambda_n^2)^{1/2}} \left(\frac{z v_i \alpha_i^* - \eta_i \pi_2^{(n)}}{p_n}, v_j \frac{\alpha_j^*}{p_n}\right)$$

$$= \frac{1 + \lambda_j^2}{(\lambda_j^2 - \lambda_n^2)^{1/2}} J$$

with

$$J = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{(z v_i \alpha_i^* - \eta_i \pi_2^{(n)})\alpha_j}{p_n p_n^*} \, d\tau$$
$$= \frac{v_i}{2\pi} \int_{-\infty}^{\infty} \frac{z \alpha_i^* \alpha_j}{p_n p_n^*} \, d\tau - \frac{\eta_i}{2\pi} \int_{-\infty}^{\infty} \frac{\pi_2^{(n)} \alpha_j}{p_n p_n^*} \, d\tau$$

From the singular value equation

$$\begin{bmatrix} \pi_1^{(n)} \\ \pi_2^{(n)} \end{bmatrix} \alpha_j = \sigma_j p_n^* \begin{bmatrix} \alpha_1^{(j)} \\ \alpha_2^{(j)} \end{bmatrix} + p_n \begin{bmatrix} \zeta_1^{(j)} \\ \zeta_2^{(j)} \end{bmatrix}$$

we get

$$\frac{\pi_2^{(n)} \alpha_j}{p_n p_n^*} = \sigma_j \frac{\alpha_2^{(j)}}{p_n} + \frac{\zeta_2^{(j)}}{p_n^*} = \frac{\mu_j}{(1 + \mu_j^2)^{1/2}} \frac{\alpha_2^{(j)}}{p_n} + \frac{\zeta_2^{(j)}}{p_n^*}$$

Hence, by integrating over the contour $\hat{\gamma}_R$, we get

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\pi_2^{(n)} \alpha_j}{p_n p_n^*} \, d\tau = \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\gamma_R} \frac{\pi_2^{(n)} \alpha_j}{p_n p_n^*} \, dz = \frac{\mu_j}{(1 + \mu_j^2)^{1/2}} \frac{\alpha_{2,n-2}^{(j)}}{p_{n,n-1}}$$

Next we compute

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{z \alpha_i^* \alpha_j}{p_n p_n^*} \, d\tau$$

From Equation (242) we have that

$$\frac{z \alpha_i^* \alpha_j}{p_n p_n^*} = \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \frac{z \omega_{ij}^*}{p_n} + \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \frac{z \omega_{ij}}{p_n^*}$$

So, for R large enough,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{z \alpha_i^* \alpha_j}{p_n p_n^*} \, d\tau = \frac{1}{2\pi i} \int_{\gamma_R} \frac{z \alpha_i^* \alpha_j}{p_n p_n^*} \, dz$$
$$= \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \frac{1}{2\pi i} \int_{\gamma_R} \frac{z \omega_{ij}^*}{p_n} \, dz = \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} (-1)^{n-3} \frac{\omega_{ij,n-3}}{p_{n,n-1}}$$

Summarizing, we have for J that

$$J = \nu_i \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} (-1)^{n-3} \frac{\omega_{ij,n-3}}{p_{n,n-1}} - \eta_i \frac{\mu_j}{(1 + \mu_j^2)^{1/2}} \frac{\alpha_{2,n-2}^{(j)}}{p_{n,n-1}} \tag{262}$$

Using the expressions in (253) and (254) we have

$$\begin{aligned}
J = {} & \frac{\varepsilon_i \lambda_i}{(\lambda_i^2 - \lambda_n^2)^{1/2}} \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \frac{(-1)^{n-3}}{p_{n,n-1}} \\
& \times (-1)^n \frac{1}{\lambda_i \lambda_j} (\lambda_i - \lambda_j)(\lambda_i - \lambda_n)(\lambda_j - \lambda_n) p_{i,n-1} p_{j,n-1} p_{n,n-1} \\
& - [-(1 + \lambda_n^2)\rho_i p_{i,n-1}] \frac{\varepsilon_j \lambda_j}{\lambda_j (1 + \lambda_j^2)^{1/2}} \frac{\varepsilon_j (\lambda_j^2 - \lambda_n^2)}{(1 + \lambda_j^2)^{1/2}(1 + \lambda_n^2)} \frac{p_{j,n-1} p_{n,n-1}}{p_{n,n-1}} \\
= {} & -\frac{\rho_i}{\lambda_i + \lambda_j} (\lambda_j - \lambda_n) p_{i,n-1} p_{j,n-1} + \rho_i p_{i,n-1} p_{j,n-1} \frac{\lambda_j^1 - \lambda_n^2}{1 + \lambda_j^2} \qquad (263) \\
= {} & -\rho_i p_{i,n-1} p_{j,n-1} (\lambda_j - \lambda_n) \left[\frac{1}{\lambda_i + \lambda_j} - \frac{\lambda_j + \lambda_n}{1 + \lambda_j^2} \right] \\
= {} & -\rho_i p_{i,n-1} p_{j,n-1} \frac{\lambda_j - \lambda_n}{1 + \lambda_j^2} \left[\frac{1 - \lambda_i \lambda_j}{\lambda_i + \lambda_j} - \lambda_n \right]
\end{aligned}$$

