[25] M. A. Naimark, *Normed Algebra*, 2nd ed. The Netherlands: Groningen, 1972.
[26] L. Pernebo, "An algebraic theory for the design of controllers for linear multivariable feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 171–194, Feb. 1981.
[27] I. Posthelwaite, J. M. Edmunds, and A. G. MacFarlane, "Principal gains and phases in the analysis of linear multivariable feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 32–46, Feb. 1981.
[28] R. Saeks and J. Murray, "Fractional representation, algebraic geometry and the simultaneous stabilization problem," *IEEE Trans. Automat. Contr.*, vol. AC-27, Aug. 1982.
[29] L. E. Sigler, *Algebra*. New York: Springer-Verlag, 1976.
[30] M. Vidyasagar, "On the use of right-coprime factorizations in distributed feedback systems containing unstable subsystems," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 916–921, Nov. 1978.
[31] M. Vidyasagar, H. Schneider, and B. Francis, "Algebraic and topological aspects of stabilization," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 880–895, Aug. 1982.
[32] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, 2nd ed. New York: Springer-Verlag, 1979.
[33] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, Jr., "Modern Wiener-Hopf design of optical controllers—Part II: The multivariable case," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 319–338, June 1976.
[34] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 301–320, Apr. 1981.

Charles A. Desoer (S'50–A'53–SM'57–F'64), for a photograph and biography, see p. 519 of the June 1984 issue of this TRANSACTIONS.

Charles L. Gustafson, for a photograph and biography, see this issue, p. 908.

Connecting Forward and Backward Autoregressive Models

BRIAN D. O. ANDERSON, FELLOW, IEEE, AND AH CHUNG TSOI, MEMBER, IEEE

Abstract—**The aim of this paper is to describe procedures for computing the matrix polynomial defining a vector backward autoregressive recursion from the matrix polynomial defining a vector forward autoregressive recursion. Direct procedures for computing the backward polynomial which do not involve a solution of a matrix Lyapunov equation are described. A novel interpretation is also included of a known procedure which involves the computation of covariance data via the matrix Lyapunov equation. This procedure depends on a standard result connecting forward and reverse time state-space models. A comparison involving operation counts is given of the algorithms.**

I. INTRODUCTION

A WELL-KNOWN problem in statistics or communication engineering is to fit an autoregressive model of prescribed order to a set of given covariance estimates. In the scalar case, Levinson [1] gave a solution to this problem by expressing the unknown coefficients of the autoregressive model in the form of a Yule–Walker equation. The unknown coefficients can then be found in a recursive fashion. In the vector case, Whittle [2] and Wiggins and Robinson [3] were among the first to provide a solution. As is now common, we term the vector or multichannel recursive determination of the matrix coefficients from the set of given covariance matrices the Levinson–Whittle–Wiggins–Robinson algorithm, or the LWR algorithm in short. It is well-known that if the covariance matrices satisfy a positive definite condition, a set of matrices known as the reflection coefficient matrices may be calculated with the aid of an auxiliary set of matrix polynomials, known as the backward polynomials (as contrasted with the coefficient matrices of the autoregressive model, which is known as the forward model). The LWR algorithm constructs the forward polynomials, the backward polynomials, and the reflection coefficient matrices recursively from the given set of covariance matrices, increasing the degree of the forward and backward polynomials at each step of the iteration. Besides serving as a tool for calculating the reflection coefficients, the backward polynomials also define a collection of reverse-time AR processes of increasing order which fit the prescribed covariance.

The LWR algorithm solves one of several related problems which can be described as follows: with A, B, R_i, and K_i denoting the forward polynomial, backward polynomial, covari-

Manuscript received November 17, 1982; revised August 3, 1983. This paper is based on a prior submission of March 15, 1982. Paper recommended by C. R. Johnson, Jr., Past Chairman of the Identification Committee.

B. D. O. Anderson is with the Department of Systems Engineering, Research School of Physical Sciences, Australian National University, Canberra, ACT, Australia.
A. C. Tsoi is with the Department of Electrical and Electronic Engineering, University of Auckland, Auckland, New Zealand.

0018-9286/84/1000-0917$01.00 ©1984 IEEE

ance, and reflection coefficients, respectively,

P1: Given the covariance matrices R_i, find A, B, and K_i.

P2a: Given the forward polynomial A, find B, R_i, and K_i.

P2b: Given the backward polynomial B, find A, R_i, and K_i.

P3: Given the reflection coefficient matrices K_i, find A, B, and R_i.

Problem P1 is solved by the standard LWR algorithm. Problems P2a and P2b are obviously closely related, although it seems that problem P2b does not have an obvious practical implication. Problem P2a may be motivated by the possibility of implementing a given forward matrix polynomial representing an autoregressive approximation of a stochastic process in terms of lattice-ladder structures which have gains closely related to the reflection coefficient matrices K_i. It is well known that the lattice-ladder structure realization is less sensitive to roundoff noises [4]–[7]. In contrast, the direct realization of the forward polynomial A may be undesirable.

Alternatively, problem P2a may be motivated from the stability theory of matrix polynomials. Several authors, e.g., [8], [9] have pointed out the relationship between a least-squares prediction problem associated with a *scalar* discrete time AR process and the problem of checking the unit-circle stability of a prescribed *scalar* polynomial via a Jury table test; the reflection coefficients correspond to certain quantities arising in the Jury table. The idea of a backward polynomial arises here in disguise, the backward polynomial basically being the forward polynomial with coefficients reversed. Related vector results can be found in [10] and [11]; thus [10] studies the stability of a matrix polynomial A with respect to the imaginary axis, and in an attempt to give a matrix Routh table type test for the stability with respect to the imaginary axis of a given matrix polynomial A, a so-called dual of the prescribed polynomial is first constructed; the relation between the dual and the initially given polynomial is like that between the forward and backward polynomials in our problem. Procedures for finding the dual polynomial are described in [10] and are relevant to our work. Reference [11] is concerned with unit circle stability of matrix polynomials and uses forward/backward polynomial pairs without explaining how one can be obtained from the other.

The solution to problem P3 is trivial. It effectively requires the running of an LWR algorithm backwards [12]. An alternative solution is given in [13]. It is shown in [13] that given the causal decomposition of a matrix AR spectrum, a set of parameters known as the Schur parameters may be extracted. This set of Schur parameters is shown to be uniquely defined by a set of so-called Szegö parameters, which turns out to be related to the reflection coefficient matrices. Hence, given the set of reflection coefficient matrices, it is possible to run the algorithms backwards to compute the covariance matrices, R_i.

In this paper, we will consider mainly the solution of problem P2a, while noting very briefly the solutions to problem P2b. We will discuss three approaches to the solution of problem P2a. The first two approaches are novel and are based on a nontrivial extension of the dual polynomial method as presented in [10]; they compute the backward polynomial directly without the intermediate step of calculating the covariance matrices R_i.

In the third approach, solution of a Lyapunov matrix equation is involved, and [14]–[16] use this idea. Our contribution is simply to interpret the schemes of [14]–[16] as an application of an important and standard idea: passing from a forward time to a reverse time state variable model of a given covariance. Connection of forward and reverse time models of continuous-time covariances using state variable equations can be found in [17]–[19], and using matrix fraction and transfer function matrix descriptions in [17]. The connections between forward and reverse time modeling of discrete-time stationary (and nonstationary) covariances using state variable equations can be found in [20], and it is [20] in particular which is most relevant for the understanding of [14]–[16].

Later in the paper, we compare the three approaches via a "big 0" estimate of the operation count of each algorithm.

The contribution of the paper has two distinct thrusts. First, our aim is to outline possible alternative algorithms for backward polynomial computation. (We establish that these are candidate competitors for the currently known algorithms via an operations count, and note that issues of numerical stability, possible parallelism of computation, and ability to incorporate existing proven software are not addressed in respect of already known algorithms, or those of this paper, here or elsewhere. Of course, the exploration of such issues is a major task, and it is obviously premature to assert that no more algorithms are necessary when existing algorithms are not fully evaluated.) Second, we aim to provide elements of a theoretical framework in which to contemplate the problem of passing from a forward to a backward polynomial, in the hope that this might assist in solving other related problems (e.g., those involving d-step ahead predictions, or predictions given missing measurements).

II. Computation of Backward Polynomial Without Intermediate Covariance Calculation

In this section, we shall describe a discrete-time version of a continuous-time procedure outlined in [10] which will allow us to obtain the backward polynomial directly from the forward polynomial.

The problem is as follows.

