Example 10: Let \((a_0, a_1, a_2, a_3, a_4) = (0, 0, 1, 1, 0)\).
Application of the above algorithm yields:
\[
\Delta_0 = 0, \quad L_0 = 0, \quad R_0 = \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}, \quad R_{-1} = \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}
\]
\[
\Delta_1 = 0, \quad L_1 = 0, \quad R_1 = \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}, \quad R_0 = \begin{bmatrix} 1 & 0 \\ 0 & s^2 \end{bmatrix}
\]
\[
\Delta_2 = 1, \quad L_2 = 2, \quad R_2 = \begin{bmatrix} 1 & -1 \\ s & 0 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 1 & -s \\ s & 0 \end{bmatrix}
\]
\[
\Delta_3 = 1, \quad L_3 = 2, \quad R_3 = \begin{bmatrix} 1 & -1 \\ 0 & s \end{bmatrix}, \quad R_2 = \begin{bmatrix} 1 & -s \\ s^2 & -s \end{bmatrix}
\]
\[
\Delta_4 = -1, \quad L_4 = 2, \quad R_4 = \begin{bmatrix} 1 & 1 \\ -s & 0 \end{bmatrix}, \quad R_3 = \begin{bmatrix} 1 & -s + s^2 \\ -s + s^2 & -s \end{bmatrix}
\]
As a result
\[
[1 - \sigma + \sigma^2 - \sigma^3] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0
\]
represents a C-MCUM for \((b)\). Taking the reciprocal row vector, we get the (unique) minimal partial realization \(1/(s^4 - s + 1)\).

V. Conclusions
The minimal partial realization problem has been considered as an instance of exact modeling of a behavior on a half-axis, as in [3]. Solutions within this framework are based on polynomials rather than Hankel matrices. A central role is played by behaviors that are the span of a finite number of trajectories and thus do not have a transfer function. It is for this reason that the notion of a behavior rather than a transfer function is essential to the approach. We put the theory to work in deriving an efficient and constructive iterative solution for the scalar case: the celebrated Berlekamp–Massey algorithm. An interesting feature of the algorithm is that its efficiency is enhanced by the update at each step of four polynomials rather than two. It is a topic of future research to put this idea to work for identification purposes, in the context of approximate modeling.

Acknowledgment
The authors would like to thank Prof. J. L. Massey for pointing out the relevance of the Berlekamp–Massey algorithm for cryptographic applications.

References

A Descriptor Solution to a Class of Discrete Distance Problems

M. M. M. Al-Husari, I. M. Jaimoukha, and D. J. N. Limebeer

Abstract—Hankel norm and Nehari-type approximation problems arise in model reduction and \(\mathcal{H}_\infty\)-control theory. Existing solutions to the discrete-time version of these problems may be derived using a standard state-space framework, but the resulting solution formulas require an invertible \(A\)-matrix. As a further complication, the \(D\)-matrix in the representation formula for all solutions becomes unbounded in the optimal case. The aim of this paper is to show that both these complications may be removed by analyzing these problems in a descriptor framework.

Index Terms—Descriptor systems, discrete-time Nehari problem, \(\mathcal{H}_\infty\) control, model reduction.

I. Introduction
It is known that many model reduction and \(\mathcal{H}_\infty\)-control problems may be transformed into the following distance problem: let \(R(z)\) be a stable real rational transfer matrix with McMillan degree \(n\). Then for any \(\gamma > 0\) and any integer \(k < n\), find all transfer matrices \(Q(z)\), with at most \(k\) poles inside the unit disc, that satisfy \(\|R(z) - Q(z)\|_\infty \leq \gamma\) [1], [6]. A necessary and sufficient condition for the existence of a solution requires \(\gamma \geq (k+1)\)st Hankel singular value of \(R(z)\) [1], [4]. The discrete-time version of this problem has received less attention than its continuous-time counterpart. Although the discrete problem can be tackled using a standard state-space approach, this approach breaks down if \(R(z)\) has poles at the origin [7], [8]. This difficulty may be traced to the fact that the conjugation operation cannot be carried out in a standard state-space framework because \(R(z)\) is

Manuscript received January 15, 1993; revised July 29, 1996.
The authors are with the Department of Electrical and Electronic Engineering, Imperial College, London SW7 2BT, U.K.
Publisher Item Identifier S 0018-9286(97)07628-9.
anticausal. As an added complication, the direct feedthrough matrix in the solution formulas becomes unbounded as \( \gamma \to \sigma_{k+1}(R) \) (even if \( R(z) \) has no poles at the origin). One can argue that the discrete-time results can be obtained from those of continuous time using a bilinear transformation. While this procedure is feasible, one would prefer to have a self-contained discrete algorithm since many control problems are inherently discrete-time problems. It is also desirable, for numerical conditioning purposes, that the solution depends in a simple and direct way on the original data of the problem [7], [8]. This paper gives a general solution in a descriptor framework which allows the model reduction of discrete-time systems containing time delays (singular \( A \)-matrix) and which avoids the problems associated with unbounded \( D \)-matrices in the solution representation formula in the optimal case. The advantage of a descriptor framework is that it is closed with respect to conjugation and system inversion as well as the usual operations of addition, subtraction, and multiplication. The solution procedure closely follows that in [4] and [5].