So

$$\begin{aligned}
\alpha_{ji} & = \frac{1 + \lambda_j^2}{(\lambda_j^2 - \lambda_n^2)^{1/2}} J \\
& = -\frac{1 + \lambda_j^2}{(\lambda_j^2 - \lambda_n^2)^{1/2}} \rho_i p_{i,n-1} p_{j,n-1} \frac{\lambda_j - \lambda_n}{1 + \lambda_j^2} \left[\frac{1 - \lambda_i \lambda_j}{\lambda_i + \lambda_j} - \lambda_n \right] \qquad (264) \\
& = -\varepsilon_j \rho_j \rho_i p_{i,n-1} p_{j,n-1} \left[\frac{1 - \lambda_i \lambda_j}{\lambda_i + \lambda_j} - \lambda_n \right] \\
& = -\varepsilon_i \rho_j \rho_i b_i b_j \left[\frac{1 - \lambda_i \lambda_j}{\lambda_i + \lambda_j} - \lambda_n \right]
\end{aligned}$$

□

In the previous theorem we have derived a state-space realization of the transfer function g_n with respect to the basis $\alpha_i^*/\pi_n^{(n)}$. The state-space realization is parametrized in terms of the leading coefficients of the polynomials p_i where the normalization of these polynomials was chosen to be the same as the one which was used in Theorem 11.2 to obtain the LQG balanced realization of g. In the following corollary we show that we can also obtain a LQG balanced realization of g_n. It is easily obtained from the above realization by a simple diagonal state-space transformation. Alternatively, the realization could also be obtained in the same way as the realization derived in the previous theorem by choosing a slightly different normalization of the basis vectors.

Corollary 12.2: *If $\mu_1 > \mu_2 > \cdots > \mu_{n-1} > \mu_n > 0$, then*

$$\left[\begin{array}{c|c} -\varepsilon_j\rho_i\rho_j b_i b_j\left[\dfrac{(1-\lambda_i\lambda_j)}{\lambda_i+\lambda_j}-\lambda_n\right] & -\rho_i(1+\lambda_n^2)^{1/2}b_i \\ \hline -(1+\lambda_n^2)^{1/2}\varepsilon_i\rho_i b_i & -\lambda_n \end{array}\right]$$

$$= \left[\begin{array}{c|c} -\varepsilon_j\dfrac{\tilde{b}_i\tilde{b}_j}{1+\lambda_n^2}\left[\dfrac{(1-\lambda_i\lambda_j)}{\lambda_i+\lambda_j}-\lambda_n\right] & \tilde{b}_i \\ \hline \varepsilon_i b_i & -\lambda_n \end{array}\right]$$

with $\tilde{b}_i = (1+\lambda_n^2)^{1/2}\rho_i b_i$, is a LQG balanced realization of g_n with LQG singular values $\mu_1 > \mu_2 > \cdots > \mu_{n-1}$.

Proof: The stated realization can be obtained from the realization in Theorem 12.2 by a state-space transformation $T = -(1+\lambda_n^2)^{1/2}I$. Comparing this with the canonical form obtained in Ober (1989) or by direct verification we conclude that the realization is LQG balanced with $\Sigma = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_{n-1})$. □

The following scheme displays the results of this and the previous sections in a schematic way. The boxes on the left-hand side contain information on the functions

$$g, \quad R^* \quad \text{and} \quad \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}$$

The boxes on the right give information on the corresponding approximants

$$g_n, \quad R_n^* \quad \text{and} \quad \begin{bmatrix} \pi_1^{(1)}/p_n \\ \pi_2^{(1)}/p_n \end{bmatrix}$$

$$g = \frac{e}{d}$$

LQG s.v. $\mu_1 > \cdots > \mu_n$

$$\left\{\frac{p_i}{d}\right\}_{i=1}^{n}$$

$$\left[\begin{array}{c|c} \left(-\varepsilon_j b_i b_j\dfrac{1-\lambda_i\lambda_j}{\lambda_i+\lambda_j}\right) & b_i \\ \hline \varepsilon_j b_j & 0 \end{array}\right]$$

$$g_n = \frac{\pi_1^{(n)}}{\pi_2^{(n)}}$$

LQG s.v. $\mu_1 > \cdots > \mu_{n-1}$

$$\left\{v_i\frac{\alpha_i}{\pi_2^{(n)}}\right\}_{i=1}^{n-1}$$

$$\left[\begin{array}{c|c} -\left(\dfrac{\varepsilon_j\rho_i\rho_j b_i b_j(1-\lambda_i\lambda_j)}{\lambda_i+\lambda_j}\right) & \rho_i b_i \\ \hline (1+\lambda_n^2)\varepsilon_j\rho_j b_j & -\lambda_n \end{array}\right]$$

$$R^* = \frac{r^*}{t^*}$$ Hankel s.v. $\mu_1 > \cdots > \mu_n$ $$\left\{\frac{p_i}{t}, \varepsilon_i \frac{p_i^*}{t^*}\right\}_{i=1}^{n}$$ $$\left[\begin{array}{c\|c} \left(\varepsilon_j b_i b_j \frac{1}{\lambda_i + \lambda_j}\right) & b_i \\ \hline -\varepsilon_j b_j & 0 \end{array}\right]$$	$$R_n^* = -\pi_- \frac{\pi_n}{p_n}$$ Hankel s.v. $\mu_1 > \cdots > \mu_{n-1}$ $$\left\{\frac{\alpha_i}{p_n^*}, \frac{\alpha_i^*}{p_n},\right\}_{i=1}^{n-1}$$ $$\left[\begin{array}{c\|c} \left(\frac{\varepsilon_j \rho_i \rho_j b_i b_j}{\lambda_i + \lambda_j}\right) & \rho_i b_i \\ \hline \varepsilon_j \rho_j b_j & 0 \end{array}\right]$$
$$\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}$$ Hankel s.v. $\sigma_1 > \cdots > \sigma_n$ $$\left\{\frac{p_i}{t}, \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}\right\}_{i=1}^{n}$$	$$\frac{1}{(1-\sigma_n^2)^{1/2}} \begin{bmatrix} \pi_1^{(n)}/p_n \\ \pi_2^{(n)}/p_n \end{bmatrix}$$ Hankel s.v. $\sigma_1 > \cdots > \sigma_{n-1}$ $$\frac{1}{(1-\sigma_n^2)^{1/2}} \left\{\frac{\alpha_i}{p_n^*}, \begin{bmatrix} \alpha_1^{(i)}/p_n \\ \alpha_2^{(i)}/p_n \end{bmatrix}\right\}_{i=1}^{n-1}$$

