Optimal Hankel-Norm Model Reductions: Multivariable Systems

SUN-YUAN KUNG, MEMBER, IEEE, AND DAVID W. LIN, STUDENT MEMBER, IEEE

Abstract—This paper represents a first attempt to derive a closed-form (Hankel-norm) optimal solution for multivariable system reduction problems. The basic idea is to extend the scalar case approach in [5] to deal with the multivariable systems. The major contribution lies in the development of a minimal degree approximation (MDA) theorem and a computation algorithm. The main theorem describes a closed-form formulation for the optimal approximants, with the optimality verified by a complete error analysis. In deriving the main theorem, some useful singular value/vector properties associated with block-Hankel matrices are explored and a key extension theorem is also developed. Imbedded in the polynomial-theoretic derivation of the extension theorem is an efficient approximation algorithm. This algorithm consists of three steps: i) compute the minimal basis solution of a polynomial matrix equation; ii) solve an algebraic Riccati equation; and iii) find the partial fraction expansion of a rational matrix.

I. Introduction

MODEL REDUCTIONS arise in many important applications for simplifying system modeling and/or controller designs. The problem has been a major attraction in the system theory literatures. Several performance criteria and many algorithms have been proposed (see, e.g., [2], [3], and the references therein). While most of them can be supported by simulation examples, the very desired error analyses and algorithm complexity studies are in general not available. Therefore, it has been very difficult to conduct an objective comparison between them.

Very recently [1]–[8], a new Hankel-norm criterion has received rapidly increasing attention. Based on the well-known Kronecker theorem and the singular value analysis (used as a robust tool for rank characterizations), Hankel-norm appears to be very natural and useful. Moreover, as has been pointed out in [5], the Hankel norm of a stable single input–output system lies between the more conventional $\mathcal{L}_2$ and $\mathcal{L}_\infty$ norms. As we shall see in Section II (after the introduction of various distance measures—norms), this is again the case for multivariable systems. Hence the Hankel-norm criterion can be viewed as a compromise between the popular least squares error criterion and the stringent maximum deviation ($\mathcal{L}_\infty$, or Chebyshev) error criterion.

Manuscript received April 10, 1980; revised December 9, 1980. Paper recommended by B. Francis, Past Chairman of the Linear Systems Committee. This work was supported by the Army Research Office under Grant DAAG 29-79-C-0054, and by the National Science Foundation under Grant ENG-7908673, and by the Joint Services Electronic Program through the Air Force Office of Scientific Research under Contract F-44620-76-C-0061.

The authors are with the Department of Electrical Engineering Systems, University of Southern California, Los Angeles, CA 90007.

For the scalar (single input–output) case, it was Adamjan *et al.* [1] who first developed a closed-form optimal solution for model reductions with respect to this criterion. As a matter of fact, it has been the only available closed-form solution for any optimality criterion. The relevance of [1] to model reductions was first mentioned by Kung [2] in 1978, while a comparison of some numerical aspects involved in [1], [2] was reported in [8]. In [3]–[7] connections between the minimal Hankel-norm approximations and rational function approximations are put into light. In [3], [4] the role of balanced realizations in state-space models is exploited to lead to an optimal approximation algorithm, requiring solving Lyapunov equations and singular value decompositions. In [5] a (one-variable) polynomial approach is taken to elucidate the singular value/vector properties of Hankel matrices, which then leads to a simple generalized eigenvalue formulation and a fast matrix-fraction description (MFD) based algorithm for the so-called *minimal degree approximation* (MDA) problems. In [6], [7] a two-variable polynomial approach is used to rederive the results of [1] and further illuminate many significant properties of the MDA problems.

Adamjan *et al.*'s work on scalar systems [1] was benefited from Nehari's work [11] related to what can be called "zeroth-order approximations." The authors of [1] also studied the "zeroth-order approximations" problem for multivariable systems [10]. However, prior to the present work, a general theory and algorithm for optimal multivariable model reductions with respect to this Hankel-norm or other criterion are still lacking. This paper aims to help close this gap.

A. Organization

Section II provides some mathematical preliminaries related to block-Hankel matrices. Section III derives the main minimal degree approximation (MDA) theorem. In the same section, some important singular value/vector properties of Hankel matrices are also explored. Section IV developes the key extension theorem, supplementing the proof of the MDA theorem and paving a way to an efficient MDA algorithm outlined in Section V. Along the way, some polynomial-theoretic results are obtained and some crucial congruence relations verified. Finally, a numerical example is presented in Section VI.

0018-9286/81/0800-0832$00.75 ©1981 IEEE

B. On Notations

Complex-Conjugate Transpose: The symbol * will be used for complex conjugate transposes. However, a difference is made between $F^*(z)$ and $[F(z)]^*$: $[F(z)]^*$ means a regular complex-conjugate transpose of $F(z)$, while $F^*(z)$ means a complex-conjugate transpose on the coefficients but *not* on the indeterminate z. In other words, $F^*(z)=[F(\bar{z})]^*$ where $\bar{z}$ stands for the complex conjugate of z.

Degree of a Polynomial Matrix and Degree (Order) of a Transfer Function: The degree of a polynomial matrix, say $P(z)$, is defined to be the degree of its highest degree entry. The degree (order) of a transfer function, say $H(z)$, is defined to be the order of its minimal state-space realizations [12]. While both of them will be symbolized by the same shorthand "deg," i.e., deg $\{P(z)\}$ and deg $\{H(z)\}$, there should be no confusion from the context they reside in.

Subscripts to Matrix Coefficients of Polynomial Matrices: The kth power term matrix coefficient of a polynomial matrix $P(z)$ will be denoted as P_k, e.g., if $P(z)$ has degree not exceeding n, then we have

$$P(z)=P_n z^n + P_{n-1} z^{n-1} + \cdots + P_0 \tag{1.1}$$

where P_i are constant matrices.

Subscripts to Identity Matrices: "I" denotes an identity matrix. The subscript, when present, denotes its dimension.

II. Basic Hankel Properties

We start with basic notations and some useful notions regarding Z-transforms.

A. Z-Transforms

The Z-transform of a square-summable[1] $q \times p$ matrix sequence $\{F_i;\ i=\cdots,-2,-1,0,1,2,\cdots\}$, where F_i are constants, is defined as

$$F(z) \triangleq \sum_{i=-\infty}^{\infty} F_i z^{-i}$$

and the inverse Z-transform as

$$F_i \triangleq \frac{1}{2\pi j}\oint_C F(z) z^{i-1}\,dz, \qquad i=0,\pm 1, \pm 2, \cdots$$

where C is the unit circle, if the integral exists. We shall refer to the half sequence $\{F_i;\ i=\cdots,-2,-1,0\}$ as the *anticausal* part, and the half-sequence $\{F_i;\ i=1,2,\cdots\}$ as the *causal* part.

Denote by $\mathcal{L}_2$ the set of all square-summable infinite sequences, and by[2] $\mathcal{L}_2^-$, $\mathcal{L}_2^+$, respectively, the set of square summable causal and anticausal sequences. We shall not distinguish explicitly between a sequence and its Z-transform representation. Therefore, we shall say $F(z)\in\mathcal{L}_2$ (or $\mathcal{L}_2^-$, $\mathcal{L}_2^+$, respectively) if $\{F_i\}\in\mathcal{L}_2$ ($\mathcal{L}_2^-$, $\mathcal{L}_2^+$, respectively), and vice versa. Note that for any $F(z)\in\mathcal{L}_2$ there exists a *unique* partition

$$F(z)=[F(z)]_- + [F(z)]_+ \tag{2.1}$$

where

$$[F(z)]_- = \sum_{i=1}^{\infty} F_i z^{-i} \in \mathcal{L}_2^-$$

and

$$[F(z)]_+ = \sum_{i=-\infty}^{0} F_i z^{-i} \in \mathcal{L}_2^+.$$

If $F(z)$ happens to be a rational function, then $[F(z)]_-$ will be strictly proper with all its poles inside the unit circle, while $[F(z)]_+$ will have all its poles outside the unit circle.

The following identity will be frequently used in later derivations.

Lemma 2.1—Truncation Property:

For any (multiplicable) matrices $F(z)\in\mathcal{L}_2$ and $G(z)\in\mathcal{L}_2^+$,

$$[F(z)G(z)]_- = [[F(z)]_- G(z)]_-.$$

Proof: $[F(z)G(z)]_- = [[F(z)]_- G(z)]_- + [[F(z)]_+ G(z)]_-$. But since $[F(z)]_+ \in \mathcal{L}_2^+$ and $G(z)\in\mathcal{L}_2^+$, $[F(z)]_+ G(z)\in\mathcal{L}_2^+$ and hence, $[[F(z)]_+ G(z)]_- = 0$. Thus, the result. □

Again consider the sequence $\{F_i\}$. For convenience, define the *antisequence* of $\{F_i\}$, denoted by $\{\check{F}_i\}$, as follows:

$$\check{F}_i \triangleq F_{1-i}, \qquad i=0,\pm 1, \pm 2, \cdots \tag{2.2a}$$

Then its Z-transform, denoted as $\check{F}(z)$, is given by

$$\check{F}(z) \triangleq \sum_{i=-\infty}^{\infty} F_i z^{i-1}. \tag{2.2b}$$

Obviously,

$$\check{F}(z)=z^{-1}F(z^{-1}), \qquad F(z)=z^{-1}\check{F}(z^{-1}). \tag{2.3}$$

Note that if $F(z)\in\mathcal{L}_2^-$, then $\check{F}(z)\in\mathcal{L}_2^+$, and vice versa.

B. Functional Representation of Block-Hankel Operators

A p-input q-output linear, discrete, time-invariant, strictly causal, and stable system can be characterized by a system transfer function matrix

$$H(z)=\sum_{i=1}^{\infty} H_i z^{-i}$$

where $\{H_i;\ i=1,2,\cdots\}$ is the impulse response (matrix) sequence. Corresponding to $H(z)$, we define an *infinite block-Hankel matrix*, denoted by $\Gamma\{H(z)\}$, as

[1] The square of a matrix F is defined to be F^*F.

[2] The space $\mathcal{L}_2^+$ is conventionally termed as Hardy space, denoted as $\mathcal{H}_2$, and $\mathcal{L}_2^- = \mathcal{L}_2 \ominus \mathcal{H}_2$.

$$\Gamma\{H(z)\}=\begin{bmatrix} H_1 & H_2 & H_3 & \cdots \\ H_2 & H_3 & & \\ H_3 & & & \\ \vdots & & & \end{bmatrix}.$$

The matrix $\Gamma\{H(z)\}$ is bounded in $\mathfrak{L}_2$.

Let $\Gamma \triangleq \Gamma\{H(z)\}$. Suppose $\Gamma\boldsymbol{\eta}=\boldsymbol{\zeta}$, where $\boldsymbol{\eta}^*=[\boldsymbol{\eta}_1^*, \boldsymbol{\eta}_2^*, \cdots]$ and $\boldsymbol{\zeta}^*=[\boldsymbol{\zeta}_1^*, \boldsymbol{\zeta}_2^*, \cdots]$. ($\boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathfrak{L}_2^-$; $\boldsymbol{\eta}_i$ are p-vectors, $\boldsymbol{\zeta}_i$ are q-vectors.) Then a simple functional representation of this equation is

$$[H(z)\check{\boldsymbol{\eta}}(z)]_-=\boldsymbol{\zeta}(z) \tag{2.4a}$$

where

$$\check{\boldsymbol{\eta}}(z)=\sum_{i=1}^{\infty}\boldsymbol{\eta}_i z^{i-1}, \qquad \boldsymbol{\zeta}(z)=\sum_{i=1}^{\infty}\boldsymbol{\zeta}_i z^{-i}. \tag{2.4b}$$

We may thus consider the functional representation in (2.4a) as a mapping which transforms an anticausal sequence into a causal one, i.e.,

$$H(z): \mathfrak{L}_2^+ \to \mathfrak{L}_2^-.$$

C. Distance Measures and Unitary Functions

Definition 2.1: Let

$$F(z)=\sum_{i=-\infty}^{\infty}F_i z^{-i}\in\mathfrak{L}_2.$$

Define

i) the sum-squares norm

$$\|F(z)\|_2 \triangleq \left\|\sum_{i=-\infty}^{\infty}F_i^*F_i\right\|_s^{1/2}$$
$$=\left\|\frac{1}{2\pi}\int_0^{2\pi}\left[F(e^{j\theta})\right]^*F(e^{j\theta})\,d\theta\right\|_s^{1/2}, \tag{2.5a}$$

ii) the Hankel norm

$$\|F(z)\|_H \triangleq \left\|\Gamma\{[F(z)]_-\}\right\|_s, \tag{2.5b}$$

and

iii) the Chebyshev norm [13]

$$\|F(z)\|_\infty \triangleq \operatorname*{ess\,sup}_{0\le\theta<2\pi}\|F(e^{j\theta})\|_s \tag{2.5c}$$

where $\|\cdot\|_s$ denotes the spectral norm of a matrix. □

Note that the Hankel norm as defined above is in fact a seminorm. But if we restrict its domain to $F(z)\in\mathfrak{L}_2^-$, it becomes a norm.

With this definition, we can write

$$\|F(z)\|_H=\sup_{\substack{\check{\boldsymbol{\eta}}(z):\ \|\check{\boldsymbol{\eta}}(z)\|_2=1\\ \text{and}\ \check{\boldsymbol{\eta}}(z)\in\mathfrak{L}_2^+}}\left\|\left[[F(z)]_-\check{\boldsymbol{\eta}}(z)\right]_-\right\|_2. \tag{2.6}$$

This equality will be used in the following lemma.

Lemma 2.2—Norm-Inequality

For any matrix $F(z)\in\mathfrak{L}_2$,

$$\|F(z)\|_H\le\|F(z)\|_\infty.$$

If, in addition, $F(z)\in\mathfrak{L}_2^-$, then

$$\|F(z)\|_2\le\|F(z)\|_H\le\|F(z)\|_\infty.$$

Proof: $\forall\check{\boldsymbol{\eta}}(z)\in\mathfrak{L}_2^+$, and $\|\check{\boldsymbol{\eta}}(z)\|_2=1$,

$$\left[[F(z)]_-\check{\boldsymbol{\eta}}(z)\right]_-=\left[F(z)\check{\boldsymbol{\eta}}(z)\right]_-$$

by truncation property. Now

$$\left\|[F(z)\check{\boldsymbol{\eta}}(z)]_-\right\|_2\le\|F(z)\check{\boldsymbol{\eta}}(z)\|_2$$
$$=\frac{1}{2\pi}\int_0^{2\pi}\left[F(e^{j\theta})\check{\boldsymbol{\eta}}(e^{j\theta})\right]^*\left[F(e^{j\theta})\check{\boldsymbol{\eta}}(e^{j\theta})\right]d\theta$$
$$\le\frac{1}{2\pi}\int_0^{2\pi}\operatorname*{ess\,sup}_\theta\|F(e^{j\theta})\|_s\cdot\|\check{\boldsymbol{\eta}}(e^{j\theta})\|_2\,d\theta$$
$$=\frac{1}{2\pi}\operatorname*{ess\,sup}_\theta\|F(e^{j\theta})\|_s\cdot\int_0^{2\pi}1\ d\theta$$
$$=\operatorname*{ess\,sup}_\theta\|F(e^{j\theta})\|_s=\|F(z)\|_\infty.$$

Thus, $\|[[F(z)]_-\check{\boldsymbol{\eta}}(z)]_-\|_2\le\|F(z)\|_\infty$, and hence, $\|F(z)\|_H\le\|F(z)\|_\infty$ by (2.6).

Now let $\boldsymbol{v}$ be a normalized singular vector corresponding to the maximum singular value of $\Sigma_{i=-\infty}^{\infty}F_i^*F_i$. Then, if $F(z)\in\mathfrak{L}_2^-$, $[F(z)\boldsymbol{v}]_-=F(z)\boldsymbol{v}$ and

$$\left\|[F(z)\boldsymbol{v}]_-\right\|_2=\|F(z)\boldsymbol{v}\|_2=\|F(z)\|_2.$$

Hence, in view of (2.6), $\|F(z)\|_2\le\|F(z)\|_H$. Q.E.D. □

Remark: It is possible now to claim a partial justification of the adoption of Hankel-norm criterion; in that the Hankel-norm lies between two other conventional norms: sum squares and Chebyshev norms. □

Definition 2.2: A square matrix $F(z)$ is said to be *unitary* [13] if $F^*(z^{-1})F(z)=I$. □

Lemma 2.3

If $F(z)$ is unitary, then

$$\|F(z)\|_2=\|F(z)\|_\infty=1.$$

Proof: If $F(z)$ is unitary, then

$$\|F(e^{j\theta})\|_s=\left\|\left[F(e^{j\theta})\right]^*F(e^{j\theta})\right\|_s^{1/2}$$
$$=\|F^*(e^{-j\theta})F(e^{j\theta})\|_s^{1/2}=1 \qquad \forall\theta$$

where the last equality comes from the above definition. Hence,

$$\|F(z)\|_\infty=\operatorname*{ess\,sup}_{0\le\theta<2\pi}\|F(e^{j\theta})\|_s=1$$

and

$$\|F(z)\|_2 = \left\| \frac{1}{2\pi} \int_0^{2\pi} [F(e^{j\theta})]^* F(e^{j\theta})\, d\theta \right\|_s^{1/2} = \left\| \frac{1}{2\pi} \int_0^{2\pi} I \cdot d\theta \right\|_s^{1/2} = 1. \quad \text{Q.E.D.} \ \square$$

D. System Order and Singular Value Analysis

A well-known result connecting the degree (order) of a system with the rank of the corresponding block-Hankel matrix is the following.