Section II describes the notation, and Section III contains some preliminaries on the representation of systems in a descriptor framework. Section IV gives a characterization of all solutions to the Nehari approximation problem. Three examples are given in Section V, and the conclusions appear in Section VI.

II. NOTATION

This is fairly standard and is reproduced here for convenience. \( \mathbb{R}^{p \times m}, \mathbb{C}^{n \times m} \) are real (complex) matrices. \( \mathcal{D}_+, \mathcal{D}_-, \mathcal{D}_0 \) are \( \{ z \in \mathbb{C} : |z| > 1 \}, \{ z \in \mathbb{C} : |z| < 1 \}, \{ z \in \mathbb{C} : |z| = 1 \} \).

\( \lambda_k(A) \), \( A^* \) is the \( k \)th largest eigenvalue of \( A \), complex conjugate transpose of \( A \).

\( \mathcal{L}_{\infty}^{p \times m} \) represents the space of \( p \times m \) transfer matrices with entries bounded on \( \mathcal{D}_0 \).

\( \| \cdot \|_\infty \) is the \( L_\infty \)-norm of matrices in \( \mathcal{L}_\infty \).

\( \mathcal{H}_{\infty}^{\mathbb{R}^{p \times m}}(\mathcal{H}_{\frac{1}{2 \pi} \mathbb{R}^{p \times m}}) \) is the subspace of \( L_\infty^{\mathbb{R}^{p \times m}} \), matrices analytic and bounded in \( \mathcal{D}_0(\mathcal{D}_-) \).

\( B, \mathcal{H}_{\frac{1}{2 \gamma} \mathbb{R}^{p \times m}} \) represent the \( \gamma \)-ball of \( \mathcal{H}_{\frac{1}{2 \pi} \mathbb{R}^{p \times m}} \), the matrices in \( \mathcal{H}_{\frac{1}{2 \gamma} \mathbb{R}^{p \times m}} \) which satisfy \( \| \cdot \|_\infty \leq \gamma \).

\( G^\sim(z) \) denotes \( G(1/z) \sigma_k(G) \) is the \( k \)th-largest Hankel singular value of \( G \).

\( \rho(A) \) is the spectral radius of (complex) matrix \( A \).

Prefix \( \mathcal{R} \) denotes real rational.

Transfer matrices will be represented by uppercase boldface type and with the dependence on \( z \) suppressed. Matrix dimensions of spaces will also be occasionally suppressed.

III. DESCRIPTOR ALLPASS SYSTEMS

Consider the descriptor system of equations [3]

\[
E x_{k+1} = Ax_k + Bu_k, \quad y_k = Cx_k + Du_k, \quad k = 0, 1, \ldots
\]

(1)

where \( E, A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, C \in \mathbb{R}^{p \times n}, \) and \( D \in \mathbb{R}^{p \times m} \).

Taking \( z \)-transforms

\[
zE x = Ax + Bu, \quad y = Cx + Du.
\]

(2)

The notations

\[
G \equiv (E, A, B, C, D) \quad \text{and} \quad G \equiv \begin{bmatrix} E & A \\ * & C \end{bmatrix} \begin{bmatrix} B \\ D \end{bmatrix}
\]

are used to represent the descriptor system in (2). A change of basis means the transformation

\[
G \equiv (E, A, B, C, D) \stackrel{T_{ET_+}}{\rightarrow} (T, ET, T, AT, TB, CT, D)
\]

for nonsingular \( T_+ \) and \( T_+ \), if \( D \) is nonsingular

\[
G^{-1} \equiv (E, A - BD^{-1}C, BD^{-1}, -D^{-1}C, D^{-1}).
\]

The expression

\[
F \left( \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \right) U = H_{11} + H_{12} U (I - H_{22} U)^{-1} H_{21}
\]

defines a lower linear fractional transformation provided \( (I - H_{22} U)^{-1} \) exists. If \( U \) is a set, \( F(H, U) \) denotes the set \( \{ F(H, U) : U \in U \} \). If \( G^{-1} \equiv \gamma^{-2} G^\sim \), then \( G \) is \( \gamma \)-allpass and satisfies \( GG^\sim = G^\sim G = \gamma^2 I \). A descriptor system is called stable if all the eigenvalues of \( (zE - A) \) are in \( \mathcal{D}_- \) and antistable if all these eigenvalues are in \( \mathcal{D}_+ \). System (1) is called causal if its state \( x \) is completely determined by the initial state \( x_0 \) and inputs \( u_0, u_1, \ldots, u_\ell \); otherwise, it is termed anticausal. Any anticausal system must have a descriptor realization [3]. These are needed since it is known that the Hankel approximation of a given stable system may be anticausal. We begin by giving an allpass lemma for the descriptor system (2). Apart from dealing with the case of a possibly singular \( E \), we cater to the case of \( (zE - A) \) singular for all \( z \).

**Lemma 3.1:** Suppose there exist \( P = P^T \) and \( X \) such that

\[
EPE' - APA' = X' A' - AX = BB', \quad EX = 0.
\]

(3)

Then we have the following.