13. Robust control

In this section we use the previously developed machinery in order to study the optimally robust stabilization problem. We identify the optimally robust controller in terms of the polynomial data of the Schmidt vectors. We re-derive, in our context, the results of Glover and McFarlane on the relation of the optimally robust controller to the Nehari complement of the normalized coprime factors and the characterization of the optimally robust stability margin. In this connection one should consult also Georgiou and Smith (1990 a, b) for another approach to the problem.

The singular vector analysis of the renormalized Nehari complement of the NCF leads to the derivation of an LQG balanced realization of the optimally robust controller. In particular we obtain the result that, given a transfer function g, the Nehari complement of the LQG symbol associated with the NCF of g is the LQG symbol of the renormalized Nehari extension of the NCF of g. All this is summed up in a scheme at the end of the section. This scheme is of course dual to the scheme at the end of the previous section.

This shows that the problems of optimally robust stabilization and model reduction via balancing and Hankel norm approximation are dual problems. This extends the duality theory developed in Fuhrmann (1991).

We will consider the standard feedback configuration shown in Fig. 7. Here, G is a $p \times m$ plant given by means of its strictly proper, rational transfer function and K is a $m \times p$ controller similarly given. Notice the full symmetry between the plant and the controller in this formulation.

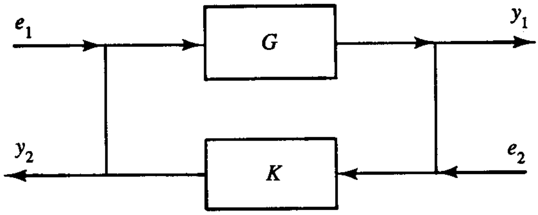

Figure 7.

The feedback configuration (G, K) is called *internally stable* if and only if

$$\begin{bmatrix} I & G \\ K & I \end{bmatrix}^{-1} = \begin{bmatrix} (I - GK)^{-1} & -(I - GK)^{-1}G \\ -K(I - GK)^{-1} & (I - KG)^{-1} \end{bmatrix} \tag{265}$$

is in H_+^∞.

Internal stability of a feedback pair (G, K) is reduced to a coprimeness condition in the following way. Let

$$G = NM^{-1} = \bar{M}^{-1}\bar{N}$$

be right and left coprime factorizations respectively over the ring H_+^∞. Also let

$$K = UV^{-1} = \bar{V}^{-1}\bar{U}$$

be right and left coprime factorizations respectively. Then the following holds (see e.g. Vidyasagar 1985).

Theorem 13.1: *The following statements are equivalent.*

(1) (G, K) *is internally stable.*

(2) $\begin{bmatrix} M & U \\ N & V \end{bmatrix}^{-1} \in H_+^\infty$

(3) $\begin{bmatrix} \bar{V} & -\bar{U} \\ -\bar{N} & \bar{M} \end{bmatrix}^{-1} \in H_+^\infty$

(4) $(\bar{V}M - \bar{U}N)^{-1} \in H_+^\infty$

(5) $(\bar{M}V - \bar{N}U)^{-1} \in H_+^\infty$

Note the conditions (4) and (5) in Theorem 13.1 are equivalent to the solvability, over H_+^∞, of the Bezout equations

$$\bar{V}M - \bar{U}N = I$$

and

$$\bar{M}V - \bar{N}U = I$$

Next we turn our attention to questions of robustness. We assume uncertainty in the plant via its normalized coprime factorizations, i.e. we consider plants of the form

$$(N + \Delta_N)(M + \Delta_M)^{-1} = (\bar{M} + \Delta_{\bar{M}})^{-1}(\bar{N} + \Delta_{\bar{N}})$$

with $\left\| \begin{bmatrix} \Delta_N \\ \Delta_M \end{bmatrix} \right\| < \varepsilon$ or $\|[\Delta_{\bar{M}} \quad \Delta_{\bar{N}}]\| < \varepsilon$ and ask whether the controller K

stabilizes all such plants. The key result has been obtained by Vidyasagar and Kimura (1986).

Theorem 13.2. (Vidyasagar and Kimura 1986): *Let (G, K) be an internally stable feedback pair. Let G and K have a doubly coprime factorization given by*

$$\begin{bmatrix} \bar{V} & -\bar{U} \\ -\bar{N} & \bar{M} \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{266}$$

Then the following statements are equivalent.

(1) $\left\| \begin{bmatrix} V \\ U \end{bmatrix} \right\|_\infty \leqslant \frac{1}{\varepsilon}$

(2) *(G', K) is an internally stable feedback pair for all G' with transfer function $(\bar{M} + \Delta_{\bar{M}})^{-1}(\bar{N} + \Delta_{\bar{N}})$ and $\|[\Delta_{\bar{M}} \quad \Delta_{\bar{N}}]\| < \varepsilon$.*

Now all stabilizing controllers of a plant G can be given via the Kucera–Youla parametrization, using the doubly coprime factorization (266), in the form

$$K = (U + MQ)(V + NQ)^{-1} = (\bar{V} + Q\bar{N})^{-1}(\bar{U} + Q\bar{M}) \tag{267}$$

with $Q \in H_+^\infty$ arbitrary, provided K is proper.

