On the Partial Stochastic Realization Problem

Christopher I. Byrnes, *Fellow, IEEE*, and Anders Lindquist, *Fellow, IEEE*

***Abstract*— In this paper, we describe a complete parameterization of the solutions to the partial stochastic realization problem in terms of a nonstandard matrix Riccati equation. Our analysis of this covariance extension equation (CEE) is based on a recent complete parameterization of all strictly positive real solutions to the rational covariance extension problem, answering a conjecture due to Georgiou in the affirmative. We also compute the dimension of partial stochastic realizations in terms of the rank of the unique positive semidefinite solution to the CEE, yielding some insights into the structure of solutions to the minimal partial stochastic realization problem. By combining this parameterization with some of the classical approaches in partial realization theory, we are able to derive new existence and robustness results concerning the degrees of minimal stochastic partial realizations. As a corollary to these results, we note that, in sharp contrast with the deterministic case, there is no generic value of the degree of a minimal stochastic realization of partial covariance sequences of fixed length.**

***Index Terms*—Covariance extension, parameterization, partial stochastic realization, positive degree, positive real.**

I. Introduction

In SIGNAL processing and speech processing, a signal is often modeled as a stationary random sequence which is the output of a linear stochastic system obtained by passing white noise through a filter with a stable transfer function and letting the system come to a statistical steady state. For example, artificial speech is synthesized by a combination of two kinds of models. Voiced sounds can be produced by passing periodic signals through a deterministic filter, while unvoiced sounds are produced by passing white noise through a shaping filter so that on a sufficiently small interval of time the unvoiced speech pattern can be regarded as a realization of a stationary random sequence. Of course, in practice only a finite string of observed data is typically available for speech synthesis (as well as for any application), in which case only a finite covariance sequence can be produced.

The need to construct stochastic models from a finite window of correlation coefficients has led to the study of several problems involving the description of classes of stationary linear stochastic systems having outputs which match a given partial covariance sequence. One of these is the *partial stochastic realization problem*, which consists of describing all such stochastic systems having the smallest possible degree, which we refer to as the *positive degree* of the partial covariance sequence. Kalman motivated the study of the partial stochastic realization problem by describing minimal realizations as being the simplest class of models capable of describing the given data. Alternatively, the *maximum entropy filter* may be interpreted as maximizing some measure of the "entropy" of the covariance window and, in this way, assumes as little as possible about the completion of the correlation sequence. The maximum entropy filter may, or may not, be minimal, but it always has degree equal to the length, n, of the covariance window.

More generally, the problem of characterizing all stationary linear stochastic systems, of degree at most n, having outputs which match a given partial covariance sequence is known as the *rational covariance extension problem*. For example, the maximal entropy solution may be characterized as the unique solution for which there are no finite zeros of the corresponding spectral density. Since the spectral zeros have intrinsic importance in speech synthesis, and since the additional memory is required by a nonminimal, an nth-order filter is both relatively cheap and available. The conjecture of Georgiou that all solutions to the rational covariance extension problem can be parameterized in terms of the partial covariance data and a choice of spectral zeros provides an attractive complement to the problem of minimal partial stochastic realization. Using the recent verification of this conjecture [10] and an integration of the various classical approaches to the partial realization problem, in this paper we prove several new results about the basic problem of parameterizing rational models for partial covariance data.

In Section II, we describe the basic problems more explicitly and introduce a Riccati-type equation, called the *covariance extension equation* (CEE), which is formulated in terms of the partial covariance data and a choice of desired modeling-filter zeros. This is a nonstandard Riccati equation, the positive semidefinite solutions of which would parameterize the solution set of the rational covariance extension problem in terms of the partial covariance sequence and the zeros of the desired modeling filter. Our first main result, Theorem 2.1, asserts that there always exists a unique positive semidefinite solution. Moreover, the rank of this solution is precisely the degree of the corresponding shaping filter, giving refined bounds for the minimal partial realization problem.

The minimal partial stochastic realization problem has three facets. It is of course related, but not equivalent, to the usual classical minimal stochastic realization problem in which complete covariance data is given. The minimal partial stochastic realization problem is also related to the partial deterministic realization problem, which consists of describing all minimal, finite-dimensional deterministic systems having Hankel parameters which match a given partial covariance sequence. The minimal degree of an interpolating deterministic system

Manuscript received July 21, 1995. Recommended by Associate Editor, V. Solo. This work was supported in part by grants from AFOSR, NSF, TFR, the Royal Swedish Academy of Sciences, and Southwestern Bell.

C. I. Byrnes is with the Department of Systems Science and Mathematics, Washington University, St. Louis, MO 63130 USA.

A. Lindquist is with the Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden.

Publisher Item Identifier S 0018-9286(97)05950-3.

is sometimes referred to as the algebraic degree of the partial covariance sequence. Naturally, the deterministic problem relaxes the constraint that the transfer function be positive real. The positivity of interpolating functions has deep historical roots as well, going back to the classical Carathéodory extension problem, which involves the parameterization of all positive real meromorphic functions which match, or interpolate, a partial sequence of Laurent coefficients. Although each of these three problems has been completely solved separately, the interrelationship between them is quite complicated, a fact which has caused some confusion in the literature and in practice. For example (see [43]), under certain conditions some popular identification procedures have been known to fail because the existence of a generic value for the positive degree of a partial covariance sequence has been implicitly assumed, something which is true for the algebraic degree of a sequence.

Nonetheless, by combining the theories underlying these three problems with the parameterization of solutions to the rational covariance extension problem, we are able to develop related existence and robustness results about minimal stochastic partial realizations which yield some rather interesting insights into the properties of the positive degree of a partial covariance sequence. As an example, Theorem 2.2 asserts that, in sharp contrast to the algebraic degree of a sequence, for each integer n^* between $\frac{1}{2}n$ and n there is a nonempty open subset of partial covariance sequences with positive degree n^*.

As we have mentioned, the techniques used in this paper are an integration of traditional, and some recent, approaches to the related problems described above. In Section III, we briefly review the status of partial realization theory beginning with some historical observations about rationality due to Euler and Kronecker, observations which play an important part in our constructions later in the paper. In Appendix A, we give a brief summary of the solution to the Carathéodory extension problem in terms of the well-known *Schur parameters*, determined by the partial covariance sequence. In addition, we describe a fast filtering algorithm which, in fact, can be viewed as a nonlinear dynamical system which propagates the Schur parameters corresponding to rational interpolants. This point of view can provide estimates for the asymptotic behavior of those Schur parameters which correspond to rational solutions to the Carathéodory extension problem (see, e.g., [10]) and is also quite useful in our analysis of the properties of the positive degree. In Appendix B, as another prerequisite to our analysis of the positive degree, we briefly relate the construction of the classical resultant to that given by Kronecker in terms of determinants of Hankel matrices. When combined with some fundamental work by Brockett on the geometry of the partial deterministic realization problem, this is extremely useful in analyzing when the algebraic, and sometimes the positive, degree can be lower or higher than expected.

Section IV is devoted to the covariance extension equation and the properties of its solutions. Our principal result regarding the CEE concerns existence and uniqueness of the positive semidefinite solution, similar in spirit to the classical existence and uniqueness theorems for the Riccati equations arising in filtering and control, and the connection of this solution to the corresponding modeling filter (40). The existence of a positive semidefinite solution is, of course, of considerable independent interest in partial stochastic realization theory.

Finally, Section V is devoted to the question of minimality. In particular we prove Theorem 2.2, which has several interesting consequences. First, it implies that the positive degree of a partial covariance sequence has no generic value. Moreover, there is an open set of partial correlation sequences for which the minimal, positive degree is n, and the minimal stochastic partial realization problem and the rational covariance extension problem are equivalent. For such sequences, then, the minimal partial stochastic realizations are parameterized by the set of Schur polynomials, i.e., by the desired zeros of the corresponding minimum phase spectral factor. Finally, the general result allows one to recast the general partial stochastic realization problem into the problem of computing the positive degree and the problem of characterizing the structure of the set of spectral zeros which yield a minimal degree realization.

II. Main Results

In signal processing and speech processing [14], [18], [23], [33], [34], [45] a signal is often modeled as a stationary random sequence $\{y(t)\}_{t\in\mathbb{Z}}$ which is the output of a linear stochastic system

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ \quad y(t) = Cx(t) + Du(t) \end{cases} \tag{1}$$

obtained by passing (normalized) white noise $\{u(t)\}_{t\in\mathbb{Z}}$ through a filter

$$\text{white noise} \xrightarrow{u} \boxed{w(z)} \xrightarrow{y}$$

with a stable transfer function

$$w(z) = C(zI - A)^{-1}B + D \tag{2}$$

and letting the system come to a statistical steady state. Here stability amounts to the matrix A having all its eigenvalues strictly inside the unit circle.

Consequently, the stationary stochastic process $\{y(t)\}_{t\in\mathbb{Z}}$ is given by the convolution

$$y(t) = \sum_{k=-\infty}^{t} w_{t-k}u(k), \qquad t = 0, 1, 2, \cdots \tag{3}$$

where $w_0 = D$ and $w_k = CA^{k-1}B$ for $k = 1, 2, 3, \cdots$, and where

$$w(z) = w_0 + w_1z^{-1} + w_2z^{-2} + w_3z^{-3} + \cdots. \tag{4}$$

The process $\{y(t)\}_{t\in\mathbb{Z}}$ has a rational spectral density

$$\Phi(z) = w(z)w(z^{-1}) \tag{5}$$

which we assume to be positive on the unit circle. In other words, $w(z)$ is a stable spectral factor of Φ which we shall take to be *minimum phase*, i.e., the rational function $w(z)$ has all its poles and zeros in the open unit disc and $w_0 = w(\infty) \neq 0$. In system-theoretical language we say that y is the output of a *shaping filter* driven by a white noise input, with the transfer function w.

It is well known that the spectral density Φ has the Fourier representation

$$\Phi(z) = c_0 + \sum_{i=1}^{\infty} c_i(z^i + z^{-i}) \tag{6}$$

where

$$c_0, c_1, c_2, c_3, \cdots \tag{7}$$

is the covariance sequence defined as

$$c_k = \mathrm{E}\{y(t+k)y(t)\}, \qquad k = 0, 1, 2, 3, \cdots. \tag{8}$$

Such a covariance sequence has the property that the infinite Toeplitz matrix

$$T_\infty = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots \\ c_1 & c_0 & c_1 & \cdots \\ c_2 & c_1 & c_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{9}$$

is positive definite. Therefore, without loss of generality we can assume that $c_0 = 1$.

The corresponding stochastic realization problem is the inverse problem of determining the stochastic system (1) given the *infinite* covariance sequence (7). The condition that $\Phi(z)$ be rational introduces a finiteness condition on the covariance sequence (7). In fact, the positive real part

$$v(z) = \frac{c_0}{2} + \sum_{i=1}^{\infty} c_i z^{-i} \tag{10}$$

of

$$\Phi(z) = v(z) + v(z^{-1}) \tag{11}$$

is rational and may be written

$$v(z) = \frac{1}{2}\frac{b(z)}{a(z)} \tag{12}$$

where

$$a(z) = z^n + a_1 z^{n-1} + \cdots + a_n \tag{13}$$

and

$$b(z) = z^n + b_1 z^{n-1} + \cdots + b_n \tag{14}$$

are monic polynomials. The property that $v(z)$ be strictly positive real is equivalent to $a(z)$ and $b(z)$ being *Schur polynomials*, i.e., having all roots in the open unit disc and satisfying

$$a(z)b(z^{-1}) + a(z^{-1})b(z) > 0 \tag{15}$$

on the unit circle. Therefore, once $a(z)$ and $b(z)$ are known, the unique stable minimum-phase spectral factor of Φ, i.e., the solution

$$w(z) = \rho\frac{\sigma(z)}{a(z)} \tag{16}$$

of (5) where $\rho \in \mathbb{R}_+$ and $\sigma(z)$ is a monic Schur polynomial

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \cdots + \sigma_n \tag{17}$$

may be determined via the polynomial spectral factorization problem

$$\tfrac{1}{2}[a(z)b(z^{-1}) + a(z^{-1})b(z)] = \rho^2\sigma(z)\sigma(z^{-1}). \tag{18}$$

In fact, identifying coefficients of nonnegative powers in z in

$$2a(z)v(z) = b(z)$$

we obtain

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = 2\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} 1 & & & \\ 2c_1 & 1 & & \\ \vdots & \vdots & & \\ 2c_{n-1} & 2c_{n-2} & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}. \tag{19}$$

Likewise, identifying coefficients of negative powers in z, we have

$$c_{n+i} = -\sum_{j=1}^{n} c_{j+i-1}a_j, \qquad i = 1, 2, 3, \cdots \tag{20}$$

so that $a := (a_1, a_2, \cdots, a_n)'$ satisfies the Hankel system

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ c_2 & c_3 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = -\begin{bmatrix} c_{n+1} \\ c_{n+2} \\ \vdots \\ c_{2n} \end{bmatrix}. \tag{21}$$

Now, Kronecker's theorem implies that

$$\begin{aligned} n^* := \deg v(z) &= \operatorname{rank}\begin{bmatrix} c_1 & c_2 & c_3 & \cdots \\ c_2 & c_3 & c_4 & \cdots \\ c_3 & c_4 & c_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} c_1 & c_2 & \cdots & c_{n^*} \\ c_2 & c_3 & \cdots & c_{n^*+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n^*} & c_{n^*+1} & \cdots & c_{2n^*-1} \end{bmatrix}. \end{aligned} \tag{22}$$

Thus, taking $n = n^*$, (21) has a unique solution a which, inserted into (19), yields b. Consequently, the spectral factor (16) is completely determined by the partial covariance sequence $\{c_1, c_2, \cdots, c_{2n^*}\}$ or, alternatively, by $\{c_1, c_2, \cdots, c_{n^*}\}$ and a.