Given a positive definite $p \times p$ symmetric matrix Q_f and a square real polynomial in z^{-1} of the same dimensions

$$A_N(z^{-1}) = I + A_{N1}z^{-1} + \cdots + A_{NN}z^{-N} \tag{2.1}$$

with $\det(A_N(z^{-1}))$ not zero for z_i, z_j such that $z_i z_j = 1$, find a real matrix polynomial

$$B_N(z^{-1}) = B_{NN} + B_{N,N-1}z^{-1} + \cdots + B_{N1}z^{-(N-1)} + Iz^{-N} \tag{2.2}$$

such that

$$A_N^*(z^{-1})Q_f^{-1}A_N(z^{-1}) = B_N^*(z^{-1})Q_b^{-1}B_N(z^{-1}) \tag{2.3}$$

for some positive definite real symmetric matrix Q_b, and

$$\det\left(A_N(z^{-1})\right) = \det\left(z^{-N}B_N(z)\right). \tag{2.4}$$

The superscript asterisk denotes transposition and replacing of z^{-1} by z.

The solution of this problem depends on the following theorem.

Theorem 2.1: Given a square polynomial $A_N(z^{-1})$, polynomial in z^{-1} with $\det(A_N(z^{-1}))$ not zero for z_i, z_j such that $z_i z_j = 1$, there exists a unique $C(z^{-1})$, polynomial in z^{-1}, such that

$$A_N^{-1}(z^{-1})Q_f\left[A_N^*(z^{-1})\right]^{-1} = A_N^{-1}(z^{-1})C(z^{-1}) + C^*(z^{-1})\left[A_N^*(z^{-1})\right]^{-1} \tag{2.5}$$

where $C(1) = \frac{1}{2}Q_f(A_N'(1))^{-1}$ (the prime denoting the transpose of a matrix), and $A_N^{-1}(z^{-1})C(z^{-1})$ is proper, i.e., is finite when $z^{-1} = 0$.

Proof: The proof is constructive in nature. Observe that

$$A_N^{-1}(z^{-1})Q_f(A_N^*(z^{-1}))^{-1}$$
$$= (z^N I + z^{N-1}A_{N1} + \cdots + A_{NN})^{-1} z^N Q_f (I + A'_{N1}z + \cdots + A'_{NN}z^N)^{-1}$$
$$= (z^N I + z^{N-1}A_{N1} + \cdots + A_{NN})^{-1} M(z) + N(z)(I + A'_{N1}z + \cdots + A'_{NN}z^N)^{-1} + T(z) \quad (2.6)$$

where $T(z)$, $M(z)$, and $N(z)$ are polynomial matrices in z, and the first two summands on the right of (2.6) are strictly proper. These two summands have disjoint poles, and hence $M(z)$ and $N(z)$ are unique. Note that $T(z)$ is straightforward to compute. In order to compute $M(z)$ and $N(z)$ (2.6) may be rewritten as

$$(z^N I + z^{N-1}A_{N1} + \cdots + A_{NN})N(z) + M(z)(I + A'_{N1}z + \cdots + A'_{NN}z^N) = V(z) \quad (2.7)$$

where

$$V(z) = z^N Q_f - (z^N I + z^{N-1}A_{N1} + \cdots + A_{NN}) \cdot T(z)(I + A'_{N1}z + \cdots + A'_{NN}z^N).$$

Since $V(z)$ is known, the unknown matrix polynomials $M(z)$ and $N(z)$ may be obtained in a number of ways [10], [21]. In Appendix A, we discuss methods for obtaining the solution of (2.7).

Now let

$$S = \tfrac{1}{2}A_N^{-1}(1)Q_f(A_N^{-1}(1))' - A_N^{-1}(1)M(1) \quad (2.8)$$

and define a polynomial in z^{-1}

$$C(z^{-1}) = z^{-N}M(z) + A_N(z^{-1})S. \quad (2.9)$$

(Note that because $[z^N A_N(z^{-1})]^{-1}M(z)$ is strictly proper, $M(z)$ as a polynomial in z has maximum degree $N-1$.) The choice of S ensures also that

$$A_N^{-1}(1)C(1) = A_N^{-1}(1)M(1) + S = \tfrac{1}{2}A_N^{-1}(1)Q_f(A_N^{-1}(1))'.$$

We note also that

$$\lim_{z\to\infty} A_N^{-1}(z^{-1})C(z^{-1}) = \lim_{z\to\infty}(z^N A_N(z^{-1}))^{-1}(M(z) + z^N A_N(z^{-1})S) = S. \quad (2.10)$$

Observe now, using (2.6) and then (2.9) that

$$A_N^{-1}(z^{-1})Q_f[A_N^*(z^{-1})]^{-1} - A_N^{-1}(z^{-1})C(z^{-1}) - C^*(z^{-1})[A_N^*(z^{-1})]^{-1}$$
$$= [z^N A_N(z^{-1})]M(z) + N(z)[A_N^*(z^{-1})]^{-1} + T(z) - A_N^{-1}(z^{-1})C(z^{-1}) - C^*(z^{-1})[A_N^*(z^{-1})]^{-1}$$
$$= [N(z) + T(z)A_N^*(z^{-1}) - SA_N^*(z^{-1}) - C^*(z^{-1})][A_N^*(z^{-1})]^{-1}. \quad (2.11)$$

The right-hand side is a matrix fraction with matrix polynomials in z. Because of the form of $A_N^*(z^{-1})$, it is clearly finite at $z = 0$. Since the left side is para-Hermitian,[1] the right side must be, and so is finite also at $z = \infty$. The right side is also equal to its para-Hermitian conjugate, viz.,

$$A_N^{-1}(z^{-1})[N^*(z) + A_N(z^{-1})T^*(z) - A_N(z^{-1})S' - C(z^{-1})]$$

and so must have the same poles (if any). Consideration of the determinantal zeros of $A_N(z^{-1})$ and $A_N^*(z^{-1})$ and the theorem hypothesis rules out any poles other than $z = 0$ or $z = \infty$, which have earlier been ruled out. Hence, the right side and thus the left side of (2.11) is constant. Set $z = 1$. The left side evaluates as zero, i.e., (2.5) holds. From (2.10) we obtain $C(1) = \frac{1}{2}Q_f^{-1}(A_N^{-1}(1))'$.

To establish the uniqueness of $C(z)$, let $C_1(z^{-1}), C_2(z^{-1})$ both satisfy the conditions of the theorem statement: Let $C_0(z^{-1}) = C_1(z^{-1}) - C_2(z^{-1})$. Then

$$A_N^{-1}(z^{-1})C_0(z^{-1}) = C_0^*(z^{-1})(A_N^*(z^{-1}))^{-1}.$$

Using the same techniques we rule out for each side of the existence of poles at $z = 0, \infty$ by one argument and poles at other points by a second argument. Thus, each side is constant. Setting $z = 1$ shows that the constant is zero, hence $C_1(z^{-1}) = C_2(z^{-1})$. ▽▽▽

We now return to the task of finding B_N and Q_b. Let

$$A_N^{-1}(z^{-1})C(z^{-1}) = D(z^{-1})E^{-1}(z^{-1}) \quad (2.12)$$

where $D(z^{-1}), E(z^{-1})$ are polynomials in z^{-1} and are right coprime [22]. Then (2.5) yields

$$E^*A_N^{-1}Q_f(A_N^*)^{-1}E = E^*D + D^*E \quad (2.13)$$

and, as the right side shows, is polynomial in z and z^{-1}. Observe that $\det(E(z^{-1}))$ is a divisor of $\det(A_N(z^{-1}))$ by (2.12), the quotient being polynomial in z^{-1}; similarly $\det(E^*(z^{-1}))$ divides $\det(A_N^*(z^{-1}))$, the quotient being polynomial in z. This means that the left-hand side of (2.13) has a determinant of the form $1/g(z)g(z^{-1})$ for $g(\cdot)$ polynomial in z. Hence, the determinant must be a nonzero constant and $\det(E(z^{-1}))$ is a constant multiple of $\det A_N$. The left-hand side of (2.13) is also para-Hermitian and positive definite almost everywhere on $|z| = 1$. (In fact, since the determinant is constant, the left-hand side of (2.13) is positive definite everywhere on $|z| = 1$.) Consequently, by elementary transformations (here the constancy of the determinant is critical [23]), we can find a W, polynomial in z^{-1} such that

$$E^*A_N^{-1}Q_f(A_N^*)^{-1}E = W^*W. \quad (2.14)$$

Furthermore, W is unimodular, i.e., W^{-1} is also polynomial in z^{-1}. Then

$$A_N^*Q_f^{-1}A_N = (EW^{-1})(EW^{-1})^*. \quad (2.15)$$

Let $N_1 = EW^{-1}|_{z^{-1}=0}$ and $Q_b^{-1} = N_1N_1'$. Because $\det A_N(z^{-1})|_{z^{-1}=0} = I$ and $\det E(z^{-1})$ is a constant multiple of $\det A_N(z^{-1})$, it follows that $\det E(z^{-1})|_{z^{-1}=0} \neq 0$. Also $\det W(z^{-1})|_{z^{-1}=0} \neq 0$ because $W(z^{-1})$ is unimodular. Hence, N_1 is nonsingular, and with $F(z^{-1}) = EW^{-1}N_1^{-1}$ we have

$$A_N^*(z^{-1})Q_f^{-1}A_N(z^{-1}) = F(z^{-1})Q_b^{-1}F^*(z^{-1}). \quad (2.16)$$

[1] A complex $n \times n$ matrix $J(z)$ is para-Hermitian if $J(z) = J^*(z^{-1})$.