Thus, from the Vidyasagar–Kimura result it is clear that the *maximum stability margin*, denoted by $\varepsilon_{\max}$ will be attained by the controller, which exists as a consequence of the commutant lifting theorem, which attains the minimum in

$$\varepsilon_{\max}^{-1} = \left\| \begin{bmatrix} U + MQ \\ V + NQ \end{bmatrix} \right\|_\infty \tag{268}$$

i.e. the H^∞ norm of the smallest solution to the Bezout equation $\bar{M}V - \bar{N}U = I$. It can be shown (Glover and McFarlane 1988, Nikolskii 1986, and references therein), that $\varepsilon_{\max} = (1 - \sigma_1^2)^{1/2}$, where σ_1 is the first Hankel singular value of $[-\bar{N} \quad \bar{M}]$.

We will proceed to interpret some of the result obtained in previous sections in terms of robust control. In the following we will assume that g is a strictly proper rational plant. No stability assumptions are made. We assume that $g = e/d$ is a polynomial coprime factorization, with the polynomial d monic. We will identify the optimally robust controller. Moreover, we will describe how to derive a LQG balanced realization for the optimally robust controller, given a LQG balanced realization of the plant. This procedure is dual to the one, developed in § 12, for the case of model reduction.

The development highlights the existing duality between problems of model reduction and those of robust control. This duality is based on the duality theory, developed in Fuhrmann (1991), between Nehari complementation and optimal Hankel norm approximation.

Note that the components of the inner function

$$\frac{1}{(1 - \sigma_1^2)^{1/2}} \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$$

are normalized coprime factors of $k = \pi_1^{(1)}/\pi_2^{(1)}$. We can therefore construct the function R_1^* associated with the LQG controller corresponding to the function k.

First, we clarify the role of the stabilizing controller obtained in Corollary 10.1.

Theorem 13.3: *Let $(e/t)(d/t)^{-1}$ be a normalized coprime factorization of the transfer function $g = e/d$. Let $\sigma_1 > \cdots > \sigma_n$ be the singular values of the Hankel operator*

$$H\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}$$

and let

$$p_i, \quad \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix}, \quad \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$$

be defined by the s.v. equations (94). Then $K = \pi_1^{(1)}/\pi_2^{(1)}$ is the optimally robust stabilizing controller for g.

Proof: By Theorem 8.1 we have, given any solution U, V of the Bezout equation $\bar{M}V - \bar{N}U = I$, that

$$\inf_{Q \in H_+^\infty} \left\| \begin{bmatrix} U \\ V \end{bmatrix} + \begin{bmatrix} M \\ N \end{bmatrix} Q \right\|_\infty = \|Z_K\| = (1 + \mu_1^2)^{1/2} = \frac{1}{(1 - \sigma_1^2)^{1/2}}$$

We show next that

$$\begin{bmatrix} U_{\text{opt}} \\ V_{\text{opt}} \end{bmatrix} = \frac{1}{1 - \sigma_1^2} \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$$

is the optimizing solution.

That it is a solution follows from Theorem 8.1. On the other hand, from Theorem 8.1.3, we know that

$$\frac{1}{(1 - \sigma_1^2)^{1/2}} \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$$

is an all-pass function. Hence

$$\left\| \begin{bmatrix} U \\ V \end{bmatrix} \right\|_\infty = \left\| \frac{1}{1 - \sigma_1^2} \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix} \right\|_\infty$$

$$= \frac{1}{(1 - \sigma_1^2)^{1/2}} \left\| \frac{1}{(1 - \sigma_1^2)^{1/2}} \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix} \right\|_\infty = \frac{1}{(1 - \sigma_1^2)^{1/2}} \qquad \square$$

We will now analyse the optimally robust controller in more detail. In particular we are going to determine the function R^* which is associated with the robust controller. Before we can do this we need more information on the normalized coprime factorization of the optimally robust controller. This information is provided in the next proposition. Here the Bezout equation corresponding to the normalized coprime factors of the optimally robust controller is examined.

Proposition 13.1: *Let e/d and let e/t, d/t be the normalized coprime factors of g. Let $\sigma_1 > \sigma_2 \geqslant \cdots \geqslant \sigma_{n-1} \geqslant \sigma_n > 0$ be the singular values of*

$$H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$$

and let

$$\begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$$

be the optimal Nehari extension corresponding to σ_1. Let $\omega_1^{(i)}$, $\omega_2^{(i)}$ be as defined by Theorem 10.3. Set

$$[\bar{N}_1 \quad \bar{M}_1] := \begin{bmatrix} N_1 \\ M_1 \end{bmatrix}^{\mathrm{T}} := \frac{1}{(1-\sigma_1^2)^{1/2}} [\pi_1^{(1)}/p_1 \quad \pi_2^{(1)}/p_1]$$

and

$$\begin{bmatrix} U_1 \\ V_1 \end{bmatrix} := \frac{1}{(1-\sigma_2^2)} \begin{bmatrix} -\dfrac{(\omega_1^{(2)})^*}{\beta_2} \\ \dfrac{(\omega_2^{(2)})^*}{\beta_2} \end{bmatrix} \in H_+^\infty$$

Then we have

(1) $(\pi_1^{(1)})^*\omega_1^{(i)} + (\pi_2^{(1)})^*\omega_2^{(i)} = (1-\sigma_1^2)^{1/2}(1-\sigma_i^2)\beta_i^* p_1^*$

(2) $\bar{M}_1 V_1 - \bar{N}_1 U_1 = I$

Proof:

(1) From the singular value equation (184), i.e.

$$\frac{1}{(1-\sigma)^{1/2}} \begin{bmatrix} \pi_1^{(1)} \\ \pi_2^{(1)} \end{bmatrix} \beta_i^* = \sigma_i p_1^* \frac{\varepsilon_1 \varepsilon_i}{(1-\sigma_i^2)^{1/2}} \begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \end{bmatrix} + p_1 \begin{bmatrix} \omega_1^{(i)} \\ \omega_2^{(i)} \end{bmatrix}$$

we get, left multiplying by $((\pi_1^{(1)})^* \quad (\pi_2^{(2)})^*)$, using (185) and the fact that $\dfrac{1}{(1-\sigma)^{1/2}} \begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$ is all-pass, that

$$\begin{aligned}
(1-\sigma_1^2)^{1/2} p_1 p_1^* \beta_i^* &= \sigma_i p_1^* \frac{\varepsilon_1 \varepsilon_i}{(1-\sigma_i^2)^{1/2}} ((\pi_1^{(1)})^* \beta_1^{(i)} + (\pi_2^{(1)})^* \beta_2^{(i)}) + p_1((\pi_1^{(1)})^* \omega_1^{(i)} \\
&\quad + (\pi_2^{(1)})^* \omega_2^{(i)}) \\
&= \sigma_i p_1^* \frac{\varepsilon_1 \varepsilon_i}{(1-\sigma_i^2)^{1/2}} \sigma_i \varepsilon_1 \varepsilon_i (1-\sigma_1^2)^{1/2} (1-\sigma_i^2)^{1/2} \beta_i^* p_1 \\
&\quad + p_1((\pi_1^{(1)})^* \omega_1^{(i)} + (\pi_2^{(1)})^* \omega_2^{(i)}) \\
&= \sigma_i^2 (1-\sigma_1^2)^{1/2} p_1 p_1^* \beta_i^* + p_1((\pi_1^{(1)})^* \omega_1^{(i)} + (\pi_2^{(1)})^* \omega_2^{(i)})
\end{aligned}$$

and the statement follows.

(2) Note that by AAK theory p_1 as well as β_2 are stable. Thus, from

$$\pi_1^{(1)}(\omega_1^{(2)})^* + \pi_2^{(1)}(\omega_2^{(2)})^* = (1-\sigma_1^2)^{1/2}(1-\sigma_2^2)\beta_2 p_1$$

we get the H_+^∞-Bezout equation

$$\frac{1}{(1-\sigma_1^2)^{1/2}} \frac{\pi_2^{(1)}}{p_1} \frac{1}{(1-\sigma_2^2)} \frac{(\omega_2^{(2)})^*}{\beta_2} - \frac{1}{(1-\sigma_1^2)^{1/2}} \frac{\pi_1^{(1)}}{p_1} \frac{-1}{(1-\sigma_2^2)} \frac{(\omega_1^{(2)})^*}{\beta_2} = 1 \quad \square$$

We proceed to identify the function R_1^* in terms of the Nehari complement of R^*. This is an analogue of Corollary 12.1.

Theorem 13.4: *The function R_1^*, i.e. the strictly proper unstable part of $M_1^*U_1 + N_1^*V_1$ is given by*

$$R_1^* = -\pi_- \frac{\pi_1^*}{p_1^*}$$

i.e. the strictly proper part of $-\dfrac{\pi_1^*}{p_1^*}$.

Proof: Let $X, Y \in H_+^\infty$ be such that

$$\bar{M}_1 Y = \bar{N}_1 X = I$$

then we know that R_1^* is the strictly proper antistable part of $M_1^*X + N_1^*Y$. Let now

$$X = \frac{1}{(1-\sigma_1^2)^{1/2}} \frac{e}{t}$$

and

$$Y = \frac{1}{(1-\sigma_1^2)^{1/2}} \frac{d}{t}.$$

Then, first note that by Equation (109)

$$\bar{M}_1 Y - \bar{N}_1 x = \frac{1}{1-\sigma_1^2}\left(\frac{\pi_2^{(1)}}{p_1}\frac{d}{t} - \frac{\pi_1^{(1)}}{p_1}\frac{e}{t}\right) = I$$

Now, using Equation (111) we have

$$\begin{aligned} M_1^*X + N_1^*Y &= \frac{1}{1-\sigma_1^2}\left(\frac{(\pi_2^{(1)})^*}{p_1^*}\frac{e}{t} + \frac{(\pi_1^{(1)})^*}{p_1^*}\frac{d}{t}\right) \\ &= \frac{1}{1-\sigma_1^2}\frac{1}{p_1^*t}\,\varepsilon_1\sigma_1\,(1-\sigma_1^2)^{1/2}\,t^*p_1 \\ &= \lambda_1 \frac{t^*p_1}{p_1^*t} \end{aligned}$$

But, by Equation (83)

$$\lambda_1 \frac{t^*p_1}{p_1^*t} = \frac{r}{t} - \frac{\pi_1^*}{p_1^*},$$

which shows that the strictly proper anti-stable part of $M_1^*X + N_1^*Y$ is given by the strictly proper part of $-\pi_1^*/p_1^*$. □

This leads to a result that is dual to Theorem 12.2. In particular, we obtain directly an LQG balanced realization of the optimally robust controller.

We need the following Lemma.

Lemma 13.1: *Let β_i, $i = 2, 3, \ldots, n$ be the polynomials defined in Theorem 10.1. Then*

$$\beta_{i,n-2} = \frac{\lambda_i - \lambda_1}{\lambda_1} p_{1,n-1} p_{i,n-1}$$

$i = 2, 3, \ldots, n$.