Kronecker's Theorem [14], [15], [12]

Let $H(z)$ be a system transfer function. Then the rank of $\Gamma\{H(z)\}$ is equal to the degree (order) of the system $H(z)$. $\square$

In practice, to determine the rank of a block-Hankel matrix is not an easy matter. The elements of the matrix are seldom given exactly, and it is unlikely that the matrix formed by these approximate values will have the same rank as the "true" one. Moreover, many rank test algorithms, e.g., Gauss elimination, tend to turn some low rank matrices into full rank. Therefore, a robust approach to characterize the rank of a matrix is needed. For this purpose, the singular value analysis on block-Hankel matrices appears to be a powerful and convenient tool [16], [17].

With $\Gamma\{H(z)\}$ bounded in ℓ_2, there exists a *singular-value decomposition* of $\Gamma\{H(z)\}$ in the form

$$\Gamma\{H(z)\} = \sum_{i=1}^{\infty} \sigma_i \zeta^{(i)} \eta^{(i)*} \tag{2.7a}$$

where the numbers $\{\sigma_i\}$ are nonnegative and are termed the singular values and $\eta^{(i)}, \zeta^{(i)}$ are such that

$$\eta^{(i)*}\eta^{(j)} = \zeta^{(i)*}\zeta^{(j)} = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{if } i \neq j \end{cases} \tag{2.7b}$$

and they are called the singular vectors. In all the following, we assume that the singular values are ordered in such a way that

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots. \tag{2.8}$$

The very first question in an approximation problem is perhaps that how close, say, a kth-order approximant $H^{(k)}(z)$ can approximate a higher order system $H(z)$. One possible closeness measure can be defined as the spectral norm of the Hankel matrix of the difference $\tilde{\tilde{H}}(z) \triangleq H(z) - H^{(k)}(z)$, i.e.,

$$\left\| \Gamma\{\tilde{\tilde{H}}(z)\} \right\|_s = \left\| \Gamma\{H(z)\} - \Gamma\{H^{(k)}(z)\} \right\|_s.$$

For convenience, we shall term this measure the *Hankel-norm distance*.

A well-known singular value analysis result is the following:

$$\inf_{A:\, \text{Rank}(A) \leq k} \|\Gamma\{H(z)\} - A\|_s = \sigma_{k+1}. \tag{2.9}$$

From this, we introduce the following lower bound properties, which will be a primary guidance on how close we can approximate $H(z)$ by a lower order model.

Lemma 2.4—Minimum-Norm Bound

For any kth-order system $H^{(k)}(z)$, the Hankel-norm distance between $H^{(k)}(z)$ and $H(z)$ is bounded from below by σ_{k+1}, i.e.,

$$\left\| \Gamma\{\tilde{\tilde{H}}(z)\} \right\|_s = \left\| \Gamma\{H(z)\} - \Gamma\{H^{(k)}(z)\} \right\|_s \geq \sigma_{k+1}.$$

Proof: By Kronecker's theorem, $\Gamma\{H^{(k)}(z)\}$ has rank k. Hence, from (2.9) we have this inequality. $\square$

This property may be stated in a slightly different version.

Lemma 2.5—Minimum-Degree Bound

Suppose ρ is a number in the interval $[\sigma_{k+1}, \sigma_k)$. Then if $H_a(z)$ meets the Hankel-norm tolerance that

$$\left\| \Gamma\{H(z) - H_a(z)\} \right\|_s \leq \rho.$$

$H_a(z)$ must be of degree greater than or equal to k. $\square$

In fact, as we progress into later sections, we shall show that for any tolerance $\rho \in (\sigma_{k+1}, \sigma_k)$, there always exists a kth-order qualified approximant such that the minimum-degree bound is achieved. The problem of finding such approximants, with a preassigned tolerance ρ, will be termed the minimal degree approximation (MDA) problem.

III. Minimal-Degree Approximations

The goal of this section is to derive a solution to the MDA problem. Our plan is to first explore the underlying algebraic framework and then induce a polynomial formulation for the solution construction. While this section deals only with the square ($p \times p$) transfer function case, we shall further show, in Section V, that any nonsquare system can be treated as a part of a square system and the same approximation scheme can be carried through naturally.

To give some background, we briefly review the scalar system approximation result by Adamjan *et al.* [1], which is the key inspiration of the present work.

Consider the singular value decomposition on a scalar system Hankel matrix:

$$\Gamma = \sum_{i=1}^{\infty} \sigma_i \zeta^{(i)} \eta^{(i)*}$$

as in (2.7). Let us treat the two infinite vectors $\eta^{(k+1)}$ and $\zeta^{(k+1)}$ as two causal sequences and denote by $\eta^{(k+1)}(z)$ and $\zeta^{(k+1)}(z)$ the Z-transforms of these sequences. Denote also $\tilde{\eta}^{(k+1)}(z) = z^{-1}\eta^{(k+1)}(z^{-1})$, as we did in (2.4). Now define $E(z) = \sum_{i=1}^{\infty} e_i z^{-i}$ with

$$e_i = \frac{1}{2\pi j}\oint_c \frac{\zeta^{(k+1)}(z)}{\check{\eta}^{(k+1)}(z)} z^{i-1}\,dz \quad (c\text{: the unit circle}).$$

In short, $E(z)=[\zeta^{(k+1)}(z)/\check{\eta}^{(k+1)}(z)]_-$. Let $\mathcal{E}$ be the infinite Hankel matrix corresponding to the sequence $\{e_i; i=1,2,\cdots\}$, i.e., $\mathcal{E}=\Gamma\{E(Z)\}$. Then according to a theorem of Nehari [11], $\|\mathcal{E}\|_s=1$. More remarkable is that the difference Hankel $\Gamma^{(k)} \triangleq (\Gamma-\sigma_{k+1}\mathcal{E})$ has rank k. This by Kronecker's theorem represents a kth-order system. And it approximates the original one with smallest possible error σ_{k+1} in Hankel-norm distance sense (Lemma 2.4).

As mentioned before, it is so far the only closed-form solution to the approximation problem with any optimality criterion. Unlike the more traditional SVD approach via $\mathcal{O}\mathcal{C}$ factorization [2], [9], this new method does not make any use of the first k singular vectors. Instead, it uses simply the $(k+1)$th singular vector to construct an error Hankel matrix. Note also that, despite its elegance in theoretical development, there is no feasible computational scheme available to numerically solve for the optimal approximants, except for finite order systems case [3], [5].

With the above background, let us now turn to the multivariable case.

A. MDA Solution for a Special Case

Suppose that $H(z)$ is the square $(p\times p)$ system function to be approximated, and that ρ is the tolerance for the error in Hankel-norm measure. For the purpose of a clean mathematical derivation, we first impose a rather artificial but heuristic assumption: Assume that $\Gamma\{H(z)\}$ has a singular value of multiplicity p and ρ is exactly that singular value; i.e., for some integer k,

$$\rho=\sigma_{k+1}=\sigma_{k+2}=\cdots=\sigma_{k+p} \tag{3.1a}$$

and

$$\sigma_k>\rho>\sigma_{k+p+1}. \tag{3.1b}$$

Such a restrictive assumption assures us exactly p independent pairs of singular vectors corresponding to (the singular value) ρ. Therefore, heuristically, the situation becomes compatible with the scalar case discussed above. More precisely, a (would be) error Hankel corresponding to a (would be) kth-order approximation can now be constructed from these p pairs of singular vectors as follows.

Let $\{(\eta^{(i)},\zeta^{(i)}); i=k+1,\cdots,k+p\}$ be a set of linearly independent singular vector pairs of Γ corresponding to the singular value ρ. Treat $\eta^{(i)}$ and $\zeta^{(i)}$ as concatenations of p-dimensional causal sequences and take Z-transforms as in Section II-B. Let the results be $\eta^{(i)}(z)$, $\zeta^{(i)}(z)$. Denote

$$X(z)\triangleq\left[\eta^{(k+1)}(z) \,\vdots\, \cdots \,\vdots\, \eta^{(k+p)}(z)\right] \tag{3.2a}$$

$$Y(z)\triangleq\left[\zeta^{(k+1)}(z) \,\vdots\, \cdots \,\vdots\, \zeta^{(k+p)}(z)\right]. \tag{3.2b}$$

Now designate (assuming $X(z)$ is nonsingular)

$$E(z)\triangleq\left[Y(z)\check{X}^{-1}(z)\right]_-, \qquad \mathcal{E}\triangleq\Gamma\{E(z)\}. \tag{3.3}$$

We shall show that

$$\|\mathcal{E}\|_s=1 \tag{3.4a}$$

and, more stimulatingly,

$$\operatorname{rank}\Gamma\{H^{(k)}(z)\}=k \tag{3.4b}$$

where $H^{(k)}(z)\triangleq H(z)-\rho E(z)$. In other words, $H^{(k)}(z)$ so constructed is indeed a kth-order optimal approximant with error, or approximation distance, being exactly $\rho=\|\rho E(z)\|_H$.

Proof of the Distance Property (3.4a): Employing the functional representation introduced in Section II-A, we can write

$$\left[H(z)\check{X}(z)-\rho Y(z)\right]_-=0 \tag{3.5a}$$

$$\left[H^*(z)\check{Y}(z)-\rho X(z)\right]_-=0 \tag{3.5b}$$

since $X(z), Y(z)$ represent the singular vectors associated with ρ. Equations (3.5) will be called the *composite singular equations* corresponding to ρ. They are in fact equivalent to

$$H(z)\check{X}(z)-\rho Y(z)=\check{K}(z) \tag{3.6a}$$

$$H^*(z)\check{Y}(z)-\rho X(z)=\check{L}(z) \tag{3.6b}$$

for some $\check{K}(z), \check{L}(z)\in\mathcal{L}_2^+$. Taking complex-conjugate transpose of (3.6b),

$$\check{Y}^*(z)H(z)-\rho X^*(z)=\check{L}^*(z). \tag{3.6b*}$$

Then $\check{Y}^*(z)\times(3.6a)-(3.6b^*)\times\check{X}(z)$ leads to

$$-\rho\left[\check{Y}^*(z)Y(z)-X^*(z)\check{X}(z)\right]=\check{Y}^*(z)\check{K}(z)-\check{L}^*(z)\check{X}(z). \tag{3.7}$$

Denote $M(z)\triangleq-\rho[\check{Y}^*(z)Y(z)-X^*(z)\check{X}(z)]$. Then it is not hard to show that $z^{-1}\check{M}^*(z)=M(z)$. However, from the right-hand side of (3.7) it can be seen that $M(z)\in\mathcal{L}_2^+$. Therefore, $z^{-1}\check{M}^*(z)$ should belong to the class $\mathcal{L}_2^-$, i.e., the only situation making $M(z)=z^{-1}\check{M}^*(z)$ is that $M(z)=0$. Hence, we have

$$\check{Y}^*(z)Y(z)-X^*(z)\check{X}(z)=0$$

or

$$\left[X^{*-1}(z)\check{Y}^*(z)\right]\left[Y(z)\check{X}^{-1}(z)\right]=I. \tag{3.8}$$

Thus, $Y(z)\check{X}^{-1}(z)$ is unitary (Definition 2.2). And hence, $\|E(z)\|_H=\|Y(z)\check{X}^{-1}(z)\|_H\leqslant\|Y(z)\check{X}^{-1}(z)\|_\infty=1$ (Lemmas 2.2, 2.3). However, we can show that 1 is indeed a singular value of $\Gamma\{E(z)\}$. This is done by considering

$$\begin{aligned}\left[E(z)\check{\eta}^{(k+1)}(z)\right]_- &= \left[\left[Y(z)\check{X}^{-1}(z)\right]_-\check{\eta}^{(k+1)}(z)\right]_- \\ &= \left[Y(z)\check{X}^{-1}(z)\check{\eta}^{(k+1)}(z)\right]_- \\ &= \left[Y(z)\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}\right]_- = \zeta^{(k+1)}(z)\end{aligned}$$

Hence, $[E(z)\check{\eta}^{(k+1)}(z)]_-=1\cdot\zeta^{(k+1)}(z)$. But since

$\|\boldsymbol{\eta}^{(k+1)}(z)\|_2=\|\boldsymbol{\zeta}^{(k+1)}(z)\|_2(=1)$, we conclude that 1 is a singular value of $\Gamma\{E(z)\}$. Thus, $\|\mathcal{E}\|_s=1$. Q.E.D.□

Therefore, the approximant $H^{(k)}(z)=H(z)-\rho E(z)$ meets the tolerance requirement on Hankel-norm distance.

Although it is possible to prove the degree property (3.4b) directly, we have found that a polynomial language can considerably simplify the proof, and provide a basis for our later algorithm derivation as well. The polynomial language calls for the following lemma.

Lemma 3.1

Suppose that $H(z)$ is of finite degree, i.e.,

$$H(z)=\frac{1}{a(z)}N(z) \tag{3.9}$$

for some polynomial $a(z)$ (assumed *monic* and of *degree n* henceforth) and some $p\times p$ polynomial matrix $N(z)$ with degree less than n. Then $X(z)$ and $Y(z)$ can be written as

$$X(z)=\frac{1}{a^*(z)}P(z) \qquad Y(z)=\frac{1}{a(z)}Q(z) \tag{3.10}$$

for some polynomial matrices $P(z)$ and $Q(z)$ with degrees less than or equal to $n-1$.

Proof: We first prove the part for $Y(z)$. From (3.5a) we have

$$\left[H(z)\check{X}(z)-\rho Y(z)\right]_-\cdot a(z)=0$$

and hence,

$$\left[\left[H(z)\check{X}(z)-\rho Y(z)\right]_-\cdot a(z)\right]_-=0.$$

By truncation property,

$$\left[\left\{H(z)\check{X}(z)-\rho Y(z)\right\}a(z)\right]_-=0$$

or

$$\left[N(z)\check{X}(z)-\rho Y(z)a(z)\right]_-=0.$$

Now since $[N(z)\check{X}(z)]_-=0$, we have $[Y(z)a(z)]_-=0$. Thus, $Y(z)=(1/a(z))Q(z)$ for some polynomial matrix $Q(z)$ of degree less than or equal to $n-1$, because $[Y(z)]_+=0$. The part for $X(z)$ can be similarly proved through working on the other singular equation (3.5b). □

Denote by $\hat{a}(z)$, $\hat{P}(z)$, and $\hat{Q}(z)$, respectively, the reciprocal polynomials of $a(z)$, $P(z)$, and $Q(z)$, i.e.,

$$\hat{a}(z)\triangleq z^n a(z^{-1}) \tag{3.11a}$$

and

$$\hat{P}(z)\triangleq z^{n-1}P(z^{-1}), \qquad \hat{Q}(z)\triangleq z^{n-1}Q(z^{-1}). \tag{3.11b}$$

Then we can write

$$\check{X}(z)=\frac{1}{\hat{a}^*(z)}\hat{P}(z), \qquad \check{Y}(z)=\frac{1}{\hat{a}(z)}\hat{Q}(z). \tag{3.12}$$

Hence, from (3.6a),

$$\check{K}(z)=\frac{N(z)\hat{P}(z)-\rho\hat{a}^*(z)Q(z)}{a(z)\hat{a}^*(z)}.$$

As $\check{K}(z)$ is in the class $\mathfrak{L}_2^+$, the roots of $a(z)$ can not be its poles and, therefore, $a(z)$ has to be cancelled by the numerator. In other words, the entries of the matrix

$$T(z)\triangleq\frac{N(z)\hat{P}(z)-\rho\hat{a}^*(z)Q(z)}{a(z)} \tag{3.13a}$$

must be polynomials. Their degrees must be less than or equal to $n-1$, as a simple result from studying the degree of each term on the right-hand side of the above expression.

Similarly, (3.6b) will lead us to conclude that

$$W(z)\triangleq\frac{N^*(z)\hat{Q}(z)-\rho\hat{a}(z)P(z)}{a^*(z)} \tag{3.13b}$$

is a polynomial matrix of degree not exceeding $n-1$. We, therefore, obtain an alternative for the singular equations (3.5), in *rational* form.

Lemma 3.2

The composite singular equations (3.5) have a rational form as

$$H(z)\hat{P}(z)-\rho\frac{\hat{a}^*(z)}{a(z)}Q(z)=T(z) \tag{3.14a}$$

$$H^*(z)\hat{Q}(z)-\rho\frac{\hat{a}(z)}{a^*(z)}P(z)=W(z) \tag{3.14b}$$

where

$$\deg T(z)\le n-1, \qquad \deg W(z)\le n-1. \tag{3.14c}$$

□

Now the degree claim (3.4b) can be stated in a polynomial setting. Note first that (3.14a) leads to

$$H(z)-\rho Y(z)\check{X}^{-1}(z)=H(z)-\rho\frac{\hat{a}^*(z)}{a(z)}Q(z)\hat{P}^{-1}(z)$$
$$=T(z)\hat{P}^{-1}(z).$$

Hence,

$$H^{(k)}(z)=\left[H(z)-\rho Y(z)\check{X}^{-1}(z)\right]_-=\left[T(z)\hat{P}^{-1}(z)\right]_-.$$

Therefore, equation (3.4b) is equivalent to saying that

$$\deg H^{(k)}(z)=\deg\left\{\left[T(z)\hat{P}^{-1}(z)\right]_-\right\}=k. \tag{3.15}$$

Proof of the Degree Property (3.4b),(3.15): For a neater language, we first define several more terms.

Definition 3.1: A unitary matrix $U(z)$ is said to be *inner* if all its poles are outside the unit circle. □

Such an inner matrix has all its zeros inside the unit circle and located at corresponding conjugate-reciprocal positions of its poles.