$\mathrm{Det}(A_N(z^{-1}))$ is clearly a constant multiple of $\det(F(z^{-1}))$. By considering $A_N(z^{-1})$ and $F(z^{-1})$ at $z^{-1}=0$, we conclude that $\det(A_N(z^{-1})) = \det(F(z^{-1}))$. Let $F(z^{-1}) = I + F_1 z^{-1} + \cdots + F_m z^{-m}$. By considering the terms in z^N and higher degree on each side of (2.16), we observe that $m = N$. Then it is straightforward to verify that

$$B_N(z^{-1}) = z^{-N} F^*(z^{-1})$$

satisfies (2.2) to (2.4). Hence, we have provided a constructive procedure for finding the backward matrix polynomial from the forward matrix polynomial.

Remark 2.1: We note that (2.5) may be regarded as providing a limited kind of partial fraction expansion of the left side of (2.5).

Remark 2.2: In the case that a stable $A_N(z^{-1})$ describes a stationary stochastic process y_t via

$$A_N(z^{-1}) y_t = u_t \tag{2.17}$$

where u_t is a Gaussian white noise process of zero mean and covariance Q_f, the expression $A_N^{-1}(z^{-1}) Q_f [A_N^*(z^{-1})]^{-1}$ may be interpreted as the spectrum $\mathbb{R}(z, z^{-1})$ of the AR process [24] y_t. Equation (2.5) then may be interpreted as an additive decomposition of this AR spectrum into rational causal and anticausal parts. Suppose, in fact we have $\mathbb{R}(z, z^{-1}) = \sum_{i=-\infty}^{\infty} R_i z^{-i}$ where $R_i = E[y_t y'_{t-i}]$. The causal part is $\mathbb{R}_+(z^{-1}) = \sum_{i=1}^{\infty} R_i z^{-i} + \frac{1}{2}R_0$ and the anticausal part is $\mathbb{R}_-(z) = \sum_{i=1}^{\infty} R_{-i} z^i + \frac{1}{2} R_0$. Evidently,

$$\mathbb{R}(z, z^{-1}) = \mathbb{R}_+(z^{-1}) + \mathbb{R}_-(z)$$

and

$$\mathbb{R}_+(z^{-1}) = A_N^{-1}(z^{-1}) C(z^{-1}). \tag{2.18}$$

This kind of decomposition is also used in stochastic realization work; see, e.g., [25].

Remark 2.3: From (2.18) it is observed that once the polynomial $C(z^{-1}) = C_{N0} + C_{N1} z^{-1} + \cdots + C_{NN} z^{-N}$ is computed, the covariance matrices R_i may be computed recursively. Since

$$[C_{N0}\ C_{N1}\ \cdots\ C_{NN}\ 0\ \cdots\ 0] = [I\ A_{N1}\ \cdots\ A_{NN}] \begin{bmatrix} \frac{R_0}{2} & R_1 & R_2 & \cdots \\ & \frac{R_0}{2} & R_1 & \cdots \\ \bigcirc & & \frac{R_0}{2} & \ddots \end{bmatrix}$$

one has

$$R_0 = 2C_{N0},$$
$$R_1 = C_{N1} - A_{N1}\frac{R_0}{2},$$
$$R_2 = C_{N2} - A_{N1}R_1 - A_{N2}\frac{R_0}{2}, \quad \text{etc.}$$

Remark 2.4: Alternatively, once $A_N(z^{-1})$ and $B_N(z^{-1})$ are known, the inverse of the block covariance matrix, i.e., $\mathbb{R}_N^{-1}$, where the $i-j$ block element of $\mathbb{R}_N$ is given by R_{j-i}, may be obtained by using the well-known Gohberg–Heinig inversion formulas for block Toeplitz matrices [26]–[29]. Then, $\mathbb{R}_N$ may be obtained by inverting the block matrix $\mathbb{R}_N^{-1}$.

Remark 2.5: Under some circumstances, $A_N(z)$ can be regarded as the Nth of a sequence of polynomials (e.g., in the fitting of an autoregressive model to a prescribed set of covariance matrices). With an appropriate definition of orthogonality the $A_N(z)$ become orthogonal, and the set of associated matrix polynomials $C(z^{-1})$ also defines a collection of orthogonal polynomials [13], [29]–[31] known as Szegö polynomials of the second kind.

III. Second MFD Based Approach to Finding the Backward Polynomial

The approach presented in this section is motivated by an idea contained in [21]; it provides a conceptually simpler (but not necessarily computationally easier) route to finding the backward polynomial, *without* also providing the polynomial $C(z^{-1})$ of the previous section. The polynomial $C(z^{-1})$ was shown to be usable in a very simple scheme to obtain the covariance matrices R_i. No such simplicity is possible here, although once forward and backward polynomials are known, the R_i can be found by other means, e.g., via a backward LWR algorithm, as noted in the Introduction.

Define

$$\Delta(z) = \det z^N A_N(z^{-1})$$

and

$$\Gamma(z) = \Delta(z)\left[z^N A_N(z^{-1})\right]^{-1} = \mathrm{adj}\left[z^N A_N(z^{-1})\right].$$

Note that $\Delta(z)$ and $\Gamma(z)$ are both polynomials in z. Then

$$A_N^*(z^{-1}) Q_f^{-1} z^N A_N(z^{-1}) \Delta^{-1}(z) = A_N^*(z^{-1}) Q_f^{-1} \Gamma^{-1}(z). \tag{3.1}$$

Excluding zeros at $z=\infty$, the determinantal zeros of $\Gamma(z)$ [which are identical to those of $\Delta(z)$] and $A_N^*(z^{-1})$ will be disjoint if $\det A_N(z^{-1})$ is not zero for z_i, z_j for which $z_i z_j = 1$; equivalently, there is no common zero of $\det z^N A_N(z^{-1})$ and $\det A_N^*(z^{-1})$. Making this assumption, (3.1) is seen to define a coprime right matrix fraction description.

Form a coprime left matrix fraction description (with matrix polynomials in z)

$$E^{-1}(z) F(z) = A_N^*(z^{-1}) Q_f^{-1} \Gamma^{-1}(z) \tag{3.2}$$

with, without loss of generality, $\det E(z) = \det \Gamma(z)$. Then we have the following.

Proposition 3.1: With quantities as defined above, and with $\det A_N(z^{-1})$ not zero for z_i, z_j with $z_i z_j = 1$, define $H(z^{-1}) = z^{-N}\Delta(z) E^{-1}(z)$. Then $H(z^{-1})$ is a matrix polynomial in z^{-1}, with the constant term H_0 possessing unity determinant, and $H(z^{-1}) = F^*(z) K$, for some constant matrix K. The expressions

$$Q_b^{-1} = H_0 K^{-1} H_0' \tag{3.3a}$$

$$B_N(z^{-1}) = z^{-N}\left(H_0^{-1}\right)' H^*(z^{-1}) \tag{3.3b}$$

define a positive definite Q_b and a polynomial $B_N(z^{-1})$ in z^{-1} such that

$$A_N^*(z^{-1}) Q_f^{-1} A_N(z^{-1}) = B_N^*(z^{-1}) Q_b^{-1} B_N(z^{-1}) \tag{3.4a}$$

and

$$\det A_N(z^{-1}) = \det\left[z^{-N} B_N(z)\right]. \tag{3.4b}$$

Proof: Since the right side of (3.2) when multiplied by $\Delta(z)$ is polynomial in z, so is the left side. Coprimeness of $E^{-1}(z)F(z)$ then implies $\Delta(z)E^{-1}(z)$ is a polynomial in z. Then using the definition of $H(z^{-1})$ together with (3.1) and (3.2)

$$H(z^{-1})F(z) = z^{-N}\Delta(z)E^{-1}(z)F(z) = A_n^*(z^{-1})Q_f^{-1}A_N(z^{-1}) \tag{3.5}$$

with

$$\det H(z^{-1}) = \det\left(z^{-N}\Delta(z)\Gamma^{-1}(z)\right) = \det A_N(z^{-1}). \tag{3.6}$$