Proof: Evaluating the leading coefficients of the polynomials in the polynomial equation (169), i.e.

$$\lambda_1 p_1^* p_i - \lambda_i p_i^* p_1 = \lambda_1 t^* \beta_i$$

we obtain the result. □

Theorem 13.5: *Assume that $\sigma_1 > \cdots > \sigma_n$ are the singular values of the Hankel operator $H_{\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}}$. Let*

$$\begin{bmatrix} -e^* \\ d^* \end{bmatrix} p_i = \sigma_i t \begin{bmatrix} \hat{p}_1^{(i)} \\ \hat{p}_2^{(i)} \end{bmatrix} + t^* \begin{bmatrix} \pi_1^{(i)} \\ \pi_2^{(i)} \end{bmatrix}$$

$$- e\hat{p}_1^{(i)} + d\hat{p}_2^{(i)} = \sigma_i t^* p_i$$

be the singular vector/singular value equations. Let μ_i and ε_i be defined as in Proposition 8.1. Then

(1) *The Hankel singular values of $\hat{H}_{R_1^*}$ are $\mu_2 > \cdots > \mu_n$, and the Schmidt pairs are $\left\{\frac{\beta_i^*}{p_1^*}, -\varepsilon_i \frac{\beta_i}{p_1}\right\}_{i=2}^{n}$.*

(2) *The shift realization of R_1^* with respect to the basis $\left\{ \eta_i \frac{\beta_i}{p_1} \right\}$, normalized so that*

$$\left\| \eta_i \frac{\beta_i^*}{p_1^*} \right\|^2 = \mu_i$$

is given by

$$\left[\begin{array}{c|c} \left(\dfrac{\varepsilon_j \tau_i \tau_j b_i b_j}{\lambda_i + \lambda_j}\right) & \tau_i b_i \\ \hline \varepsilon_j \tau_j b_j & 0 \end{array}\right] \tag{269}$$

We assume that the $\{p_i/t\}$ are normalized so that $\|p_i/t\|^2 = \mu_i$, $i = 2, 3, \ldots, n$. Here $b_i = \varepsilon_i p_{i,n-1}$ and $\eta_i = \left(\dfrac{\mu_1^2}{\mu_1^2 - \mu_i^2}\right)^{1/2}$.

(3) *Given the parameters of the realization of R_1^*, $k = g_1 = \pi_1^{(1)}/\pi_2^{(1)}$ has the following LQG balanced realization*

$$\left[\begin{array}{c|c} -\left(\dfrac{\varepsilon_j \tilde{b}_i \tilde{b}_j (1 - \lambda_i \lambda_j)}{(1 + \lambda_1^2)(\lambda_i + \lambda_j)}\right) & \tilde{b}_i \\ \hline \varepsilon_j \tilde{b}_j & -\lambda_1 \end{array}\right] \tag{270}$$

Here

$$\tilde{b}_i = \tau_i (1 + \lambda_1^2 b_i)^{1/2} \quad i = 2, \ldots, n \tag{271}$$

and

$$\tau_i = \left(\frac{\lambda_1 - \lambda_i}{\lambda_1 + \lambda_j}\right)^{1/2} \tag{272}$$

(4) *k has LQG singular values* $\mu_2 > \cdots > \mu_n$

Proof:

(1) This follows since $-R_1$ is the strictly proper part of the Nehari extension of R^*.

(2) From the (171)

$$\left\|\frac{\beta_i}{p_1^*}\right\|^2 = \left(1 - \frac{\mu_i^2}{\mu_1^2}\right)\left\|\frac{p_i}{t}\right\|^2$$

we can determine the constant η_i which is such that

$$\mu_i = \eta_i^2 \left\|\frac{\beta_i}{p_1^*}\right\|^2 = \eta_i^2 \left(1 - \frac{\mu_i^2}{\mu_1^2}\right)\left\|\frac{p_i}{t}\right\|^2 = \eta_i^2\left(1 - \frac{\mu_i^2}{\mu_1^2}\right)\mu_i$$

Therefore

$$\eta_i = \left(\frac{\mu_1^2}{\mu_1^2 - \mu_i^2}\right)^{1/2}$$

By Proposition 11.2 we have to determine the leading coefficients of the numerator polynomials $\hat{\beta}_{i,n-2}$ of the Schmidt vectors assuming that the denominator polynomial is monic. We have that

$$\begin{aligned}\hat{\beta}_{i,n-2} &= \eta_i \frac{\beta_{i,n-2}}{(-1)^{n-1} p_{1,i}} \\ &= \left(\frac{\mu_1^2}{\mu_1^2 - \mu_i^2}\right)^{1/2} \frac{1}{(-1)^{n-1} p_{1,i}} \frac{\lambda_{-}\lambda_1}{\lambda_1} p_{1,n-1} p_{i,n-1} \\ &= (-1)^n \tau_i p_{i,n-1}\end{aligned}$$

This, together with the fact that the Hankel singular values of the system are given by $\mu_2 > \cdots > \mu_n > 0$ and the signs are given by $-\varepsilon_i$ proves the result.

(3) In Theorem 11.2 (Theorem 12.2 and Corollary 12.2) we have seen how, given a Lyapunov balanced realization of R (R_n^*) corresponding to a function g (g_n), we can construct a LQG balanced parametrization of g (g_n) using the parameters that were used to parametrize the Lyapunov balanced realization of R (R_n^*). Those results were independent of the particular situation. They can therefore be applied to any situation in which the Lyapunov balanced parameters are known of a function R and where it is required to find a LQG balanced realization of the g. We can therefore also apply those results to our situation.