It is known [13] that for any $p\times p$ rational function matrix $F(z)$ having all its poles outside the unit circle, there exists a left factorization

$$F(z)=U_L(z)F_L(z) \tag{3.16a}$$

and also a right factorization

$$F(z)=F_R(z)U_R(z) \tag{3.16b}$$

where $U_L(z)$, $F_L(z)$, $F_R(z)$, and $U_R(z)$ are $p\times p$ matrices, and they have the following properties.

i) $U_L(z)$ and $U_R(z)$ are inner. Their zeros are those of $F(z)$ that locate inside the unit circle.

ii) $F_L(z)$ and $F_R(z)$ are maximal phase, i.e., all their zeros (and poles) are not inside the unit circle.

Obviously, $U_L(z)$ and $U_R(z)$ share the same zeros and poles, and hence the same degree. So do $F_L(z)$ and $F_R(z)$.

Definition 3.2: $U_L(z)$ and $U_R(z)$ are called the *left inner factor* and the *right inner factor* of $F(z)$, respectively. The degree of the inner factors is called the *inner degree*. □

We are to show that $\deg\{[T(z)\hat{P}^{-1}(z)]_-\}=k$. Note first that the system poles of $[T(z)\hat{P}^{-1}(z)]_-$ are these "inner zeros" (zeros locating inside the unit circle) of $\det\{\hat{P}(z)\}$ that are not cancelled by $T(z)$. *Assume* $\det\{P(z)\}$ *has* m "*inner zeros*." Then the system $[T(z)\hat{P}^{-1}(z)]_-$ will have a degree $m'\leq m$. Note, however, since

$$\check{X}(z)=\frac{1}{\hat{a}^*(z)}\hat{P}(z),$$

the number of inner zeros of $\hat{P}(z)$ in fact gives the inner degree of $\check{X}(z)$, i.e., *the inner degree of* $\check{X}(z)$ *is equal to* m. The following lemma, related to the inner degree of $\check{X}(z)$, will help in establishing an equality between m, m', and k.

Lemma 3.3

Let $\check{Y}(z)=U_{yL}(z)\check{Y}_L(z)$ be a left inner factorization of $\check{Y}(z)$, and let $\{\sigma_i'\}$ denote the decreasingly ordered singular values of $\Gamma\{[H^*(z)U_{yL}(z)]_-\}$.

i) *Dominance Property*: $\sigma_i'\leq\sigma_i\ \forall i$.

ii) *Multiplicity Property*: If the inner degree of $\check{X}(z)$ is m, then

$$\sigma_1'=\sigma_2'=\cdots=\sigma_{m+p}'=\rho$$

Proof: See Appendix A. □

From the multiplicity property, we have $\rho=\sigma_{m+p}'$. From the dominance property, we have $\sigma_{m+p}'\leq\sigma_{m+p}$. Combining with $\rho=\sigma_{k+p}>\sigma_{k+p+1}$ (3.1), $\sigma_{m+p}\geq\sigma_{m+p}'=\rho=\sigma_{k+p}>\sigma_{k+p+1}$. Hence, $\sigma_{m+p}>\sigma_{k+p+1}$. This implies $m<k+1$, or $m'\leq k$. But m' cannot be smaller than k, for otherwise it would violate the minimum degree bound lemma as $\|\Gamma\{H(z)-H^{(k)}(z)\}\|_s=\|\rho\mathcal{E}\|_s=\rho=\sigma_{k+1}$. Thus, $m=m'=k$, and

$$\deg H^{(k)}(z)=\deg\{[T(z)\hat{P}^{-1}(z)]_-\}=k$$

(Note also that the above result also proves that no cancellation can happen between the "inner zeros" of $T(z)$ and $\det\{\hat{P}(z)\}$.) Q.E.D. □

Summarizing, we have the following.

Lemma 3.4—The MDA Lemma

Let the singular values of $\Gamma\{H(z)\}$ be such that $\sigma_k>\rho=\sigma_{k+1}=\sigma_{k+2}=\cdots=\sigma_{k+p}>\sigma_{k+p+1}$, and let $\{P(z),Q(z),T(z),W(z)\}$ be as defined in Lemmas 3.1 and 3.2. Denote

$$H^{(k)}(z)\triangleq[T(z)\hat{P}^{-1}(z)]_-$$

Then

i) $\|\Gamma\{H(z)-H^{(k)}(z)\}\|_s=\rho$

ii) $\deg H^{(k)}(z)=k$

i.e., the claim (3.4) given earlier in this section is justified. □

Remark: A straightforward dual argument will show that $H^{(k)}(z)$ can also be derived as

$$H^{(k)}(z)=[\hat{Q}^{*-1}(z)W^*(z)]_-$$ □

B. Regular Situation

Clearly, the major difficulty yet to resolve is that it is very unlikely that any singular value of Γ will have multiplicity p. Besides, ρ, the assigned tolerance on approximation error (Hankel-norm distance), can be other than one of the singular values. Motivated by the scalar case results [1], [5], we propose a solution toward such situation by a proper "extension" of the original system.

Definition 3.3: An *extension* of the matrix $\Gamma=\Gamma\{H(z)\}$ (where $H(z)=\sum_{i=1}^{\infty}H_iz^{-i}$), denoted as $\underline{\Gamma}$, is a block-Hankel matrix generated by the sequence $\{H_0,H_1,H_2,\cdots\}$, i.e.,

$$\underline{\Gamma}=\begin{bmatrix}H_0 & H_1 & H_2 & \cdots\\ H_1 & H_2 & H_3 & \\ H_2 & H_3 & & \\ \vdots & & & \ddots\end{bmatrix}$$

where H_0 is a $p\times p$ constant matrix. □

Denote by $\{\underline{\sigma}_i\}$ $(i=1,2,\cdots)$ the decreasingly ordered singular value set for $\underline{\Gamma}$. We have the following.

The Extension Theorem

Given Γ with complex (real) entries, then for every number $\rho\in(\sigma_{k+1},\sigma_k)$, $(k=1,2,\cdots)$, there exists at least one complex (real) extension $\underline{\Gamma}$ for Γ such that ρ is a singular value of $\underline{\Gamma}$ with multiplicity p. More precisely, when such extension exists,

$$\underline{\sigma}_k>\rho=\underline{\sigma}_{k+1}=\underline{\sigma}_{k+2}=\cdots=\underline{\sigma}_{k+p}>\underline{\sigma}_{k+p+1}. \tag{3.17}$$

Proof: Provided in next section. □

Treat $\{H_0,H_1,H_2,\cdots\}$ as a strictly causal sequence, and let

$$\underline{H}(z)=\sum_{i=1}^{\infty}H_{i-1}z^{-i}=H_0z^{-1}+H(z)z^{-1}.$$

If $\tilde{\Gamma}$ is a "properly extended Hankel" for $\rho\in(\sigma_{k+1},\sigma_k)$ as stated in the extension theorem, then we can apply the MDA lemma to obtain a kth order approximant $\tilde{H}^{(k)}(z)$ of $\tilde{H}(z)$:

$$\tilde{H}^{(k)}(z)=\left[\tilde{T}(z)\tilde{P}^{-1}(z)\right]_-=\left[\hat{\tilde{Q}}^{*-1}(z)\tilde{W}^*(z)\right]_-$$

where $\tilde{T}(z)$, $\tilde{P}(z)$, $\tilde{Q}(z)$, and $\tilde{W}(z)$ are defined parallelly to their "untilded" counterparts in Lemmas 3.1 and 3.2.

Now we are ready to present the main conclusion of the above study.

Main Theorem—The Minimal Degree Approximation Theorem

Given a complex (real) coefficient system $H(z)$ and a tolerance ρ for approximation error (Hankel-norm distance). Then a complex (real) *minimal degree approximant* for $H(z)$ is given by

$$H_{\text{mda}}(z)=\left[z\tilde{T}(z)\hat{\tilde{P}}^{-1}(z)\right]_-=\left[\hat{\tilde{Q}}^{*-1}(z)\tilde{W}^*(z)z\right]_-$$

where $\{\tilde{T}(z),\tilde{P}(z),\tilde{Q}(z),\tilde{W}(z)\}$ is a solution to the rational form singular equations (3.14) corresponding to a properly extended system $\tilde{H}(z)$ (i.e., ρ is a singular value of $\Gamma\{\tilde{H}(z)\}$ with multiplicity p). In other words, we have

i) $\|H(z)-H_{\text{mda}}(z)\|_H\leq\rho$

ii) $\deg H_{\text{mda}}(z)=k$, if $\rho\in(\sigma_{k+1},\sigma_k)$.

Proof: Notice that $\Gamma\{H(z)-H_{\text{mda}}(z)\}$ is a submatrix of $\Gamma\{\tilde{H}(z)-\tilde{H}^{(k)}(z)\}$. Hence, $\|\Gamma\{H(z)-H_{\text{mda}}(z)\}\|_s\leq\|\Gamma\{\tilde{H}(z)-\tilde{H}^{(k)}(z)\}\|_s=\rho$, in view of Part i) of the MDA lemma. As to the degree of $H_{\text{mda}}(z)$, note that it cannot be greater than that of $\tilde{H}^{(k)}(z)$ since $\Gamma\{H_{\text{mda}}(z)\}$, being a submatrix of $\Gamma\{\tilde{H}^{(k)}(z)\}$, has a rank less than or equal to that of $\Gamma\{\tilde{H}^{(k)}(z)\}$. On the other hand, the minimum degree bound lemma sets k as the lower bound for $\deg\{H_{\text{mda}}(z)\}$. Thus, the result. □

The above results also imply that, theoretically speaking, we can attain a tolerance $\rho=\sigma_{k+1}+\varepsilon$ with a kth degree approximant for any small $\varepsilon>0$. And thus the minimum norm bound is *essentially* also achievable. However, the numerical aspects of this issue deserves a closer attention. Preliminary studies seems to indicate that the tolerance ρ can be a singular value *without* jeopardizing the existence of a proper extension. The mathematical analysis for this situation is currently under our investigation.

IV. Proof of the Extension Theorem

We choose to adopt a polynomial function, which will not only facilitate the proof but also benefit the derivation of an efficient algorithm (detailed in next section) numerically solving for minimal degree approximations.

Note that the extended system $\tilde{H}(z)$ can be written as

$$\tilde{H}(z)=H_0z^{-1}+H(z)z^{-1}=\frac{1}{za(z)}\{H_0a(z)+N(z)\}\tag{4.1}$$

where the denominator $za(z)$ has degree $n+1$. Suppose that a proper extension has been found (so that ρ is a p-multiple singular value of $\tilde{\Gamma}$) and that $(\tilde{P}(z)/za^*(z),\tilde{Q}(z)/za(z))$ represents a set of p linearly independent singular vector pairs [cf. (3.10)] associated with the singular value ρ. Then, using an argument parallel to that in Lemmas 3.1 and 3.2, we get

$$\tilde{H}(z)\hat{\tilde{P}}(z)-\rho\frac{\hat{a}^*(z)}{za(z)}\tilde{Q}(z)=\tilde{T}(z)\tag{4.2a}$$

$$\tilde{H}^*(z)\hat{\tilde{Q}}(z)-\rho\frac{\hat{a}(z)}{za^*(z)}\tilde{P}(z)=\tilde{W}(z)\tag{4.2b}$$

where $\tilde{P}(z)$, $\tilde{Q}(z)$, $\tilde{T}(z)$, and $\tilde{W}(z)$ are all polynomial matrices with

$$\deg\tilde{P}(z)\leq n,\qquad \hat{\tilde{P}}(z)\triangleq z^n\tilde{P}(z^{-1})\tag{4.3a}$$

$$\deg\tilde{Q}(z)\leq n,\qquad \hat{\tilde{Q}}(z)\triangleq z^n\tilde{Q}(z^{-1})\tag{4.3b}$$

$$\deg\tilde{T}(z)\leq n-1,\qquad \deg\tilde{W}(z)\leq n-1.\tag{4.3c}$$

(The trivial verification is left to the reader.) By (4.1), equations (4.2) are equivalent to

$$H(z)\hat{\tilde{P}}(z)-\rho\frac{\hat{a}^*(z)}{a(z)}\tilde{Q}(z)=z\tilde{T}(z)-H_0\hat{\tilde{P}}(z)\triangleq R(z)\tag{4.4a}$$

$$H^*(z)\hat{\tilde{Q}}(z)-\rho\frac{\hat{a}(z)}{a^*(z)}\tilde{P}(z)=z\tilde{W}(z)-H_0^*\hat{\tilde{Q}}(z)\triangleq S(z).\tag{4.4b}$$

Backward tracing of the above discussion leads us to the following polynomial formulation of the problem.

Lemma 4.1

Let $\tilde{P}(z)$, $\tilde{Q}(z)$, $R(z)$, and $S(z)$ be $p\times p$ polynomial matrices with degrees not exceeding n that satisfy the equations

$$\frac{N(z)}{a(z)}\hat{\tilde{P}}(z)-\rho\frac{\hat{a}^*(z)}{a(z)}\tilde{Q}(z)=R(z)\tag{4.5a}$$

$$\frac{N^*(z)}{a^*(z)}\hat{\tilde{Q}}(z)-\rho\frac{\hat{a}(z)}{a^*(z)}\tilde{P}(z)=S(z)\tag{4.5b}$$

If $\tilde{P}(z)$ or $\tilde{Q}(z)$ is nonsingular and there exists a constant $p\times p$ matrix H_0 such that

$$R(0)+H_0\hat{\tilde{P}}(0)=0\tag{4.6a}$$

$$S(0)+H_0^*\hat{\tilde{Q}}(0)=0\tag{4.6b}$$

then ρ is a singular value of the extended Hankel matrix $\tilde{\Gamma}$ (with H_0 being the extending block as in Definition 3.3) and the columns of $(\tilde{P}(z)/za^*(z),\tilde{Q}(z)/za(z))$ represent p *linearly independent* pairs of singular vectors associated with ρ.

Proof: If (4.5) and (4.6) are true, then we define

$$\underset{\sim}{T}(z) \triangleq z^{-1}\left[R(z)+H_0\hat{\underset{\sim}{P}}(z)\right],$$

$$\underset{\sim}{W}(z) \triangleq z^{-1}\left[S(z)+H_0^*\underset{\sim}{Q}(z)\right].$$

And then (4.2)–(4.4) can be seen to be true. By Lemma 3.2, it is clear that $(\underset{\sim}{P}(z)/za^*(z), \underset{\sim}{Q}(z)/za(z))$ represents p independent pairs of singular vectors associated with a singular value ρ of $\underset{\sim}{\Gamma}$. Q.E.D. □

Based on this lemma, the study on the existence of a proper extension will be carried out in two steps. *First*, we shall obtain a special set of solutions $\{\underset{\sim}{P}(z), \underset{\sim}{Q}(z), R(z), S(z)\}$ to (4.5). *Second*, we look, among these solutions, for one that satisfies (4.6).

For the first step, note that by defining

$$\hat{N}(z) \triangleq z^n N(z^{-1}), \quad \hat{S}(z) \triangleq z^n S(z^{-1}), \tag{4.7}$$

(4.5) can be written as

$$\begin{bmatrix} -\rho\hat{a}^*(z)I_p & N(z) \\ \hat{N}^*(z) & -\rho a(z)I_p \end{bmatrix}\begin{bmatrix} \underset{\sim}{Q}(z) \\ \hat{\underset{\sim}{P}}(z) \end{bmatrix} = \begin{bmatrix} a(z)I_p & 0 \\ 0 & \hat{a}^*(z)I_p \end{bmatrix}\begin{bmatrix} R(z) \\ \hat{S}(z) \end{bmatrix} \tag{4.8}$$

which is equivalent to

$$\begin{bmatrix} a(z)I_p & -N(z) & \rho\hat{a}^*(z)I_p & 0 \\ 0 & \rho a(z)I_p & -\hat{N}^*(z) & \hat{a}^*(z)I_p \end{bmatrix}\begin{bmatrix} R(z) \\ \hat{\underset{\sim}{P}}(z) \\ \underset{\sim}{Q}(z) \\ \hat{S}(z) \end{bmatrix} = 0. \tag{4.9}$$

For convenience, denote

$$V(z) \triangleq \begin{bmatrix} a(z)I_p & N(z) & \rho\hat{a}^*(z)I_p & 0 \\ 0 & \rho a(z)I_p & -\hat{N}^*(z) & \hat{a}^*(z)I_p \end{bmatrix}. \tag{4.10}$$

Clearly, the solution space of (4.9) is $2p$-dimensional. It is also known that [18] this space is spanned by a *minimal basis*, whose key properties are summarized below for later reference. See [18] for proof.

Lemma 4.2

Let $Z(z)$ be a minimal basis for the solution space of (4.9). Then $Z(z)$ is a $4p\times 2p$ polynomial matrix such that

$$V(z)Z(z)=0 \tag{4.11}$$

and

i) $Z(z)$ is column-proper, i.e., the highest degree coefficient vectors, one from each column's highest degree term, combine to yield a linearly independent set;

ii) the column-degree-sum (sum of the degrees of the columns) of $Z(z)$ is no greater then the row-degree-sum (sum of the degrees of the rows) of $V(z)$. □

Based on the special structure of $V(z)$ in (4.10), we can further derive several other definitive properties of $Z(z)$. They are presented in the following two lemmas.

Lemma 4.3

Let $\rho\neq\sigma_i\ \forall i$ and $V(z)$ be given as in (4.10). If $Z(z)$ is a minimal basis solution to (4.11), then

i) the $2p$ columns of $Z(z)$ all have an identical degree n; and

ii) Z_n, the highest degree term (z^n term) coefficient matrix of $Z(z)$, is of full rank $2p$.