(Note that we have not yet shown that $H(z^{-1})$ is polynomial in z^{-1}, although it is clearly polynomial in z and z^{-1}.) Now from (3.5) we have $H(z^{-1})F(z) = F^*(z)H^*(z^{-1})$ and so $H^{-1}(z^{-1})F^*(z) = F(z)[H^*(z^{-1})]^{-1}$. Now the poles of the left side are the zeros of $\det(H(z^{-1}))$ and (since F^* is polynomial in z^{-1}), possibly $z=0$. The poles of the right side are the zeros of $\det(H^*(z^{-1}))$, i.e., the zeros of $\det(A_N^*(z^{-1}))$. These zeros are disjoint from those of $\det(H(z^{-1}))$ [and hence $\det(A_N(z^{-1}))$ from (3.6)], and cannot include zero since $\det A_N^*(0) = I$. Hence, the poles of $H^{-1}(z^{-1})F^*(z)$ and $F(z)[H^*(z^{-1})]^{-1}$ are disjoint. Hence, $F(z)[H^*(z^{-1})]^{-1}$ is polynomial in z. Since $F(z)[H^*(z^{-1})]^{-1} = H^{-1}(z^{-1})F^*(z)$, this polynomial must be its para-Hermitian conjugate, i.e., it must be constant. Hence,

$$F(z) = KH^*(z^{-1}) \tag{3.7}$$

for some constant K. Since $\det H(z^{-1}) = \det A_N(z^{-1}) \not\equiv 0$ and $F(z)$ satisfies (3.2), so that $\det F(z) \not\equiv 0$, K must be nonsingular.

Because $F(z)$ is polynomial in z, $H^*(z^{-1})$ is polynomial in z. Because $\det H(z^{-1})|_{z^{-1}=0} = \det A_N(z^{-1})|_{z^{-1}=0} = I$, the constant term H_0 of $H(z^{-1})$ has unity determinant. Now (3.7) and (3.5) yield

$$H(z^{-1})KH^*(z^{-1}) = A_N^*(z^{-1})Q_f^{-1}A_N(z^{-1}).$$

Consideration of the highest power of z on both sides of this equality shows that $H(z^{-1})$ is a polynomial in z^{-1} of degree N. Then the definition (3.3) yields (3.4) immediately. That $Q_b > 0$ follows on setting $z=1$. (Note that $A_N(1)$ is nonsingular by assumption.)

IV. Computation of the Backward Polynomial Via the Solution of a Lyapunov Equation

In this section, we will show briefly how the concepts of forward/reverse discrete time models as detailed in [20] may be applied, with a modest extension to find the backward polynomial via the solution of a Lyapunov equation. The Lyapunov equation approach is not new. However, the material in this section may clarify why the Lyapunov equation approach works.

Assume that we are given a forward AR model

$$y_t + \sum_{i=1}^{N} A_{Ni}y_{t-i} = u_t \tag{4.1}$$

with $E[u_t]=0$, $E[y_t u_s'] = Q_f\delta_{t,s}$. Equation (3.1) may be rewritten in the form

$$x_{t+1} = Fx_t + Gu_t$$
$$y_t = Hx_t + Ju_t \tag{4.2}$$

where

$$F = \begin{bmatrix} 0 & I & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & I \\ -A_{NN} & & \cdots & -A_{N1} \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \end{bmatrix}$$

$$H = -[A_{NN} \;\cdots\; A_{N1}], \qquad J = I$$

and

$$x_t' = \left[y_{t-N}' \; y_{t-N+1}' \; \cdots \; y_{t-1}'\right]. \tag{4.3}$$

Note that $[F, G]$ forms a controllable pair.

Now, with $\Pi = E[x_t x_t']$, a block Toeplitz matrix, then, if we assume that y_t is stationary, we have

$$\Pi = F\Pi F' + GQ_f G'. \tag{4.4}$$

We note that with the definition of x_t in (4.3) $\Pi \equiv \mathbb{R}_N$. Following [30] and assuming for the moment that F is nonsingular or equivalently that A_{NN} is nonsingular, we may define a reverse time model associated with (4.2) as

$$x_t^b = F_b x_{t+1}^b + G_b u_t^b$$
$$y_t^b = H_b x_{t+1}^b + J_b u_t^b \tag{4.5}$$

where

$$\begin{aligned} F_b &= F^{-1}\left(I - GQ_fG'\Pi^{-1}\right) = \Pi F'\Pi^{-1} \\ G_b &= -F^{-1}G \\ J_b &= J - HF^{-1}G \\ H_b &= HF^{-1} + (J - HF^{-1}G)Q_fG'\Pi^{-1} \\ x_t^b &= x_t \\ y_t^b &= y_t \end{aligned} \tag{4.6}$$

and

$$E\left[u_t^b\right] = 0,$$
$$E\left[u_t^b u_s^{b\prime}\right] = \left(Q_f - Q_fG'\Pi^{-1}GQ_f\right)\delta_{t,s} = Q_b\delta_{t,s}$$
$$E\left[x_T u_t^{b\prime}\right] = 0 \qquad \forall t < T.$$

Note that nonsingularity of Π is guaranteed by the controllability of $[F, G]$. For F nonsingular, we have

$$F^{-1} = \begin{bmatrix} -A_{NN}^{-1}A_{N,N-1} & -A_{NN}^{-1}A_{N,N-2} & \cdots & -A_{N,N}^{-1}A_{N1} & -A_{NN}^{-1} \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & & I & 0 \end{bmatrix}.$$

Then, for some B_{Ni}, $i=1,\cdots,N$,

$$F_b = F^{-1}(I - GG'\Pi^{-1}) = \begin{bmatrix} -B_{N1} & -B_{N2} & \cdots & -B_{N,N-1} & -B_{N,N} \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}.$$

In fact

$$B_{Ni} = A_{NN}^{-1}(A_{NN-i} - P_{1,N-i+1}) \qquad i = 1, 2, \cdots, N-1$$
$$B_{NN} = A_{NN}^{-1}(I - P_{11})$$

where P_{ij} is the block i, j entry of Π^{-1}. We also obtain

$$H_b = [0 \ \cdots \ 0 \ I]$$
$$J_b = 0$$
$$G_b = \begin{bmatrix} A_{NN}^{-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Then, the backward model, which would have the same state-vector as the forward model, is given by

$$\sum_{i=0}^{N} B_{N,N-i} y_{t-i} = A_{NN}^{-1} u_t^b \qquad (B_{N,0} = I). \tag{4.7}$$

It is noted that for F singular, then the representation as given in (4.6) is no longer suitable. However, as shown in [30], there is little adjustment required, and one still has $F_b = \Pi F' \Pi^{-1}$ from which the backward AR model follows. We omit the details

Remark 4.1: The above description of computing $B_N(z^{-1})$ from $A_N(z^{-1})$ presents an application of the forward/reverse discrete time model connection as presented in [20] to yield a novel interpretation of [14]–[16].

Remark 4.2: The above procedure may still be used even if $\det A_N(z^{-1})$ does not have all its zeros in $|z| < 1$, as long as no two zeros of $\det(A_N(z^{-1}))$ have a product of 1, this being the solvability condition for Π. The matrix $\bar{Q}_b^{1/2}$ may be obtained as

$$\bar{Q}_b = (\bar{Q}_b^{1/2}) \Sigma (\bar{Q}_b)^{-1/2}$$

where Σ is a signature matrix. Of course, in this case, we can no longer impose a stochastic process interpretation.

Remark 4.3: In [14]–[16] Hadidi, Morf, and Porat studied the problem of obtaining the backward matrix polynomial from a given forward matrix polynomial. They do not explicitly formulate the backward matrix polynomial as we did using the forward/reverse discrete time model connection. In [15], [16], Porat and Morf began with the Yule–Walker normal equation. Then it is possible to rewrite the normal equation as a linear matrix equation involving the covariance matrices R_i, $i = 0, 1, \cdots N$ as unknown variables, the other matrices being of Hankel and Toeplitz type. By stacking up the columns of R_i, this equation may be reformulated as a linear matrix equation while preserving the Hankel and Toeplitz matrix form. The dimension of this linear matrix equation may be reduced by noting the special form of Hankel and Toeplitz matrices, and the resulting matrix equation may be solved using standard numerical methods for linear equation solution or with special methods appropriate to Hankel and Toeplitz matrices. In [14], [15] Hadidi and Morf formulated the forward polynomial in a forward discrete time model as depicted in (4.2). Then the Lyapunov equation (4.4) is obtained. Again by stacking up the columns of the covariance matrices R_i, the Lyapunov equation may be reformulated as a linear matrix equation. It is found that this linear matrix equation may be solved efficiently by noting the special nature of the Lyapunov equation. The immediate contribution of this section is then to exhibit a natural formulation of the problem of passing from the forward to the backward polynomial, in terms of forward and reverse time state variable models, rather than to set out a new computational procedure.