(4) This follows either by verification or from the general results in § 6. □

The following scheme shows, on the left-hand side information on the functions

$$g, \quad R^* \quad \text{and} \quad \begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}$$

The right-hand side displays information on the functions associated through the Nehari extension and optimal control problem, i.e.

$$k, \quad R_1^* \quad \text{and} \quad \frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$$

$$g = \frac{e}{d}$$

$$\text{LQG s.v. } \mu_1 > \cdots > \mu_n$$

$$\left\{\frac{p_i}{d}\right\}_{i=1}^n$$

$$\left[\begin{array}{c|c} -\left(\varepsilon_j b_i b_j \dfrac{1-\lambda_i\lambda_j}{\lambda_i+\lambda_j}\right) & b_i \\ \hline \varepsilon_j b_j & 0 \end{array}\right]$$

$$k = g_1 = \frac{\pi_1^{(1)}}{\pi_2^{(1)}}$$

$$\text{LQG s.v. } \mu_2 > \cdots > \mu_n$$

$$\left\{\nu_i \frac{\beta_i}{\pi_2^{(1)}}\right\}_{i=2}^n$$

$$\left[\begin{array}{c|c} -\left(\dfrac{\varepsilon_j \tau_i \tau_j b_i b_j (1-\lambda_i\lambda_j)}{\lambda_i+\lambda_j}\right) & \dfrac{\tau_i b_i}{(1-\sigma_1^2)^{1/2}} \\ \hline \dfrac{\varepsilon_j \tau_j b_j}{(1-\sigma_1^2)^{1/2}} & -\lambda_1 \end{array}\right]$$

$$R^* = \frac{r^*}{t^*}$$

$$\text{H.s.v. } \mu_1 > \cdots > \mu_n$$

$$\left\{\frac{p_i}{t}, \frac{p_i^*}{t^*}\right\}_{i=1}^n$$

$$\left[\begin{array}{c|c} \left(\varepsilon_j b_i b_j \dfrac{1}{\lambda_i+\lambda_j}\right) & b_i \\ \hline \varepsilon_j b_j & 0 \end{array}\right]$$

$$R_1^* = -\pi_- \frac{\pi_1}{p_1}$$

$$\text{H.s.v. } \mu_1 > \cdots > \mu_{n-1}$$

$$\left\{\frac{\beta_i^*}{p_1^*}, -\varepsilon_i \frac{\beta_i}{p_1},\right\}_{i=2}^n$$

$$\left[\begin{array}{c|c} -\left(\dfrac{\varepsilon_j \tau_i \tau_j b_i b_j}{\lambda_i+\lambda_j}\right) & \tau_i b_i \\ \hline \varepsilon_j \tau_j b_j & 0 \end{array}\right]$$

$$\begin{bmatrix} -e^*/t^* \\ d^*/t^* \end{bmatrix}$$

$$\text{H.s.v. } \sigma_1 > \cdots > \sigma_n$$

$$\left\{\frac{p_i}{t}, \begin{bmatrix} \hat{p}_1^{(i)}/t^* \\ \hat{p}_2^{(i)}/t^* \end{bmatrix}\right\}_{i=1}^n$$

$$\frac{1}{(1-\sigma_1^2)^{1/2}}\begin{bmatrix} \pi_1^{(1)}/p_1 \\ \pi_2^{(1)}/p_1 \end{bmatrix}$$

$$\sigma_2 > \cdots > \sigma_n$$

$$\left\{\frac{\beta_i^*}{p_1^*}, \frac{\varepsilon_1\varepsilon_i}{(1-\sigma_i^2)^{1/2}}\begin{bmatrix} \beta_1^{(i)}/p_1 \\ \beta_2^{(i)}/p_1 \end{bmatrix}\right)\right\}_{i=2}^n$$

Acknowledgment

Professor Fuhrmann's research was partially supported by the Israeli Academy of Sciences under Grant No. 249/90.

References

Adamjan, V. M., Arov, D. Z., and Krein, M. G. 1968 a, Infinite Hankel matrices and generalized problems of Caratheodory–Fejer and F. Riesz, *Functions Analysis Applications* **2**, 1–18; 1968 b, Infinite Hankel matrices and generalized problems of Caratheodory-Fejer and I. Schur. *Ibid.* **2**, 269–281; 1971, Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. *Mathematics USSR Sbornik*, **15**, 31–73; 1978, Infinite Hankel block matrices and related extension problems, *American Mathematical Society Translationals*, series 2, **111**, 133–156.

Antoulas, A. C. 1985, A new approach to synthesis problems in linear system theory. *IEEE Transactions on Automatic Control*, **30**, 465–473.

Coppel, W. A., 1972, Linear Systems, *Notes in Pure Mathematics*, Australian National University, Canberra, vol. 6.

Enns, D., 1984, Model reduction for control systems. Ph.D. dissertation, Department of Astronautics and Aeronautics, Stanford University.

Foias, C., Georgiou, T. T., and Smith, M. C., 1990, Geometric techniques for robust stabilization of linear time-varying systems. *Proceedings of the Conference on Decision and Control*, Honolulu, December 1990.