1) If \( (AP + X') \) is nonsingular and

\[
DD' + (AP + X')C' = 0, \quad DD' + C PC' = \gamma^2 I
\]

(2) defines a unique \( \gamma \)-allpass transfer function \( H = D + C(zE - A)^\# B \). The symbol \( \langle \# \rangle \) denotes a generalized inverse (defined in (6) below).

2) If \( X = 0 \), and \( (E + A) \) and \( P \) are nonsingular, \( \nu_d(E, A) \leq \pi_d(P) \) and \( \pi_d(E, A) \leq \nu_d(P) \).

3) Suppose that the descriptor system has the special form

\[
\begin{bmatrix} E & A \\ * & C \end{bmatrix} D = \begin{bmatrix} 0 & A_{11} & A_{12} \\ 0 & 0_{n_2} & A_{21} & A_{22} \\ * & * & C_1 & C_2 \\ B_1 & B_2 \end{bmatrix} D
\]

(6)

then the dimension of the system can be reduced to \( n_1 + [n_2 - \text{rank}(A_{22})] \). Furthermore if \( A_{22} \) is nonsingular, then the system has a standard realization of order \( n_1 \).

**Proof:**

1) Let \( zE - A = NFM \), be a Smith diagonalization where \( F = \text{diag}(F_1, 0) \), \( N \) and \( M \) are unimodular polynomial matrices and \( F_1 \) has full normal rank. Define

\[
(zE - A)^\# := M^{-1} \text{diag}(F_1^{-1}, 0) N^{-1}
\]

(6)

Now, (3) gives

\[
(zE - A)P(z^{-1}E' - A') + (zE - A)PA' + X
\]

(7)

\[
+ (AP + X')(z^{-1}E' - A') = BB'
\]

\[
\Rightarrow FMFM^F z^{-1} E' - FM(PA' + X)(N^{-1})^{-1}
\]

(8)

\[
\Rightarrow JN^{-1} = 0
\]

(9)
where $J = [0 \ J]$. This shows that range $(B) \subset$ range $(z E - A)$, and so (2) has a solution for $x$ given any $u$. A simple verification shows that $x$ solves (2) if and only if $x = (z E - A)\# B u + M^{-1} J w$ for some $w$. Multiplying (8) from the left by $J$ and using (9)

$$
J N^{-1} (A P + X') M^{-1} F' N^{-1} = 0 \Rightarrow N^{-1} (A P + X') M^{-1} = \begin{bmatrix} \ast & \ast \\ 0 & Z' \end{bmatrix}
$$

since $J F = 0$ and $F_1$ has full normal rank. The $\ast$ denotes an expression irrelevant for the present purpose. $Z$ is full rank since $N, M,$ and $A P + X'$ are full rank by assumption. Hence, (4) implies

$$
DB'(N^{-1}) J' + CM^{-1} J' Z = 0 \Rightarrow CM^{-1} J' = 0
$$

by (9) and the nonsingular nature of $Z$. Thus

$$
y = Du + C (z E - A)\# B u + C M^{-1} J' w
$$

$$
= \left( \mathbf{D} + C (z E - A)\# B \right) u = \mathbf{H} u.
$$

Equations (6) and (10) imply that $C (z E - A)\# (z E - A) = C M^{-1} \text{diag}(I, 0) M = C$. Now we conclude from (4) and (7) that

$$
\mathbf{H} \mathbf{H}^{-1} = DB' + C PC' + C (z E - A)\#
$$

$$
\cdot (A PC' + X'C + B D')
$$

$$
+ (C PA' + CX + D B')(z E - A') = \gamma^2 I.
$$

2) Since $X = 0$, (3) can be written as

$$
(E + A) P(E - A') + (E - A) P(E' + A')
$$

$$
= 2 B B' \Rightarrow \hat{A} + \hat{A}' = 2 B B' \geq 0
$$

where $\hat{A} = (E - A) P(E' + A')$. Since $\delta_e (P) = 0$, [4, Corollary 3.2] implies that $\pi_e (\hat{A} H) \leq \pi_e (\hat{H}) \leq \nu_e (\hat{H})$, where $\hat{H} = (E' - A')^{-1} P - (E + A)^{-1}$. Since $\pi_e (\hat{H}) = \pi_e (P)$

$$
\pi_e [(E - A)(E + A)^{-1}] \leq \pi_e (P)
$$

$$
\nu_e [(E - A)(E + A)^{-1}] \leq \nu_e (P).
$$

(11)

A calculation using the bilinear transformation $s = (z - 1) z + 1)$ shows that $\pi_e (E, A) = \pi_e [(E - A)(E + A)^{-1}]$ and $\nu_e (E, A) = \nu_e [(E - A)(E + A)^{-1}]$, which, together with (11), proves the result.