Remark 4.4: It is important to note the various choices suggested in this section. First one can obtain Π by several different ways, including an infinite sum doubling algorithm [32, p. 67] in case $\det A(z^{-1})$ is stable. Second, one can proceed via $A_N(z^{-1}) \to \Pi \to \Pi^{-1} \to B_N(z^{-1})$ as our formulas have suggested. Alternatively, one can employ the LWR algorithm, using the entries of Π, to find $B_N(z^{-1})$. Even the step $\Pi \to \Pi^{-1}$ could use a specialized Toeplitz algorithm, e.g., [26], [33], [34] or the LWR algorithm.

V. Comparison of the Three Methods

In earlier sections, we have explained what is involved in the various methods. Here we will discuss computational aspects. By way of general comments, we note that a full appreciation of the computational issues involved demands much more than a "big 0" theory, with a possibly underestimated scaling constant, and sometimes the use of analytically useful but computationally conservative assumptions in generating the argument of the $0[\cdot]$ operation. It is also relevant to consider for example the question of numerical stability [35], the possibilities for parallel computation, and the availability of existing proven software.

The scheme of [15] referred to earlier is characterized to the point of a "big 0" theory with an estimate of the scaling constant. The operation count is $0[p^6 n^2] + 0[p^4 n^2]$.

In Appendix B, we show that the compoutation of the Hermite form of a $p \times p$ polynomial matrix of degree n involves $0[p^6 n^2] + 0[p^4 n \log^2 pn]$ operations. It is not hard to verify that the Hermite form computation is the major burden of the scheme of Section II. We shall argue below that it provides a computational bound also for the scheme of Section III. For embedded within the Hermite form calculations are subalgorithms (of lesser complexity) for passing from a right to a left matrix fraction description, and for finding a row proper matrix equivalent to a given matrix under unimodular multiplication on the left. Of course, the Hermite form calculation also provides an approach to evaluating a determinant of a matrix.

Let us consider how the procedure of Section III can be bounded. In outline:

1) The calculation of $\det z^N A_N(z^{-1})$ can be bounded by the Hermite calculation.

2) The calculation of $\Gamma(z)$ can be proved by finding a left MFD equivalent to a right MFD (or vice versa: $\Gamma^{-1}(z) = z^N A_N(z^{-1})[\Delta(z) I]^{-1}$).

3) The calculation of E, F proceeds by finding a left MFD equivalent to a right MFD.

4) The calculation of $H(z^{-1})$ is like that of $\Gamma(z)$.

5) The construction of Q_b and $B_N(z^{-1})$ is obviously over-bounded by the Hermite bound.

Consequently, the bound $0[p^6 n^2] + 0[p^4 n \log^2 pn]$ applies.

VI. Conclusions

In this paper, we have described two procedures and provided a new interpretation of a third procedure for passing from a forward polynomial to a backward polynomial associated with AR processes or stationary spectra. The first two procedures involve the construction of the backward polynomial directly without the intermediate step of evaluating the covariance matrices. The third procedure involves the solution of a Lyapunov equation which would yield the covariance matrices as an intermediate step. In either case, the reflection coefficient matrices may be obtained.

As foreshadowed in the Introduction, we comment presently on the task of passing from a backward polynomial to a forward polynomial. This is virtually the same problem as we have solved, noting the following observation: $[A_N, B_N]$ is a forward–backward pair if and only if $[z^{-N} B_N(z), z^{-N} A_N(z)]$ is a forward–backward pair. So, given B_N, we can find A_N using the forward to backward algorithms of the paper taking $z^{-N}B_N(z)$ as

As commented earlier, the problem of finding the forward polynomial, the backward polynomial, and the covariance matrices from the reflection coefficient matrices K_i is straightforward. The forward and backward polynomials may be obtained by running the LWR algorithm backwards [12]. Alternatively, one may make use of the equivalence of the Schur parameters and the Szegö parameters (which may be regarded as a scaled version of the reflection coefficient matrices) as shown in [13]. Once the Schur parameters are obtained, then, the covariance matrices may be obtained. Then the forward and backward polynomials may be obtained accordingly using the LWR algorithm.

We would also note that efficient solutions of the Lyapunov equation exist [14]–[16] and may be employed in the third approach. However, we note that the first two approaches solve a problem posed but not solved at the conclusions of both [14] and [16]. The computational burden is of the same order of magnitude for the three approaches while more detailed numerical properties are not yet available.

Our new understanding of the third approach is believed to have relevance in defining a solution to a related problem under examination by the authors: that of relating d-step ahead forward and backward predictors.

Appendix A
Solution of $z^N A_N(z^{-1})N(z)+M(z)A_N^*(z^{-1})=V(z)$

In this Appendix, we will indicate ways of solving the equation $z^N A_N(z^{-1})N(z)+M(z)A_N^*(z^{-1})=V(z)$, where $[z^N A_N(z^{-1})]^{-1}M(z)$ and $N(z)[A_N^*(z^{-1})]^{-1}$ are strictly proper. There are a number of methods for solving this equation [10], [21] with [21] containing a useful summary.

1) Use a property of matrix Bezout identities (the original approach in [21]).

2) Reduce $z^N A_N$ and A_N^* to Smith form, and solve the resulting equation term by term.

3) Reduce $z^N A_N$ and A_N^* to Hermite form and solve the resulting equation term by term.

4) In case $z^N A_N, A_N^*$ are regular, replace the equation by a linear matrix equation (over the reals).

5) By equating coefficients of each power in each entry, form a vector matrix equation ($Ax=b$ type).

In order to discuss the computational burden of the scheme of Section II, we shall consider approach 3) in more detail [10]. One transforms the matrix polynomial $(z^N I+z^{N-1}A_{N1}+\cdots+A_{NN})$ into an upper triangular Hermite form [22] by finding some postmultiplying unimodular matrix U_1. Similarly, the matrix polynomial $(I+A'_{N1}z+\cdots+A'_{NN}z^N)$ may be transformed into a lower triangular Hermite form by another premultiplying unimodular matrix U_2. Equation (2.8) becomes

$$\left(z^N I+z^{N-1}A_{N1}+\cdots+A_{NN}\right)U_1(z)U_1^{-1}(z)N(z) + M(z)U_2^{-1}(z)U_2(z)\left(I+A'_{N1}z+\cdots+A'_{NN}z^N\right)=V(z)$$

or

$$\begin{bmatrix} \bar a_{11} & & \cdots & \bar a_{1p} \\ & \bar a_{22} & & \\ & & \ddots & \vdots \\ 0 & & & \bar a_{pp}\end{bmatrix}\bar N(z)+\bar M(z)\begin{bmatrix} \bar b_{11} & & & 0 \\ & \bar b_{22} & & \\ \vdots & & \ddots & \\ \bar b_{p1} & & & \bar b_{pp}\end{bmatrix} = V(z). \tag{A.1}$$

The (p,p) term of (A.1) becomes $\bar a_{pp}\bar n_{pp}+\bar b_{pp}\bar m_{pp}=v_{pp}$, where $\bar a_{pp}, \bar b_{pp}$ are relatively prime. Hence, $\bar m_{pp}, \bar n_{pp}$ may be obtained. The $(p-1,p)$ term gives

$$\bar a_{p-1,p-1}\bar n_{p-1,p}+\bar m_{p-1,p}\bar b_{pp}=-\bar a_{p-1,p}\bar n_{pp}+v_{p-1,p}$$

where $\bar a_{p-1,p-1}, \bar b_{pp}$ are relatively prime. Hence, $\bar m_{p-1,p}$ and $\bar n_{p-1,p}$ may be calculated. By proceeding successively with terms $(p-2,p)$, $(p-3,p),\cdots$ the last column of $\bar M(z)$ and $\bar N(z)$ may be found. Proceed then to examine the entries in positions $(p,p-1)$, $(p-1,p-1)$, $(p-2,p-1)\cdots$ and so on. In this manner, all entries of $\bar M(z)$ and $\bar N(z)$ are obtained. The polynomial equations encountered at each stage are all solvable as each diagonal entry of $(z^N I+z^{N-1}A_{N1}+\cdots+A_{NN})U_1$ is coprime with each diagonal entry of $U_2(I+A'_{N1}z+\cdots+A'_{NN}z^N)$. We also demand that at each stage $\bar m_{ij}$ or $\bar n_{ij}$ can be chosen to have the least degree. Once $\bar M(z)$ and $\bar N(z)$ are found, one can find $N(z)=U_1(z)\bar N(z)$ and $M(z)=\bar M(z)U_2(z)$ as required, save that $[z^N A_N(z^{-1})]^{-1}M(z)$ and $N(z)[A_N^*(z^{-1})]^{-1}$ may not be strictly proper. Divide $z^N A_N(z^{-1})$ into $M(z)$ to determine a remainder $\tilde M(z)$ for which $[z^N A_N(z)^{-1})]^{-1}\tilde M(z)$ is strictly proper. Then replace $M(z), N(z)$ by $\tilde M(z)$ and $N+[z^N A_N(z^{-1})]^{-1}[M(z)-\tilde M(z)]A_N^*(z^{-1})$ to fulfill the properness requirement.