Proof: First, we note that not any column of $Z(z)$ can have degree less than n. This is proved by contradiction as follows. If a particular column of $Z(z)$ is of degree less than n, we partition it as

$$Z(z)=\begin{bmatrix} \boldsymbol{t}(z) \\ \hat{\boldsymbol{p}}(z) \\ \boldsymbol{q}(z) \\ \hat{\boldsymbol{w}}(z) \end{bmatrix}$$

where $\boldsymbol{t}(z)$, $\boldsymbol{p}(z)$, $\boldsymbol{q}(z)$, and $\boldsymbol{w}(z)$ are all p-vectors and $\hat{\boldsymbol{p}}(z)\triangleq z^{n-1}\boldsymbol{p}(z^{-1})$, $\hat{\boldsymbol{w}}(z)\triangleq z^{n-1}\boldsymbol{w}(z^{-1})$. Compared to (3.14) in Lemma 3.2, it can be seen that $(\boldsymbol{p}(z)/a^*(z), \boldsymbol{q}(z)/a(z))$ represents a pair of singular vectors of $\Gamma\{H(z)\}$ corresponding to the "singular value" ρ. This contradicts the assumption that ρ is *not* a singular value of $\Gamma\{H(z)\}$.

Now since $V(z)$ has a row-degree-sum $2pn$, the minimal basis $Z(z)$ has to have a *uniform* column-degree n; for otherwise it would have a column-degree-sum exceeding $2pn$, violating Lemma 4.2.ii). The rank property of Z_n follows trivially from Lemma 4.2.i). □

Lemma 4.4

Given $V(z)$ as in (4.10), there exists a special minimal basis solution $Z(z)$ for (4.11), such that

$$Z_n=\begin{bmatrix} Z_n^u \\ -- \\ I_{2p} \end{bmatrix} \tag{4.12}$$

where Z_n^u is a $p\times p$ constant matrix. (In the sequel, the polynomial matrix $Z(z)$ of this structure will be termed as in *polynomial echelon form*.)

Proof: Consider the coefficient matrix of the highest degree term of the product $V(z)Z(z)$

$$\left[\begin{array}{cc|c} I_p & 0 & \boxtimes \\ 0 & \rho I_p & \end{array}\right]\begin{bmatrix} Z_n^u \\ -- \\ Z_n^l \end{bmatrix}=0.$$

Note that Z_n^u is related to Z_n^l by

$$Z_n^u = -\begin{bmatrix} I_p & 0 \\ 0 & (1/\rho)I_p \end{bmatrix}\left[\boxtimes\right] Z_n^l.$$

Thus, Z_n^l must be nonsingular, for otherwise Z_n will not have full rank and hence contradicts Lemma 4.3 ii. Therefore, without loss of generality we can assume a normalized form, i.e., $Z_n^l = I_{2p}$. Q.E.D. □

It is easily seen that, for any $2p \times p$ constant matrix M, $Z(z)M$ will be a polynomial solution to (4.9) (In fact, as a result of the minimal basis properties, we can prove that every solution to (4.9) can be expressed as $Z(z)M$.) Now, we partition the *polynomial echelon form* $Z(z)$ as

$$Z(z) = \begin{bmatrix} R'(z) & R''(z) \\ \underset{\sim}{\hat{P}}'(z) & \underset{\sim}{\hat{P}}''(z) \\ \underset{\sim}{Q}'(z) & \underset{\sim}{Q}''(z) \\ \hat{S}'(z) & \hat{S}''(z) \end{bmatrix} \tag{4.13}$$

where the submatrices are all $p \times p$ polynomial matrices.

For a reason to become clear later, we shall concentrate on a subset of solutions to (4.9) in which

$$M = \begin{bmatrix} I_p \\ \hline -K^* \end{bmatrix} \tag{4.14}$$

where K is a $p \times p$ constant matrix. Namely, we concentrate on the solution set of (4.9) with the form

$$\begin{bmatrix} R(z) \\ \underset{\sim}{\hat{P}}(z) \\ \underset{\sim}{Q}(z) \\ \hat{S}(z) \end{bmatrix} = Z(z)\begin{bmatrix} I_p \\ \hline -K^* \end{bmatrix} = \begin{bmatrix} R'(z) - R''(z)K^* \\ \underset{\sim}{\hat{P}}'(z) - \underset{\sim}{\hat{P}}''(z)K^* \\ \underset{\sim}{Q}'(z) - \underset{\sim}{Q}''(z)K^* \\ \hat{S}'(z) - \hat{S}''(z)K^* \end{bmatrix} \tag{4.15}$$

Our second step is to show that there exists a matrix K such that $\{R(z), \underset{\sim}{P}(z), Q(z), S(z)\}$ as given in (4.15) will satisfy (4.6) for some $\tilde{H}_0$, and therefore the conditions in Lemma 4.1 can all be fulfilled.

Noting that $\hat{S}(z) = z^n S(z^{-1})$ and $\underset{\sim}{\hat{Q}}(z) = z^n \underset{\sim}{Q}(z^{-1})$, we can write (4.6) as

$$R_0 + H_0 \underset{\sim}{\hat{P}}_0 = 0$$

$$\hat{S}_n + H_0^* \underset{\sim}{Q}_n = 0$$

where we used the convention (1.1) for subscripts of $\underset{\sim}{P}$, $\underset{\sim}{Q}$, R, and S. By (4.15) these are equivalent to

$$(R_0' - R_0''K^*) + H_0(\underset{\sim}{\hat{P}}_0' - \underset{\sim}{\hat{P}}_0''K^*) = 0 \tag{4.16a}$$

$$(\hat{S}_n' - \hat{S}_n''K^*) + H_0^*(\underset{\sim}{Q}_n' - \underset{\sim}{Q}_n''K^*) = 0. \tag{4.16b}$$

Notice that the coefficients in (4.16a) are from the upper half of Z_0 [constant term in $Z(z)$]

$$\begin{bmatrix} R_0' & R_0'' \\ \underset{\sim}{\hat{P}}_0' & \underset{\sim}{\hat{P}}_0'' \end{bmatrix}$$

and the coefficients in (4.16b) are from the lower half of Z_n

$$\begin{bmatrix} \underset{\sim}{Q}_n' & \underset{\sim}{Q}_n'' \\ \hat{S}_n' & \hat{S}_n'' \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_p \end{bmatrix}$$

where the equality comes from the polynomial echelon form requirement (4.12). Therefore, (4.16b) is exactly[3]

$$-K^* + H_0^* = 0. \tag{4.17}$$

Hence, (4.16a) is equivalent to

$$R_0' - R_0''H_0^* + H_0\underset{\sim}{\hat{P}}_0' - H_0\underset{\sim}{\hat{P}}_0''H_0^* = 0 \tag{4.18a}$$

or

$$\begin{bmatrix} I_p & H_0 \end{bmatrix}\begin{bmatrix} R_0' & -R_0'' \\ \underset{\sim}{\hat{P}}_0' & -\underset{\sim}{\hat{P}}_0'' \end{bmatrix}\begin{bmatrix} I_p \\ H_0^* \end{bmatrix} = 0 \tag{4.18b}$$

or, with an obvious denotation,

$$\begin{bmatrix} I_p & H_0 \end{bmatrix} U \begin{bmatrix} I_p \\ H_0^* \end{bmatrix} = 0. \tag{4.18c}$$

Therefore, the existence of a proper extension hinges upon the existence of a solution to the matrix quadratic equation (4.18). Indeed, we can show [Appendix B.1, (B.14)] that U is Hermitian, i.e., $U = U^*$; and, therefore, we shall term (4.18) as an *algebraic Riccati equation* (*ARE*) [12]. More importantly, we have the following result.

Lemma 4.5

There exists a solution H_0 for the ARE (4.18).

Proof: See Appendixes B.2 and B.3. □

Remark: Note that by the echelon form constraint (4.12) we have

$$\underset{\sim}{Q}(z) = \underset{\sim}{Q}'(z) - \underset{\sim}{Q}''(z)H_0^*$$

$$= I_p z^n + (\text{lower power terms}).$$

Hence, the nonsingularity of $\underset{\sim}{Q}(z)$ is guaranteed (to fulfill the condition in Lemma 4.1). Therefore, the columns of $(\underset{\sim}{P}(z)/za^*(z), \underset{\sim}{Q}(z)/za(z))$ represent *p linearly independent* pairs of singular vectors corresponding to ρ. That is, ρ is a singular value of $\underset{\sim}{\Gamma}$ with multiplicity *at least p*. Further, the nonsingularity of $\underset{\sim}{Q}(z)$ implies that of $P(z)$. □

[3]If the upper half of M is some matrix K' other than I_p, then instead of (4.17), (4.6b) will lead to $-K^* + H_0^*K' = 0$. Hence, the rank of M depends on the rank of K' as $K^* = H_0^*K'$. However, M has to have full rank to make possible p linearly independent columns in $P(z)$ or $Q(z)$ [see (4.15)]. Therefore, K' needs to be nonsingular. Without loss of generality we can let $K' = I_p$, which justifies the choice (4.14).

We thus have shown the existence of a proper extension. What is left to prove in the extension theorem is the inequality (3.17), which gives the relative position of ρ among the ordered singular value set of $\underset{\sim}{\Gamma}$. This is accomplished by the following lemma.

Lemma 4.6

Let $\rho \in (\sigma_{k+1}, \sigma_k)$, and $\underset{\sim}{\Gamma}$ be a properly extended Hankel matrix as discussed above. Then ρ is a singular value (of $\underset{\sim}{\Gamma}$) with multiplicity exactly p, and

$$\underset{\sim}{\sigma}_k > \rho = \underset{\sim}{\sigma}_{k+1} = \underset{\sim}{\sigma}_{k+2} = \cdots = \underset{\sim}{\sigma}_{k+p} > \sigma_{k+p+1} \quad (4.19)$$

where $\{\underset{\sim}{\sigma}_i\}$ denotes the decreasingly ordered singular value set of $\underset{\sim}{\Gamma}$.

Proof: We need only to show that

$$\underset{\sim}{\sigma}_k > \rho > \underset{\sim}{\sigma}_{k+p+1}$$

for the fact that ρ is a singular value of multiplicity at least p (see the previous remark of Lemma 4.1) will immediately lead to (4.19).

Note that

$$\underset{\sim}{\Gamma}^*\underset{\sim}{\Gamma} = \Gamma^*\Gamma + (\underset{\sim}{\Gamma}^*\mathcal{E})(\mathcal{E}^*\underset{\sim}{\Gamma})$$

where $\mathcal{E} \triangleq [I_p \; 0 \; 0 \; \cdots]^*$. It is clear that $(\underset{\sim}{\Gamma}^*\mathcal{E})(\mathcal{E}^*\underset{\sim}{\Gamma})$ is Hermitian and nonnegative definite with rank not exceeding p. Therefore, by a perturbation analysis result [16, pp. 102–103]

$$\underset{\sim}{\sigma}_k \geqslant \sigma_k \quad (4.20a)$$

and

$$\sigma_{k+1} \geqslant \underset{\sim}{\sigma}_{k+p+1}. \quad (4.20b)$$

In view that $\sigma_k > \rho > \sigma_{k+1}$, the two relations (4.20) trivially lead to the inequality $\underset{\sim}{\sigma}_k > \rho > \underset{\sim}{\sigma}_{k+p+1}$. Q.E.D. □

V. A Minimal-Degree-Approximation Algorithm

A. Square Systems Case

The discussion in the previous section in fact provides a method which can derive an optimal approximant while at the same time explicitly give rise to a proper extending block H_0. (This can be seen by a look at (4.18), (4.17), (4.15), and the MDA theorem.) The complete procedure for (square) multivariable systems approximation can now be summarized as follows.

Step 1—Find an Echelon-Form Minimal Basis Solution Z(z): From the original system function $H(z) = N(z)/a(z)$ and the approximation tolerance ρ, form (4.11)

$$V(z)Z(z) = 0 \quad (5.1a)$$

where (4.10)

$$V(z) = \begin{bmatrix} a(z)I_p & -N(z) & \rho\hat{a}^*(z)I_p & 0 \\ 0 & \rho a(z)I_p & -\hat{N}^*(z) & \hat{a}^*(z)I_p \end{bmatrix} \quad (5.1b)$$

and find the polynomial echelon form minimal basis solution for $Z(z)$ (see Lemma 4.4).

Step 2—Solve the ARE (4.18): Partition $Z(z)$ as (4.13)

$$Z(z) = \begin{bmatrix} R'(z) & R''(z) \\ \underset{\sim}{\hat{P}}'(z) & \underset{\sim}{\hat{P}}''(z) \\ \underset{\sim}{Q}'(z) & \underset{\sim}{Q}''(z) \\ \hat{S}'(z) & \hat{S}''(z) \end{bmatrix} \quad (5.2)$$

(where all submatrices are $p \times p$) and single out the upper half of the constant term to form the ARE (4.18)

$$\begin{bmatrix} I_p & H_0 \end{bmatrix} \begin{bmatrix} R_0' & -R_0'' \\ \underset{\sim}{\hat{P}}_0' & -\underset{\sim}{\hat{P}}_0'' \end{bmatrix} \begin{bmatrix} I_p \\ H_0^* \end{bmatrix} = 0. \quad (5.3)$$

Solve for H_0.

Step 3—Compute the Optimal Approximant:

$$H_{\text{mda}}(z) = \left[R(z)\underset{\sim}{\hat{P}}^{-1}(z)\right]_- = \left[\{R'(z) - R''(z)H_0^*\}\{\underset{\sim}{\hat{P}}'(z) - \underset{\sim}{\hat{P}}''(z)H_0^*\}^{-1}\right]_- \quad (5.4a)$$

(See (4.15), (4.17) and noting that $[R(z)\underset{\sim}{\hat{P}}^{-1}(z)]_- = [\{z\underset{\sim}{T}(z) + H_0\underset{\sim}{\hat{P}}(z)\}\underset{\sim}{\hat{P}}^{-1}(z)]_- = [z\underset{\sim}{T}(z)\underset{\sim}{\hat{P}}^{-1}(z)]_-$.) Or,

$$H_{\text{mda}}(z) = \left[\underset{\sim}{\hat{Q}}^{*-1}(z)S^*(z)\right]_- = \left[\{\underset{\sim}{\hat{Q}}'^*(z) - H_0\underset{\sim}{\hat{Q}}''^*(z)\}^{-1}\{S'^*(z) - H_0S''^*(z)\}\right]_- \quad (5.4b)$$

The projection operation $[\;]_-$ can be done by, say, partial fraction expansion.

Remark: The equation (5.1a) was originated from a more basic form

$$F(z)D(z) = C(z) \quad (5.5)$$

where

$$F(z) = \begin{bmatrix} -\rho\dfrac{\hat{a}^*(z)}{a(z)}I_p & \dfrac{N(z)}{a(z)} \\ \dfrac{\hat{N}^*(z)}{\hat{a}^*(z)} & -\rho\dfrac{a(z)}{\hat{a}^*(z)}I_p \end{bmatrix} \quad (5.6)$$

and

$$C(z)=\begin{bmatrix} R'(z) & R''(z) \\ \hat{S}'(z) & \hat{S}''(z) \end{bmatrix}, \qquad D(z)=\begin{bmatrix} \tilde{Q}'(z) & \tilde{Q}''(z) \\ \tilde{\hat{P}}'(z) & \tilde{\hat{P}}''(z) \end{bmatrix} \tag{5.7}$$

(Compare these to (4.5) and (4.8).) The $2p\times 2p$ rational matrix $F(z)$ is termed the *adjoint system matrix* [5]. It is seen that $C(z)D^{-1}(z)$ is a right matrix-fraction description (MFD) [12] of $F(z)$. This offers a different viewpoint to the minimal basis problem (5.1), viz. a viewpoint of *minimal design problems* [5], [19]–[21], in which a minimal degree MFD (in this case $C(z)D^{-1}(z)$) for a system (in this case $F(z)$) is the objective. On the other hand, we note the striking similarity between the *Hamiltonian system* [12] and the adjoint system matrix $F(z)$, which strongly suggests the mathematical relevance between this optimal reduction problem and optimal control/estimation problems. □

In the following, we briefly comment on the actual implementation of the steps outlined above. It is not intended to be elaborate and the reader is suggested to consult proper literatures for details.

For Step 1, a brute-force approach could be a Gauss-elimination type procedure. However, a fast projection method has been devised to solve such problems [20], [21]. This fast method takes $\mathcal{O}(n^2p^3)$ operations as opposed to $\mathcal{O}(n^3p^3)$ needed for a Gauss-elimination type of method. In addition, it also saves storage to a large extent if n and p are large ($\mathcal{O}(np^2)$ versus $\mathcal{O}(n^2p^2)$).

As to Step 2, much effort has been devoted to the study of ARE like (5.3) (see, e.g., [22], [23]). (Note that the ARE's conventionally encountered in least-squares optimal control/estimation problems form a subclass of (5.3), as the former has some positive-definiteness requirement while the latter does not. See Lemma B.7.) Various methods solving this type of equations exist. Finite algorithms give one option, the underlying principle being to block-diagonalize the $2p\times 2p$ coefficient matrix in (5.3) as discussed in Appendix B.3.[4] However, as is well known, iterative methods are generally a better choice for such problems, as they are relatively free from the curse of ill-conditionedness.[5] Laub [25] has an interesting discussion on some iterative methods. The time complexity for finite and iterative algorithms are both $\mathcal{O}(p^3)$, which is the least costly among the three steps (and hence we can afford an iterative procedure without painfully weighing the tradeoffs).