3) Let $A_{22}$ have a singular value decomposition, $U' A_{22} V = \text{diag}(A_{221}, 0_{n_2 - r})$, where $A_{221}$ is nonsingular and $U, V$ are orthogonal. Applying the change of basis $T_1 = \text{diag}(I, U')$, $T_2 = \text{diag}(I, V)$ to (5) and partitioning conformably gives, after an appropriate adjustment of indexes

$$
\mathbf{H} = \begin{bmatrix}
I & 0 & 0 & 0 \\
0 & 0 & A_{11} & A_{12} \\
0 & 0 & A_{21} & A_{22} \\
0 & 0 & A_{31} & 0 & 0
\end{bmatrix}
$$

$$
B_1 \\
B_2 \\
B_2 \\
B_3
\end{bmatrix}
$$

Applying the basis change

$$
T_1 = \text{diag}\left( \begin{bmatrix} I & -A_{12} A_{22}^{-1} \\ 0 & I \end{bmatrix} \right)
$$

and

$$
T_2 = \text{diag}\left( \begin{bmatrix} I & 0 \\ -A_{22}^{-1} A_{21} & I_2 \end{bmatrix} \right)
$$

$$
\mathbf{H} = \begin{bmatrix}
I & 0 & 0 & 0 & 0 \\
0 & A_{11} & A_{12} & A_{13} & B_1 \\
0 & 0 & A_{21} & A_{22} & 0 & B_2 \\
0 & 0 & A_{31} & 0 & 0 & B_3 \\
0 & 0 & 0 & C_1 & C_2 & C_3
\end{bmatrix}
$$

$$
D
$$

If $A_{22}$ in (5) is nonsingular, $n_2 = r$, and $\mathbf{H}$ has a standard realization of order $n_1$.

Remark 3.1: The (numerically ill-conditioned) Smith form is used in the proof of Lemma 3.1 to establish some system theoretic properties (existence and all pass character). This form will, however, not be used to calculate the solution of the Nehari approximation problem considered in this paper.

Remark 3.2: Equations (3) and (4) will be referred to as the descriptor allpass equations. If $E = I$, (3) implies $X = 0$. In this case, these equations reduce to the familiar allpass equations

$$
P - A P A' = B B', B D' + A P C' = 0, D D' + C P C' = \gamma^2 I.
$$

IV. THE NEHARI APPROXIMATION PROBLEM

This section solves the suboptimal and optimal discrete-time Nehari approximation problem.

Problem 4.1: Suppose $R \in \mathcal{H}_{\infty}^{+p \times m}$ has McMillan degree $n$. Then for all integer $k < n$ and all $\sigma_k (R) \geq \gamma \geq \sigma_{k+1} (R)$, find the set of all error systems, $S := \{ E = R + Q : Q \in \mathcal{H}_{\infty}^{n \times m} (k) : ||E||_{\infty} \leq \gamma \}$.

Our first result gives conditions under which an allpass embedding of $R$ will act as a generator of the set of all error systems $S$ in the suboptimal case $\sigma_k (R) > \gamma > \sigma_{k+1} (R)$.

Lemma 4.1 [5]: For $R$ and $k$ defined in Problem 4.1, let $s_k (R) > \gamma > s_{k+1} (R)$ and suppose there exists a $\gamma$-allpass embedding $\mathbf{H} \in \mathcal{R} L_{\infty}$ such that

$$
\mathbf{H} = \begin{bmatrix}
0 & Q_{11} & Q_{12} \\
Q_{21} & Q_{22}
\end{bmatrix}
$$

$$
Q_{11} \in \mathcal{H}_{\infty} (- \gamma), Q_{12} \in \mathcal{H}_{\infty} - \gamma, Q_{21} \in \mathcal{H}_{\infty} \gamma.
$$

Then the set of all suboptimal error systems associated with $R$ is given by

$$
S = \mathcal{F} (\mathbf{H}, B, C, z \mathcal{H}_{\infty}^{n \times m}).
$$

The next theorem gives such an allpass embedding of $R$ in a descriptor framework.

Theorem 4.2: Let $R \in \mathcal{H}_{\infty}^{+p \times m}$ have a minimal realization given by $\mathbf{R} = (A, B, C, 0)$ where $P (A) < 1$, $A \in \mathcal{R}^{n \times n}$, $B \in \mathcal{R}^{p \times n}$, $C \in \mathcal{R}^{p \times m}$, and $D \in \mathcal{R}^{p \times m}$. Let $P_1 = P_1^T > 0$ and $Q_1 = Q_1^T > 0$ be the controllability and the observability Grammians satisfying

$$
P_1 - A P_1 A' - B B' = 0, \quad Q_1 - A' Q_1 A - C' C = 0.
$$

(14)
Then \( \sigma_k(R) = \sqrt{\lambda_1(P_1Q_1)} \). If \( \sigma_k(R) > \gamma > \sigma_{k+1}(R) \), Problem 4.1 is solved by (13) with

\[
Q_a = \begin{bmatrix}
E_Q & A_Q & B_{Q_1} & B_{Q_2} \\
0 & C_{Q_1} & D_{11} & D_{12} \\
0 & C_{Q_2} & D_{21} & D_{22} \\
I & 0 & -I & -Z' \\
0 & 0 & I & A_0 \\
& & & Q_{11}A_0B + \gamma A_0C' \gamma^{-1} \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0 \\
& & & + \gamma^{-1} B' A_0
\end{bmatrix}
\] (15)

where \( A_0 := (I + A)^{-1}, A_0 := Q_1A_0P_1 = \gamma^{-2}A'A'_0 \) and \( Z := P_1Q_1 - \gamma^2 I \). Furthermore, the order of the realization of \( Q_a \) can be made equal to \( 2n - \text{rank}(A_0) \). Finally if \( A_0 \) is nonsingular, \( Q_a \) has a standard realization of order \( n \).