Appendix B
Computation of the Hermite Form: An Operation Count

We begin with some preliminary calculations.

Lemma B1: The computation of

$$g=\sum_{i=1}^{p} a_i x_i$$

where x_i, $i=1,2,\cdots,p$ are polynomials in z, with highest degree at most d, g, i, the greatest common divisor of the polynomial set $(X_1, X_2,\cdots,X_p)$ a_i, $i=1,2,\cdots,p$ are polynomials in z, with highest degree at most d, can be found in $0[p^2 d\log^2 d]$ operations.

Proof: We shall use the fact [36, p. 308] that we can find α_{12}, β_{12}, and $g_{12}=\gcd(x_1,x_2)$ such that $\alpha_{12}x_1+\beta_{12}x_2=g_{12}$ in $0[d\log^2 d]$ operations. Without loss of generality, assume that $\deg x_i \leqslant \deg x_2 \leqslant \cdots$. Then, construct $g_{12}=\alpha_{12}x_1+\beta_{12}x_2$, then

$$\begin{aligned} g_{13}&=\gcd(g_{12},x_3)=a_{13}g_{12}+c_{13}x_3\\ &=(a_{13}\alpha_{12})x_1+(a_{13}\beta_{12})x_2+c_{13}x_3\\ &=\alpha_{13}x_1+\beta_{13}x_2+\gamma_{13}x_3.\end{aligned}$$

Note that $a_{13}\alpha_{12}$ and $a_{13}\beta_{12}$ are evaluated with two polynomial multiplications—taking $0[d\log d]$ operations [36, p. 269], and α_{13}, β_{13} are the remainder of $a_{13}\alpha_{12}$ and $a_{13}\beta_{12}$ divided by x_3 again, $0[d\log d]$ operations for each division [36, p. 289]. Continuing in this vein, we find a total operation count of

$$\begin{aligned}&0[d\log^2 d]\\&+0[d\log^2 d]+4\ 0[d\log d]\\&+0[d\log^2 d]+6\ 0[d\log d]\\&\quad\vdots\\&+0[d\log^2 d]+2(p-1)\ 0[d\log d]\end{aligned}$$

which can be crudely overbounded by $0[p^2 d\log^2 d]$.

Next we recall the following result from [37].

Lemma B2: Consider a proper right matrix fraction description ND^{-1} with $\deg D=d$, $q=$ number of rows of $N+$ number of rows of D, $m=$ number of rows of D, $\nu=$ maximum row degree in a row proper minimal left MFD $A^{-1}B$ with $A^{-1}B=ND^{-1}$. Then A, B can be found in $3q^2\nu md$ operations. We shall now show the following.

Lemma B3: Suppose that $x_1,\cdots,x_p$ are prescribed polynomials with degree $\leqslant d$. Then a unimodular matrix U with degree at most d such that

$$U\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=\begin{bmatrix}g\\0\\\vdots\\0\end{bmatrix}\tag{B.1}$$

where $g=\gcd(x_1,\cdots,x_p)$ can be found in $0[3p^2d^2]+0[p^2d\log^2 d]$ operations.

Proof: Without loss of generality, let $\deg x_1\leqslant\deg x_2\leqslant\cdots$. Find the first row of U by the procedure of Lemma B1. Use Lemma B2 with $N=[x_1,x_2,\cdots,x_{p-1}]'$, $D=x_p$ and set

$$U=\begin{bmatrix}a_1 & \cdots & a_p\\ -A & & B\end{bmatrix}.$$

Note that the last $(p-1)$ rows of (B.1) are equivalent to $-AN+BD=0$ or $A^{-1}B=ND^{-1}$. Thus, (B.1) indeed holds.

To show that U is unimodular, let u_i' denote the ith row of U. If U is not unimodular, there exist constants, α_i, $i=1,2,\cdots,p$ and a z_0 such that $\Sigma_i\alpha_i u_i'(z_0)=0$. Also since the last $(p-1)$ rows of U are linearly independent for all z_0 (by the minimality of $A^{-1}B$, see Lemma B.2), we must have $\alpha_1\neq 0$. But

$$U\begin{bmatrix}x_1/g\\x_2/g\\\vdots\\x_p/g\end{bmatrix}=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}.$$

Set $z=z_0$ and premultiply by $[\alpha_1\ \cdots\ \alpha_p]$ to obtain $\alpha_1=0$—a contradiction.

The operation count for the first row of U is immediate from Lemma B1. The quantities q, ν, m, d of Lemma B2 become p, $\leqslant d$, 1, d in Lemma B3, giving a bound $3p^2d^2$.

The other tool we need also makes use of Lemma B2.

Lemma B4: Let $M(z)$ be $p\times p$ of degree d and determinantal degree d'. Then $M(z)$ can be reduced to row proper form in at most $0[24p^3d'(d'+d)]$ operations.

Proof: Apply Lemma B2 with $N=I$, $D=M'(z)$. Then we can find A, B with $A^{-1}B=(M^{-1})'$ in approximately $3(2p)^2\cdot d'\cdot p\cdot d=12p^3dd'$ operations. Note that $\nu\leqslant d'$. Now apply Lemma B2 with $N=B'$, $D=A'$. Then we obtain coprime $\bar{A}$, $\bar{B}$ with $\bar{A}^{-1}\bar{B}=M^{-1}I$ and $\bar{A}$ row proper. (The second operation count is $3(2p)^2d'\cdot p\cdot d'=12p^2d'$.)^2 Notice that $\bar{A}^{-1}\bar{B}$ and $M^{-1}I$ are both coprime MFD's, so that M and $\bar{A}$ are related by a unimodular transformation.

The main result is as follows.

Theorem B5: Let $X(z)$ be a $p\times p$ polynomial matrix with each entry of degree at most n. Then a unimodular matrix reducing $X(z)$ to Hermite form and the Hermite form of $X(z)$ can be found in $0[p^4n\log^2 pn]+0[p^6n^2]$ operations.

Proof:

Step 1: Find a unimodular $U(z)$ so that

$$U_1(z)X(z)=\begin{bmatrix}h_{11} & x_{12}' & \cdots & x_{1p}'\\ 0 & x_{22}' & \cdots & x_{2p}'\\ \vdots & & & \\ 0 & x_{p2}' & \cdots & x_{pp}'\end{bmatrix}=\begin{bmatrix}h_{11} & X_{12}'\\ 0 & X_{22}'\end{bmatrix}$$

using the ideas of Lemma B3. Note that because the entries of $U_1(z)$ have degree at most n, the x_{ij}' have degree at most $2n$. There are $0(3p^2n^2)+0(p^2n\log^2 n)$ operations involved in obtaining $U_1(z)$ and $h_{11}(z)$ a further $[pn\log n]$ operations in evaluating each entry of X_{12}' and X_{22}', or $0[p^2(p-1)n\log n]$ operations for both matrices. Throw away U_1.

Step 2: Reduce X_{22}' to row proper form (using the ideas of Lemma B4). Note that $X(z)$ has maximum determinantal degree pn (as has X_{22}'). The operation count is $0[24(p-1)^3pn(pn+2n)]=0[p^5n^2]$ crudely. Call the row reduced matrix X_{22}. It has maximum determinantal degree pn, and being row reduced, must have maximum row degree pn. The operation count is now overbounded by $0[p^5n^2]$.