The partial fraction expansion in Step 3 is the most costly in the whole algorithm. It may require as many as $\mathcal{O}(n^3p^3)$ operations simply in computing the poles of $R(z)\tilde{\hat{P}}^{-1}(z)$ or $\tilde{\hat{Q}}^{*-1}(z)S^*(z)$, since $\det\{\tilde{\hat{P}}(z)\}$ (and $\det\{\tilde{\hat{Q}}^*(z)\}$) has degree np. It is possible that the projection operation $[\cdot]_-$ can be achieved via a spectral factorization on $R(z)\tilde{\hat{P}}^{-1}(z)\tilde{\hat{P}}^{*-1}(z^{-1})R^*(z^{-1})$, which may result in some saving for computation. This route is being examined.

[4]See Lemma B.2 and especially (B.28)–(B.31), assuming $\Delta_p=0$. The dimension of Δ_p is zero if $\hat{P}_0''$ is nonsingular, and Λ_p is now a matrix not necessarily of eigenvalues of $\hat{P}_0''$. For a survey of methods for symmetrically decomposing Hermitian matrices, see [24].

[5]For example, when $\hat{P}_0''$ is nearly singular, and hence ill-conditioned.

B. Nonsquare Systems Case

Suppose the given system $H(z)$ has p-inputs and q-outputs with $q\leq p$. We can augment it to square by adding extra zeros, e.g.,

$$H^{\#}(z)\triangleq\begin{bmatrix} 0 \\ \hline H(z) \end{bmatrix} \;\}\; (p-q) \text{ rows} \tag{5.8}$$

Then the Hankel matrix $\Gamma\{H^{\#}(z)\}$ will be the same as $\Gamma\{H(z)\}$ except for some extra zero rows. Clearly, the extra zero rows will not affect the singular values. Hence, $\Gamma\{H^{\#}(z)\}$ and $\Gamma\{H(z)\}$ have exactly the same singular values; or, $\sigma_i^{\#}=\sigma_i\ \forall i$.

Let ρ be the tolerance of approximation-error Hankel-norm for $H(z)$. Conceptually, we can first use the previous algorithm to find a minimal degree approximant for $H^{\#}(z)$, say $H^{\#}_{\mathrm{MDA}}(z)$, satisfying the tolerance requirement. Then chopping off the first $(p-q)$ rows from $H^{\#}_{\mathrm{MDA}}(z)$, we have the remaining $q\times p$ matrix $H_{\mathrm{MDA}}(z)=[0 \mid I_q]H^{\#}_{\mathrm{MDA}}(z)$. It will be a minimal degree approximant for $H(z)$. This can be easily verified by noting that

$$\left\|\Gamma\{[0 \mid I_q][H^{\#}(z)-H^{\#}_{\mathrm{MDA}}(z)]\}\right\|_s \leq \left\|\Gamma\{H^{\#}(z)-H^{\#}_{\mathrm{MDA}}(z)\}\right\|_s \leq \rho$$

and that

$$\deg H_{\mathrm{MDA}}(z)\leq \deg H^{\#}_{\mathrm{MDA}}(z)=k$$

whereas the minimum degree bound lemma assures that $\deg H_{\mathrm{MDA}}(z)\geq k$. Therefore, $\deg H_{\mathrm{MDA}}(z)=k$ and the assertion is verified.

Now we have demonstrated that the method originally for square systems approximation can be trivially extended to deal with general cases. We would like to further note that the computation procedure can be simplified as presented below (proof omitted).

Step 1—Find an Echelon-Form Minimal Basis Solution Z(z) to the Equation:

$$\begin{bmatrix} a(z)I_q & -N(z) & \rho\hat{a}^*(z)I_q & 0 \\ 0 & \rho a(z)I_p & -\hat{N}^*(z) & \hat{a}^*(z)I_p \end{bmatrix} Z(z)=0. \tag{5.9a}$$

$Z(z)$ is now an nth degree $(2p+2q)\times(p+q)$ polynomial matrix with its z^n term coefficient matrix in the form

$$Z_n=\begin{bmatrix} \times \\ I_{p+q} \end{bmatrix}. \tag{5.9b}$$

Step 2—Solve the Algebraic Riccati Equation

$$\begin{bmatrix} I_p & \vdots & H_0^{\#} \end{bmatrix} \left[\begin{array}{c|cc} -\rho I_{p-q} & & 0 \\ \hline & R_0' & -R_0'' \\ 0 & & \\ & \hat{P}_0' & -\hat{P}_0'' \end{array}\right] \begin{bmatrix} I_p \\ \hline H_0^{\#*} \end{bmatrix} = 0 \tag{5.10}$$

for $H_0^{\#}$, where the coefficient matrices R_0', R_0'' and $\hat{P}_0'$, $\hat{P}_0''$ are defined according to the partition on $Z(z)$

$$Z(z) = \begin{bmatrix} R'(z) & R''(z) \\ \hat{P}'(z) & \hat{P}''(z) \\ Q'(z) & Q''(z) \\ \hat{S}'(z) & \hat{S}''(z) \end{bmatrix} \begin{matrix} \} \; q \text{ rows} \\ \} \; p \\ \} \; q \\ \} \; p \end{matrix} \tag{5.11}$$

$$\underbrace{}_{q} \quad \underbrace{}_{p}$$

columns

Step 3—Compute the Optimal Approximant

$$H_{\mathrm{MDA}}(z) = \left[\hat{Q}^{*-1}(z) S^*(z)\right]_- \tag{5.12a}$$

where

$$\hat{Q}(z) = \hat{Q}'(z) - \hat{Q}''(z) H_0^* \tag{5.12b}$$

$$S(z) = S'(z) - S''(z) H_0^* \tag{5.12c}$$

with

$$H_0 = \begin{bmatrix} 0 & \vdots & I_q \end{bmatrix} H_0^{\#}, \tag{5.12d}$$

i.e., H_0 consists of the last q rows of $H_0^{\#}$.

VI. A Numerical Example

The sole purpose of the following simple example is to demonstrate the procedure of computing the minimal degree approximation as outlined in Section V. No attempt is made here to discuss the numerical aspects or the practicality consideration of the algorithm. For more real-data simulations, we refer to the scalar case paper [5b].

Let

$$H(z) = \begin{bmatrix} \dfrac{z+1}{z^2 - z + \frac{1}{4}} & \dfrac{1}{z - \frac{1}{2}} \\[2ex] \dfrac{-z^2 + z + 1}{z^3 + \frac{1}{2}z^2 - \frac{1}{4}z - \frac{1}{8}} & \dfrac{z - \frac{1}{4}}{z^2 + z + \frac{1}{4}} \end{bmatrix}$$

Then

$$a(z) = z^4 - \tfrac{1}{2}z^2 + \tfrac{1}{16}$$

and

$$N(z) = \begin{bmatrix} z^3 + 2z^2 + \frac{5}{4}\cdot z + \frac{1}{4} & z^3 + \frac{1}{2}\cdot z^2 - \frac{1}{4}\cdot z - \frac{1}{8} \\ -z^3 + \frac{3}{2}\cdot z^2 + \frac{1}{2}\cdot z - \frac{1}{2} & z^3 - \frac{5}{4}\cdot z^2 + \frac{1}{2}\cdot z - \frac{1}{16} \end{bmatrix}.$$

The system has four poles located at 1/2, 1/2, −1/2, and −1/2, respectively. Its four nonzero singular values were found to be $\sigma_1 = 5.56$, $\sigma_2 = 3.83$, $\sigma_3 = 1.33$, and $\sigma_4 = 1.04$. We took $\rho = 2$. According to the theory, the algorithm should yield a second order approximation since $\sigma_2 > \rho > \sigma_3$.

Step 1—An Echelon-Form Minimal Basis Solution to (5.1): Note that

$$\hat{a}(z) = \tfrac{1}{16}z^4 - \tfrac{1}{2}z^2 + 1$$

and

$$\hat{N}(z) = \begin{bmatrix} \frac{1}{4}z^4 + \frac{5}{4}z^3 + 2z^2 + z & -\frac{1}{8}z^4 - \frac{1}{4}z^3 + \frac{1}{2}z^2 + z \\ \frac{1}{2}z^4 + \frac{1}{2}z^3 + \frac{3}{2}z^2 - z & -\frac{1}{16}z^4 + \frac{1}{2}z^3 - \frac{5}{4}z^2 + z \end{bmatrix}.$$

The echelon-form solution $Z(z)$ is shown in Table I. It is seen that the upper half of Z_0 shows certain symmetry as it should be (B.14).

TABLE I
POLYNOMIAL ECHELON FORM MINIMAL BASIS SOLUTION

R_0'	2.7178444E+00	2.7247064E+00	1.2912721E+00	3.2960047E-01	R_0''	
	2.7247064E+00	1.3323256E+00	8.6558115E-01	-7.0248389E-01		
$\hat{P}_0'$	-1.2912721E+00	-8.6558115E-01	-3.3044918E-01	-1.4970277E-01	$\hat{P}_0''$	
	-3.2960047E-01	7.0248365E-01	-1.4975277E-01	-3.7900683E-01		
Q_0'	-2.2574163E-01	-2.5724936E-01	-7.2256115E-02	-5.3305547E-03	Q_0''	Z_0
	2.4797097E-01	1.5280745E-01	6.0242409E-02	1.1247591E-02		
$\hat{S}_0'$	1.6140982E-01	1.0619726E-01	4.1301859E-02	1.8724096E-02	$\hat{S}_0''$	
	4.1200059E-02	-8.7810426E-02	1.8724056E-02	4.7387111E-02		
R_1'	4.2616495E+00	1.9991548E+00	7.4820664E-01	2.9708167E-01	R_1''	
	1.1523705E-01	-1.8854423E+00	-5.4806170E-01	4.2405910E-01		
$\hat{P}_1'$	-2.6310599E-01	-3.5715705E-01	1.2326260E-01	2.5112735E-01	$\hat{P}_1''$	
	8.6088221E-01	3.1517216E-01	2.9726028E-01	-1.5452981E-01		
Q_1'	-9.6821471E-01	-7.5498768E-01	-2.1434126E-01	-1.4463187E-02	Q_1''	Z_1
	-3.6494528E-01	9.6748420E-02	-1.4972322E-01	-2.0342641E-01		
$\hat{S}_1'$	-4.5653286E-02	-3.4591532E-01	-1.4795964E-01	-1.0796706E-01	$\hat{S}_1''$	
	-8.5380924E-02	-1.2358635E-01	-4.5202743E-02	6.9253163E-02		
R_2'	2.0228646E+00	1.4753147E-01	-5.2467702E-02	-1.0701352E-01	R_2''	
	-7.6694970E-01	5.7885172E-01	-1.7976943E-01	5.6554974E-02		
$\hat{P}_2'$	1.3919409E+00	1.4030052E+00	3.6574456E-01	5.6111920E-02	$\hat{P}_2''$	
	2.0967281E-01	-6.9659174E-01	-4.7272572E-02	3.2874583E-01		
Q_2'	-9.7395927E-01	-1.7295118E-01	1.0187802E-02	6.0930096E-03	Q_2''	Z_2
	-1.4342999E-01	-1.0358044E+00	1.0807216E-01	-2.8575187E-02		
$\hat{S}_2'$	-2.0873519E+00	-2.0859025E+00	-4.7655704E-01	1.5272381E-01	$\hat{S}_2''$	
	-2.1112047E+00	-2.2215655E-01	-6.0992852E-01	-7.8611583E-01		
R_3'	-1.7875908E-02	-3.7825608E-01	-6.6715469E-02	-3.4677016E-02	R_3''	
	-1.5765947E-01	2.2549661E-01	4.3699327E-02	-6.8859412E-02		
$\hat{P}_3'$	8.9624655E-01	3.5704522E-01	8.6466719E-02	-6.9773627E-02	$\hat{P}_3''$	
	-1.5385690E-01	2.4469225E-01	-5.1752135E-02	-1.5853151E-02		
Q_3'	6.5900726E-01	7.7604962E-01	2.5957222E-01	2.7416129E-02	Q_3''	Z_3
	-2.5712426E-01	-3.7624852E-02	-2.5534699E-02	2.0087523E-01		
$\hat{S}_3'$	-4.1566642E+00	-1.2978526E+00	-7.7267570E-01	-6.4581767E-02	$\hat{S}_3''$	
	-1.5746951E-01	-1.3507276E+00	5.2290507E-01	1.5159339E-01		
R_4'	-1.2500000E-01	0.0000000E+00	0.0000000E+00	0.0000000E+00	R_4''	
	0.0000000E+00	-1.2500000E-01	0.0000000E+00	0.0000000E+00		
$\hat{P}_4'$	1.2500000E-01	-2.5000000E-01	-3.1250000E-02	0.0000000E+00	$\hat{P}_4''$	
	-6.2500000E-02	-3.1250000E-02	0.0000000E+00	-3.1250000E-02		
Q_4'	1.0000000E+00	0.0000000E+00	0.0000000E+00	0.0000000E+00	Q_4''	Z_4
	0.0000000E+00	1.0000000E+00	0.0000000E+00	0.0000000E+00		
$\hat{S}_4'$	0.0000000E+00	0.0000000E+00	1.0000000E+00	0.0000000E+00	$\hat{S}_4''$	
	0.0000000E+00	0.0000000E+00	0.0000000E+00	1.0000000E+00		

Step 2—Solve the Algebraic Riccati Equation (5.3): Equation (5.3) now takes the form

$$\begin{bmatrix} I_p & \vdots & H_0 \end{bmatrix} \left[\begin{array}{cc|cc} 2.718 & 2.725 & -1.291 & 0.330 \\ 2.725 & 1.332 & -0.866 & 0.702 \\ \hline -1.291 & -0.866 & 0.330 & 0.150 \\ -0.330 & 0.702 & 0.150 & 0.379 \end{array}\right] \begin{bmatrix} I_p \\ H_0^* \end{bmatrix} = 0.$$

One solution is

$$H_0 = \begin{bmatrix} 1.343 & 1.062 \\ 3.463 & 0.275 \end{bmatrix}.$$

As a side remark, note that

$$R_0' + R_0''\hat{P}_0''^{-1}R_0''^* = \begin{bmatrix} -2.532 & -1.553 \\ -1.553 & -4.784 \end{bmatrix}.$$

It is not hard to verify that both $\hat{P}_0$ and $(R_0' + R_0''\hat{P}_0''^{-1}R_0''^*)$ have two negative eigenvalues, and hence they are congruent, as has been theoretically proved (Lemma B.3). To check that the extended Hankel $\underline{\Gamma}$ does have the desired

TABLE II

$$R(z) = \begin{bmatrix} 6.5538672E-01 & -1.6570850E+00 \\ 2.3080858E+00 & -1.4717343E+00 \end{bmatrix} z^0 + \begin{bmatrix} 2.9411777E+00 & -4.7324763E-01 \\ 1.1900027E-01 & -8.3545526E-01 \end{bmatrix} z^1 + \begin{bmatrix} 2.1456402E+00 & 2.0799225E-01 \\ -6.2894007E-01 & 1.1714215E+00 \end{bmatrix} z^2 + \begin{bmatrix} 1.0918941E-01 & -1.3093465E-01 \\ -1.4342878E-01 & 9.1106360E-02 \end{bmatrix} z^3 + \begin{bmatrix} -1.2500000E-01 & 0.0000000E+00 \\ 0.0000000E+00 & -1.2500000E-01 \end{bmatrix} z^4$$

$$\hat{P}(z) = \begin{bmatrix} -6.8834129E-01 & 3.1941019E-01 \\ 2.7419238E-01 & 1.5253751E+00 \end{bmatrix} z^0 + \begin{bmatrix} -7.1549529E-01 & -8.4935106E-01 \\ 6.2579103E-01 & -6.7343825E-01 \end{bmatrix} z^1 + \begin{bmatrix} 7.7174298E-01 & 4.0855550E-02 \\ -7.5947367E-02 & -6.2347544E-01 \end{bmatrix} z^2 + \begin{bmatrix} 8.5420359E-01 & 1.1682656E-01 \\ -9.4120686E-02 & 3.5984377E-01 \end{bmatrix} z^3 + \begin{bmatrix} 1.6697478E-01 & -1.4179318E-01 \\ -2.4931521E-02 & -2.2659165E-02 \end{bmatrix} z^4$$

multiplicity $(p-2)$ for ρ, its (nonzero) singular values were computed: 7.125, 4.890, 2.000, 2.000, 0.622, 0.173. Note that $\sigma_3=\sigma_4=2$, exactly as predicted.

Step 3—Calculate the Minimal Degree Approximation (5.4): The matrices $R(z)$ and $\hat{P}(z)$ corresponding to this extension are given in Table II. The "inner zeros" of $\det\{\hat{P}(z)\}$ were found to be -0.662 and 0.768. Hence, $H_{\mathrm{MDA}}(z)=[R(z)\hat{P}^{-1}(z)]_-$ is indeed a second-order system, as expected. We actually have

$$H_{\mathrm{MDA}}(z)=\frac{1}{z+0.662}\begin{bmatrix} 0.437 & -1.497 \\ -0.264 & 0.905 \end{bmatrix} + \frac{1}{z-0.768}\begin{bmatrix} 1.697 & 0.896 \\ 0.870 & 0.459 \end{bmatrix}.$$

The (nonzero) singular values of the error Hankel $\Gamma\{H(z)-H_{\mathrm{MDA}}(z)\}$ were also computed. They are 2.00, 2.00, 1.99, 1.99, 0.62, 0.39. Therefore, we do have $\|H(z)-H_{\mathrm{MDA}}(z)\|_H \leq \rho$.