As an illustration from speech synthesis, recall that artificial speech is produced by a combination of two kinds of models, one kind for voiced sounds (such as vowels) and one kind for unvoiced sounds (for consonants such as "s" or "t"), the transfer functions of which vary on different small intervals of time. Voiced speech can be produced by passing periodic signals through a deterministic filter, while unvoiced signals can be produced by passing white noise through a shaping filter. On a sufficiently small interval of time, the unvoiced speech pattern can be regarded as a realization of a stationary random sequence y with covariances

$$c_k := E\{y(t+k)y(t)\} \tag{23}$$

where $E\{\cdot\}$ denotes mathematical expectation and with a spectral density

$$\Phi(z) = c_0 + \sum_{k=1}^{\infty} c_k(z^k + z^{-k}). \tag{24}$$

Given a (infinite) string of observed data

$$y_0, y_1, y_2, y_3, \cdots \tag{25}$$

satisfying a certain ergodicity property, the covariance sequence $(c_0, c_1, c_2, c_3, \cdots)$ can be determined as

$$c_k = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} y_{t+k}y_t \tag{26}$$

which defines a unique spectral density and hence a unique shaping filter.

However, in practice only a finite string of observed data

$$y_0, y_1, y_2, \cdots, y_N \tag{27}$$

is typically available for speech synthesis (as well as for most applications). If N is sufficient large, there is a $T < N$ such that

$$\frac{1}{T} \sum_{t=0}^{T} y_{t+k}y_t \tag{28}$$

is a good approximation of c_k, but now only a finite covariance sequence

$$c_0, c_1, c_2, \cdots, c_n \tag{29}$$

where $n \ll N$, can be produced.

The need to construct stochastic models from a finite window of correlation coefficients has led to the study of two fundamental, but related, problems: the *partial* stochastic realization problem and the rational covariance extension problem. Both problems begin with a partial list of correlation coefficients (29) with the property that the Toeplitz matrix

$$T_n = \begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & \cdots & c_0 \end{bmatrix} \tag{30}$$

is positive definite.

The partial stochastic realization problem involves both characterizing the minimal degree of a stochastic realization, which generates the given window (29), and parameterizing all such minimal realizations. Kalman motivated this problem by describing a minimal realization as being the simplest model for the given data.

Another interpretation of simplicity can be given in terms of the *maximum entropy filter* determined by the data (29). (See Section III for more details.) The maximum entropy filter may be interpreted as maximizing some measure of the "entropy" of the covariance window (29) and, in this way, assumes as little as possible about the completion of the correlation sequence.

The maximum entropy filter has degree n and thus may not be minimal. Nevertheless, its construction shows that there always exist models for (29) having degree less than or equal to n. Moreover, with the increased availability and cheaper cost of memory, since the time when the partial realization problem was originally formulated, the related problem of parameterizing all positive rational extensions having degree less than or equal to n becomes considerably more relevant. This is the rational covariance extension problem.

Of course, at a purely set-theoretic level, there are many *ad hoc* ways to parameterize a solution set, not all of which are so useful. For this reason, we shall formulate these two problems more precisely. Denote by $\mathcal{C}_n$ the subset of $\mathbb{R}^n$, consisting of partial covariance sequences (29). First, recalling that c_0 has been normalized to one, we will suppress c_0 from the partial covariance sequence c and denote by c the point $c = (c_1, c_2, \cdots, c_n)$ in $\mathcal{C}_n$. Next, let n^* be any integer satisfying $0 \leq n^* \leq n$. We define the subset, $S(n^*)$ of $\mathbb{R}^n$ as the set of those partial covariance sequences c having a minimal stochastic realization of degree n^*. Finally, we define the subset $\Sigma(n^*)$ of $\mathbb{R}^n$ as the set of those partial covariance sequences c having a minimal stochastic realization of degree less than or equal to n^*. For example, zero is contained in $\Sigma(n^*)$, for every n^* but is contained in $S(n^*)$ only for $n^* = 0$.

In this notation, the partial stochastic realization problem consists of two parts.

1) Describe $S(n^*)$ as explicitly as possible as a subset of $\mathbb{R}^n$.
2) For c in $S(n^*)$ parameterize those realizations of degree n^* in terms of familiar, or useful, system-theoretic objects.

Similarly, the basic mathematical problem underlying the rational covariance extension problem is the following.

3) Given a c in $\mathcal{C}_n$, in terms of familiar, or useful, system-theoretic objects, parameterize all infinite extensions

$$c_{n+1}, c_{n+2}, c_{n+3}, \cdots \tag{31}$$

of (29) such that

$$v(z) = c_0 + c_1z^{-1} + c_2z^{-2} + c_3z^{-3} + \cdots \tag{32}$$

defines a function which is

a) rational of degree at most n;
b) strictly positive real, i.e., analytic for $|z| > 1$ and satisfying the positivity condition

$$v(e^{i\theta}) + v(e^{-i\theta}) > 0, \qquad \text{for all } \theta \in [0, 2\pi). \tag{33}$$

To each such extension there is a unique *modeling filter*, i.e., the minimum phase spectral factor (16) of the spectral density (33).

We shall first illustrate what we mean by a parameterization in terms of "familiar, or useful, system-theoretic objects." Problem 3) combines two requirements, positivity and rationality. Such extension problems have a long history. Suppressing rationality, we obtain the *Carathéodory extension problem*, i.e., the problem of finding all positive real functions v, analytic outside the unit disc, which satisfy (32). This was posed by Carathéodory and was solved by Schur in terms of an associated sequence of parameters, equivalent to (29) and now known as the *Schur parameters*; see Appendix A. However, the basic question of which Schur sequences correspond to rational solutions remains open.

On the other hand, dropping the positivity Condition a), one obtains another well-known problem, namely the *partial realization problem*, which is presented in more detail in Section III. The partial realization problem without positivity has been extensively studied, and there exist explicit parameterizations of the set of rational, but not necessarily positive real, functions which satisfy (32). Thus, in contrast to the Schur parameterization, such parameterizations guarantee that v will be rational of degree at most n but leave open the rather challenging problem of characterizing positivity in terms of the remaining parameters.

In this setting, there was a long-standing conjecture of Georgiou [23] that, for any desired choice of spectral density zero structure, there is one and only one positive extension, i.e., one and only one modeling filter. The existence question had already been settled by Georgiou in [23]. In [10], we not only proved injectivity, but also that the bijection is actually a diffeomorphism and that the problem of determining the appropriate modeling filter is well posed. This result was obtained by viewing a certain fast filtering algorithm as a nonlinear dynamical system defined on the space of positive real rational functions of degree less than or equal to n. It is observed that filtering and interpolation induce complementary, or "dual" decompositions (or foliations), of this space. From this assertion about the geometry of positive real functions follows a result [10] which itself answers Georgiou's conjecture in the affirmative and provides the first complete parameterization of all positive rational extensions.

This solution to the rational covariance extension problem expresses the choice of free parameters in familiar systems theoretic terms, viz. the numerator of the resulting modeling filter. While the numerator can be any Schur polynomial, the resulting pole polynomial, which is determined by this choice of zeros and by the interpolation conditions, must be obtained by solving a system of nonlinear equations which gives little *a priori* information about the degree of the resulting realization. As we shall demonstrate in this paper, it turns out that these parameters can also be expressed in terms of a new Riccati-type equation which we shall call the CEE [8], [9]. Thus, in the *partial* stochastic realization problem, the algebraic Riccati equation (87) of stochastic realization theory needs to be replaced by a nonstandard, quadratic matrix equation of another type, containing certain indefinite terms. Moreover, the rank of the unique positive semidefinite solution of the CEE is the degree of the associated partial stochastic realization.

In this problem the given covariance data may also be represented in terms of the first n coefficients obtained from the expansion

$$\frac{z^n}{z^n + c_1 z^{n-1} + \cdots + c_n} = 1 - u_1 z^{-1} - u_2 z^{-2} - u_3 z^{-3} - \cdots \tag{34}$$

about infinity, in terms of which we define

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \qquad U = \begin{bmatrix} 0 & & & & \\ u_1 & 0 & & & \\ u_2 & u_1 & & & \\ \vdots & \vdots & \ddots & & \\ u_{n-1} & u_{n-2} & \cdots & u_1 & 0 \end{bmatrix}. \tag{35}$$

Next, for any Schur polynomial

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \cdots + \sigma_n \tag{36}$$

we define

$$\sigma = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_n \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} -\sigma_1 & 1 & 0 & \cdots & 0 \\ -\sigma_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\sigma_{n-1} & 0 & 0 & \cdots & 1 \\ -\sigma_n & 0 & 0 & \cdots & 0 \end{bmatrix}$$

and

$$h = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{37}$$

We can now formulate a nonstandard Riccati equation which, as we shall see below, parameterizes solutions to the rational covariance extension problem in terms of the partial covariance sequence and the auxiliary Schur polynomial σ corresponding to desired zeros. This CEE has the form

$$P = \Gamma(P - Phh'P)\Gamma' + g(P)g(P)' \tag{38}$$

where $'$ denotes transposition and the function g: $\mathbb{R}^{n\times n} \to \mathbb{R}^n$ is defined as

$$g(P) = u + U\sigma + U\Gamma Ph. \tag{39}$$

Our principal result concerning the CEE concerns the existence and uniqueness of the positive semidefinite solution, similar in spirit to the classical existence and uniqueness theorems for the Riccati equations arising in filtering and control and the connection of this solution to the corresponding modeling filter (40). This, of course, is of considerable independent interest in partial stochastic realization theory.

Theorem 2.1: Let $(1, c_1, \cdots, c_n)$ be a given positive partial covariance sequence. For every Schur polynomial $\sigma(z)$, there exists a unique positive semidefinite solution P of the CEE satisfying $h'Ph < 1$, to which in turn there corresponds a unique modeling filter

$$w(z) = \rho \frac{\sigma(z)}{a(z)} \tag{40}$$

for which the denominator polynomial

$$a(z) = z^n + a_1 z^{n-1} + \cdots + a_n \tag{41}$$

is given by

$$a = (I - U)(\Gamma Ph + \sigma) - u \tag{42}$$

and $\rho \in (0, 1]$ is a real number given by

$$\rho = (1 - h'Ph)^{1/2}. \tag{43}$$

All modeling filters are obtained in this way. Moreover, the degree of $w(z)$, and hence that of $v(z)$, equals the rank of P.

The results stated so far relate to the parameterization aspects of our problems for a fixed c in $\mathcal{C}_n$. While memory constraints might not make the minimality of the degree as essential, we will now turn to Problem 1) since it is also the case that certain matrices which are ubiquitous in linear

systems theory will become singular, or that certain numerical algorithms may become ill-conditioned, when computed for nonminimal realizations. As we shall show in Section V, there is a sharp contrast between Problem 1) and the corresponding problem for deterministic partial realization theory; e.g., for $n \geq 2$ there is no generic value of the minimal degree $n^*(c)$ for $c \in \mathcal{C}_n$. In fact, even $S(n)$ has a nonempty interior. Existence results such as these follow from the general existence results inherent in the solution of the rational covariance extension problem and provide for the use of geometric approach to describing properties of the level sets of the function $n^*(c)$.

Recall that a subset of $\mathbb{R}^n$ is semialgebraic, provided it can be defined by a finite number of polynomial equations, inequations, and inequalities. For example, $\mathcal{C}_n$ is a semialgebraic subset of $\mathbb{R}^n$, being defined by polynomial inequalities. A subset of $\mathbb{R}^n$ is algebraic, provided it can be defined by a finite number of polynomial equations. Finally, a property of points in $\mathbb{R}^n$ is said to be *generic* if the set of points which enjoy this property is nonempty, with its complement being contained in an algebraic set.

Theorem 2.2: Let n^* be any integer satisfying $0 \leq n^* \leq n$. Then the subsets $S(n^*)$ of $\mathbb{R}^n$, consisting of partial covariance sequences $(c_1, c_2, \cdots, c_n)$ having a minimal stochastic realization of degree n^*, are a nonempty semialgebraic set. The subset $\Sigma(n^*)$ of those partial covariance sequences c having a minimal stochastic realization of degree less than or equal to n^* is also semialgebraic. Moreover, $S(n^*)$ and $\Sigma(n^*)$ have nonempty interiors if and only if $\frac{1}{2}n \leq n^* \leq n$.

Further results concerning these sets and their properties can be found in Section V. However, we remark that Theorem 2.2 has several interesting consequences. First, it implies that the *positive degree* of a partial covariance sequence $(1, c_1, \cdots, c_n)$, i.e., the minimal dimension of any partial stochastic realization of $(1, c_1, \cdots, c_n)$, has no generic value. Moreover, there is an open set of partial correlation sequences for which the positive degree is n and for which, therefore, Problems 2) and 3) are equivalent. For such sequences, then, the minimal partial stochastic realizations are parameterized by the set of Schur polynomials. Alternatively, the minimal partial stochastic realizations of such sequences are in one-to-one correspondence with an arbitrary choice of zeros for the associated minimum phase spectral factor. Finally, the general result allows one to recast the general partial stochastic realization problem into Problem 1) and the problem of characterizing the structure of the set of spectral zeros which yield a minimal degree realization.

III. A Review of Partial Realization Theory

One approach to the partial stochastic realization problem is to suppress rationality and to first obtain the solutions to the *Carathéodory extension problem*, the problem of finding all meromorphic positive real functions which interpolate the first n Laurent coefficients. As described in Appendix A, these functions can be parameterized in terms of an associated sequence of Schur parameters $(\gamma_0, \gamma_1, \gamma_2, \cdots)$, which are equivalent to the correlation coefficients. Not surprisingly, characterizing which Schur sequences correspond to rational solutions is apparently quite difficult.

While the Schur parameters and an associated family of orthogonal polynomials, the Szegö polynomials, are nevertheless very useful in constructing rational modeling filters, rationality is so central to the circle of problems considered in this paper that it is appropriate and useful to begin a review of partial realization theory with the question: When is a proper meromorphic function v on $\mathbb{C}$ rational? By proper, we mean that v takes on a finite value at infinity. In particular, by replacing v by f, where $f(z) = v(z) - v(\infty)$, we may assume that the meromorphic function f is strictly proper; i.e., that f vanishes at infinity. There are no doubt several classical approaches to determining whether such an f is rational. In this section, we will rely on two: the method of continued fractions pioneered by Euler and a technique involving quadratic forms developed by Kronecker in his study of the elimination theory of two or more polynomials [35].