Step 3: Find $U_2(z)$ so that

$$U_2(z)X_2(z)=\begin{bmatrix}h_{22} & x_{23}'' & \cdots & x_{2p}''\\ 0 & x_{33}'' & \cdots & x_{3p}''\\ \vdots & \vdots & & \\ 0 & x_{p3}'' & \cdots & x_{pp}''\end{bmatrix}=\begin{bmatrix}h_{22} & X_{23}''\\ 0 & X_{33}''\end{bmatrix}.$$

The entries of $U_2(z)$ may have degree pn and the x_{ij}'' may have degree $2pn$. There are

$$\begin{aligned}0\left[3p^2(pn)^2\right]&+0\left[p^2(pn)\log^2(pn)\right]\\&=0\left[3p^4n^2\right]+0\left[p^3n\log^2(pn)\right]\end{aligned}$$

operations, by Lemma B3, together with a lesser number for computing X_{23}'' and X_{33}''.

Step 4: Reduce X_{33}'' to row proper form, using Lemma B4. Note that X_{33}'' has maximum determinantal degree pn and maximum degree $2pn$. The operation count is $0[24(p-1)^3pn(pn+2pn)]=0[p^5n^2]$. Call the resulting row reduced matrix X_{33}. It has maximum determinantal degree pn and maximum row degree pn.

Steps of this nature through Step 2 $(p-1)$ are carried through (throwing away the unimodular matrices) until the following is obtained:

$$\begin{bmatrix}h_{11} & k_{12} & \cdots & k_{1p}\\ 0 & h_{22} & \cdots & k_{2p}\\ & 0 & & \\ \vdots & \vdots & & \vdots\\ 0 & 0 & & h_{pp}\end{bmatrix}.$$

(The total operation count to this point is $0[p^4n\log^2 pn]+0[p^6n^2]$.) Step $(2p-1)h_{22}$ is divided into k_{12} and k_{12} replaced by the remainder. Then h_{33} is divided into k_{13} and k_{23}, and each is replaced by its remainder. The process continues through with h_{pp} being divided into each of $k_{1p},\cdots,k_{p-1,p}$, followed by replacement of k_{ip} with the remainder. (These operations are equivalent to a series of subtractions of multiples of one row from another row, and could be achieved through premultiplication by a unimodular matrix. However, we do not compute this matrix.) The number of operations associated with each h_{ii} and

k_{ij} is overbounded by $0[p^2 n \log pn]$ and there $0(p^2)$ such operations. So this step is overbounded by $0[p^4 n \log pn]$ operations.

At this stage, the Hermite matrix $H(z)$ has been found.

Last, it remains to reconstruct the unimodular matrix $U(z)$ linking $X(z)$ to $H(z)$ via $X(z) = U(z)H(z)$. Observe that

$$\begin{bmatrix} u_{11} \\ \vdots \\ u_{p1} \end{bmatrix} = h_{11}^{-1} \begin{bmatrix} x_{11} \\ \vdots \\ x_{p1} \end{bmatrix}$$

$$\begin{bmatrix} u_{12} \\ \vdots \\ u_{p2} \end{bmatrix} = h_{22}^{-1} \left\{ \begin{bmatrix} x_{12} \\ \vdots \\ x_{p2} \end{bmatrix} - h_{12} \begin{bmatrix} u_{11} \\ \vdots \\ u_{p1} \end{bmatrix} \right\}$$

and so on. Since the u_{ij} are polynomials $\deg h_{11} \leqslant \deg[x_{i1}]$, $i = 1, 2, \cdots$, $p \leqslant n$. So the u_{j1} are found in $0[pn \log n]$ operations. Their degree is overbounded by n. Next, h_{12} has degree at most $(pn - 1)$. Hence, u_{j2} are found in $0[2p(pn-1)\log(pn-1)]$ operations, and their degree is also overbounded by n. More generally, for fixed i, the u_{ji} are found in $[0ip(pn-1)\log(pn-1)]$ operations and have degree overbounded by n. The total operation count is overbounded by $0[p^3 n \log(pn)]$ operations. Overall then, the operation count is bounded as described in the theorem statement, even when $U(z)$ is to be found.

In case the inverse of $U(z)$ is required, it can be found by applying the Hermite reduction algorithm to $[U(z) I]$ for which the Hermite form is $[IU^{-1}(z)]$. Since the degree of entries of $U(z)$ are bounded by n, the determination of $U^{-1}(z)$ will clearly be overbounded by $0[p^4 n \log^2(pn)] + 0[p^6 n^2]$—in fact it will be significantly less.

References

[1] N. Levinson, "The Wiener rms (root-mean square) error criterion in filter design and prediction," *J. Math Phys.*, vol. 25, pp. 261–278, 1947.

[2] P. Whittle, "On the fitting of multivariable autoregressions and the approximate canonical factorization of a spectral density matrix," *Biometrika*, vol. 50, pp. 129–134, 1963.

[3] R. A. Wiggins and E. A. Robinson, "Recursive solution to the multichannel filtering problem," *J. Geophys. Res.*, vol. 70, pp. 1885–1891, 1965.

[4] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, 1975.

[5] L. R. Rabiner and R. Schafer, *Digital Processing of Speech Signals.* Englewood Cliffs, NJ: Prentice-Hall, 1978.

[6] R. W. Schafer and L. R. Rabiner, "Digital representation of speech signals," *Proc. IEEE*, vol. 63, no. 4, pp. 662–677, 1975.

[7] A. H. Gray and J. D. Markel, "A normalized digital filter structure," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, no. 3, pp. 268–277, 1975.

[8] A. J. Berkout, "Stability and least square estimation," *Automatica*, vol. 11, pp. 637–638, 1975.

[9] A. Vieira and T. Kailath, "On another approach to the Schur–Cohn criterion," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 218–220, 1977.

[10] B. D. O. Anderson and R. R. Bitmead, "Stability of matrix polynomials," *Int. J. Contr.*, vol. 26, pp. 235–247, 1977.

[11] Y. Monden and S. Arimoto, "Generalized Rouché theorem and its application to multivariate autoregressions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 6, pp. 733–738, Dec. 1980.

[12] M. Morf, A. Vieira, and T. Kailath, "Covariance characterization by partial auto correlation matrices," *Annals Stat.*, vol. 6, no. 3, pp. 643–648, 1978.

[13] P. Delsarte, Y. V. Genin, and Y. G. Kamp, "Schur parametrization of positive definite block-Toeplitz systems," *SIAM J. Appl. Math.*, vol. 36, p. 34, 1979.

[14] M. T. Hadidi and M. Morf, "Efficient solution of the inverse Levinson problem in the multichannel case via the Lyapunov equation," in *Proc. 1981 Conf. Decision and Contr.*, San Diego, CA, Dec. 1981, pp. 1075–1076.

[15] M. T. Hadidi, M. Morf, and B. Porat, "Efficient construction of canonical ladder forms for vector autoregressive processes," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 1222–1233, 1982.

[16] B. Porat and M. Morf, "Efficient solution of Lyapunov equation for matrix autoregressive model and its applications to the inverse Levinson problem," in *Proc. 1981 Conf. Decision and Contr.*, San Diego, CA, Dec. 1981, pp. 1070–1074.

[17] B. D. O. Anderson and T. Kailath, "Forwards, backwards and dynamically reversible Markovian models of second-order processes," *IEEE Trans. Circuits Syst.*, vol. CAS-26, pp. 956–965, Nov. 1979.

[18] A. Lindquist and G. Picci, "On the stochastic realization problem," *SIAM J. Contr. Optimiz.*, vol. 17, pp. 361–389, May 1979.

[19] L. Ljung and T. Kailath, "Backwards Markovian models for second-order stochastic processes," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 488–491, July 1976.

[20] G. Verghese and T. Kailath, "A further note on backwards Markovian models," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 121–124, Jan. 1979; correction vol. IT-25, p. 501, July 1979.

[21] W. A. Wolovich, "Skew prime polynomial matrices," *IEEE Trans. Automat. Contr.*, vol. AC-23, no. 5, pp. 880–887, 1978.

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

[23] P. R. Motyka and J. A. Cadzow, "The factorization of discrete-process spectral matrices," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 698–706, 1967.

[24] M. Morf, A. Vieira, D. T. L. Lee, and T. Kailath, "Recursive multichannel maximum entropy spectral estimation," *IEEE Trans. Geosci. Electron.*, vol. GE-16, pp. 86–94, Apr. 1978.

[25] J. Rissanen and T. Kailath, "Partial realization of random systems," *Automatica*, vol. 8, pp. 389–396, 1972.

[26] R. R. Bitmead and B. D. O. Anderson, "Asymptotically fast solution of Toeplitz and related systems of linear equations," *Linear Algebra Appl.*, vol. 34, pp. 103–116, 1980.