**Proof:** Since \( \pi_k(R) > \gamma > \pi_{k+1}(R) \), \( Z \) is nonsingular and

\[
\pi_k(P, Z^{-1}) = \pi_k(Q_1, Z) = k. \tag{16}
\]

Also, since \( \rho(A) < 1 \), \( A_0 \) exists. Define \( H \in \mathcal{RH}_\infty \) by

\[
H = D_H + C_H (E_{H} - A_H)^{-1} B_H
\]

\[
\begin{bmatrix}
I & 0 & 0 & 0 \\
0 & E_Q & 0 & A_Q \\
0 & B_{Q_1} & B_{Q_2} \end{bmatrix}
\] (17)

We use Lemma 3.1 to prove that \( H \) satisfies the conditions of Lemma 4.1. Let

\[
T_r = \begin{bmatrix} 0 & 0 & I & 0 \\
0 & I & 0 & 0 \\
0 & I & 0 & T_{33} \end{bmatrix}
\]

where \( T_{33} = A_0P_1 \). A routine calculation verifies that

\[
\det[T_r A_0P_1 + X_H T_{33}] = -\gamma^{-2} \det(Q_1, Z) \neq 0,
\]

since \( Z \) and \( Q_1 \) are nonsingular. Hence \( A_0P_1 + X_H \) is nonsingular. It can be verified that the descriptor allpass equations (3) and (4) are satisfied so that (17) defines a unique \( \gamma \)-allpass transfer function from Lemma 3.1. We show that (15) has the correct stability properties by demonstrating that \( \nu_d(E_Q, A_Q) \leq k \). Setting

\[
T_r = \begin{bmatrix} A_0 & -I & 0 & 0 \\
& & & \end{bmatrix},
\]

\[
\nu_d(E_Q, A_Q) = \nu_d(E_0T_r, A_Q T_r) = \nu_d(A_0, A_1)
\]

where \( A_1 := Z' - A_0 \). Using the (3, 3) block of (3) partitioned conformably with \( P_1, A_1, P_1Z' + A_0, -A_1, P_1Z^{-1}A_1' = Q_1A_0B'A_0Q_1 + \gamma^2 A_0C \gamma^{-1} A_1 \). Hence, it follows from (16) and part 2) of Lemma 3.1 that \( \nu_d(E_Q, A_0) \leq k \), since \( A_0 + A_1 = Z' \) and \( P_1, Z^{-1} \) are nonsingular. Next, we establish the properties in (12). Observe that \( Q_{12} = (E_0, A_Q, B_{Q_2}, C_{Q_1}, \gamma, I) \). It follows that

\[
Q_{12}^{-1} = \begin{bmatrix}
\hat{E}_r & \hat{A} & \hat{B} \\
* & \hat{C} & \hat{D} \\
0 & 0 & I \\
& & & \end{bmatrix}
\]

\[
\begin{bmatrix}
I & 0 & -I & -Z' \\
0 & 0 & I & A_0' A_0' \gamma^{-1} A_1 \\
& & & + \gamma^{-1} A_0' A_0' \gamma^{-1} A_1 \\
& & & + \gamma^{-1} A_0' A_0' \gamma^{-1} A_1
\end{bmatrix}
\]

Multiplying \( \hat{E} \) and \( \hat{A} \) from the right by

\[
\hat{T}_r = \begin{bmatrix} A_0' A_0' & -I & 0 \end{bmatrix}
\]

implies

\[
\nu_d(E, \hat{A}) = \nu_d(E, \hat{T}_r, \hat{A}_r') = \nu_d(A_0' A_0' Z', A_0' Z') = \nu_d(A', I).
\]

Hence, it follows from (14) and part 2) of Lemma 3.1 that

\[
\nu_d(A', I) \leq \pi_k(Z) = 0.
\]

This proves that \( \nu_{12} - 1 \in \mathcal{RH}_{\infty} \). A similar argument proves that \( \nu_{21} - 1 \in \mathcal{RH}_{\infty} \). Finally, the order of the realization of \( Q_a \) follows from part 3) of Lemma 3.1.

Next, we treat the optimal case when \( \gamma = \sigma_{k+1}(R) \). Complications arise because \( Z \) in Theorem 4.2 becomes singular. The next result gives conditions under which an allpass embedding of \( R \) will act as a generator of \( S \).