VII. Conclusion

A multivariable systems model reduction problem with Hankel-norm error criterion has been studied. Through an algebraic analysis, it is shown that the optimal solutions adopt a closed-form formulation and then accordingly an MDA algorithm is developed.

While an *exact* error analysis (not existing for any other model reduction scheme) is provided in the theoretical domain, the suitability of this Hankel-norm criterion remains to be carefully tested by comparing its performance with other alternatives. However, a useful fact justifying the choice of Hankel-norm criterion is that this measure stands between two conventionally used measures, the sum-squares ($\mathcal{L}_2$) and the Chebyshev ($\mathcal{L}_\infty$) norms. More ambitiously, the Hankel-norm is now being examined as a possible measure of stability margin in closed-loop systems analysis.

Turning to the numerical aspects of the MDA algorithm, we note that the polynomial formulation offers an attractive computation speed (see Section V). Unfortunately, the important numerical stability analysis is admittedly very much lacking. It is a priority task to address this issue. Research is also underway to compare this new algorithm with other numerically tested approximation methods, [2], [9], [29].

The polynomial-theoretic interpretation of the algorithm is rather stimulating itself. It basically amounts to solving a Hamiltonian-type system [see (5.5)–(5.7)], just like what arises in the optimal control/estimation context. The tight relationship between optimality and Hamiltonian systems solution surfaces again in this approximation problem. This mathematical overtone will hopefully shed some light on future theoretical research.

Finally, we remark that the extension step serves to link a generic situation to an artificial one where we have p pairs of singular vectors corresponding to a singular value designating the approximation error tolerance. Further, it has a computational advantage that the artificial singular vectors introduced through extension can be computed with some fast algorithm (see Section V). The question now is whether such an extension step is absolutely necessary. Our conjecture [28] is that the p pairs of singular vectors corresponding to p consecutive singular values can serve equally well to construct a desired solution in a similar manner. Of course, it may not enjoy the fast algorithm and therefore may need more computations. However, this is certainly an interesting and important problem from a mathematical standpoint.

Appendix A—Proof of Lemma 3.3—Dominance Property and Multiplicity Property

Proof of Dominance Property

Let $H'(z)\triangleq[U_{yI}^*(z)H(z)]_-$ and $\Gamma'\triangleq\Gamma\{H'(z)\}$. Then all we need to show is that $\Gamma'^*\Gamma'\leq\Gamma^*\Gamma$. This then, by a perturbation theory result in [16, pp. 102–103], implies $\sigma_i'\leq\sigma_i$. To show $\Gamma'^*\Gamma'\leq\Gamma^*\Gamma$, we need only to prove that $\|\Gamma'\boldsymbol{\eta}\|_2\leq\|\Gamma\boldsymbol{\eta}\|_2\ \forall\boldsymbol{\eta}\in\mathcal{L}_2^-$, or in terms of equivalent functional representation, $\|[H'(z)\check{\boldsymbol{\eta}}(z)]_-\|_2\leq\|[H(z)\check{\boldsymbol{\eta}}(z)]_-\|_2\ \forall\check{\boldsymbol{\eta}}(z)\in\mathcal{L}_2^+$.

Note that

$$\begin{aligned}[H'(z)\check{\boldsymbol{\eta}}(z)]_- &= \big[[U_{yI}^*(z)H(z)]_-\check{\boldsymbol{\eta}}(z)\big]_- \\ &= [U_{yI}^*(z)H(z)\check{\boldsymbol{\eta}}(z)]_- \\ &= \big[U_{yI}^*(z)[H(z)\check{\boldsymbol{\eta}}(z)]_-\big]_-\end{aligned}$$

by repeatedly applying the truncation property, noting that $U_{yI}^*(z)\in\mathcal{L}_2^+$. Therefore,

$$\begin{aligned}\|[H'(z)\check{\boldsymbol{\eta}}(z)]_-\|_2 &= \big\|\big[U_{yI}^*(z)[H(z)\check{\boldsymbol{\eta}}(z)]_-\big]_-\big\|_2 \\ &\leq \|U_{yI}^*(z)[H(z)\check{\boldsymbol{\eta}}(z)]_-\|_2 \\ &= \|[H(z)\check{\boldsymbol{\eta}}(z)]_-\|_2\end{aligned}$$

where the last equality is because that $U_{yI}^*(z)$ is unitary (Lemma 2.3). Q.E.D. □

Proof of Multiplicity Property

The proof is separated into two steps. First, we shall show that $\|\Gamma\{[H^*(z)U_{yL}(z)]_-\}\|_s \le \rho$. Then we shall show that σ_1', the largest singular value of $\Gamma\{[H^*(z)U_{yL}(z)]_-\}$, is precisely equal to ρ, and that ρ is in fact a singular value of multiplicity at least $m+p$.

i) *To show that* $\|\Gamma\{[H^*(z)U_{yL}(z)]_-\}\|_s \le \rho$.

The singular equation (3.5a) gives

$$\left[H^*(z)\check{Y}(z)-\rho X(z)\right]_- = \left[H^*(z)U_{yL}(z)\check{Y}_L(z)-\rho X(z)\right]_- = 0.$$

Hence, noting that $\check{Y}_l(z)$ is maximal phase, by truncation property[6] we obtain

$$\left[\left[H^*(z)U_{yL}(z)\check{Y}_L(z)-\rho X(z)\right]_-\check{Y}_L^{-1}(z)\right]_- = \left[H^*(z)U_{yL}(z)-\rho X(z)\check{Y}_L^{-1}(z)\right]_- = 0,$$

i.e., $[H^*(z)U_{yL}(z)]_- = \rho[X(z)\check{Y}_L^{-1}(z)]_-$. Now,

$$\begin{aligned}\left\|\left[X(z)\check{Y}_L^{-1}(z)\right]_-\right\|_H &\le \|X(z)\check{Y}_L^{-1}(z)\|_\infty\\ &= \|X(z)\check{Y}_L^{-1}(z)U_{yL}^{-1}(z)\|_\infty\\ &= \|X(z)\check{Y}^{-1}(z)\|_\infty = 1\end{aligned}$$

where the first equality holds because $U_{yL}(z)$ is unitary and the last equality comes from that $X(z)\check{Y}^{-1}(z)$ is unitary (3.8). Therefore,

$$\|\Gamma\{[H^*(z)U_{yL}(z)]_-\}\|_s \le \|[H^*(z)U_{yL}(z)]_-\|_\infty \le \rho.$$

ii) *To prove that* $\sigma_1'=\sigma_2'=\cdots=\sigma_{m+p}'=\rho$.

Let $\check{X}(z)=\check{X}_R(z)U_{xR}(z)$ be a right inner factorization of $\check{X}(z)$. Denote $X_{aR}(z) \triangleq Z^{-1}\check{X}_R(z^{-1})$ and $U_{axR}(z) \triangleq U_{xR}(z^{-1})$. Then

$$X(z)=z^{-1}\check{X}(z^{-1})=X_{aR}(z)U_{axR}(z).$$

$U_{axR}(z)$ is still unitary and $X_{aR}(z)$ becomes minimal phase [13]. Let $\Psi(z)$ be any unitary right divisor of $U_{axR}(z)$, i.e.,

$$U_{axR}(z)=\Phi(z)\Psi(z)$$

for some unitary function matrix $\Phi(z)$ with $\Phi(z^{-1})$ and $\Psi(z^{-1})$ being inner. Again we start with the composite singular equation

$$\left[H^*(z)U_{yL}(z)\check{Y}_L(z)-\rho X(z)\right]_- = 0.$$

Multiplying from the right by $\Psi^{-1}(z)$ and after some manipulations, we obtain

$$\begin{aligned}&\left[\left[H^*(z)U_{yL}(z)\right]_-\check{Y}_L(z)\Psi^{-1}(z)\right]_-\\ &=\rho\left[X(z)\Psi^{-1}(z)\right]_- -\rho\left[X_{aR}(z)\Phi(z)\right]_- = \rho X_{aR}(z)\Phi(z)\\ &=\rho X(z)\Psi^{-1}(z).\end{aligned}$$

Note that $\check{Y}_l(x)\Psi^{-1}(z)$ has all its poles outside the unit circle.

We claim that for any constant p-vector v,

$$\|\check{Y}_L(z)\Psi^{-1}(z)v\|_2 = \|X(z)\Psi^{-1}(z)v\|_2 \qquad \text{(A.1)}$$

and therefore, ρ is a, indeed the maximum, singular value of $\Gamma\{[H^*(z)U_{yL}(z)]_-\}$, with $(Y_L(z)\Psi^*(z)v, X(z)\Psi^{-1}(z)v)$ representing a corresponding pair of singular vectors.

To prove the claim, observe that

$$\check{Y}_L(s)\Psi^{-1}(z)v = U_{yL}^{-1}(z)\left[\check{Y}(z)X^{-1}(z)\right]\cdot X(z)\Psi^{-1}(z)v$$

whereas $U_{yL}^{-1}(x)[\check{Y}(z)X^{-1}(z)]$ is unitary. Hence, (A.1).

Next, we prove that the multiplicity of ρ as a singular value is no less than $m+p$.

By assumption, $U_{axR}(z)$ has degree m. Let $U_{axR}(z)=\Phi_m(z)\Phi_{m-1}(z)\cdots\Phi_1(z)$, where $\Phi_i(z)$ are first order unitary factors of $U_{axR}(z)$. Define

$$\Psi_k(z) \triangleq \prod_{j=k}^{1}\Phi_j(z).$$

We claim that there exists a set of $m+p$ linearly independent vectors formed by the p columns of $\check{Y}_l(z)$ and one column from each of the m matrices $\{\check{Y}_l(z)\Psi_k^{-1}(z);\ k=1,2,\cdots,m\}$. This is proved by contradiction as follows.

Suppose that no such independent set can be formed. This implies that for certain $l\le m$, all the columns of the matrix $\check{Y}_L(z)\Psi_l^{-1}(z)$ are linearly dependent on the columns of $\check{Y}_L(z)$ and the matrices $\{\check{Y}_L(z)\Psi_k^{-1}(z);\ 1<k\le l-1\}$, i.e.,

$$\check{Y}_L(z)\Psi_l^{-1}(z)=\check{Y}_L(z)\Theta+\sum_{k=1}^{l-1}\check{Y}_L(z)\Psi_k^{-1}(z)\Xi_k$$

or

$$\Psi_l^{-1}(z)=\Theta+\sum_{k=1}^{l-1}\Psi_k^{-1}(z)\Xi_k \qquad \text{(A.2)}$$

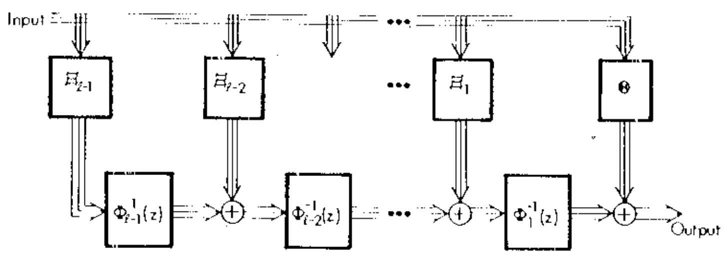
for some $p\times p$ constant matrices Θ and Ξ_k. But this is impossible, for the right-hand side of (A.2) can be modeled as the transfer function of a $(l-1)$th-order realization (Fig. 1), while the left-hand side has order l.

Appendix B—Existence of Solution to the Algebraic Riccati Equation (4.18)

A. Coefficients Expressed in Terms of Infinite Hankel Matrix

To start, we have [cf. (4.11)]

[6]Since $\check{Y}_l(z)$ may have zeros on the unit circle, $\check{Y}_L^{-1}(z)$ may have poles right on the unit circle. Hence, we have to generalize the truncation property to take care of this case. For any rational function matrix $F(z)$, we can define a partition $F(z)=[F(z)]_- + [F(z)]_+$ similar to (2.1) where $[F(z)]_-$ is still in the class $\mathfrak{L}_2^-$ (i.e., it is strictly proper with all its poles inside the unit circle) but $[F(z)]_+$ now allows poles on the unit circle. Then the modified truncation property says the following: for any multiplicable rational function matrices $F(z)$ and $G(z)$, if $[G(z)]_+ = G(z)$, then $[F(z)G(z)]_- = [[F(z)]_- G(z)]_-$.



Transfer Function $= \Theta + \sum_{k=1}^{l-1} \Psi_k^{-1}(z)\Xi_k$

Fig. 1

$$V(z)Z(z)=0 \tag{B.1}$$

where

$$V(z)=\begin{bmatrix} a(z)I_p & -N(z) & \rho\hat{a}^*(z)I_p & 0 \\ 0 & \rho a(z)I_p & -\hat{N}^*(z) & \hat{a}^*(z)I_p \end{bmatrix} \tag{B.2}$$

and

$$Z(z)=\begin{bmatrix} R'(z) & R''(z) \\ \underset{\sim}{\hat{P}}'(z) & \underset{\sim}{\hat{P}}''(z) \\ \underset{\sim}{Q}'(z) & \underset{\sim}{Q}''(z) \\ \hat{S}'(z) & \hat{S}''(z) \end{bmatrix}. \tag{B.3}$$

This is equivalent to (cf. (4.2), (3.14); for brevity, only the single-primed part is given)

$$\{z^{-1}H(z)\}\left\{\frac{\underset{\sim}{\hat{P}}'(z)}{\hat{a}^*(z)}\right\}-\rho\left\{\frac{\underset{\sim}{Q}'(z)}{za(z)}\right\}=\frac{R'(z)/z}{\hat{a}^*(z)} \tag{B.4a}$$

$$\{z^{-1}H^*(z)\}\left\{\frac{\underset{\sim}{\hat{Q}}'(z)}{\hat{a}(z)}\right\}-\rho\left\{\frac{\underset{\sim}{P}'(z)}{za^*(z)}\right\}=\frac{S'(z)/z}{\hat{a}(z)}. \tag{B.4b}$$

We then obtain [cf. (3.5)]

$$\left[\{z^{-1}H(z)\}\frac{\underset{\sim}{\hat{P}}'(z)}{\hat{a}^*(z)}-\rho\frac{\underset{\sim}{Q}'(z)}{za(z)}\right]_-=R'_0z^{-1} \tag{B.5a}$$

$$\left[\{z^{-1}H^*(z)\}\frac{\underset{\sim}{\hat{Q}}'(z)}{\hat{a}(z)}-\rho\frac{\underset{\sim}{P}'(z)}{za^*(z)}\right]_-=S'_0z^{-1}. \tag{B.5b}$$

Expanded into infinite matrix multiplication form (a reverse procedure of that in Section II-B which obtained the functional form (2.4) from $\Gamma\boldsymbol{\eta}=\boldsymbol{\zeta}$),

$$\begin{bmatrix} 0 & H_1 & H_2 & \cdots \\ \hline & \Gamma & & \end{bmatrix}\begin{bmatrix} \mathcal{P}'_1 & \mathcal{P}''_1 \\ \mathcal{P}'_2 & \mathcal{P}''_2 \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}-\rho\begin{bmatrix} \mathcal{Q}'_1 & \mathcal{Q}''_1 \\ \mathcal{Q}'_2 & \mathcal{Q}''_2 \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}=\begin{bmatrix} R'_0 & R''_0 \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \end{bmatrix} \tag{B.6a}$$

$$\left[\begin{array}{c|c} 0 & \\ H_1^* & \\ H_2^* & \Gamma^* \\ \vdots & \end{array}\right]\begin{bmatrix} \mathcal{Q}'_1 & \mathcal{Q}''_1 \\ \mathcal{Q}'_2 & \mathcal{Q}''_2 \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}-\rho\begin{bmatrix} \mathcal{P}'_1 & \mathcal{P}''_1 \\ \mathcal{P}'_2 & \mathcal{P}''_2 \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}=\begin{bmatrix} S'_0 & S''_0 \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \end{bmatrix} \tag{B.6b}$$

where $\{\mathcal{P}'_i\}$, $\{\mathcal{P}''_i\}$, $\{\mathcal{Q}'_i\}$, and $\{\mathcal{Q}''_i\}$ are the sequences obtained from inverse Z-transforming $\underset{\sim}{P}'(z)/za^*(z)$, $\underset{\sim}{P}''(z)/za^*(z)$, $\underset{\sim}{Q}'(z)/za(z)$, and $\underset{\sim}{Q}''(z)/za(z)$, respectively; and the double-primed part is reinserted. It is easy to show that

$$\mathcal{P}'_1=\underset{\sim}{P}'_n=\underset{\sim}{\hat{P}}'_0, \qquad \mathcal{P}''_1=\underset{\sim}{P}''_n=\underset{\sim}{\hat{P}}''_0 \tag{B.7a}$$

$$\mathcal{Q}'_1=\underset{\sim}{Q}'_n=\underset{\sim}{\hat{Q}}'_0, \qquad \mathcal{Q}''_1=\underset{\sim}{Q}''_n=\underset{\sim}{\hat{Q}}''_0 \tag{B.7b}$$

by monicity of $a(z)$. As noted in Section IV already, the echelon form requirement (4.12) implies that

$$\mathcal{Q}'_1=\underset{\sim}{Q}'_n=I_p, \qquad \mathcal{Q}''_1=\underset{\sim}{Q}''_n=0 \tag{B.8a}$$

$$S'_0=\hat{S}'_n=0, \qquad S''_0=\hat{S}''_n=I_p. \tag{B.8b}$$

We can then obtain the following expressions for the coefficients in (4.18).