In his first published work on continued fractions, Euler [20] studied the very basic question as to whether the number e was rational, appreciating that rationality of a real number α would be equivalent to the finiteness of the continued fraction expansion

$$\alpha = n + \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_2 + \cfrac{1}{\alpha_3 + \cdots}}}. \tag{44}$$

In fact, Euler shows that for $\alpha = e$ one has

$$(\alpha_1, \alpha_2, \alpha_3, \cdots) = (1, 2, 1, 1, 4, 1, 1, 6, \cdots). \tag{45}$$

Euler's proof was actually based on expressing the proper meromorphic function (on the punctured complex plane) $v(z) = e^{1/z}$ as a continued fraction, obtaining instead of the constants α_i a sequence of polynomial functions $\alpha_i(z)$ which he proposes to evaluate at $z = 1$. In fact, Euler computes these functions by first computing the numerators and denominators of the partial sums of the continued fraction, for which he finds a two-dimensional linear differential recurrence equation which is solved (as we would today) in terms of an associated Riccati equation [20]. This remarkable method also gives, of course, a proof that $v(z)$ is irrational.

Following Kronecker, we begin by fashioning the infinite (Hankel) matrix

$$H_f = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots \\ c_2 & c_3 & c_4 & \cdots \\ c_3 & c_4 & c_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{46}$$

from the Laurent series of f

$$f(z) = c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3} + \cdots. \tag{47}$$

There are two principal points we shall need to review here about this construction. We shall later make use of its relationship to the resultant of two polynomials, also discovered by Kronecker, in Section V.

Kronecker's first, now widely appreciated, observation was that if f were in fact rational, say

$$f(z) = \frac{g(z)}{a(z)} \tag{48}$$

then by multiplying each side of (47) by $a(z)$ one obtains a recurrence of length $\deg a(z)$ among the Laurent coefficients of f. Therefore, the $(\deg a(z)+j)$th column of H_f is linearly dependent on the preceding $\deg a$ columns.

Conversely, Kronecker's theorem asserts that f is rational if and only if $\operatorname{rank} H_f$ is finite. We need to phrase this observation more carefully for an analysis of the partial realization theorem. More precisely, then, for any infinite sequence

$$c_1, c_2, c_3, \cdots \tag{49}$$

of real numbers, consider the family

$$H_{ij} = \begin{bmatrix} c_1 & c_2 & \cdots & c_j \\ c_2 & c_3 & \cdots & c_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_i & c_{i+1} & \cdots & c_{i+j-1} \end{bmatrix} \tag{50}$$

of rectangular Hankel matrices. Define the *degree indexes* $n_0, n_1, n_2, \cdots$ of (49) in the following way. Set $n_0 = 0$, and, for $k = 1, 2, 3, \cdots$, let n_k be the smallest integer greater than n_{k-1} such that $H_{n_{k-1}+1, n_k}$ has full rank. According to Kronecker's theorem, the Laurent series (47) defines a rational function of degree n^* if and only if (49) has a finite number of degree indexes, n^* being the largest. In this case

$$\begin{aligned} n^* &= \operatorname{rank} H_f \\ &= \operatorname{rank} \begin{bmatrix} c_1 & c_2 & \cdots & c_{n^*} \\ c_2 & c_3 & \cdots & c_{n^*+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n^*} & c_{n^*+1} & \cdots & c_{2n^*-1} \end{bmatrix} \end{aligned} \tag{51}$$

and there are matrices (A, B, C) of dimensions $n^* \times n^*$, $n^* \times 1$, and $1 \times n^*$, respectively, such that

$$CA^{k-1}B = c_k, \qquad \text{for } k = 1, 2, 3, \cdots \tag{52}$$

so that $C(zI-A)^{-1}B$ is the minimal realization of the rational function f defined by (47). We call n^* the McMillan degree or the *algebraic degree* of (49).

Next, following Kalman [31], consider the problem of finding an infinite extension of a finite sequence

$$c_1, c_2, c_3, \cdots, c_m \tag{53}$$

having the smallest possible algebraic degree, i.e., such that (47) is a rational function of smallest degree. This is the *partial realization problem* [26], [31], [32], [49], and the infinite sequence, which may or may not be unique, is called a *minimal rational extension* of (53). The degree indexes of a finite sequence (53) are constructed precisely as for an infinite one, with the exception that the process stops when there are no more data. Clearly (53) has the same degree indexes as any of its minimal rational extensions.

In order to compute the minimal algebraic degree of the partial sequence (53), and to parameterize those minimal partial realizations, it is convenient to turn to the method of continued fractions for determining whether a strictly proper meromorphic function f is rational. Just as in Euler's treatise on the irrationality of e in the application of continued fractions to the partial realization problem, the equivalence between rationality and finiteness and the computation of the fractional partial sums play an equally important role [26], [32]. In fact, let $f_0, f_1, f_2, \cdots, f_\nu$ be a sequence of strictly proper rational functions defined recursively in the following way. Given f_{k-1}, apply the Euclidean division algorithm to obtain

$$\frac{\beta_k}{f_{k-1}(z)} = \pi_k(z) + f_k(z) \tag{54}$$

where β_k is a normalizing coefficient chosen so that $\pi_k(z)$ is a monic polynomial

$$\pi_k(z) = z^{d_k} - \pi_{k1} z^{d_k-1} - \cdots - \pi_{kd_k} \tag{55}$$

and $f_{k+1}(z)$ is the remainder in the form of a strictly proper rational function. Now, taking $f_0 = f$, it can be shown that $f_{\nu+1} = 0$ for some finite ν if and only if f is rational. Then a simple calculation gives the following continued fraction expansion:

$$f(z) = \cfrac{\beta_1}{\pi_1(z) - \cfrac{\beta_2}{\pi_2(z) - \cfrac{\beta_3}{\pi_3(z) - \cdots}}} \tag{56}$$

with $f_{\nu+1} = 0$. This is the *principal part continued fraction* of Magnus [44].

It was pointed out in [32] and further elaborated upon in [26] that the family of minimal rational extensions can be parameterized via such a finite continued fraction expansion. Indeed, the degree indexes of the sequence (49) or, equivalently, that of (47), is then given by the recursion

$$n_k = n_{k-1} + d_k, \qquad n_0 = 0 \tag{57}$$

and the algebraic degree of these sequences is $n^* = n_\nu$.

This suggests [26] that the class of rational functions (48) is parameterized by the sequence

$$\rho := (\rho_1, \rho_2, \cdots, \rho_m) = (s_1, s_2, \cdots, s_\nu, 0, \cdots, 0) \tag{58}$$

of m real numbers, where

$$s_k = (0, \cdots, 0, \beta_k, \pi_{k1}, \cdots, \pi_{kd_k}) \tag{59}$$

is a subsequence of $2d_k$ parameters, the first $d_k - 1$ being all zero. We shall call s_k the *kth section* of the parameter sequence ρ, and the corresponding subsequence of c is the kth section of c. More exactly, ρ has the form (58) in the case $2n_\nu \leq m$. If $2n_\nu > m$, the last section, s_ν, will not be completely filled, so one or several of the parameters $\pi_{n_\nu 1}, \pi_{n_\nu 2}, \cdots, \pi_{n_\nu d_\nu}$ will be arbitrary and will not appear in the parameter sequence ρ. Whenever $2n_\nu < m$, there are $m - 2n_\nu$ zeros after the last section s_ν, indicating that the last $m - 2n_\nu$ elements in the sequence (53) are automatically matched. A "generic section" has the form (β_k, ϕ_{k1}), consisting of only two parameters. The following is a statement of [26, Th. 6].

Theorem 3.1 [26]: The function f: $\mathbb{R}^m \to \mathbb{R}^m$ sending the sequence (53) to the parameter sequence (58) is a bijection. Moreover, if $c, \hat{c} \in \mathbb{R}^m$, $\rho = f(c)$, and $\hat{\rho} = f(\hat{c})$, then $c_i = \hat{c}_i$ for $i = 1, 2, \cdots, k < m$ if and only if $\rho_i = \hat{\rho}_i$ for $i = 1, 2, \cdots, k$.

The proof of this theorem relies on a particular form of the converse of the continued fraction expansion (54), which will also play an important role in some of the constructions needed for our main results. Namely, as shown in [26], the rational function (48) can be reconstructed from $\beta_1, \beta_2, \cdots, \beta_\nu$ and $\pi_1(z), \pi_2(z), \cdots, \pi_\nu(z)$ via the three-term recursion

$$\begin{cases} P_{k+1}(z) = \pi_{k+1}(z)P_k(z) - \beta_{k+1}P_{k-1}(z), \\ \qquad P_0 = 0,\ P_{-1} = -1, \\ Q_{k+1}(z) = \pi_{k+1}(z)Q_k(z) - \beta_{k+1}Q_{k-1}(z), \\ \qquad Q_1 = 1,\ Q_{-1} = 0 \end{cases} \tag{60}$$

where the polynomials P_k, Q_k are actually the *Lanczos polynomials* used in block tridiagonalization [37]. In fact, the polynomials $g(z)$ and $a(z)$ are given by

$$g(z) = P_\nu(z) \quad \text{and} \quad a(z) = Q_\nu(z) \tag{61}$$

which are coprime polynomials [26].

Returning to the partial stochastic realization problem, recall that in addition to the rationality requirement on the interpolating filter which we have just discussed, it is also required that the rational filter be positive real. The additional difficulty imposed by positivity can be illustrated by comparing the algebraic and positive degrees as c varies over the set $\mathcal{C}_n$ of positive sequences. For c in $\mathcal{C}_n$, it is perhaps tempting to proceed as above by taking n^* to be the algebraic degree r of the partial covariance sequence (29), i.e., the maximum rank of the Hankel matrices H_{ij} with $i+j = n+1$, which, as pointed out in [42] and [43], is common in applications to identification [3], [51]. For this algebraic partial realization problem, it is well known [5] that the algebraic degree has the generic value of r if $n = 2r$, or if $n = 2r-1$. However, this will in general not lead to a positive real $v(z)$, or even a stable $a(z)$ for that matter; cf., [6]. Indeed, by Theorem 2.2, the smallest degree p that will preserve the positive realness of $v(z)$, the *positive degree* of (29), can be any integer between zero and n. Moreover, in contrast to the purely algebraic problem, for each $\nu = [(n+1)/2], [(n+1)/2]+1, \cdots, n$ there is even an open subset of the n-dimensional set of covariance data for which $p = \nu$. In fact, as we shall show in Section V, examples of interior points in $S(n)$ are given by the maximum entropy filter for partial covariance sequences satisfying $c_i = 0$, for $i = 1, \cdots n-1$ and $c_n \neq 0$.

In general, for any c in $\mathcal{C}_n$ the maximum entropy filter is obtained by setting

$$\gamma_i = 0, \qquad \text{for } i = n, n+1, n+2, \cdots.$$

This corresponds to taking $a(z) = \varphi_n(z)$ and $b(z) = \psi_n(z)$, where $\{\varphi_t(z)\}$ and $\{\psi_t(z)\}$ are the Szegö polynomials of the first and second kind, respectively, as defined in Appendix A. It can be shown that

$$\frac{1}{2}\frac{\psi_n(z)}{\varphi_n(z)} = \frac{1}{2} + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_n z^{-n} + \cdots \tag{62}$$

and that

$$\varphi_n(z)\psi_n(z^{-1}) + \varphi_n(z^{-1})\psi_n(z) = r_n > 0. \tag{63}$$

Consequently, since $\varphi_n(z)$ and $\psi_n(z)$ are Schur polynomials

$$v(z) = \frac{1}{2}\frac{\psi_n(z)}{\varphi_n(z)} \tag{64}$$

is strictly positive real and

$$v(z) + v(z^{-1}) = \frac{r_n}{\varphi_n(z)\varphi_n(z^{-1})} \tag{65}$$

yielding the modeling filter

$$w(z) = \frac{\sqrt{r_n}z^n}{\varphi_n(z)}. \tag{66}$$

This is the *maximum entropy solution*, which in general has the property that the corresponding spectral density (65) lacks finite zeros [28]. However, in many applications, such as the speech processing example described above, zeros are desired, and the question arises whether it is possible to assign zeros arbitrarily and still satisfy (24). To this end Georgiou [23] and Kimura [34] observed that the formula (62) could be generalized to

$$v(z) = \frac{1}{2}\frac{\psi_n(z) + \alpha_1\psi_{n-1}(z) + \cdots + \alpha_n\psi_0(z)}{\varphi_n(z) + \alpha_1\varphi_{n-1}(z) + \cdots + \alpha_n\varphi_0(z)} \tag{67}$$

$$= \tfrac{1}{2} + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_n z^{-n} + \cdots \tag{68}$$

thus expressing $v(z)$ in terms of the $2n$ parameters (α, γ), where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)' \in \mathbb{R}^n$ and $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{n-1})' \in \mathbb{R}^n$. In fact, it was shown in [23] and also later in [12] that the interpolation condition in (67) holds for all α, and consequently the *Kimura–Georgiou parameterization* (67) characterizes rationality but not positivity. We denote by $\mathcal{P}_n$ the subset of $\mathbb{R}^{2n}$ for which $v(z)$ is strictly positive real and let

$$\mathcal{P}_n(\gamma) = \{(\alpha, \gamma) \in \mathcal{P}_n \mid \gamma \text{ fixed}\} \subset \mathbb{R}^n$$

be the positive real region for fixed covariance data.

Given the fact that there is an open set of partial correlation sequences c which have the maximum positive degree, n, possible, and that memory is both available and cheap, the Kimura–Georgiou parameterization is an attractive parameterization of those deterministic partial realizations of degree n which should also play an important role in the stochastic partial realization theorem. Of course, given the partial covariance data γ, the choice $\alpha = 0$ is the maximum entropy solution, but in general it is very complicated to characterize those other α for which $v(z)$ is positive real.

To illustrate our point let us give some low-dimensional examples. For $n = 1$ the representation (67) takes the form

$$v(z) = \frac{1}{2}\frac{z + \gamma_0 + \alpha_1}{z - \gamma_0 + \alpha_1}.$$

The strictly positive real region is the diamond depicted in Fig. 1, and fixing the partial covariance data γ_0, the admissible α are the ones on the open interval in the figure.

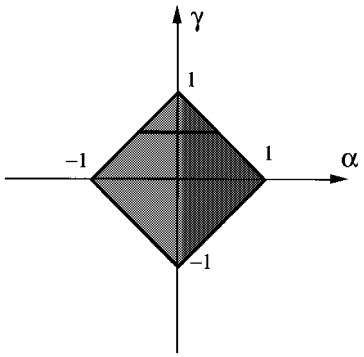


Fig. 1. The positive region P_1; the marked interval is $P_1(\frac{3}{5})$.