**Lemma 4.3 [5]:** For \( R \) and \( k \) defined in Problem 4.1, suppose that \( \gamma = \sigma_{k+1}(R) \) has multiplicity \( r \) and that there exists a \( \gamma \)-allpass embedding \( H^* \in \mathcal{RL}_\infty \) such that

\[
H^* = m \begin{bmatrix} R + Q_{11} & \nu_{12} \\
Q_{21} & Q_{22} \end{bmatrix}
\]

\[
Q_{12} = \begin{bmatrix} Q_{11} & Q_{12} \end{bmatrix} \in \mathcal{RH}_{\infty}(k), |Q_{12}^{-1}|, (Q_{22})' \in \mathcal{RH}_{\infty}
\]

where \( l \leq r \) and where \( (\cdot)' \) and \( (\cdot)^r \) denote left and right inverses, respectively. Then

\[
S = \mathcal{F}(H^*, B_{\delta}, \mathcal{H}_\infty(l \times (m-1)) \times (m-1)). \tag{18}
\]

In order to give a construction for such an allpass embedding of \( R \), we need the following result which gives some properties of balanced realizations of discrete-time systems.

**Lemma 4.4:**

1. For \( R \) and \( \gamma \) defined in Lemma 4.3 there exists a balanced realization

\[
R = \begin{bmatrix} A & B \\
C & D \end{bmatrix}
\]

\[
\begin{bmatrix} A_{11} & A_{12} & B_1 \\
A_{21} & A_{22} & B_2 \\
C_1 & C_2 & 0 \end{bmatrix}
\]

such that

\[
P_1 = Q_1 = \text{diag}(\Sigma_1, \gamma I_r)
\]

\[
Z = \text{diag}(Z_1, Z_2) = \text{diag}(\Sigma_1' - \gamma^2 I_{n-r}, 0_r)
\]

with \( \delta_1(\Sigma_1 - \gamma I_{n-r}) = 0 \). Furthermore, \( \rho(A_{11}) < 1 \) and \( \rho(A_{22}) < 1 \).

2. Assume that part 1) is satisfied and introduce the following partitions into (15):

\[
A_s = Q_1A_1P_1 = \gamma^{-2} A_0A_0' = A_0' A_0' + C'C A_0 P_1
\]

\[
B_s = m \begin{bmatrix} Q_1A_1B & \gamma A_0A_0' \gamma^{-1} A_1 \end{bmatrix} = m \begin{bmatrix} B_{11} & B_{12} \end{bmatrix}
\]

\[
C_s = m \begin{bmatrix} C_0A_1P_1 & \gamma B'A_0' \gamma^{-1} A_1 \end{bmatrix} = m \begin{bmatrix} C_{11} & C_{12} \end{bmatrix}
\] (21)
Then
\[
\gamma A_1 B B' A_1' = \gamma^{-1} \begin{bmatrix} C_2 \bar{C}_1 & C_2 \bar{C}_{22} \\
C_2 \bar{C}_{21} & C_{22} \bar{C}_{22} 
\end{bmatrix}
\]
\[
= \begin{bmatrix} * & A_{21}' \\
* & A_{22}' \n\end{bmatrix} \geq 0
\]
\[
\gamma A_1' C' C A_1 = \gamma^{-1} \begin{bmatrix} B_21 B_{12}' & B_21 B_{22}' \\
B_21 B_{12} & B_{22} B_{22}' \n\end{bmatrix}
\]
\[
= \begin{bmatrix} * & A_{12}' \\
* & A_{12}' \n\end{bmatrix} \geq 0
\] (22)

\[
\gamma^{-1} B_{21} B_{21}' = \gamma^{-1} C_2 C_{12} = \gamma^{-1} B_{22} B_{22}' = \gamma^{-1} C_{22} C_{22} = A_{22}' \geq 0.
\] (23)

**Proof:** See [9] for part 1). Part 2) follows from direct calculations.

**Remark 4.1:** It can be easily shown, using the semidefinite character of (22) to (23), that without loss of generality we can assume that they are parts of orthogonal matrices. Hence there introduces pole-zero cancellations at corresponding to equation (22).

**Theorem 4.5:** Let \( \mathbf{R}, k, \gamma \) and all symbols be as defined in Lemma 4.4. Assume that \( A_{22} \) is nonsingular. Then Problem 4.1 has the solution (18) with (w), as shown at the bottom of the page. The order of the realization of \( Q'_n \) can be chosen equal to \( 2(n - r) - \text{rank}(A_1) - 2 A_{12} A_{22}^{-1} A_{22}' \). Finally, if \( (A_1 - A_{12} A_{22}^{-1} A_{22}') \) is nonsingular, \( Q'_n \) has a standard realization of order \( n - r \).

**Proof:** Since \( \delta_{n}(\Sigma_1 - \gamma J_{n-r}) = 0 \), \( Z_1 \) is nonsingular and

\[
\pi_{\gamma}(\Sigma_1 Z_1^{-1}) = \pi_{\gamma}(\Sigma_1 Z_1) = k.
\] (24)

Substituting (19)-(21) into (15) and removing the nonminimal modes corresponding to \( z = -1 \)

\[
Q_n = \begin{bmatrix} I & 0 & 0 & -I & -Z_1' & 0 & 0 & 0 \\
0 & 0 & 0 & I & A_{11}' & A_{12}' & B_{11} & B_{12} \\
0 & 0 & 0 & 0 & A_{21}' & A_{22}' & 0 & 0 \\
* & * & 0 & C_{11} & C_{12} & D_{11} & D_1 \gamma I \\
* & * & 0 & C_{21} & C_{22} & D_{21} & D_2 \gamma I \n\end{bmatrix}
\]

Using a procedure similar to that employed in part 3) of Lemma 3.1 and applying a basis change

\[
T_1 = \text{diag} \begin{bmatrix} I & 0 & -A_1 A_{22}^{-1} \\
0 & I \n\end{bmatrix}
\]
\[
T_2 = \text{diag} \begin{bmatrix} I & 0 & -A_{22}^{-1} A_{21} \\
0 & I \n\end{bmatrix}
\]

and we have (x), also shown at the bottom of the page.