Fuhrmann, P. A., On the corona problem and its application to spectral problems in Hilbert space. *Transactions American Mathematical Society*, **132**, 55–67; 1976, Algebraic system theory: An analyst's point of view. *Journal of the Franklin Institute*, **301**, 521–540; 1977, On strict system equivalence and similarity. *International Journal of Control*, **25**, 5–10; 1979, Linear feedback via polynomial models. *International Journal of Control*, **30**, 363–377; 1981, *Linear Systems and Operators in Hilbert Space.* (New York: McGraw-Hill); 1984, On Hamiltonian transfer functions. Linear Algebra Applications, 1–93; 1985, The algebraic Riccati equation—a polynomial approach. *Systems and Control Letters*, 369–376; 1986, Block Hankel matrix inversion—the polynomial approach, In *Proceedings of Amsterdam Workshop on Operator Theory and Applications*, Birkhauser Verlag, 207–230; 1991, A polynomial approach to Hankel norm and balanced approximation. *Linear Algebra and Applications*, **146**, 133–220.

Fuhrmann, P. A., and Hautus, M. L. J., 1980, On the zero module of rational matrix functions. *Proceedings of the Conference on Decision and Control Albuquerque*.

Fuhrmann, P. A., and Krishnaprasad, P. S., 1986, Towards a cell decomposition for rational functions. *IMA Journal of Mathematical Control and Information*, **3**, 137–150.

Georgiou, T. T., and Smith, M. C., 1990 a Optimal robustness in the gap metric. *IEEE Transactions on Automatic Control*, **35**, 673–686; 1990 b, Robust stabilization in the gap metric; controller design for distributed plants. *1990 American Control Conference*, San Diego, May 1990; 1991, Upper and lower bounds for approximation in the gap metric. To be published in *Proceedings of the 1991 CDC*.

Glover, K. 1984, All optimal Hankel-norm approximations and their L^∞-error bounds. *International Journal Control* **39**, 1115–1193.

Glover, K. and McFarlane, D. C., 1988, Robust stabilization of normalized coprime factors: an explicit H^∞ solution. *Proceedings Automatic Control Conference*, 1988, Atlanta, Georgia, USA.

Hammer, J., 1985, Nonlinear systems, stabilization, and coprimeness. *International Journal of Control*, **42**, 1–20.

Hautus, M. L. J. and Heymann, M. 1978, Linear feedback-an algebraic approach. *SIAM Journal of Control*, **16**, 83–105.

Helmke, U., Hinrichsen, D., and Manthey, W., 1988, A cell decomposition of the space of real Hankels to rank $\leqslant n$ and some applications. Report, Institut für dynamische Systeme, Universität Bremen, Number 183.

Jonckheere, E., and Silverman, L. 1983, A new set of invariants for linear systems—application to reduced order compensator design. *IEEE Transactions on Automatic Control* **28**, 953–964.

Kimura, H., 1984, Robust stabilization for a class of transfer functions. *IEEE Transaction on Automatic Control* **29**, 788–793.

Kucera, V., 1979, *Discrete Linear Control* (Wiley).

McFarlane, D., and Glover, K., 1990, *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions*. Lecture Notes in Control and Information Sciences, Vol. 10

(Springer Verlag).

MEYER, D., and FRANKLIN, G., 1987, A connection between normalized coprime factorizations and linear quadratic regulator theory. *IEEE Transactions on Automatic Control* **32**, 227–228.

MOORE, B. C., 1981, Principal component analysis in linear systems: controllability, observability and model reduction. *IEEE Transactions on Automatic Control* **26**, 17–32.

NIKOLSKII, N. K., 1986, *Treatise of the Shift Operator* (Springer Verlag).

OBER, R. J., 1987, Balanced realizations for finite and infinite dimensional linear systems. Ph.D. thesis, Engineering Department, Cambridge University; 1988, Balanced parametrization of classes of linear systems. Report Cambridge University Engineering Department. *SIAM Journal of Optimization and Control*, to be published; 1989, Connectivity properties of classes of linear systems. *International Journal of Control*, **50**, 2049–2073.

OBER, R. J. and MCFARLANE, D. C., 1989, Balanced canonical forms for minimal systems: normalized coprime factor approach. *Linear Algebra and its Applications*, 122/123/124: 23–64.

OBER, R. J. and SEFTON, J. A., 1990, Stability of control systems and graphs of linear systems. *Report Programs in Mathematical Sciences*, August 1990, submitted for publication.

PERNEBO, L. and SILVERMAN, L., 1982, Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control* **27**, 382–387.

RANTZER, A., 1989, Equivalence between stability of partial realizations and feedback stabilization—applications to reduced order stabilization. *Linear Algebra and its Applications*, **122–124**, 641–653.

ROSENBROCK, H. H. 1970, *State Space ad Multivariable Theory* (New York: Wiley).

SEFTON, J., 1991, Ph.D. thesis, Cambridge University, in preparation.

SZ.-NAGY, B. and FOIAS, C., 1970, *Harmonic Analysis of Operators on Hilbert Space* (Amsterdam: North Holland).

VERMA, M. S., 1988 Coprime fractional representations and stability of nonlinear feedback systems. *International Journal of Control*, **48**, 897–918.

VIDYASAGAR, M., 1985 *Control System Synthesis: a Coprime Factorization Approach* (Cambridge, Mass: MIT Press).

VIDYASAGAR, M. and KIMURA, H., 1986, Robust controllers for uncertain linear multivariable systems. *Automatica*, **22**, 85–94.

WEILAND, S., 1989, Balanced representations and approximation of linear systems, to be published.

YEH, F.-B. and WEI, L.-F., 1990, Inner-outer factorizations of right invertible real rational matrices. *Systems and Control Letters*, **14**, 31–36.

YOULA, D. C., BONGIORNO, J. J. and JABR, H. A., 1976, Modern Wiener–Hopf design of optimal controllers. Part: 2 The multivariable case. *IEEE Transactions on Automatic Control*, **21**, 319–338.

ZAMES, G., 1981, Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms and approximative inverses. *IEEE Transactions on Automatic Control*, **23**, 301–320.