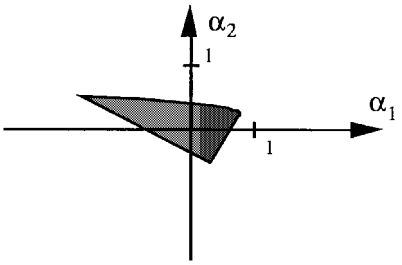


Fig. 2. The positive region $P_2(\frac{1}{2}, \frac{1}{3})$.

Next, let us consider the case $n = 2$. Fixing the covariance data at $\gamma_0 = \frac{1}{2}$ and $\gamma_1 = \frac{1}{3}$, we obtain

$$v(z) = \frac{z^2 - \frac{2}{3}z + \frac{1}{3} + \alpha_1(z + \frac{1}{2}) + \alpha_2}{z^2 + \frac{1}{3}z + \frac{1}{3} + \alpha_1(z - \frac{1}{2}) + \alpha_2}$$

and the region of positive real $\alpha = (\alpha_1, \alpha_2)$ is as depicted in Fig. 2.

The higher-dimensional cases become much more complicated. While it is true that $\mathcal{P}_n(\gamma)$ is always diffeomorphic to Euclidean space [7], any good solution to Problem 2) or 3) would give a parameterization in terms of familiar system theoretic objects. In this direction, the possibility of parameterizing those filters which are positive real by arbitrarily prescribing the zeros of a modeling filter was suggested earlier by Georgiou [23]. Indeed, using a very innovative application of topological degree theory Georgiou proved that to each choice of zeros there corresponds some modeling filter.

Recently we proved an amplification of a long-standing conjecture of Georgiou that, for any desired choice of spectral density zero structure, there is one and only one positive extension, i.e., one and only one modeling filter. This result was obtained by viewing a certain fast filtering algorithm as a nonlinear dynamical system defined on the space of positive real rational functions of degree less than or equal to n. It is then observed that filtering and interpolation induce complementary, or "dual" decompositions (or foliations) of this space. From this assertion about the geometry of positive real functions follows a result [10] which itself answers Georgiou's conjecture in the affirmative and provides the first complete parameterization of all positive rational extensions.

Theorem 3.2 [10]: Let $(1, c_1, \cdots, c_n)$ be a given positive partial covariance sequence. Then given any Schur polynomial

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \cdots + \sigma_n$$

there exists a unique monic Schur polynomial $a(z)$ of degree n and a unique $\rho \in (0, 1]$ such that

$$w(z) = \rho \frac{\sigma(z)}{a(z)}$$

is a minimum phase spectral factor of a spectral density $\Phi(z)$ satisfying

$$\Phi(z) = 1 + \sum_{i=1}^{\infty} \hat{c}_i(z^i + z^{-i}), \qquad \hat{c}_i = c_i \text{ for } i = 1, 2, \cdots, n.$$

In particular, the solutions of the rational positive extension problem are in one-to-one correspondence with self-conjugate sets of n points (counted with multiplicity) lying in the open unit disc, i.e., with all possible zero structures of modeling filters. Moreover, the modeling filter $w(z)$ depends analytically on the covariance data and the choice of zeros of the spectral density.

As an example, Fig. 3 depicts the connected open submanifolds $\mathcal{P}_2(\gamma)$ and $\mathcal{S}_2$, consisting of the monic polynomials in $\mathcal{S}_2$, for $\gamma = (\frac{1}{2}, \frac{1}{3})$. These sets form the domain and codomain of the diffeomorphism, described in Theorem 3.2, sending α to σ. Theorem 3.2 states that to any point σ in $\mathcal{S}_2$, there is one and only one α such that $(\alpha, \gamma) \in \mathcal{P}_2(\gamma)$. This α defines a modeling filter $w(z)$ having the zeros of $\sigma(z)$. Conversely, any α such that $(\alpha, \gamma) \in \mathcal{P}_2(\gamma)$ determines a Schur polynomial $\sigma(z)$. We remark that $\sigma(z)$ can also be computed via the convergence of the dynamical system (155) with the initial condition determined by (α, γ); see Appendix A.

IV. The Covariance Extension Equation

Recall that the problem under consideration is as follows. Given a partial covariance sequence $\{1, c_1, \cdots, c_n\}$ and a monic stable polynomial σ, representing the required zeros, find monic Schur polynomials $a(z)$ and $b(z)$ such that:

1) the rational function

$$v(z) = \frac{1}{2}\frac{b(z)}{a(z)} \tag{69}$$

with Laurent expansion

$$v(z) = \tfrac{1}{2} + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_n z^{-n} + \cdots \tag{70}$$

about infinity satisfies the interpolation condition

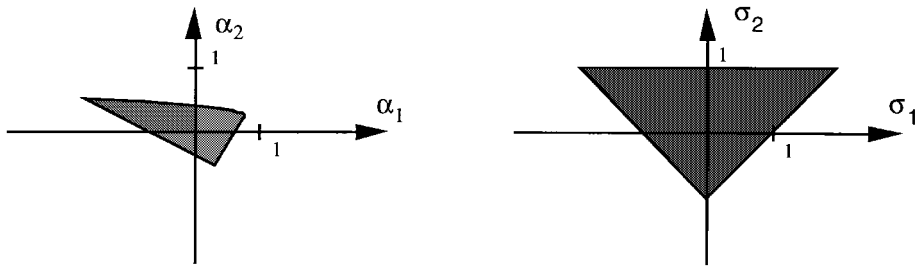
$$\hat{c}_i = c_i, \qquad \text{for } i = 1, 2, \cdots, n; \tag{71}$$

2)

$$\tfrac{1}{2}[a(z)b(z^{-1}) + a(z^{-1})b(z)] = \rho^2 \sigma(z)\sigma(z^{-1}). \tag{72}$$

for some positive real number ρ.

We shall first relax the problem by temporarily dropping the requirement that $a(z)$ and $b(z)$ both be Schur polynomials. In fact, our first result parameterizes the set of all pairs $(a(z), b(z))$ of monic, not necessarily Schur, polynomials in terms of symmetric solutions of the CEE (38).



(a) (b)

Fig. 3. The domain and codomain of the diffeomorphism of Theorem 3.2 in the case $n = 2$ and $\gamma = (\frac{1}{2}, \frac{1}{3})$.

Theorem 4.1: There is a one-to-one correspondence between symmetric solutions P of the CEE (38) such that $h'Ph < 1$ and pairs of monic polynomials

$$a(z) = z^n + a_1 z^{n-1} + \cdots + a_n \tag{73}$$

$$b(z) = z^n + b_1 z^{n-1} + \cdots + b_n \tag{74}$$

satisfying the interpolation condition (71) and the positivity condition (72). Under this correspondence

$$a = (I - U)(\Gamma Ph + \sigma) - u \tag{75}$$

$$b = (I + U)(\Gamma Ph + \sigma) + u \tag{76}$$

$$\rho = (1 - h'Ph)^{1/2} \tag{77}$$

and P is the unique solution of the Lyapunov equation

$$P = JPJ' - \tfrac{1}{2}(ab' + ba') + \rho\sigma\sigma' \tag{78}$$

where

$$J = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{79}$$

is the upward shift matrix. Moreover, the following conditions are equivalent.

1) $P \geq 0$.
2) $a(z)$ is a Schur polynomial.
3) $b(z)$ is a Schur polynomial.

If they are fulfilled

$$\deg v(z) = \operatorname{rank} P. \tag{80}$$

The proof of this theorem, as well as the first of its corollaries, will be deferred to the end of the section. The following example of polynomial spectral factorization illustrates Theorem 4.1.

Example 4.2: Let us consider the case

$$c_1 = c_2 = \cdots = c_n = 0.$$

Then, in view of (19), $a = b$ in (72) so we must have $v(z) = \frac{1}{2}$. Consequently, (72) yields

$$\rho^{-2} a(z)a(z^{-1}) = \sigma(z)\sigma(z^{-1}). \tag{81}$$

Moreover, $u = 0$ and $U = 0$ so the covariance extension equation (38) becomes

$$P = \Gamma(P - Phh'P)\Gamma' \tag{82}$$

and a and ρ are given by

$$a = \sigma + \Gamma Ph, \qquad \rho = (1 - h'Ph)^{1/2}. \tag{83}$$

Now, by Theorem 4.1, there is a one-to-one correspondence between symmetric solutions P of (82) and polynomial spectral factors $\rho^{-1}a(z)$ of the pseudo-polynomial $\sigma(z)\sigma(z^{-1})$, and this correspondence is described by (83). The stable solution to (81) corresponds to $P = 0$, the only positive semidefinite solution to (82), and in this case (83) yields $a = \sigma$ and $\rho = 1$ as expected.

As a corollary of Theorem 4.1 we have that P is also a solution to a certain algebraic Riccati equation related to the rational function $v(z)$, defined by (69). In fact, it is elementary to check that $v(z)$ has a minimal realization

$$v(z) = 1/2 + h'(zI - F)^{-1}g \tag{84}$$

where F is the companion matrix

$$F = J - ah' \tag{85}$$

and

$$g = \tfrac{1}{2}(b - a) \tag{86}$$

to which representation there corresponds an *algebraic Riccati equation*

$$P = FPF' + (g - FPh)(1 - h'Ph)^{-1}(g - FPh)'. \tag{87}$$

We shall say that a symmetric solution P of (87) is *stabilizing* if $h'Ph < 1$ and

$$F_+ = F - (1 - h'Ph)^{-1}(g - FPh)h' \tag{88}$$

has all its eigenvalues in the closed unit disc. It can be shown (see, for example, [48]) that (87) has a unique stabilizing solution if and only if

$$v(e^{i\theta}) + v(e^{-i\theta}) > 0, \qquad \text{for all } \theta \tag{89}$$

which follows from (72).

Corollary 4.3: Let P be a symmetric solution of the CEE (38) such that $h'Ph < 1$, and let $v(z)$ be the corresponding rational function (69) defined via (75) and (76). Then P is the unique stabilizing solution to the algebraic Riccati equation (87) corresponding to $v(z)$.

Now, restricting our attention to positive semidefinite solutions of the CEE (38), $a(z)$ and $b(z)$ become Schur polynomials and $v(z) := \frac{1}{2}(b(z)/a(z))$ is strictly positive real. Also

$v(z)$ is analytic for $|z| \geq R$ for some $R < 1$, and hence the Laurent series (70) is valid there. Consequently

$$\Phi(z) = 1 + \sum_{i=1}^{\infty} \hat{c}_i(z^i + z^{-i}) \tag{90}$$

is defined in an annulus containing the unit circle and is therefore a bona fide spectral density.

Corollary 4.4: There is a one-to-one correspondence between positive semidefinite solutions of the covariance extension equation (38) satisfying $h'Ph < 1$ and monic Schur polynomial $a(z)$ of degree n such that for some $\rho \in (0, 1]$

$$w(z) = \rho \frac{\sigma(z)}{a(z)} \tag{91}$$

satisfies

$$w(z)w(1/z) = 1 + \sum_{i=1}^{\infty} \hat{c}_i(z^i + z^{-i}) \tag{92}$$

on the unit circle, where

$$\hat{c}_i = c_i, \qquad \text{for } i = 1, 2, \cdots, n. \tag{93}$$

Under this correspondence $a(z)$ and ρ are given by (75) and (79), respectively. The degree of $w(z)$ equals the rank of P.

Proof: It remains to show that there is a one-to-one correspondence between $w(z)$ and $v(z)$. The linear operator $S(a)$, from the space of polynomials of degree less than or equal to n to the space of symmetric pseudo-polynomials, defined by

$$S(a)p = a(z)p(z^{-1}) + a(z^{-1})p(z)$$

is invertible if and only if $a(z)$ has reciprocal roots, as is the case if $a(z)$ is a Schur polynomial. This follows from [19], noting that the Jury matrix of $a(z)$ is a matrix representation of $S(a)$; also see [13, Lemma 5.5]. Therefore (72) can be uniquely solved for $b(z)$ and hence $v(z)$ is uniquely determined by $w(z)$; the reverse is trivial. Finally, $\rho \leq 1$ follows from $P \geq 0$. □

One of our main results, namely Theorem 2.1, is now an immediate consequence of Corollary 4.4 and Theorem 3.2. We note that the minimum-phase spectral factor $w(z)$ is precisely the modeling filter corresponding to P. Passing white noise through this modeling filter and letting it come to statistical steady state, we obtain a linear stochastic system (1). In view of Corollary 4.3 and the classical stochastic realization theory [2], [21], [41], P is actually the state covariance matrix of this system, i.e.,

$$P = \mathrm{E}\{x(t)x(t)'\}.$$

Theorem 4.1 and Corollary 4.3 are a consequence of the following chain of lemmas.

Lemma 4.5: The interpolation condition (71) holds if and only if

$$g = (I - U)^{-1}u + (I - U)^{-1}Ua \tag{94}$$

where U, u, a, and g are defined by (35), (75), and (86).

Proof: To prove that (94) is equivalent to the interpolation condition (71), first note that (19) can be written

$$b = 2c + (2C_n - I)a \tag{95}$$

where

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \qquad C_n = \begin{bmatrix} 1 & & & & \\ c_1 & 1 & & & \\ c_2 & c_1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ c_{n-1} & c_{n-2} & c_{n-3} & \cdots & 1 \end{bmatrix}.$$

Now, identifying coefficients in (34), we see that u is the unique solution of

$$C_n u = c. \tag{96}$$

This equation may also be written

$$C_{n+1}\begin{bmatrix} 1 \\ -u \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and therefore a simple inspection shows that

$$C_n(I - U) = I. \tag{97}$$

Consequently, (95) takes the form

$$b = 2(I - U)^{-1}u + 2(I - U)^{-1}a - a$$

which is equivalent to $g = \frac{1}{2}(b - a)$ satisfying (86). □

Lemma 4.6: Let f be the function, defined by (75) and (76), sending symmetric solutions P of the covariance extension equation (38) to points $(a, b) \in \mathbb{R}^{2n}$. Then f is injective and maps onto the set of (a, b) satisfying the interpolation condition (72) and the positivity condition (73). Its inverse $f^{-1}(a, b)$ is the unique solution of the Lyapunov equation (78).