We will now show that this defines a unique \( \gamma \)-allpass transfer function. Let

\[
P_n = \begin{bmatrix} \Sigma_1 & 0 & Z_1 & -I \\
0 & \gamma & 0 & 0 \\
Z_1' & 0 & \Sigma_1 Z_1 & 0 \\
-\gamma & 0 & 0 & 0 \n\end{bmatrix}
\]
\[
X_n = \begin{bmatrix} 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
-\gamma & 0 & 0 & -\Sigma_1 \n\end{bmatrix}.
\] (25)

A simple calculation verifies that \( \det(A_n P_n + X_n') = -\gamma^2 \det(\Sigma_1 Z_1) \neq 0 \), since, \( \Sigma_1, Z_1 \) and \( (I + A_1) \) are nonsingular. Let

\[
H = \begin{bmatrix} \mathbf{R} + Q_{11} & Q_{12} \\
Q_{21} & Q_{22} \n\end{bmatrix}
\]

\[
= \begin{bmatrix} I & 0 & A & 0 & 0 & -I & -Z_1' & 0 \\
0 & E_Q & 0 & A_Q & 0 & B_0 & B_{02} & B_{02} \\
* & * & 0 & C_{Q1} & C_{Q1} & D_{11} & D_{12} \\
* & * & 0 & C_{Q2} & C_{Q2} & D_{21} & D_{22} \\
I & 0 & 0 & -I & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & -I & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -I & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -I \n\end{bmatrix}
\]

It can be verified that the descriptor allpass (3) and (4) are satisfied so that this defines a unique \( \gamma \)-allpass transfer function. Let

\[
\nu(\mathbf{E}_Q, A_Q) = \nu_{\gamma}(\mathbf{A}^{n} - \gamma I) \begin{bmatrix} A_1 & I \\
0 & -I \n\end{bmatrix} = \nu_{\gamma}(A^n, A_1)
\]

where \( A_1 \equiv Z_1' - A^n. \) Using the (3, 3) block of (3) verifies that \( A^n \Sigma_1 Z_1^{-1} A^n - A_1 \Sigma_1 Z_1^{-1} A_1 = B_1 B_1' + B_2 B_2'. \) It follows
from (24) and part 2) of Lemma 3.1 that \( \nu_d(E_Q, A_Q) \leq k \), since \( A^* + A_1 = Z_1^* \) and \( \Sigma_1 Z_1^{-1} \) are nonsingular. A simple calculation using (20) and (21) shows that \( B_{21}Q_{22}C_{12} = -\gamma I \). Consider the identity

\[
\begin{bmatrix}
I & 0 \\
0 & B_{21}
\end{bmatrix}
\begin{bmatrix}
R + Q_{11} & Q_{12} \\
Q_{21} & Q_{22}
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
0 & C_{12}
\end{bmatrix}
= \begin{bmatrix}
R + Q_{11} & Q_{12}C_{12} \\
B_{21}Q_{21} & B_{21}Q_{22}C_{12}
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
0 & C_{12}
\end{bmatrix}
\]

Since the left-hand side is \( \gamma \)-allpass, it follows that

\[
H^* = \begin{bmatrix}
R + Q_{11} & Q_{12}C_{12} \\
B_{21}Q_{21} & B_{21}Q_{22}C_{12}
\end{bmatrix} = \begin{bmatrix}
R + Q_{11}' & Q_{12}' \\
Q_{21}' & Q_{22}'
\end{bmatrix}
\tag{26}
\]

is \( \gamma \)-allpass. Note that (3) and (4) are still satisfied with \( B_{21} \) and \( X \) defined in (25). It only remains for us to show that \( Q_{12}' \) and \( Q_{21}' \) have left and right inverses, respectively, in \( \mathcal{RH}_\infty \). It can be verified, using (26) and the allpass equations, that we have \( (y) \), shown at the bottom of the page, since \( C_{12} \) is nonsingular. It now follows from part 1) of Lemma 4.4 that \( (Q_{12}')^\dagger \in \mathcal{RH}_\infty \). A similar procedure can be used to prove that \( (Q_{21}')^\dagger \in \mathcal{RH}_\infty \). The order of the realization \( Q_* \) follows from part 3) of Lemma 3.1.