[27] B. Friedlander, M. Morf, T. Kailath, and L. Ljung, "New inversion formulas for matrices classified in terms of their distances from Toeplitz matrices," *Linear Algebra Appl.*, vol. 27, pp. 31–60, 1979.

[28] I. C. Gohberg and G. Heinig, "Inversion of finite Toeplitz matrices with entries being element from a non-commutative algebra," *Rev Roumaine Math Pures Appl.*, vol. XIX, pp. 623–665, 1974.

[29] T. Kailath, A. Vieira, and M. Morf, "Inverses of Toeplitz operators, innovations, and orthogonal polynomials," *SIAM Rev.*, vol. 20, pp. 106–119, Jan. 1978.

[30] P. Delsarte, Y. V. Genin, and Y. G. Kamp, "Orthogonal polynomial matrices on the unit circle," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 3, pp. 149–160, 1978.

[31] D. C. Youla and N. N. Kazanjian, "Bauer-type factorization of positive definite matrices and the theory of matrix polynomials orthogonal on the unit circle," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 57–69, 1978.

[32] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

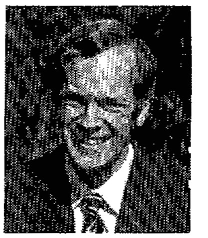
[33] F. G. Gustavson and D. Y. Y. Yun, "Fast algorithm for rational Hermite approximation and solution of Toeplitz systems," *IEEE Trans. Circuits Syst.*, vol. CAS-26, no. 9, pp. 750–755, Sept. 1979.

[34] M. Morf, "Doubling algorithms for Toeplitz and related equations," in *Proc. IEEE Conf. Acoustics, Speech, Signal Processing*, Denver, CO, 1980, pp. 954–959.

[35] G. Cybenko, "The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations," *SIAM J. Sci. Stat. Comput.*, vol. 1, no. 3, pp. 303–319, 1980.

[36] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms.* Reading, MA: Addison-Wesley, 1974.

[37] S. Kung and T. Kailath, "Fast projection methods for minimal design problem in linear system theory," *Automatica*, vol. 16, pp. 399–403, 1980.



Brian D. O. Anderson (S'62–M'66–SM'74–F'75) was born in Sydney, Australia, in 1941. He received the B.S. degrees in pure mathematics and electrical engineering from the University of Sydney, Sydney, Australia, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1966.

He is currently Professor and Head of the Department of Systems Engineering at the Australian National University, Canberra, Australia. From 1967 through 1981, he was Professor of Electrical Engineering at the University of Newcastle, New South Wales, Australia. He has also held appointments as a Visiting Professor at a number of universities in the United States, Australia, and Europe. He is coauthor of several books and his research interests are in control and signal processing.

Dr. Anderson is a Fellow of the Australian Academy of Science, Australian Academy of Technological Sciences, and the Institution of Engineers, Australia.

Ah Chung Tsoi (S'70–M'72–M'84) was born in Hong Kong in 1947. He received the Diploma of Technology in electrical engineering in 1969, the M.Sc. and the Ph.D. degrees in control engineering from the University of Salford, Salford, England, in 1970 and 1972, respectively, and the B.D. (extra-mural) degree in theology from Otago University, Dunedin, New Zealand, in 1980.

He is currently a Senior Lecturer in Electrical Engineering at the University of Auckland, Auckland, New Zealand. From 1972 to 1974, he was a Senior Research Fellow in Distributed Parameter Systems at the Inter-University Institute of Engineering Control, University College of North Wales, Bangor. From 1974 to 1977 he was a Lecturer at Paisley College of Technology, Scotland, and in 1977 he joined the University of Auckland. He has held visiting positions at the Department of Systems Engineering, Australian National University, Canberra. His research interests are in control and digital signal processing.

Dr. Tsoi is an associate member of the IEE, England.

Technical Notes and Correspondence

A LINPACK-Style Condition Estimator for the Equation $AX - XB^T = C$

RALPH BYERS

***Abstract*—Estimation of the condition number of $AX - XB^T = C$ requires an approximation of the smallest l_2 singular value of the linear transformation $X \to AX - XB^T$. The approximation is obtained from a step of inverse iteration applied to a heuristically chosen starting matrix. In the style of LINPACK, the cost of the estimator is kept down by using the same matrix factorizations to make the estimate as are used to solve $AX - XB^T = C$. Computational experiments verify its reliability.**

Introduction

Consider the equation

$$AX - XB^T = C \tag{1}$$

where $A \in R^{m\times m}$, $B \in R^{n\times n}$, $C \in R^{m\times n}$, and $X \in R^{m\times n}$. This is a system of mn equations in the mn unknowns x_{ij}. Expressed in standard matrix vector notation, it becomes

$$\begin{bmatrix} A - b_{11}I_m & -b_{12}I_m & \cdots & -b_{1n}I_m \\ -b_{21}I_m & A - b_{22}I_m & \cdots & -b_{2n}I_m \\ & \cdots & & \\ -b_{n1}I_m & -b_{n2}I_m & \cdots & A - b_{nn}I_m \end{bmatrix} \begin{bmatrix} X_{*1} \\ X_{*2} \\ \cdots \\ X_{*n} \end{bmatrix} = \begin{bmatrix} C_{*1} \\ C_{*2} \\ \cdots \\ C_{*n} \end{bmatrix}. \tag{2}$$

The notation I_m indicates the m-by-m identity matrix. The jth columns of X and C are represented by X_{*j} and C_{*j}. The mn-by-mn matrix on the left-hand side of (2) is the Kronecker sum $K = A \otimes I_m - I_n \otimes B$. For convenience, we will assume that A and B have distinct eigenvalues, so that K is nonsingular.

The sensitivity of the solution X to perturbations in the data A, B, and C is related to the condition number of (1)

$$K(A, B) = \|K\| \|K^{-1}\|.$$

Throughout this correspondence, we use the Euclidean norm $\|Z\| = [\text{trace}(Z^TZ)]^{(1/2)}$. A rough estimate of $K(A, B)$ is useful when assessing the accuracy of a computed solution X [3]. The LINPACK [7] condition estimator DGECO is not attractive to use with (1) because it would require explicit construction of the mn-by-mn matrix K and $O(m^3n^3)$ work. This is an excessive amount of work and storage compared to solving (1) by the method of [1]. This correspondence describes a LINPACK-like way to estimate $K(A, B)$ during the solution of (1) at little extra cost.

Condition Estimation

It is not difficult to calculate $\|K\|$ from (2) without forming an nm-by-nm array. Some care must be taken to avoid overflows and underflows. An estimate of $\|K^{-1}\|$ is more difficult.

Following [10], define the separation of $A \in R^{m\times m}$ and $B \in R^{n\times n}$ $\text{SEP}(A, B^T)$, by

$$\text{SEP}(A, B^T) = \min \|AZ - ZB^T\| / \|Z\|$$

where $Z \neq 0$ varies over $R^{m\times n}$. The l_2 norm of K^{-1} is $1/\text{SEP}(A, B^T)$, so

$$(mn)^{(-1/2)} \|K^{-1}\| \leqslant 1/\text{SEP}(A, B^T) \leqslant \|K^{-1}\|.$$

An approximation to $\text{SEP}(A, B^T)$ gives an approximation to $\|K^{-1}\|$.

The relation of $\text{SEP}(A, B^T)$ to (1) has been known for some time [5]. It also appears in the perturbation theory of algebraic Riccati equations [2] and invariant subspaces [9], [10].

Among the most successful solution methods for solving (1) are those of [1] and [6]. Both methods begin by using an orthogonal similarity transformation to reduce A and B to quasi-triangular form, i.e., to block-triangular form with 1-by-1 and 2-by-2 blocks along the diagonal. We have

$$\begin{aligned} A &:= Q^TAQ \\ B &:= U^TBU \end{aligned} \tag{3}$$

where $Q^{-1} = Q^T$, $U^{-1} = U^T$. It is easy to show [10] that (3) does not change $\text{SEP}(A, B^T)$.

For ease of explication, suppose for the moment that A and B are triangular. Then, K is an upper triangular mn-by-mn matrix.

Cline, Moler, Stewart, and Wilkinson [3] suggest that $\|K^{-1}\|$ be estimated by $\|y\|/\|z\|$ where y, z, and $w \in R^{nm}$ satisfy

$$\begin{aligned} K^Tz &= w \\ Ky &= z. \end{aligned} \tag{4}$$

Manuscript received October 12, 1983; revised May 21, 1984.

The author is with the Department of Mathematical Sciences, Northern Illinois University, DeKalb, IL 60115.

0018-9286/84/1000-0926$01.00 ©1984 IEEE