Proof: Let P be a symmetric solution of the covariance extension equation (38) such that $1 - h'Ph < 1$. Then a straightforward reformulation of (38) yields

$$\begin{aligned} P = {} & (\Gamma + \sigma h')P(\Gamma + \sigma h')' - (\Gamma Ph + \sigma) \\ & \cdot (\Gamma Ph + \sigma)' + \rho^2 \sigma\sigma' + gg' \end{aligned} \tag{98}$$

where

$$g = u + U(\Gamma Ph + \sigma) \tag{99}$$

and ρ is given by (77). Now, if a and b are defined in terms of P by (75) and (76), we have

$$\tfrac{1}{2}(a + b) = \Gamma Ph + \sigma \tag{100}$$

and

$$\tfrac{1}{2}(b - a) = g. \tag{101}$$

Therefore, since

$$\Gamma = J - \sigma h' \tag{102}$$

P must satisfy

$$p_{ij} - p_{i+1,\,j+1} = -\tfrac{1}{2}(a_i b_j + b_i a_j) + \rho^2 \sigma_i \sigma_j \tag{103}$$

where $p_{ij} = 0$ when i or j is greater than one. Multiplying (103) by $z^{j-i} = z^{n-i}z^{n-j}$ and summing over all $i, j = 1, 2, \cdots, n$, we obtain

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}z^{j-i} - \sum_{i=2}^{n}\sum_{j=2}^{n} p_{ij}z^{j-i} \\ = -\tfrac{1}{2}\{[a(z)-z^n][b(z^{-1})-z^{-n}] \\ + [a(z^{-1})-z^{-n}][b(z)-z^n]\} \\ + \rho^2[\sigma(z)-z^n][\sigma(z^{-1})-z^{-n}] \end{aligned} \tag{104}$$

the left member of which may be written

$$\text{LM} = \sum_{i=1}^{n} p_{i+1,1}(z^i + z^{-i}) + p_{11}. \tag{105}$$

But, in view of (102) and (77), (100) is the same as

$$JPh = \tfrac{1}{2}(a+b) - \rho^2\sigma \tag{106}$$

which may also be written

$$p_{i+1,1} = \tfrac{1}{2}(a_i + b_i) - \rho^2\sigma_i. \tag{107}$$

Therefore, since $p_{11} = 1 - \rho^2$, the left member of (104) becomes

$$\begin{aligned}\text{LM} = \tfrac{1}{2}\{z^n[a(z^{-1}) - z^{-n} + b(z^{-1}) - z^{-n}] \\ + z^{-n}[a(z) - z^n + b(z) - z^n]\} \\ - \rho^2\{[z^n[\sigma(z^{-1}) - z^{-n}] \\ + z^{-n}[\sigma(z) - z^n]\} + 1 - \rho^2\end{aligned}$$

and consequently

$$\tfrac{1}{2}[a(z)b(z^{-1}) + a(z^{-1})b(z)] = \rho^2\sigma(z)\sigma(z^{-1}). \tag{108}$$

This establishes (72). It remains to show that the interpolation condition (71) also holds. To this end note that

$$\Gamma Ph + \sigma = a + g \tag{109}$$

which inserted into (99) yields (94), which in turn is equivalent to (71) (Lemma 4.5). We have thus established that the function f maps into the set of (a, b) satisfying the interpolation condition (71) and the factorization condition (72). To prove that f actually maps onto, choose any pair (a, b) satisfying these conditions. Let ρ be the unique positive number satisfying (108), obtained by identifying coefficients of like powers in z, and let P be the unique symmetric solution of the Lyapunov equation (78). Here uniqueness is a consequence of the fact that the eigenvalues of J are all zero and hence in the open unit disc. It remains to show that a, b, and P satisfy (75) and (76) and that P satisfies the covariance extension equation (38). To this end, write the Lyapunov equation (78) in the form (105). Together with the factorization condition (108) this yields

$$\begin{aligned}\sum_{i=1}^{n} p_{i+1,1}(z^i + z^{-i}) + p_{11} \\ = \tfrac{1}{2}\{z^n[a(z^{-1}) + b(z^{-1})] + z^{-n}[a(z) + b(z)]\} \\ - \rho^2[z^n\sigma(z^{-1}) + z^{-n}\sigma(z)] - 1 + \rho^2\end{aligned}$$

from which (107), or equivalently (106), and $p_{11} = 1 - \rho^2$ are obtained by identifying coefficients of like powers in z. Consequently, (77) and (100) hold.

We now invoke the interpolation condition (71), which by Lemma 4.5, is equivalent to $g := \frac{1}{2}(b - a)$ satisfying (94) or, equivalently

$$g = u + U(a + g). \tag{110}$$

But, since $b = a + 2g$, (100) is the same as (109), which together with (110) yields (99), which in turn, together with (100) and (101), yields (75) and (76). Now, inserting (99)–(101) and (77) into (78), a simple calculation yields (38), showing that P is a symmetric solution of the covariance extension equation. Hence we have proved that f maps onto the set of (a, b) which satisfies (71) and (72).

To prove that f is injective, let (a, b) be any point in the range of f. Then ρ^2 is uniquely defined by (108). Any P such that $f(P) = (a, b)$ must satisfy the Lyapunov equation (78) which has a unique solution. This establishes both injectiveness and the last statement of Lemma 4.3. □

Lemma 4.6 shows that there corresponds a unique rational function $v(z)$, defined via (69), to any solution P to the CEE. Next we shall show that P is also a solution to the algebraic Riccati equation (87) corresponding to (84). Since (F', h) is reachable, the existence of a unique stabilizing solution follows, for example, from [48, Th. 1]. If, in addition, (F, g) is reachable so that $(F, g, h, \frac{1}{2})$ is a minimal realization of $v(z)$, then it is well known [2], [21] and immediately seen from the Kalman–Yakubovich–Popov Lemma that $P > 0$ if and only if $v(z)$ is strictly positive real. In general, we have the following result.

Lemma 4.7: The rational function $v(z)$, defined by (84) and satisfying (89), is strictly positive real if and only if the unique stabilizing solution P of the algebraic Riccati equation is positive semidefinite. In this case, the degree of $v(z)$ equals the rank of P.

Proof: Setting $k := (g - FPh)(1 - h'Ph)^{-1/2}$, we may write the algebraic Riccati equation (87) in the Lyapunov form

$$P = FPF' + kk'. \tag{111}$$

To proceed we shall need some properties of such equations, namely the following.

1) If $P > 0$, then $|\lambda(F)| \leq 1$.
2) If $|\lambda(F)| \leq 1$, then $P \geq 0$.
3) If F has no pair of eigenvalues λ_1, λ_2 which are reciprocal, i.e., such that $\lambda_1 = 1/\lambda_2$, then

$$\text{rank}\, P = \text{rank}\,(k, Fk, \cdots, F^{n-1}k).$$

The continuous-time versions of these statements follow from [24, Th. 3.3]: i) from (2), ii) from (7), and iii) from (5) in that theorem. The discrete-time results are obtained by applying the usual linear fractional transformations—see, for example [24, Sec. 2.2] or in [21, Sec. 3.3]—keeping in mind that the left half-plane is transformed into the unit disc as concerns the spectrum of F, while P remains the same in the two settings.

The rational function $v(z) := \frac{1}{2}(b(z)/a(z))$ satisfying (91) is equivalent to the existence of a Schur polynomial $\sigma(z)$ and a positive number ρ such that (108) holds. Then $v(z)$ is strictly positive real if and only if $a(z)$ is a Schur polynomial.

First suppose that $a(z)$ is a Schur polynomial. Then, since $a(z)$ is the characteristic polynomial of F, $|\lambda(F)| < 1$, so it follows from ii) that $P \geq 0$. Conversely, suppose that $P \geq 0$. Setting $r := \operatorname{rank} P$, there is a nonsingular linear transformation T and a positive definite symmetric $r \times r$ matrix P_1 such that

$$TPT' = \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Transforming (111) accordingly yields

$$\begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} F_{11}' & F_{21}' \\ F_{12}' & F_{22}' \end{bmatrix} + \begin{bmatrix} k_1 k_1' & k_1 k_2' \\ k_2 k_1' & k_2 k_2' \end{bmatrix} \tag{112}$$

where

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = TFT^{-1} \quad \text{and} \quad \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = Tk.$$

Since therefore

$$F_{21} P_1 F_{21}' + k_2 k_2' = 0$$

we must have $F_{21} = 0$ and $k_2 = 0$, and hence

$$a(z) = \det(zI - F) = \det(zI - F_{11}) \det(zI - F_{22}). \tag{113}$$

Also a straightforward calculation shows that

$$w(z) := \rho \frac{\sigma(z)}{a(z)} = h_1 (zI - F_{11})^{-1} k_1 + \rho$$

where h_1 is the r-vector $(1, 0, \cdots, 0)'$, so $\det(zI - F_{22})$ must be a common factor of $a(z)$ and $\sigma(z)$ which is being canceled. Since $\sigma(z)$ is a Schur polynomial, then so is $\det(zI - F_{22})$. Therefore, it only remains to prove that $\det(zI - F_{11})$ is a Schur polynomial, i.e., that $|\lambda(F_{11})| < 1$. To this end, note that, in view of (112)

$$P_1 = F_{11} P_1 F_{11}' + k_1 k_1'.$$

Then, since $P_1 > 0$, ii) implies that $|\lambda(F_{11}| \leq 1$. But $a(z)$ cannot have any zeros on the unit circle and hence, in view of (113), we must have $|\lambda(F_{11})| < 1$ as claimed. In fact, if λ_0 is a zero of $a(z)$ on the unit circle, then so is $1/\lambda_0$. Therefore, in view of (108), either λ_0 or $1/\lambda_0$ is a zero of $\sigma(z)$ contradicting the assumption that $\sigma(z)$ is a Schur polynomial.

To prove the last statement, observe that if F is stable, there are no reciprocal eigenvalues, so iii) implies that

$$\operatorname{rank} P = \operatorname{rank}(k, Fk, \cdots, F^{n-1}k)$$

which, in view of the fact that (h, F) is observable, equals the degree of

$$w(z) = h'(zI - F)^{-1} k + \rho. \tag{114}$$

However, in view of (108) and the fact that $a(z)$, $b(z)$, and $\sigma(z)$ are all Schur polynomials, any common polynomial factor of $a(z)$ and $b(z)$ is a common factor also of $a(z)$ and $\sigma(z)$ and vice versa. Hence, $w(z)$ and $v(z)$ have the same degree, namely rank P. □

It is well known that the rational function (69) is strictly positive real if and only if (89) holds and either $a(z)$ or $b(z)$ is a Schur polynomial, in which case both $a(z)$ and $b(z)$ are Schur polynomials. Therefore, the last two statements of Theorem 4.1 follow from Lemma 4.7 and Corollary 4.3, which we prove next.

Proof of Corollary 4.3: Let P be a symmetric solution to the covariance extension equation (38) such that $h'Ph < 1$, and let g be defined by (39). Then, as demonstrated above, (109) holds, i.e.,

$$g = \Gamma P h + \sigma - a. \tag{115}$$

Inserting (115) into (38) yields

$$P = FPF' + \rho^2 (\sigma - a)(\sigma - a)' \tag{116}$$

where $\rho^2 = 1 - h'Ph$ and

$$F = \Gamma + (\sigma - a) h' \tag{117}$$

which, in view of (102), is the matrix (85) defined in Lemma 4.5. Now, from (114) and (117) we have

$$\sigma - a = \rho^{-2}(g - FPh) \tag{118}$$

which inserted into (116) yields the algebraic Riccati equation (87). Hence P is a symmetric solution of (87), which, by Lemma 4.5, clearly is the algebraic Riccati equation corresponding to $v(z)$. However, it remains to show that it is the unique stabilizing solution of (87). To this end, observe that (117) and (118) imply that $\Gamma = F_+$, where F_+ is the feedback matrix (88). Hence, it follows from the fact that the characteristic polynomial $\sigma(z)$ of Γ is a Schur polynomial and the fact that $1 - h'Ph > 0$ that P is the stabilizing solution of (87). □

V. Minimal Partial Stochastic Realizations

The question of minimality of the dimension of partial stochastic realizations will now be studied in more detail. In this direction, Theorem 2.1 gives some information about the minimal partial stochastic realization problem. In fact, for each choice of zero polynomial there is a unique solution which we may denote $P(\sigma)$. In this setting, the minimal partial realization problem could also be phrased as finding the zero polynomial $\hat{\sigma}$ minimizing the function

$$r(\sigma) = \operatorname{rank} P(\sigma)$$

over the region $\mathcal{S}_n$ of Schur polynomials (36). The optimal zero structure $\hat{\sigma}$ is in general not unique, and the structure of the optimizing set of $\hat{\sigma}$ depends on $(c_1, \cdots, c_n)$. In harmony with Example 4.2, all σ are optimizing and $r(\sigma)$ is identically zero if and only if $c = (c_1, c_2, \cdots, c_n) = 0$. It can be further seen that all σ are optimizing if $c = (0, \cdots, 0, c_n)$ with $c_n \neq 0$, in which case $r(\sigma)$ is identically n. In this section, we investigate how the positive degree of a partial covariance sequence depends on the values of the covariance

data $c_0, c_1, \cdots, c_n$. In general, according to Theorem 2.2, for $n > 1$, the sets of such n-tuples for which $r(\hat{\sigma}) = n$ and $r(\hat{\sigma}) < n$ are both open sets in $\mathbb{R}^n$.

We shall first consider the "codimension one" case, i.e., the situation where the minimal dimension of a partial stochastic realization can be reduced to at least $n-1$. This, of course, occurs when there are "extensions" (α, γ) for which the corresponding choice of polynomials (a, b) has a common root. Classically, this may be tested by computing whether the resultant, $R(a, b)$ of the pair (a, b), vanishes; see Appendix B. Regarding the sequence γ as being fixed, the resultant is then a polynomial $R_\gamma(\alpha)$ in α which defines an affine hypersurface H_γ in $\mathbb{R}^n$ as its zero locus. A better understanding of the real hypersurface yields a simple, but powerful, geometric criterion for a codimension one reduction for partial stochastic realizations. In particular, we shall prove the following result, which is a special case of Theorem 2.2, but which already illustrates the profound difference between deterministic and stochastic partial realization theory.