Lemma B.1

The coefficients in the quadratic equation (4.18) can be expressed in terms of ρ and Γ as

$$\underset{\sim}{\hat{P}}'_0=\rho\mathcal{E}^*(\rho^2I_\infty-\Gamma^*\Gamma)^{-1}\mathcal{T}\Gamma^*\mathcal{E} \tag{B.9a}$$

$$R'_0=\rho\mathcal{E}^*\Gamma\mathcal{T}^*(\rho^2I_\infty-\Gamma^*\Gamma)^{-1}\mathcal{T}\Gamma^*\mathcal{E}-\rho I_p \tag{B.9b}$$

$$\underset{\sim}{\hat{P}}''_0=-\rho\mathcal{E}^*(\rho^2I_\infty-\Gamma^*\Gamma)^{-1}\mathcal{E} \tag{B.9c}$$

$$R_0'' = -\rho\mathcal{E}^*\Gamma\mathcal{T}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathcal{E} \tag{B.9d}$$

where

$$\mathcal{E} \triangleq [I_p \quad 0 \quad 0 \quad \cdots]^* \qquad \mathcal{T} \triangleq \left[\begin{array}{c} 0 \\ \hline I_\infty \end{array}\right] \} \ p \text{ rows}$$

($\mathcal{T}$ is a shift-one-block-down operator. $\mathcal{T}^*$ is, on the contrary, shift-one-block-up operator.)

Proof: For convenience, let

$$\mathcal{P}' \triangleq [\mathcal{P}_1'^* \mid \mathcal{P}_2'^* \mid \cdots]^*, \qquad \mathcal{Q}' \triangleq [\mathcal{Q}_1'^* \mid \mathcal{Q}_1'^* \mid \cdots]^* \tag{B.10a}$$

$$\mathcal{P}'' \triangleq [\mathcal{P}_1''^* \mid \mathcal{P}_2''^* \mid \cdots]^*, \quad \mathcal{Q}'' \triangleq [\mathcal{Q}_1''^* \mid \mathcal{Q}_2''^* \mid \cdots]^*. \tag{B.10b}$$

Then (B.6) can be written as

$$\left[\begin{array}{c} \mathcal{E}^*\Gamma\mathcal{T}^* \\ \hline \Gamma \end{array}\right]\left[\mathcal{P}' \mid \mathcal{P}''\right] - \rho\left[\begin{array}{c|c} \mathcal{Q}_1' & \mathcal{Q}_1'' \\ \hline \mathcal{T}^*\mathcal{Q}' & \mathcal{T}^*\mathcal{Q}'' \end{array}\right] = \left[\begin{array}{c|c} R_0' & R_0'' \\ \hline 0 & 0 \end{array}\right] \tag{B.11a}$$

$$\left[\mathcal{T}\Gamma^*\mathcal{E} \mid \Gamma^*\right]\left[\begin{array}{c|c} \mathcal{Q}_1' & \mathcal{Q}_1'' \\ \hline \mathcal{T}^*\mathcal{Q}' & \mathcal{T}^*\mathcal{Q}'' \end{array}\right] - \rho\left[\mathcal{P}' \mid \mathcal{P}''\right] = \left[\mathcal{E}S_0' \mid \mathcal{E}S_0''\right]. \tag{B.11b}$$

Focus on the single-primed part first. Equation (B.11b) implies

$$\mathcal{T}\Gamma^*\mathcal{E}\mathcal{Q}_1' + \Gamma^*\mathcal{T}^*\mathcal{Q}' - \rho\mathcal{P}' = \mathcal{E}S_0' \tag{B.12}$$

while (B.11a) implies $\Gamma\mathcal{P}' - \rho\mathcal{T}^*\mathcal{Q}' = 0$ and hence $\mathcal{T}^*\mathcal{Q}' = (1/\rho)\Gamma\mathcal{P}'$. Substituting into (B.12) yields

$$\rho\mathcal{T}\Gamma^*\mathcal{E}\mathcal{Q}_1' + \Gamma^*\Gamma\mathcal{P}' - \rho^2\mathcal{P}' = \rho\mathcal{E}S_0'. \tag{B.13}$$

By the facts that (B.8) $\mathcal{Q}_1' = I_p$ and $S_0' = 0$, we have

$$\mathcal{P}' = \rho(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathcal{T}\Gamma^*\mathcal{E}$$

where $(\rho^2 I_\infty - \Gamma^*\Gamma)$ is invertible because $\rho \in (\sigma_{k+1}, \sigma_k)$, i.e., $\rho \neq \sigma_i\ \forall i$, and hence $(\rho^2 I_\infty - \Gamma^*\Gamma)$ is nonsingular. Hence,

$$\hat{P}_0' = \mathcal{P}_1' = \rho\mathcal{E}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathcal{T}\Gamma^*\mathcal{E}.$$

From (B.11a), we also find that

$$R_0' = \mathcal{E}^*\Gamma\mathcal{T}^*\mathcal{P}' - \rho\mathcal{Q}_1'$$

which is exactly equal to what is given in (B.9b). As for the double-primed part, we can similarly derive a parallel to (B.13):

$$\rho\mathcal{T}\Gamma^*\mathcal{E}\mathcal{Q}_1'' + \Gamma^*\Gamma\mathcal{P}'' - \rho^2\mathcal{P}'' = \rho\mathcal{E}S_0''.$$

Imposing the dual conditions (B.8) $\mathcal{Q}_1'' = 0$ and $S_0'' = I_p$ gives

$$\mathcal{P}'' = -\rho(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathcal{E}.$$

Hence,

$$\hat{P}_0'' = \mathcal{P}_1'' = -\rho\mathcal{E}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathcal{E}$$

Again from (B.11a) we obtain that

$$R_0'' = \mathcal{E}^*\Gamma\mathcal{T}^*\mathcal{P}'' - \rho\mathcal{Q}_1'',$$

which equals to (B.9d). Q.E.D. □

Note that R_0' and $\hat{P}_0''$ are self-Hermitian, while $\hat{P}_0'$ and $-R_0''$ are mutually Hermitian, i.e.,

$$(R_0')^* = R_0', \quad (\hat{P}_0'')^* = \hat{P}_0'', \quad (\hat{P}_0')^* = -R_0''. \tag{B.14}$$

Hence, the quadratic equation (4.18) is Hermitian (algebraic Riccati equation).

B. Existence of Solution when $\hat{P}_0''$ Nonsingular

The goal is to solve the ARE

$$[I_p \quad H_0]U\begin{bmatrix} I_p \\ H_0^* \end{bmatrix} = 0 \tag{B.15a}$$

where

$$U = \begin{bmatrix} R_0' & -R_0'' \\ -R_0''^* & -\hat{P}_0'' \end{bmatrix}. \tag{B.15b}$$

The study can be easier accomplished by transforming the equation into a structure easier to manipulate. For this the following lemma will be useful (proof omitted).

Lemma B.2

Let

$$L = \begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix} \tag{B.16}$$

where L_{11}, L_{12}, and L_{22} are $p \times p$ constant matrices with L_{11} and L_{22} being nonsingular. Then (B.15) has a solution H_0 *if and only if* the following equation has a solution $\tilde{H}_0$:

$$[I_p \quad \tilde{H}_0]LUL^*\begin{bmatrix} I_p \\ \tilde{H}_0^* \end{bmatrix} = 0 \tag{B.17}$$

In fact, $H_0 = L_{11}^{-1}\tilde{H}_0 L_{22} + L_{11}^{-1}L_{12}$. □

Letting $L_{11} = L_{22} = I_p$ and $L_{12} = -R_0''\hat{P}_0''^{-1}$, we have

$$LUL^* = \begin{bmatrix} R_0' + R_0''\hat{P}_0''^{-1}R_0''^* & 0 \\ 0 & -\hat{P}_0'' \end{bmatrix} \tag{B.18}$$

Hence, by the above lemma, (B.15) has a solution H_0 if the following equation has a solution $\tilde{H}_0$:

$$\tilde{H}_0\hat{P}_0''\tilde{H}_0^* = R_0' + R_0''\hat{P}_0''^{-1}R_0''^*. \tag{B.19}$$

It can be seen that (B.19) will have a solution if $\hat{P}_0''$ and $(R_0' + R_0''\hat{P}_0''^{-1}R_0''^*)$ are congruent [26].

Lemma B.3

If $\hat{P}_0''$ is nonsingular, then $\hat{P}_0''$ and $(R_0' + R_0''\hat{P}_0''^{-1}R_0''^*)$ are congruent. □

The proof will be carried out in two steps. We shall first show that $\hat{P}_0''$ is congruent to Π where [cf. (B.9c)]

$$\Pi \triangleq -\rho\mathfrak{E}^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\mathfrak{E}. \tag{B.20}$$

Then as a second step we shall show that, as a matter of fact, $\Pi^{-1} = R_0' + R_0''\hat{P}_0''^{-1}R_0''^*$. The congruence between $\hat{P}_0''$ and $(R_0' + R_0''\hat{P}_0''^{-1}R_0''^*)$ is thus established.

For convenience in later derivations, let us first obtain an alternative expression for $\hat{P}_0''$ by using the well-known matrix inversion formula

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}. \tag{B.21}$$

Then we have

$$\begin{aligned}\hat{P}_0'' &= -\rho\mathfrak{E}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathfrak{E}\\ &= -\rho\mathfrak{E}^*\left[\rho^{-2}I_\infty + \rho^{-2}\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\right]\mathfrak{E}, \quad \text{(by (B.21))}\\ &= -\rho^{-1}\left[I_p + \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}\right]. \end{aligned}\tag{B.22}$$

As a dual case,

$$\Pi = -\rho^{-1}\left[I_p + \mathfrak{E}^*\Gamma(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\Gamma^*\mathfrak{E}\right] \tag{B.23}$$

Lemma B.4

$\hat{P}_0''$ and Π are congruent.

Proof: Let "~" denote the congruence relation. Define

$$L_1 = \left[\begin{array}{c|c} I_\infty & 0 \\ \hline \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1} & I_p\end{array}\right]$$

$$L_2 = \left[\begin{array}{c|c} I_\infty & \Gamma\mathfrak{E} \\ \hline 0 & I_p\end{array}\right].$$

Then

$$\begin{aligned} L_2 L_1 &\left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\Gamma^* & 0 \\ \hline 0 & \rho\hat{P}_0''\end{array}\right] L_1^* L_2^* \\ &= L_2\left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\Gamma^* & \Gamma\mathfrak{E} \\ \hline \mathfrak{E}^*\Gamma^* & \rho\hat{P}_0'' + \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}\end{array}\right] L_2^* \\ &= L_2\left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\Gamma^* & \Gamma\mathfrak{E} \\ \hline \mathfrak{E}^*\Gamma^* & -I_p\end{array}\right] L_2^*, \qquad \text{[by (B.22)]} \\ &= \left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\Gamma^* + \Gamma\mathfrak{E}\mathfrak{E}^*\Gamma^* & 0 \\ \hline 0 & -I_p\end{array}\right] \\ &= \left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\mathfrak{T}\mathfrak{T}^*\Gamma^* & 0 \\ \hline 0 & -I_p\end{array}\right]. \end{aligned}$$

Hence,

$$\left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\Gamma^* & 0 \\ \hline 0 & \rho\hat{P}_0''\end{array}\right] \sim \left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma\mathfrak{T}\mathfrak{T}^*\Gamma^* & 0 \\ \hline 0 & -I_p\end{array}\right] \tag{B.24a}$$

Similarly, we can prove that

$$\left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma^*\Gamma & 0 \\ \hline 0 & \rho\Pi\end{array}\right] \sim \left[\begin{array}{c|c} \rho^2 I_\infty - \Gamma^*\mathfrak{T}\mathfrak{T}^*\Gamma & 0 \\ \hline 0 & -I_p\end{array}\right]. \tag{B.24b}$$

In the next, we shall show that i) $(\rho^2 I_\infty - \Gamma^*\Gamma) \sim (\rho^2 I_\infty - \Gamma\Gamma^*)$, and ii) $(\rho^2 I_\infty - \Gamma^*\mathfrak{T}\mathfrak{T}^*\Gamma) \sim (\rho^2 I_\infty - \Gamma\mathfrak{T}\mathfrak{T}^*\Gamma^*)$. And hence by comparing (B.24a) with (B.24b), we shall have $\hat{P}_0'' \sim \Pi$.

To prove i), note that $\Gamma^*\Gamma$ and $\Gamma\Gamma^*$ share the same eigenvalues $\{\sigma_i^2\}$. Hence, $(\rho^2 I_\infty - \Gamma^*\Gamma)$ and $(\rho^2 I_\infty - \Gamma\Gamma^*)$ share the same eigenvalues $\{\rho^2 - \sigma_i^2\}$. And thus, $(\rho^2 I_\infty - \Gamma^*\Gamma) \sim (\rho^2 I_\infty - \Gamma\Gamma^*)$. As for ii), note that $(\Gamma^*\mathfrak{T})(\mathfrak{T}^*\Gamma) = (\mathfrak{T}^*\Gamma^*)(\Gamma\mathfrak{T})$. Since $(\Gamma\mathfrak{T})$ is again block-Hankel, a similar argument as that for i) concludes that $(\rho^2 I_\infty - \Gamma^*\mathfrak{T}\mathfrak{T}^*\Gamma) \sim (\rho^2 I_\infty - \Gamma\mathfrak{T}\mathfrak{T}^*\Gamma^*)$. Q.E.D. □

Lemma B.5

$$R_0' + R_0''\hat{P}_0''^{-1}R_0''^* = \Pi^{-1}.$$

Proof: This is derived via repeated application of the matrix inversion formula (B.21). Consider Π first. From (B.23) we have

$$\begin{aligned}\Pi^{-1} &= -\rho\left[I_p + \mathfrak{E}^*\Gamma(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\Gamma^*\mathfrak{E}\right]^{-1}\\ &= -\rho\left\{I_p - \mathfrak{E}^*\Gamma\left[(\rho^2 I_\infty - \Gamma^*\Gamma) + \Gamma^*\mathfrak{E}\mathfrak{E}^*\Gamma\right]^{-1}\Gamma^*\mathfrak{E}\right\}, \quad \text{(by (B.21))}\\ &= -\rho\left\{I_p - \mathfrak{E}^*\Gamma\left[\rho^2 I_\infty - \mathfrak{T}^*\Gamma^*\Gamma\mathfrak{T}\right]^{-1}\Gamma^*\mathfrak{E}\right\}. \end{aligned}\tag{B.25}$$

Expanding the inverse of bracketed expression above by (B.21),

$$\left[\rho^2 I_\infty - \mathfrak{T}^*\Gamma^*\Gamma\mathfrak{T}\right]^{-1} = \rho^{-2}I_\infty + \rho^{-2}\mathfrak{T}^*\Gamma^*(\rho^2 I_\infty - \Gamma\mathfrak{T}\mathfrak{T}^*\Gamma^*)^{-1}\Gamma\mathfrak{T} \tag{B.26}$$

Again calling for (B.21) to expand the last parenthesized expression,

$$\begin{aligned}(\rho^2 I_\infty - \Gamma\mathfrak{T}\mathfrak{T}^*\Gamma^*)^{-1} &= \left[(\rho^2 I_\infty - \Gamma\Gamma^*) + \Gamma\mathfrak{E}\mathfrak{E}^*\Gamma^*\right]^{-1}\\ &= (\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\\ &\quad - (\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}\left[I_p + \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}\right]^{-1}\\ &\quad \cdot \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\end{aligned}$$

$$= (\rho^2 I_\infty - \Gamma\Gamma^*)^{-1} - (\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}$$
$$\cdot \Gamma\mathfrak{E}\left[-\rho\hat{P}_0''\right]^{-1}\mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}$$

using (B.22). Substituting the last expression into (B.26) and then back into (B.25), we arrive at the following messy equality:

$$\Pi^{-1} = -\rho I_p + \rho\mathfrak{E}^*\Gamma\Big\{\rho^{-2}I_\infty + \rho^{-2}\mathfrak{T}^*\Gamma^*\Big[(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}$$
$$+ \rho^{-1}(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}\hat{P}_0''^{-1}$$
$$\cdot \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Big]\Gamma\mathfrak{T}\Big\}\Gamma^*\mathfrak{E}. \tag{B.27}$$

We shall show that, by expanding $(R_0' + R''\hat{P}_0''^{-1}R_0''^*)$, the same expression can be obtained. First note that

$$R_0'' = -\rho\mathfrak{E}^*\Gamma\mathfrak{T}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathfrak{E}$$
$$= -\rho\mathfrak{E}^*\Gamma\mathfrak{T}^*\left[\rho^{-2}I_\infty + \rho^{-2}\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\right]\mathfrak{E}$$
$$= -\rho^{-1}\mathfrak{E}^*\Gamma\mathfrak{T}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}$$

where the last equality is obtained in view that $\mathfrak{T}^*\mathfrak{E} = 0$. Also note that

$$R_0' = \rho\mathfrak{E}^*\Gamma\mathfrak{T}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}\mathfrak{T}\Gamma^*\mathfrak{E} - \rho I_p$$
$$= \rho\mathfrak{E}^*\Gamma\mathfrak{T}^*\left[\rho^{-2}I_\infty + \rho^{-2}\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\right]\mathfrak{T}\Gamma^*\mathfrak{E} - \rho I_p$$
$$= \rho^{-1}\mathfrak{E}^*\Gamma\Gamma^*\mathfrak{E} + \rho^{-1}\mathfrak{E}^*\Gamma\mathfrak{T}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}$$
$$\cdot \Gamma\mathfrak{T}\Gamma^*\mathfrak{E} - \rho I_p$$

where in the last equality we used the fact that $\mathfrak{T}^*\mathfrak{T} = I_\infty$. Now we have

$$R_0''\hat{P}_0''^{-1}R_0'' + R_0' = \rho^{-2}\mathfrak{E}^*\Gamma\mathfrak{T}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{E}\hat{P}_0''^{-1}$$
$$\cdot \mathfrak{E}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}\Gamma\mathfrak{T}\Gamma^*\mathfrak{E}$$
$$+ \rho^{-1}\mathfrak{E}^*\Gamma\Gamma^*\mathfrak{E} + \rho^{-1}\mathfrak{E}^*\Gamma\mathfrak{T}^*\Gamma^*(\rho^2 I_\infty - \Gamma\Gamma^*)^{-1}$$
$$\cdot \Gamma\mathfrak{T}\Gamma^*\mathfrak{E} - \rho I_p.$$

A comparison of this to (B.27) shows that they are equal. Hence, $\Pi^{-1} = R_0' + R_0''\hat{P}_0''^{-1}R_0''^*$. □

By the discussion before Lemma B.3, this concludes the proof for solution existence when $\hat{P}_0''$ is nonsingular.