V. EXAMPLES

This section gives three examples to illustrate the algorithms presented earlier. The first shows that a causal system with poles at the origin does not necessarily lead to an anticausal generator. Let \( R = \frac{3}{z(z - \frac{3}{2})} \). It can be checked that \( A_* = \frac{\sqrt{3}z(z - \frac{3}{2})}{z - \frac{3}{2}} \) is singular. Applying the algorithm of Theorem 4.2

\[
Q_* = \begin{bmatrix}
-6/(z + 4) & -2\sqrt{3}(z - 2)/(z + 4) \\
-2\sqrt{3}(z - 2)/(z + 4) & -12(z + 1)/(z + 4)
\end{bmatrix}
\]

In the final example, we illustrate the solution of the optimal problem. Let \( R = \sqrt{3}/(z - 0.5) \). Here

\[
A_* = \begin{bmatrix}
24 & 0 & 0 \\
0 & 8/3 & -2/3 \\
0 & -32 & 8
\end{bmatrix}
\]

is singular. Applying the algorithm of Theorem 4.5, we have \( (z) \), shown at the top of the page.

VI. CONCLUSIONS

A general algorithm is presented for the solution of the one-block discrete-time Nehari approximation problem. The form of the solution depends on the parameters of the problem in a simple and direct way. An approach using properties of descriptor systems is used, and it is shown that this approach lends itself naturally to the solution of the discrete-time Nehari approximation problem. In the development of the solution we have made no assumptions regarding the poles of \( R \).
In particular, the algorithm applies even if $R$ has poles at the origin. As Theorems 4.2 and 4.5 show, the generator will have standard state-space realization if $A_n$ is nonsingular.

REFERENCES


ON AN OPEN PROBLEM RELATED TO THE STRICT LOCAL MINIMA OF MULTILINEAR OBJECTIVE FUNCTIONS

Xue-Bin Liang and Li-De Wu

Abstract—This paper gives a combinatorial proof of a “yes” answer to an open question presented in [1], stated as follows: “given a multilinear polynomial $E(x): \{0,1\}^n \rightarrow \mathbb{R}$, is it true that $E(x) = E(\mathbf{x}) - \mathbf{b}^T \mathbf{x}$ has a strict local minimum over the discrete set $\{0,1\}^n$ for almost all $\mathbf{b}$ of sufficiently small norm?” As was pointed out in [1], the answer to this question as stated is “yes,” which can be proved from the obtained analysis results in [1]. The proof, however, is very indirect and unsatisfactory. The question is basically combinatorial in nature, and the answer should therefore have a combinatorial proof [1]. In this paper, we will give a combinatorial proof as expected. Interestingly, the given combinatorial proof can be completed directly by providing a sufficient condition for a conjecture on the strict local minima of multilinear polynomials also postulated in Appendix B of [1] to hold, which was regarded as going slightly beyond the above open question in [1]. It will be demonstrated by a simple counterexample that the conjecture may be not true if the provided sufficient condition is not satisfied.

Index Terms—Analog neural networks, discrete optimization problems, multilinear polynomials, objective functions.

I. INTRODUCTION

In [1] the author has given a strict and elegant analysis of minimum-seeking properties of analog neural networks with multilinear objective functions over the discrete set $\{0,1\}^n$. These analysis results provided a theoretical foundation of the analog and neural approach to discrete optimization problem. A typical discrete optimization problem is to minimize $E(x)$ as $x = (x_1, x_2, \ldots, x_n)$.

Manuscript received December 29, 1995; revised August 26, 1996.

II. A COMBINATORIAL PROOF OF THE OPEN QUESTION

We need the definitions and lemmas stated as follows.

Definition 2.1: A function $E: \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be a multilinear polynomial if it can be written as the form

$$E(x) = a_0 + \sum_{i_1=1}^n a_{1i_1} x_{i_1} + \sum_{i_2=1}^n a_{12i_2} x_{i_2} x_{1i_2} + \cdots + \sum_{i_k=1}^n a_{1\cdots 2i_k} x_{1i_k} x_2 x_{i_{k-1}} x_k + \cdots \sum_{i_1=1}^n a_{1\cdots i_n} x_{1i_n} x_2 \cdots x_n$$

where $a_0, a_{1i_1}, a_{12i_2}, \ldots, a_{1\cdots i_n}$ and $a_1, \ldots, a_n$ are real numbers.

Definition 2.2: A vector $x_0 \in \{0,1\}^n$ is said to be a local minimum of the objective function $E$ if

$$E(x) \leq E(y), \quad \text{for all } y \in N(x)$$

where $N(x)$ denotes the set of all vectors in $\{0,1\}^n$, lying at a Hamming distance of one from $x$. $x$ is said to be a strict local minimum of $E$ if

$$E(x) < E(y), \quad \text{for all } y \in N(x)$$

Definition 2.3: A sequence $x_{i_1}, x_{i_2}, \ldots, x_k$ is said to be a chain in $\{0,1\}^n$ if $x_{i_{i+1}} \in N(x_{i_i})$ for $i = 1, \ldots, k - 1$. A set $S \subseteq \{0,1\}^n$ is said to be connected if there is a chain between every pair of points in $S$.

Lemma 2.1: Let $M \subseteq \{0,1\}^n$ be the set of local minima of the objective function $E$. Then $M$ can be divided into its connected