Theorem 5.1: Let $\Sigma(n-1)$, $S(n)$, and $\mathcal{C}_n$ be defined as in Section II. Then

$$\mathcal{C}_n = \Sigma(n-1) \cup S(n)$$

is a decomposition of $\mathcal{C}_n$ into two semialgebraic subsets with nonempty interiors. In fact, $\Sigma(n-1) \setminus \{0\}$ is an open, semialgebraic subset of $\mathcal{C}_n \setminus \{0\}$. That is

$$\Sigma(n-1) = \mathcal{O}(n-1) \cup \{0\}$$

where $\mathcal{O}(n-1)$ is an open, semialgebraic subset of $\mathcal{C}_n$. If $n \geq 2$, $\mathcal{O}(n-1)$ is nonempty.

Remark 5.2: In order to prove results guaranteeing the structural stability, with respect to γ, of the intersection of H_γ with $\mathcal{P}_n(\gamma)$, we shall need to prove a separation theorem, a result which is not at all immediate from the definitions. Indeed, not every algebraic set defined by a single equation is a hypersurface in what is called the geometric sense. For example, in $\mathbb{R}^2$ the equation $x^2+y^2=0$ defines an algebraic hypersurface, but the zero locus does not separate $\mathbb{R}^2$ into two or more open sets, nor is it dimension one, in any sense but the purely algebraic sense of counting equations. This example, however, illustrates exactly what can go wrong for real hypersurfaces. Very briefly, in $\mathbb{C}^2$, the equation $x^2+y^2=0$ defines a one-dimensional algebraic curve, with just one *singular point*, (0, 0)—that is, a point at which the total derivative (or gradient) of the defining equation vanishes. As it turns out, the only real point on this complex curve is a singular point, which is precisely why the real locus has "algebraic" dimension one but "geometric" dimension zero in a sense we shall now make precise.

Recall [47] that an algebraic subset of $\mathbb{R}^n$ defined by a single polynomial equation is called an algebraic hypersurface. An algebraic hypersurface is a geometric hypersurface if and only if it contains a *regular point*, i.e., a point at which the gradient of the defining equation is nonzero. Geometric hypersurfaces have dimension $n-1$, in the sense that a geometric hypersurface is always an $(n-1)$-dimensional manifold in a neighborhood of any regular point. Moreover, the complement of a geometric hypersurface is a union of the two disjoint open sets where the defining equation is positive and negative.

According to a theorem of Whitney [47], the complement of an algebraic subset of $\mathbb{R}^n$ defined by a single polynomial equation of degree d has at most $d+1$ connected components. In preparation for the proof of Theorem 5.1, we need the following separation criterion, the essence of which is that H_γ contains a real, regular point and therefore separates $\mathbb{R}^n$ itself into at least two disjoint open subsets. In this language, we first wish to characterize for which covariance sequences the hypersurface H_γ is a proper algebraic subset, when it is nonempty, and when it is a geometric hypersurface.

Theorem 5.3: Consider a partial covariance sequence $c = (c_1, c_2, \cdots, c_n)$ in $\mathbb{R}^n$. There exists an α such that $R_\gamma(\alpha) \neq 0$ if and only if $c_i \neq 0$ for some $i = 1, \cdots, n$. A necessary and sufficient condition for $c \neq 0$ to admit a partial stochastic realization of dimension less than or equal to $n-1$ is that the hypersurface H_γ intersects $\mathcal{P}_n(\gamma)$ nontrivially; i.e.,

$$H_\gamma \cap \mathcal{P}_n(\gamma) \neq \emptyset. \tag{119}$$

In this case, H_γ separates $\mathcal{P}_n(\gamma)$ into at least two open subsets.

Of course, for low-dimensional problems, the separation criterion provides for a complete analysis of the minimal partial realization problem. Before turning to the higher-codimension cases, for the sake of completeness we describe these lower-dimensional examples. If $n=1$, then $\mathcal{O}(n-1)$ is empty, since H_γ reduces instead to the constraint $\gamma_0 = 0$. In particular, for $n=1$ we have $S(0) = \{0\}$ and $S(1) = \mathcal{C}_1 \setminus \{0\}$.

If $n=2$, we still have $S(0) = \{0\}$, and, as pointed out by Georgiou [23], it is easy to see that $c \in S(1)$ if and only if

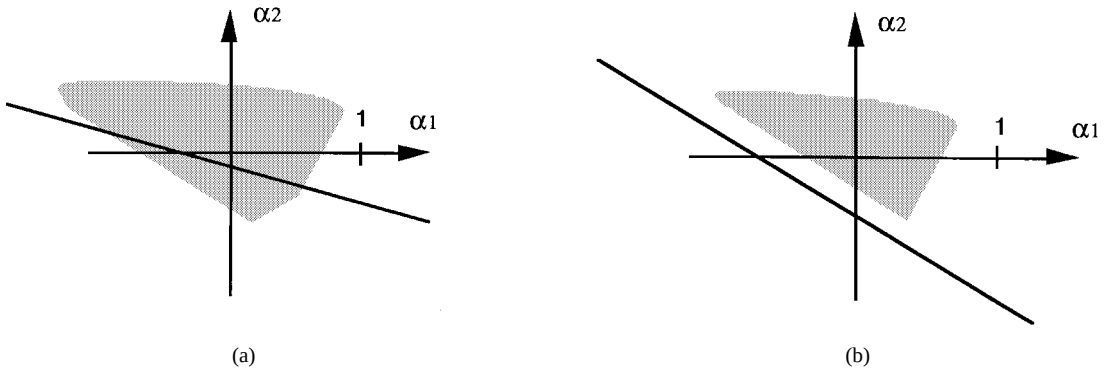
$$|\gamma_1| < \frac{|\gamma_0|}{1+|\gamma_0|}. \tag{120}$$

For such a γ, $\gamma_0 \neq 0$, and, as depicted in Fig. 4, only points α in the intersection between the (shaded) positive real region and the line will have $\deg v(z) = 1$. We note that in this case, for γ fixed the resultant hypersurface H_γ is linear and is defined by

$$\alpha_2 = \gamma_1\left(1 - \frac{1}{\gamma_0^2}\right)(\gamma_0\alpha_1 + \gamma_1). \tag{121}$$

These are precisely the points for which $a(z)$ and $b(z)$ have a common factor. All other α correspond to $v(z)$ of degree two. For example, if $\gamma_0 = 0.5$ and $\gamma_1 = 0.2$, (120) is satisfied, and the line (121) intersects the positive real region as depicted to the left in Fig. 4. All $v(z)$ corresponding to points on the interval defined by this intersection have degree one. The set $S(2)$ consists of those c for which $\gamma \neq 0$ and condition (120) is violated. The situation corresponding to such a point, $\gamma_0 = 0.5$ and $\gamma_1 = 0.4$, is illustrated to the right in Fig. 4. Here the intersection between $\mathcal{P}_2$ and the hypersurface H_γ is empty.

For $n > 2$ the situation is more complicated, but we have a sufficient condition for the positive degree n^* to be strictly less than one, which is similar to (120).



(a) (b)

Fig. 4. The hypersurface H_γ (line) and $P_2(\gamma)$ (shaded region) for two different choices of γ.

Corollary 5.4: Suppose $n \geq 2$. Any partial covariance sequence satisfying the condition

$$|\gamma_{n-1}| < \frac{|\gamma_{n-2}|}{1+|\gamma_{n-2}|} \tag{122}$$

is contained in $\mathcal{O}(n-1)$. In particular, (122) is a sufficient condition for n^* to be less than n. If $n = 2$, the condition is also necessary.

Corollary 5.4 gives an independent proof that $\Sigma(n-1)$ has a nonempty interior in $\mathcal{C}_n$. Of course, Theorem 2.2 also asserts that the semialgebraic sets $S(n^*)$ and $\Sigma(n^*)$ have nonempty interiors if and only if $\frac{1}{2}n \leq n^* \leq n$. In this direction, one can very easily see that for n even, the subset $S(\frac{1}{2}n)$ is open. Indeed, to say $c_1, \cdots, c_n$ lies in $S(\frac{1}{2}n)$ is to say that both the Hankel matrix

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_{(1/2)n} \\ c_2 & c_3 & \cdots & c_{(1/2)n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{(1/2)n} & c_{(1/2)n+1} & \cdots & c_{n-1} \end{bmatrix}$$

is invertible and that the Toeplitz matrix

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & \cdots & c_0 \end{bmatrix}$$

is positive definite. Since these conditions define open sets of sequences, $S(\frac{1}{2}n)$ is open but, of course, not dense as it is for the deterministic partial realization problem. Our proof is that in the intermediate cases $\frac{1}{2}n \leq n^* \leq n-1$, $S(n^*)$ still contains interior points is constructive. Indeed, we shall show that the maximum entropy filters corresponding to certain choices of partial covariance sequences are interior points by using a minimality criterion derived from the fast filtering algorithm described in Appendix A. From this construction, and the results stated above, Theorem 2.2 follows.

We now turn to the proofs. Our first result characterizes when H_γ is a proper algebraic subset of $\mathbb{R}^n$. Recall that $c = (c_1, c_2, \cdots, c_n) = 0$ if and only if $a(z) = b(z)$ and therefore if and only if $n^* = 0$.

Proposition 5.5: Suppose that $\gamma \neq 0$. Then there exists an α such that $R_\gamma(\alpha) \neq 0$.

Proof: Given the partial covariance sequence c, or equivalently γ, as seen in Section III, another n parameters are needed to parameterize all partial *stochastic* realizations of degree at most n. These could be α, as in the Kimura–Georgiou parameterization, or a, or, equivalently, $c_{n+1}, c_{n+2}, \cdots, c_{2n}$, as explained in Section II. Consequently, we consider a sequence

$$(c_1, c_2, \cdots, c_{2n}) \tag{123}$$

where $c = (c_1, c_2, \cdots, c_n)$ is fixed, and $c_{n+1}, c_{n+2}, \cdots, c_{2n}$ are free parameters to be selected. Now, if

$$0 = n_0 < n_1 < n_2 < \cdots < n_\nu \tag{124}$$

are the degree indexes of (123), the condition that (123) has McMillan degree at most n is equivalent to $n_\nu \leq n$, and in this case the last section of (123) is completed. If $n_\nu < n$, the parameter sequence $\rho = (\rho_1, \rho_2, \cdots, \rho_{2n})$ will end with $2(n - n_\nu)$ zeros.

Therefore, the statement of the proposition is equivalent to saying that for any fixed choice of $c_1, c_2, \cdots, c_n$ there exists an extension $c_{n+1}, c_{n+2}, \cdots, c_{2n}$ such that the square Hankel matrix of (123) is nonsingular, and hence $n_\nu = n$. Define ℓ to be the integer with the property that $c_1 = c_2 = \cdots = c_{\ell-1} = 0$ and $c_\ell \neq 0$. If $\ell = n$, then there is only one section and $n = n_1$. If $\ell < n$, then any section initiated in the first fixed part of c will end with c_{2k}, where $k \leq n-1$, if the arbitrary c_i are chosen appropriately. If $2k \geq n$, $c_{n+1}, c_{n+2}, \cdots, c_{2n}$ can be chosen so that there are, say, $n-k$ generic sections at the end, then the highest degree index equals n. □

Our next result ensures that the hypersurface H_γ in $\mathbb{R}^n$ is nonempty, except when $c = (0, \cdots, 0, c_n)$ with $c_n \neq 0$.

Proposition 5.6: The algebraic degree of a sequence c is n if, and only if, $c = (0, \cdots, 0, c_n)$ with $c_n \neq 0$.

Proof: By definition, the sequence c has degree indexes $(n_0, n_1) = (0, n)$, and hence the algebraic degree is n. To prove the converse, first note that if $c = 0$, the algebraic degree is zero. Next suppose that c_ℓ, $1 \leq \ell < n$ is the first nonzero element in c. Then the first section of c has 2ℓ elements, so if $\ell \geq n/2$, the algebraic degree is ℓ. If $\ell < n/2$, $n-2\ell$ elements remain for further sections, so the algebraic degree is at most $\ell + (n - 2\ell) = n - \ell < n$. □

Theorem 5.7: The hypersurface H_γ is the closure of the set of all regular points on H_γ. In particular, H_γ is a geometric hypersurface except when $c = (c_1, c_2, \cdots, c_n) = 0$ or when $c = (0, \cdots, 0, c_n)$ with $c_n \neq 0$.

Proof: The first step is to characterize the regular points on H_γ.

Proposition 5.8: Suppose that $\gamma \neq 0$. The set of regular points on H_γ consists of those α for which the corresponding polynomials a and b have a simple (real) common factor over $\mathbb{R}$.

Proof: We shall begin our analysis on the space $\mathbb{R}^{2n}$ of pairs (a, b) of monic polynomials of degree n. As we have noted in the proof of Proposition 5.5, this space can also be parameterized in terms of the partial sequence c (or equivalently γ), together with the parameters α in the Kimura–Georgiou parameterization. Following Kronecker, as described in Section III (see also Appendix B), the zero locus of the resultant $R(a, b)$ in $\mathbb{R}^{2n}$ can be identified with the zero locus Z_n of the determinant of the Hankel matrix

$$\begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ c_2 & c_3 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n-1} \end{bmatrix}.$$

Since

$$H_\gamma = \{(a, b) \in Z_n | \ \gamma \text{ is fixed}\} \tag{125}$$

H_γ is the intersection of Z_n with an n-dimensional manifold M_γ with tangent space given by (see [10])

$$T_{(a,b)}(M_\gamma) = \{(p, q) | \ aq - bp = r, \deg r \leq n-1\}. \tag{126}$$

From the representation of H_γ as such an intersection, one sees that every singular point of the hypersurface Z_n remains a singular point of H_γ. In particular (Appendix B), every α corresponding to a pair (a, b) having a greatest common divisor with degree larger than one is a singular point of H_γ.

Conversely, from the representations

$$T_{(a,b)}(Z_n) = \{(u_0\theta, v_0\theta) + \rho(a_0, b_0) | \ \deg u_0 \leq n-2, \deg u_0 \leq n-2, \rho \in \mathbb{R}\} \tag{127}$$

of the tangent space of Z_n (Proposition B.3) and (126) of the tangent space of M_γ and H_γ, we conclude that all other pairs are regular points on H_γ. That is, those α corresponding to a pair (a, b) having a greatest common divisor with degree one are precisely the regular points, as claimed in the Proposition.