C. Existence of Solution when $\hat{P}_0''$ Singular

The special kind of congruence transform as represented by (B.16) will often be appealed to. Therefore, we deliberately denote it as "$\underset{s}{\sim}$", contrasting to "$\sim$" which stands for general congruence transforms.

The two simple results below will also be important in the proof.

Lemma B.6

U is nonsingular $\forall\rho \in (\sigma_{k+1}, \sigma_k)$.

Proof: Recall that U is obtained from the upper half of Z_0 (4.13) and that U is nonsingular if and only if the upper half of Z_0 is nonsingular. Define $\hat{Z}(z) \triangleq z^n Z(z^{-1})$ and

$$\hat{V}(z) \triangleq z^n V(z^{-1})$$
$$= \begin{bmatrix} \hat{a}(z)I_p & -\hat{N}(z) & \rho a^*(z)I_p & 0 \\ 0 & \rho\hat{a}(z)I_p & -N^*(z) & a^*(z)I_p \end{bmatrix}$$

Note that Z_0 is now the z^n term (highest degree term) coefficient matrix of $\hat{Z}(z)$. Since $\hat{V}(z)\hat{Z}(z) = 0$, by a similar minimal basis argument as in Lemmas 4.3 and 4.4 it can be shown that the upper half of Z_0 is nonsingular. □

Lemma B.7

For $\rho \in (\sigma_{k+1}, \sigma_k)$, the eigenvalues of $\hat{P}_0''$ are strictly increasing analytic functions of ρ and $\hat{P}_0''$ can be singular at no more than p values of ρ.

Proof:

$$\frac{d}{d\rho}\hat{P}_0'' = 2\rho^2\mathfrak{E}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-2}\mathfrak{E}.$$

Since $[(\rho^2 I_\infty - \Gamma^*\Gamma)^{-1}]^2$ is positive definite, as a leading diagonal block $\mathfrak{E}^*(\rho^2 I_\infty - \Gamma^*\Gamma)^{-2}\mathfrak{E}$ must also be positive define. Hence, $(d/d\rho)\hat{P}_0''$ is positive definite. Let $\{\lambda_i, i = 1, 2, \cdots, p\}$ be the eigenvalues of $(d/d\rho)\hat{P}_0''$. λ_i are analytic functions of ρ because $\hat{P}_0''$ is (see (B.9c)) [27]. By the perturbation analysis [16, pp. 102–103], $(d/d\rho)\hat{P}_0'' > 0$ implies that $(d/d\rho)\lambda_i > 0$ $\forall i = 1, 2, \cdots, p$. Therefore, any λ_i can be zero at no more than one value of $\rho \in (\sigma_{k+1}, \sigma_k)$. And hence, $\hat{P}_0''$ can be singular at no more than p values of $\rho \in (\sigma_{k+1}, \sigma_k)$. □

Suppose that there exists $\rho_0 \in (\sigma_{k+1}, \sigma_k)$ such that $\hat{P}_0''$, an analytic function of ρ, becomes singular at $\rho = \rho_0$. Then by Lemma B.7 there is an interval $(\rho_0 - \epsilon, \rho_0 + \epsilon)$ for ρ $(\epsilon > 0)$ such that $\hat{P}_0'' \sim \mathrm{diag}(\Delta_\rho, \Lambda_\rho)$ where Δ_ρ and Λ_ρ are diagonal matrices of the eigenvalues of $\hat{P}_0''$ with $\Delta_\rho \to 0$ as $\rho \to \rho_0$ while Λ_ρ remains nonsingular throughout the interval. Now, by an obvious transform we have

$$U = \begin{bmatrix} R_0' & -R_0'' \\ -R_0''^* & \hat{P}_0'' \end{bmatrix} \underset{s}{\sim} \left[\begin{array}{c|c|c} R_0' & \bar{F}_\rho & 0 \\ \hline \bar{F}_\rho^* & \Delta_\rho & 0 \\ \hline 0 & 0 & \Lambda_\rho \end{array}\right] \tag{B.28}$$

where $\bar{F}_\rho$ is a bounded analytic function of $\rho \in (\rho_0 - \epsilon, \rho_0 + \epsilon)$. Note that $\bar{F}_{\rho_0}$ is of full rank by Lemma B.6. By continuity there exists a matrix (function of ρ) L_F such that

$$L_F\bar{F}_\rho = \begin{bmatrix} 0 \\ F_\rho \end{bmatrix}$$

with F_ρ being nonsingular, for ρ in the vicinity of ρ_0. Therefore, the transform

$$\begin{bmatrix} L_F & 0 \\ 0 & I_p \end{bmatrix}$$

can be used to derive the congruence relation

$$\left[\begin{array}{c|c|c} R_0' & \bar{F}_\rho & 0 \\ \hline \bar{F}_\rho^* & \Delta_\rho & 0 \\ \hline 0 & 0 & \Lambda_\rho \end{array}\right] \underset{s}{\sim} \left[\begin{array}{ccc|c} E_\rho & E_2 & 0 & \\ E_2^* & E_1 & F_\rho & 0 \\ 0 & F_\rho^* & \Delta_\rho & 0 \\ \hline \multicolumn{3}{c|}{0} & 0 \quad \Lambda_\rho \end{array}\right] \triangleq \bar{U} \tag{B.29}$$

for some E_1, E_2, and E_ρ. In the following we shall show first that at $\rho=\rho_0$, the quadratic equation (4.18), or (B.15), has a solution H_0 if $E_{\rho_0}\sim(-\Lambda_{\rho_0})$. Then we shall show that, indeed $E_{\rho_0}\sim(-\Lambda_{\rho_0})$, by a continuity argument. This completes the proof.

Lemma B.8

If $E_{\rho_0}\sim(-\Lambda_{\rho_0})$, then (B.15) has a solution for H_0.

Proof: Note that at $\rho=\rho_0$, $\Delta_\rho=0$ and

$$\bar{U} \underset{s}{\sim} \left[\begin{array}{ccc|c} E_{\rho_0} & 0 & 0 & \\ 0 & 0 & F_{\rho_0} & 0 \\ 0 & F_{\rho_0}^* & 0 & 0 \\ \hline \multicolumn{3}{c|}{0} & 0 \quad \Lambda_{\rho_0} \end{array}\right] \triangleq \tilde{U},$$

$$\left(\text{by a transform} \left[\begin{array}{cc|c|c} I & 0 & -E_4F_2^{*-1} & \\ 0 & I & -\tfrac{1}{2}E_3F_2^{*-1} & 0 \\ \hline \multicolumn{2}{c|}{0} & I & 0 \\ & & 0 & I \end{array}\right]\right) \tag{B.30}$$

By Lemma B.6, $E_{\rho 0}$ must be nonsingular. If one can show that $E_{\rho 0}\sim(-\Lambda_{\rho 0})$, then by inspection we can construct a solution $\tilde{H}_0$ for the equation

$$\begin{bmatrix} Ip & \tilde{H}_0 \end{bmatrix} \tilde{U} \begin{bmatrix} Ip \\ \tilde{H}_0^* \end{bmatrix} = 0 \tag{B.31a}$$

as

$$\tilde{H}_0 = \left[\begin{array}{c|c} 0 & G \\ \hline 0 & 0 \end{array}\right] \tag{B.31b}$$

where G is such that

$$E_{\rho_0} = -G\Lambda_{\rho_0}G^*. \tag{B.31c}$$

From Lemma B.2, the existence of solution to (B.31a) implies the existence of solution to (B.15). □

Finally, with the following lemma the proof will be complete.

Lemma B.9

E_{ρ_0} and $(-\Lambda_{\rho_0})$ are congruent.

Proof: Note that for $\rho\in(\rho_0-\epsilon, \rho_0)$,

$$\bar{U} \underset{s}{\sim} \left[\begin{array}{ccc|c} E_\rho & E_2 & 0 & \\ E_2^* & E_3 & 0 & 0 \\ 0 & 0 & \Delta_\rho & 0 \\ \hline \multicolumn{3}{c|}{0} & 0 \quad \Lambda_\rho \end{array}\right], \quad \left(\text{by a transform} \left[\begin{array}{cc|c|c} I & 0 & 0 & \\ 0 & I & -F_2\Delta_\rho^{-1} & 0 \\ \hline \multicolumn{2}{c|}{0} & I & 0 \\ & & 0 & I \end{array}\right]\right)$$

$$\underset{s}{\sim} \left[\begin{array}{cc|c} E_4 & 0 & \\ 0 & E_3 & 0 \\ \hline \multicolumn{2}{c|}{0} & \begin{array}{cc} \Delta_\rho & 0 \\ 0 & \Lambda_\rho \end{array} \end{array}\right], \quad \left(\text{by a transform} \left[\begin{array}{c|c|c} I & -E_4E_5^{-1} & \\ \hline 0 & I & 0 \\ \hline \multicolumn{2}{c|}{0} & I \end{array}\right]\right) \tag{B.32}$$

where

$$E_3 = E_1 - F_\rho\Delta_\rho^{-1}F_\rho^*, \quad E_4 = E_\rho - E_2E_3^{-1}E_2^*.$$

As $\rho\to\rho_0$, $\Delta_\rho\to 0$. Hence,

$$E_3\sim(-\Delta_\rho) \tag{B.33a}$$

and

$$E_\rho\sim E_4 \tag{B.33b}$$

since $E_3\to 0$. (For rigorosity, note that E_1, E_2, and E_ρ are bounded continuous function of ρ in the vicinity of ρ_0, and that E_{ρ_0} is nonsingular.) On the other hand, note that $\hat{P}_0''$ is nonsingular for $\rho\in(\rho_0-\epsilon, \rho_0)$ and hence,

$$\begin{bmatrix} E_4 & 0 \\ 0 & E_3 \end{bmatrix} \sim \begin{bmatrix} \Delta_\rho & 0 \\ 0 & -\Lambda_\rho \end{bmatrix}$$

(cf. (B.18) and Lemma B.3). Now since $E_3\sim(-\Delta_\rho)$, we have $E_4\sim(-\Lambda_\rho)$ and thus $E_\rho\sim(-\Lambda_\rho)$ by (B.33b), for $\rho\in(\rho_0-\epsilon, \rho_0)$. By boundedness and continuity of E_ρ and Λ_ρ we get $E_{\rho_0}\sim(-\Lambda_{\rho_0})$. Q.E.D. □

ACKNOWLEDGMENT

It is our great pleasure to thank Profs. T. Kailath and L. M. Silverman for their encouragements and valuable discussions. We also wish the thank Dr. Y. Genin and M. Bettayeb for many of their suggestions. The very helpful comments of the reviewers are also highly appreciated.

References

[1] V. M. Adamjan, D. Z. Arov and M. G. Krein, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem," *Math. USSR Sbornik*, vol. 15, no. 1, pp. 31–73, 1971.

[2] S. Kung, "A new identification and model reduction algorithm via singular value decompositions," in *Proc. 12th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, pp. 705–714, Nov. 1978; also, "A new low-order approximation algorithm via singular value decompositions," submitted to *IEEE Trans. Automat. Contr.*

[3] L. M. Silverman and M. Bettayeb, "Optimal approximation of linear systems," in *Proc. 1980 Joint Automat. Contr. Conf.*, San Francisco, CA, Paper FA8.A, Aug. 1980; also submitted to *IEEE Trans. Automat. Contr.*

[4] A. Bultheel and P. Dewilde, "On the Adamjan–Arov–Krein approximation, identification, and balanced realization," submitted to *IEEE Trans. Circuits Syst.*

[5] (a) S. Kung, "Optimal Hankel-norm model reductions: scalar systems," in *Proc. 1980 Joint Automat. Contr. Conf.*, San Francisco, CA, Paper FA8.D; Aug. 1980.
(b) ——, "Optimal Hankel-norm model reductions: scalar systems," submitted to *IEEE Trans. Automat. Contr.*

[6] Y. V. Genin and S. Kung, "Rational approximation with Hankel-norm criterion," in *Proc. 19th IEEE Conf. Decision and Control*, Albuquerque, NM, pp. 486–487, Dec. 1980.

[7] Y. V. Genin and S. Kung, "A two-variable approach to the model reduction problem with Hankel norm criterion," submitted to *IEEE Trans. Circuits Syst.*

[8] A. Bultheel, "Numerieke aspekten van de Adamjan- Arov–Krein benadering," Second part of doctoral dissertation, Faculty of Sciences, Leuven University, Leuven, Belgium, 1979.

[9] H. P. Zeiger and A. J. McEwen, "Approximate linear realizations of given dimension via Ho's algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-19, p. 153, Apr. 1974.

[10] V. M. Adamjan, D. Z. Arov, and M. G. Krein, "Infinite Hankel block matrices and related extension problems," *Amer. Math. Soc. Transl.*, series 2, vol. 111, pp. 133–156, 1978.

[11] Z. Nehari, "On bounded bilinear forms," *Ann. Math.*, series 2, vol. 65, no. 1, pp. 153–162, Jan. 1957.

[12] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[13] E. A. Robinson, *Multichannel Time Series Analysis with Digital Computer Programs* (revised edition). San Francisco, CA: Holden-Day, 1967.

[14] F. R. Gantmacher, *The Theory of Matrices*. New York: Chelsea, 1960, chs. 10, 15.

[15] B. L. Ho, "An effective construction of realizations from input/output descriptions," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1966.

[16] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. London, England: Oxford Univ. Press, 1965.

[17] G. W. Stewart, *Introduction to Matrix Computations*. New York: Academic, 1973.

[18] G. D. Forney, Jr., "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Contr.*, vol. 13, no. 3, pp. 493–520, May 1975.

[19] S. H. Wang and E. J. Davison, "A minimization algorithm for the design of linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 220–225, June 1973.

[20] S. Kung and T. Kailath, "Fast projection methods for minimal design problems in linear system theory," *Automatica*, vol. 16, pp. 399–403, 1980; also presented at 4th IFAC Symp. Multivariable Technol. Syst., Fredericton, N.B., Canada, July 1977.

[21] S. Kung, "Multivariable and multidimensional systems: Analysis and design," Ph.D. dissertation, Dep. Electrical Engineering, Stanford Univ., Stanford, CA, 1977.

[22] J. E. Potter, "Matrix quadratic solutions," *SIAM J. Appl. Math.*, vol. 14, no. 3, pp. 496–501, May 1966.

[23] W. A. Coppel, "Matrix quadratic equations," *Bull. Austral. Math. Soc.*, vol. 10, pp. 377–401, 1974.

[24] J. R. Bunch and B. N. Parlett, "Direct methods for solving symmetric indefinite systems of linear equations," *SIAM J. Numer. Anal.*, vol. 8, no. 4, pp. 639–655, Dec. 1971.

[25] A. J. Laub, "A Schur method for solving algebraic Riccati equations," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 913–921, Dec. 1979.

[26] C. C. MacDuffee, *The Theory of Matrices*. New York: Chelsea, 1946.

[27] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1966.

[28] S. Kung and D. W. Lin, "Optimal Hankel-norm model reductions: multivariable systems," in *Proc. 19th IEEE Conf. Decision Contr.*, Albuquerque, NM, pp. 187–193, Dec. 1980.

[29] B. C. Moore, "Singular value analysis of linear systems, Parts I and II," Univ. Toronto, Systems Control Reports 7801 and 7802, Apr. 1978; also in *Proc. 1978 IEEE Conf. Decision Contr.*, pp. 66–73.

Sun-Yuan Kung (M'77) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1971, the M.S.E.E. degree from the University of Richester, NY, in 1974, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1977.

In 1973, he had held a fellowship at the University of Rochester. In 1974, he joined the Amdahl Corporation, Sunnyvale, CA, as an Associate Engineer. From 1974 to 1977 he was a Research Assistant in the Information System Laboratories at Stanford University. Since July 1977, he has joined the faculty of the Department of Electrical Engineering, University of Southern California, Los Angeles, as an Assistant Professor, teaching courses in the areas of control system theory, and digital signal processing. Since July 1979, he has also been a consultant with the Stanford University, Stanford, CA, and General Electric Company, Syracuse, NY.

Dr. Kung's research interests are in the areas of digital signal processing, multivariable and two-dimensional system theory, and VLSI parallel processing.

David W. Lin (S'78) was born in Taipei, Taiwan, China, in 1953. He received the B.S. degree in 1975 from National Chiao Tung University, Taiwan, China, and the M.S. degree in 1979 from the University of Southern California, Los Angeles, both in electrical engineering. He is currently a Ph.D. candidate at the University of Southern California.

Since 1977, he has been holding a Teaching/Research Assistantship in the Electrical Engineering Department of the University of Southern California. During summer 1979, he worked at California Institute of Technology, Pasadena, on VLSI design-automation data base. His current research interests include systems model reduction, digital signal processing and high resolution spectrum estimation.