To see this, it suffices to prove that for any α for which the corresponding polynomials a and b have a simple (real) common factor over $\mathbb{R}$, there exists a tangent vector (p, q) to (a, b) in $T_{(a,b)}(M_\gamma)$ which is not in $T_{(a,b)}(Z_n)$. Alternatively, it suffices to check that the subspace $T_{(a,b)}(Z_n)$ intersects $T_{(a,b)}(M_\gamma)$ in a codimension-one subspace. To say that (p, q) lies in the intersection

$$T_{(a,b)}(Z_n) \cap T_{(a,b)}(M_\gamma) \tag{128}$$

is to say that

$$a_0\theta(v_0\theta + \rho b_0) - b_0\theta(u_0\theta + \rho a_0) = r \tag{129}$$

where $r = r_0\theta$ and r_0 is an arbitrary polynomial satisfying $\deg r_0 < n-1$. However, for (p, q) lying in (128) we have (129) or, equivalently

$$(a_0v_0 - b_0u_0)\theta^2 = r \tag{130}$$

so that r_0 is not arbitrary but rather is itself also divisible by θ so that the corresponding set of polynomial r is each divisible by θ^2. In particular (see Remark B.2), the subspace (128) has codimension one in $T_{(a,b)}(M_\gamma)$, which concludes the proof of the proposition. □

Having characterized the regular points, we now show that they exist, except in the two cases delineated above.

Lemma 5.9: Suppose that $\gamma \neq 0$ and that there exists an α such that $R_\gamma(\alpha) = 0$. Then there exists an α such that $(a, b) = \theta$, where $\deg \theta = 1$, i.e., there exists a regular point.

Proof: This is equivalent to saying that for any fixed choice of $c_1, c_2, \cdots, c_n$ there exists an extension $c_{n+1}, c_{n+2}, \cdots, c_{2n}$ such that the largest degree index n_ν of $c_1, c_2, \cdots, c_{2n}$ equals $n-1$, i.e., the last section s_ν should end with ρ_{2n-2}, and $\rho_{2n-1} = \rho_{2n} = 0$. Let ℓ be defined as in the proof of Lemma 5.5. Then we must have $\ell < n$, so any section initiated in the first (fixed) half of c will end with c_{2k}, where $k \leq n-1$, if the arbitrary c_i in that section is chosen appropriately. We must have $\rho_{2n-1} = \rho_{2n} = 0$. If $k < n-1$, we fill the gap with $n-1-k$ generic sections. Then $n_\nu = n-1$, as claimed. □

Lemma 5.10: The hypersurface H_γ coincides with the closure of all regular points on H_γ.

Proof: Let $C_k \subset \mathbb{R}^n$ be the set

$$C_k = \{(c_1, c_2, \cdots, c_{2n}) | \ (a, b) = \theta, \deg \theta = k\}.$$

If we can show that

$$C_{k+1} \subset \overline{C_k}, \qquad k = 1, 2 \cdots, n \tag{131}$$

then $H_\gamma = \overline{C_1}$ as claimed. To this end, choose an arbitrary point $\hat{c}$ in C_{k+1}. Then $\hat{c}$ has a parameter sequence $\hat{\rho}$ such that

$$\hat{\rho}_{2n-2k-2} = \hat{\rho}_{2n-2k-1} = \cdots = \hat{\rho}_{2n} = 0$$

and $\hat{\rho}_{2n-2k-1}$ is the last element in a section. We want to find a sequence $\hat{c}^{(1)}, \hat{c}^{(2)}, \hat{c}^{(3)}, \cdots$ in C_k which converges to $\hat{c}$. Clearly, such a sequence is obtained by taking $\hat{c}^{(j)}$ to have the parameter sequence $\hat{\rho}^{(j)}$, where

$$\hat{\rho}_i^{(j)} = \begin{cases} \dfrac{1}{j}, & \text{for } i = 2n-2k-2 \\ \hat{\rho}_i, & \text{otherwise.} \end{cases} \tag{132}$$

In fact, if ν is the number of sections in $\hat{c}$, then $\hat{c}^{(j)}$ has $\nu+1$ sections and

$$\begin{cases} \hat{g}^{(j)}(z) = P_{\nu+1}(z) = zP_\nu(z) - \dfrac{1}{j}P_{\nu-1}(z) \\ \hat{a}^{(j)}(z) = Q_{\nu+1}(z) = zQ_\nu(z) - \dfrac{1}{j}Q_{\nu-1}(z) \end{cases}$$

so that

$$\frac{\hat{g}^{(j)}(z)}{\hat{a}^{(j)}(z)} = \frac{z\hat{g}(z) - \dfrac{1}{j}P_{\nu-1}(z)}{z\hat{a}(z) - \dfrac{1}{j}Q_{\nu-1}(z)} \to \frac{\hat{g}(z)}{\hat{a}(z)}$$

as $j \to \infty$, where $\hat{g}(z)$ and $\hat{a}(z)$ correspond to $\hat{c}$. Consequently, $\hat{c}^{(j)} \to \hat{c}$ as $j \to \infty$, as claimed. □

This concludes the proof of Theorem 5.7. □

Clearly, Theorem 5.7 implies Theorem 5.3. Theorem 5.1 then follows from these results and the following corollary.

Corollary 5.11: The subset $\Sigma(n-1)\backslash\{0\}$ is an open, semialgebraic subset of $\mathcal{C}_n\backslash\{0\}$. That is

$$\Sigma(n-1) = \mathcal{O}(n-1) \cup \{0\}$$

where $\mathcal{O}(n-1)$ is an open, semialgebraic subset of $\mathcal{C}_n$. If $n \geq 2$, $\mathcal{O}(n-1)$ is nonempty.

Proof: We first prove that $\Sigma(n-1)$ and hence $\Sigma(n-1)\backslash\{0\}$ is a semialgebraic set. Since the change of coordinates between $(c_1, c_2, \cdots, c_n)$ and $(\gamma_1, \gamma_2, \cdots, \gamma_n)$ is birational without poles on $\mathcal{C}_n$, it suffices to prove this claim in the γ coordinates. To this end, consider the space

$$\mathbb{R}^{2n} = \{((\alpha_1, \alpha_2, \cdots, \alpha_n), (\gamma_1, \gamma_2, \cdots, \gamma_n))\}.$$

The subset $\mathcal{P}_n$ of those sequences for which v is strictly positive real is, of course, a semialgebraic set, since it is definable in terms of polynomial inequalities [13]. We also consider the algebraic set Z_n, the zero set of the resultant R_γ introduced in the proof of Proposition 5.8. In particular, $Z_n \cap \mathcal{P}_n$ is semialgebraic. Finally, we define the projection p_n: $\mathbb{R}^{2n} \to \mathbb{R}^n$ via

$$p_n((\alpha_1, \alpha_2, \cdots, \alpha_n), (\gamma_1, \gamma_2, \cdots, \gamma_n)) = (\gamma_1, \gamma_2, \cdots, \gamma_n).$$

According to the Tarski–Seidenberg theorem [28], $\Sigma(n-1) = p_n(M_{n-1} \cap \mathcal{P}_n)$ is also semialgebraic.

To see that $\Sigma(n-1)\backslash\{0\}$ is open, we recall that H_γ is a proper hypersurface (Proposition 5.5), i.e., a codimension one, provided at least one γ_i differs from zero. Suppose then that $\hat{c}$ lies in $\Sigma(n-1)\backslash\{0\}$. Denoting by $\hat{\gamma}$ the corresponding partial Schur sequence, according to Theorem 5.3, the proper hypersurface $H_{\hat{\gamma}}$ meets the open set $\mathcal{P}_n(\hat{\gamma})$ at a regular point. Then, there exist points $(\alpha_1, \hat{\gamma})$ and $(\alpha_2, \hat{\gamma})$ in $\mathcal{P}_n(\hat{\gamma})$ such that

$$R_{\hat{\gamma}}(\alpha_1) > 0 \quad \text{and} \quad R_{\hat{\gamma}}(\alpha_2) < 0.$$

Since the set of strictly positive real, degree n, transfer functions is open, for γ sufficiently near $\hat{\gamma}$, we must have $(\alpha_1, \hat{\gamma}) \in \mathcal{P}_n(\gamma)$ and $(\alpha_2, \hat{\gamma}) \in \mathcal{P}_n(\gamma)$. Moreover, we must also have

$$R_\gamma(\alpha_1) > 0 \quad \text{and} \quad R_\gamma(\alpha_2) < 0.$$

Since $\mathcal{P}_n(\gamma)$ is connected and f_1 is continuous, there is a point $(\alpha, \gamma) \in \mathcal{P}_n(\gamma)$ such that $f_1(\alpha_2, \gamma) = 0$. Therefore, $\Sigma(n-1)\backslash\{0\}$ is open.

To see that $\mathcal{O}(n-1)$ is nonempty, we need only construct $(\hat{\alpha}, \hat{\gamma})$ corresponding to a modeling filter having degree n^* satisfying $1 \leq n^* \leq n-1$. One such choice

$$\hat{\alpha} = (0, 0, \cdots, 0) \quad \text{and} \quad \hat{\gamma} = (0, \cdots, 0, \hat{\gamma}_{n-2}, 0)$$

with $\hat{\gamma}_{n-2} \neq 0$, corresponds to a maximum entropy filter

$$\hat{w}(z) = \rho \frac{z^{n-1}}{\varphi_{n-1}(z)}$$

of degree $n-1$; see Section III. □

We have just seen that the separation criterion implies that $\Sigma(n-1)\backslash\{0\}$ is open. We now prove Corollary 5.4, which gives an explicit sufficient condition (122) for membership in $\Sigma(n-1)\backslash\{0\}$ in terms of the Schur parameters.

Proof of Corollary 5.4: Given a sequence of Schur parameters $\gamma_0, \gamma_1, \cdots, \gamma_{n-1}$ satisfying (122), we want to find a positive real function

$$v(z) = \frac{1}{2}\frac{b(z)}{a(z)} \tag{133}$$

of degree at most $n-1$; the first n Schur parameters are precisely $\gamma_0, \gamma_1, \cdots, \gamma_{n-1}$. We shall demonstrate that there is a real number α_1 such that

$$\begin{cases} a(z) = \varphi_{n-1}(z) + \alpha_1 \varphi_{n-2}(z) \\ b(z) = \psi_{n-1}(z) + \alpha_1 \psi_{n-2}(z) \end{cases} \tag{134}$$

defines such a function, where φ_{n-1}, φ_{n-2}, $\psi_{n-1}(z)$, and ψ_{n-2} are the appropriate Szegö polynomials defined as in Appendix A from $\gamma_0, \gamma_1, \cdots, \gamma_{n-2}$. Consider the dynamical system (155) with $2(n-1)$ equations and initial condition $\alpha(0) = (\alpha_1, 0, \cdots, 0)' \in \mathbb{R}^{n-1}$ and $\gamma(0) = (\gamma_0, \gamma_1, \cdots, \gamma_{n-2})' \in \mathbb{R}^{n-1}$. Then, to match the remaining Schur parameter γ_{n-1}, we must choose α_1 so that

$$\gamma_{n-1} = \gamma_{n-2}(1) = -\frac{\gamma_{n-2}}{1-\gamma_{n-2}^2}\,\alpha_1. \tag{135}$$

Now, in view of (122), $\gamma_{n-2} \neq 0$. Consequently, it remains to show that (134) with α_1 given by

$$\alpha_1 = -\frac{\gamma_{n-1}}{\gamma_{n-2}}(1-\gamma_{n-2}^2) \tag{136}$$

defines a positive real function (133). To prove this we need to show that $a(z)$ is a Schur function and that (15) holds on the unit circle.

Let us start with the last requirement. To this end, first note that, in view of (136), (122) is equivalent to

$$|\alpha_1| < 1 - |\gamma_{n-2}|. \tag{137}$$

From the recursions (149) and (150) it is not hard to see that

$$\varphi_{n-1}(z)\psi_{n-1}(z^{-1}) + \psi_{n-1}(z)\varphi_{n-1}(z^{-1}) = r_{n-1} \tag{138}$$

and that

$$\varphi_{n-1}(z)\psi_{n-2}(z^{-1}) + \psi_{n-1}(z)\varphi_{n-2}(z^{-1}) = r_{n-2}z \tag{139}$$

and therefore, since $r_{n-1} = r_{n-2}(1-\gamma_{n-2}^2)$

$$\tfrac{1}{2}[a(z)b(z^{-1}) + a(z^{-1})b(z)] = \tfrac{1}{2}(1-\gamma_{n-2}^2 + 2\alpha_1 \cos\theta + \alpha_1^2)$$

for $z = e^{i\theta}$. This is positive for all θ if and only if (137) holds.

Next we prove that $a(z)$ is a Schur polynomial. In view of (149), we have

$$a(z) = (z+\alpha_1)\varphi_{n-2}(z) - \gamma_{n-2}\varphi_{n-2}^*(z). \tag{140}$$

Since, by (137), $|\alpha_1| < 1$, the function

$$f(z) = (z+\alpha_1)\varphi_{n-2}(z) \tag{141}$$

has all its $n-1$ zeros inside the unit circle. We want to show that the same is true for $a(z)$. To this end, observe that, in view of (137)

$$|a(z_0)-f(z_0)|<(1-|\alpha_1|)|\varphi^*_{n-2}(z_0)|$$

for any z_0 on the unit circle. But, since $|\varphi_{n-2}(z_0)| = |\varphi^*_{n-2}(z_0)|$ (see [4, p. 119]) and $|z_0|=1$, this implies that

$$|a(z_0)-f(z_0)|<|f(z_0)| \tag{142}$$

and hence by Rouche's theorem, $a(z)$ is a Schur polynomial as required. If $n=2$, all $a(z)$ and $b(z)$ have the form (134), and therefore we have also necessity, as claimed. □

In the course of determining properties (e.g., nonvacuousness) of the sets $S(n^*)$ and $\Sigma(n^*)$, we will find it useful to apply not just the solution to the rational covariance extension problem, but also certain of the tools which played an important role in its resolution. One of these is a nonlinear dynamical system (155), which is a reformulation of a fast algorithm for Kalman filtering [38], [39] and is also related to the Schur algorithm. We note that for initial conditions $(\alpha,\gamma)\in\mathcal{P}_n$ it is known that the trajectory remains in $\mathcal{P}_n$ and converges to $(\alpha_\infty,0)$, where $\alpha_\infty\in\mathcal{S}_n=\mathcal{P}_n(0)$ [13]. Of course to say $(\alpha,\gamma)\in\mathcal{P}_n$ is to say that to (α,γ) there corresponds a positive real, rational function v. It is important to note that

$$v(z)=\tfrac{1}{2}+\sum_{i=1}^{\infty}c_iz^i \tag{143}$$

where the covariance sequence $(1,c_1,c_2,\cdots)$ has as its corresponding sequence of Schur parameters the components of the partial state $\gamma(t)$ propagated by this dynamical system.

This observation provides a useful alternative for an analysis of the minimal partial stochastic realization problem. Indeed, in this language, we note that to say that a modeling filter corresponding to a pair (α,γ) has degree less than or equal to n^* is to say that $f(\alpha,\gamma)=0$, where $f\colon \mathbb{R}^{2n}\to\mathbb{R}^n$ is defined by

$$\begin{cases} f_1(\alpha,\gamma)=\gamma_{n^*}+\alpha_1(1)\gamma_{n^*-1}+\alpha_2(1)\gamma_{n^*-2} \\ \qquad\qquad -\cdots+\alpha_{n^*}(1)\gamma_0 \\ f_2(\alpha,\gamma)=\gamma_{n^*+1}+\alpha_1(2)\gamma_{n^*}+\alpha_2(2)\gamma_{n^*-1} \\ \qquad\qquad +\cdots+\alpha_{n^*}(2)\gamma_1 \\ \qquad\qquad \vdots \\ f_{n-n^*}(\alpha,\gamma)=\gamma_{n-1}+\alpha_1(n-n^*)\gamma_{n-2} \\ \qquad\qquad +\alpha_2(n-n^*)\gamma_{n-3} \\ \qquad\qquad +\cdots+\alpha_{n^*}(n-n^*)\gamma_{n-n^*-1}. \end{cases} \tag{144}$$

More precisely, if the degree of (143) is less than or equal to n^*, there is a recursion

$$\begin{aligned}\gamma_{t+n^*}=&-\alpha_1(t+1)\gamma_{t+n^*-1}-\alpha_2(t+1)\gamma_{t+n^*-2}\\ &-\cdots-\alpha_{n^*}(t+1)\gamma_t\end{aligned} \tag{145}$$

of type (158), where $\alpha(t)=(\alpha_1(t),\cdots,\alpha_{n^*}(t))'$ and $\gamma(t)=(\gamma_t,\gamma_{t+1},\cdots,\gamma_{t+n^*-1})'$ are generated by a reduced-order dynamical system (155) of dimension $2n^*$. In order to also match the remaining $n-n^*$ Schur parameters $\gamma_{n^*},\cdots,\gamma_{n-1}$, the constraints $f(\alpha,\gamma)=0$ are therefore required. We note that, for fixed γ, $f_1(\alpha,\gamma)=0$ is of course an alternative expression for the resultant of the pair (a,b) of polynomials corresponding to (α,γ). We are now in a position to complete our proof of Theorem 2.2 with the following sequence of lemmas.

Lemma 5.12: Let n^* be any integer satisfying $0\le n^*\le n$. Then the subsets $S(n^*)$ of $\mathbb{R}^n$ consisting of partial covariance sequences $(c_1,c_2,\cdots,c_n)$ having a minimal stochastic realization of degree n^* are a nonempty semialgebraic set. The subset $\Sigma(n^*)$ of those partial covariance sequences c having a minimal stochastic realization of degree less than or equal to n^* is also semialgebraic.

Proof: We begin by proving that $S(n^*)$ and $\Sigma(n^*)$ are semialgebraic. As in the proof of Theorem 5.3 it suffices to prove this claim in the γ coordinates. For codimension greater than one, we consider the algebraic subset M_{n^*} of $\mathbb{R}^{2n}$ defined by $M_{n^*}=f^{-1}(0)$, where f is defined by (144). Since $M_{n^*}\cap\mathcal{P}_n$ is semialgebraic, $\Sigma(n^*)=p_n(M_{n^*}\cap\mathcal{P}_n)$ is also semialgebraic by the Tarski–Seidenberg theorem [29]. In this notation

$$\begin{aligned}S(0)&=\Sigma(0)\\ S(n^*)&=\Sigma(n^*)\backslash\Sigma(n^*-1).\end{aligned}$$

Since the complement of a semialgebraic set is semialgebraic, it follows from induction that $S(n^*)$ is semialgebraic.

We next show that $S(n^*)$, and hence $\Sigma(n^*)$, is nonempty. As before, consider the choice $(\hat\alpha,\hat\gamma)=((0,0,\cdots,0),(0,\cdots,\hat\gamma_{n^*-1},\cdots,0))$ corresponding to a maximum entropy filter of degree n^*

$$\hat w(z)=\rho\frac{z^{n^*}}{a(z)}$$

where $a(z)$ is the n^*th Szegö polynomial φ_{n^*}. We shall further assume that $\hat\gamma_{n^*-1}\neq 0$ so that $\varphi_{n^*}(0)\neq 0$, and therefore $\hat w$ has minimal degree n^*. We now note that, for all choices of α, whenever n^* is replaced by n^*-1, we must have

$$\begin{aligned}f_1(\alpha,\hat\gamma)=&\gamma_{n^*-1}+\alpha_1(1)\gamma_{n^*-2}+\alpha_2(1)\gamma_{n^*-3}\\ &-\cdots+\alpha_{n^*-1}(1)\gamma_0=\hat\gamma_{n^*-1}\neq 0.\end{aligned}$$

In particular, for no choice of α will the modeling filter corresponding to $(\alpha,\hat\gamma)$ have degree less than n^*. That is, if $\hat c$ is the partial covariance sequence corresponding to $\hat\gamma$, then $\hat c\in S(n^*)$. □

Lemma 5.13: Let n^* be any integer satisfying $0\le n^*\le n$. Then, there is an open neighborhood of $\hat\gamma$ such that, for all γ in this neighborhood, there does not exist an $\alpha\in\mathcal{P}_n(\gamma)$ for which the modeling filter corresponding to (α,γ) has degree less than n^*.

Proof: Suppose the contrary, so that there exists a sequence $\{\gamma^i\}$ approaching $\hat\gamma$ such that for each i there exists α^i for which the degree of the system determined by (α^i,γ^i) is less than or equal to n^*-1. In particular, we must have that $f_1(\alpha,\gamma)=0$ holds with n^* replaced by n^*-1. More

explicitly, we must have

$$f_1(\alpha^i, \gamma^i) = \gamma^i_{n^*-1} + \alpha^i_1(1)\gamma^i_{n^*-2} + \alpha^i_2(1)\gamma^i_{n^*-3} - \cdots + \alpha^i_{n^*-1}(1)\gamma^i_0 = 0.$$

Since $\mathcal{P}_n$ is relatively compact, the sequence (α^i, γ^i) has a cluster point and, by choosing a subsequence if necessary, we may assume that $(\alpha^i, \gamma^i) \to (\alpha^0, \hat{\gamma})$ where $(\alpha^0, \hat{\gamma}) \in \overline{\mathcal{P}_n}$. This, however, leads to a contradiction since

$$0 = f_1(\alpha^i, \gamma^i) \to f_1(\alpha^0, \hat{\gamma}) = \hat{\gamma}_{n^*-1} \neq 0.$$

□

Lemma 5.14: Let n^* be any integer satisfying $\frac{1}{2}n \leq n^* \leq n$. Then, $\Sigma(n^*)$ has a nonempty interior which contains $\hat{\gamma}$.

Proof: We shall show that there exists an open neighborhood $\Omega(n^*)$ of $\hat{\gamma}$ in $\mathbb{R}^n$ such that for all $\gamma \in \Omega(n^*)$, there exists an $\alpha \in \mathcal{P}_n(\gamma)$ for which the modeling filter corresponding to (α, γ) has degree less than or equal to n^*. Existence of $\Omega(n^*)$ follows from an application of the implicit function theorem to the equation $f(\alpha, \gamma) = 0$. For the parameter choice $(\hat{\alpha}, \hat{\gamma})$, recalling that the matrix $A(\gamma)$ in the fast filtering algorithm is upper triangular and that $\hat{\alpha} = 0$, an interesting but routine calculation shows that

$$\frac{\partial f}{\partial \alpha}(\hat{\alpha}, \hat{\gamma}) = \begin{bmatrix} \hat{\gamma}_{n^*-1} & * & \cdots & * & \cdots \\ 0 & \hat{\gamma}_{n^*-1} & \cdots & * & \cdots \\ \vdots & & \ddots & & \\ 0 & \cdots & 0 & \hat{\gamma}_{n^*-1} & \cdots \end{bmatrix}$$

which, under our hypothesis, has rank $n - n^*$. Therefore, augmenting the equation $f(\alpha, \gamma) = 0$ by adding the slack equations

$$\begin{cases} \alpha_{n-n^*+1} = 0 \\ \alpha_{n-n^*+2} = 0 \\ \quad \vdots \\ \alpha_n = 0 \end{cases} \tag{146}$$

we obtain a system of n equations $g(\alpha, \gamma) = 0$ for which the Jacobian matrix $(\partial g/\partial \alpha)(\hat{\alpha}, \hat{\gamma})$ has full rank. Therefore, for γ sufficiently near $\hat{\gamma}$, there exists an analytic function F such that $\hat{\alpha} = F(\hat{\gamma})$ and $(\alpha, \gamma) = (F(\gamma), \gamma)$ is a solution to $f_1(\alpha, \gamma) = 0$. Since, for such γ, $\mathcal{P}_n$ is a nonempty open set and F continuous, it follows that $\Omega(n^*)$ is a nonempty open neighborhood of $\hat{\gamma}$ in $\Sigma(n^*)$. □

These lemmas imply that $S(n^*)$ has a nonempty interior, for any integer n^* satisfying $\frac{1}{2}n \leq n^* \leq n$. We now show that if n^* satisfies $0 \leq n^* < \frac{1}{2}n$, then the interior of $S(n^*)$ is empty. To this end, consider that to each point c in $\mathbb{R}^n$ we assign the standard Hankel matrix

$$C = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \\ c_2 & c_3 & \cdots & c_{m+1} \\ \vdots & \vdots & \ddots & \\ c_m & c_{m+1} & \cdots & c_{2m-1} \end{bmatrix}$$

where either $n = 2m$ is even or $n = 2m+1$ is odd. We also consider the algebraic set

$$M_{n^*} = \{(c_1, c_2, \cdots, c_n) | \operatorname{rank} C \leq n^*\}.$$

Then (see [5] or Appendix B), M_{n^*} is an algebraic subset of $\mathbb{R}^n$ having geometric (and algebraic) dimension $2n^*$. In particular, if n^* satisfies $0 \leq n^* < \frac{1}{2}n$, then M_{n^*} has empty interior. Since M_{n^*} contains $S(n^*)$, if n^* satisfies $0 \leq n^* < \frac{1}{2}n$, then the subset $S(n^*)$ of $\mathcal{P}_n$ has empty interior as well. This concludes the proof of Theorem 2.2.

From Theorem 2.2 and from Corollary 5.11, we can deduce the following properties of $S(n)$, differing substantially from the deterministic case.

Corollary 5.15: The subset $S(n)$ is a closed semialgebraic subset of $\mathcal{C}_n \backslash \{0\}$ having a nonempty interior.

Appendix A
Positivity of Meromorphic and Rational Covariance Extensions

There are three principal constraints in the partial stochastic realization problem: rationality, positivity, and minimality of the (positive) degree. In Section III, we discussed classical and recent approaches to rationality, and the minimality of the algebraic degree. Positivity also has deep historical roots, going back to the classical Carathéodory extension problem [15], [16]. It is well known [27], [50] that to any sequence $(1, c_1, c_2, \cdots, c_n)$ one can bijectively assign a sequence $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ of *Schur parameters* defined in terms of the Szegö polynomials of the first kind $\{\varphi_0(z), \varphi_1(z), \varphi_2(z), \cdots\}$, a sequence of monic Schur polynomials

$$\varphi_t(z) = z^t + \varphi_{t1}z^{t-1} + \cdots + \varphi_{tt}$$

which are orthogonal on the unit circle [1], [27]. The Schur parameters are then given by

$$\gamma_t = \frac{1}{r_t}\sum_{k=0}^{t} \varphi_{t,\,t-k}c_{k+1} \tag{147}$$

where $r_0, r_1, r_2, \cdots$ and the coefficients $\{\varphi_{ti}\}$ can be determined recursively [1] by

$$r_{t+1} = (1 - \gamma_t^2)r_t, \qquad r_0 = 1 \tag{148}$$

and the Szegö–Levinson equations

$$\begin{cases} \varphi_{t+1}(z) = z\varphi_t(z) - \gamma_t\varphi_t^*(z), & \varphi_0(z) = 1 \\ \varphi_{t+1}^*(z) = \varphi_t^*(z) - \gamma_t z\varphi_t(z), & \varphi_0^*(z) = 1 \end{cases} \tag{149}$$

with $\varphi_t^*(z)$ being the reversed polynomials

$$\varphi_t^*(z) = \varphi_{tt}z^t + \varphi_{t,\,t-1}z^{t-1} + \cdots + 1.$$

For the sake of completeness, we also define the Szegö polynomials of the second kind, generated by the recursion

$$\begin{cases} \psi_{t+1}(z) = z\psi_t(z) + \gamma_t\psi_t^*(z), & \psi_0(z) = 1 \\ \psi_{t+1}^*(z) = \psi_t^*(z) + \gamma_t z\psi_t(z), & \psi_0^*(z) = 1. \end{cases} \tag{150}$$

Clearly $\{\psi_t(z)\}$ are obtained from $\{\varphi_n(z)\}$, by merely switching the signs of the Schur parameters.

These recursive schemes show that if $(\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$ are the Schur parameters of $(1, c_1, c_2, \cdots, c_m)$, then, for any $k < m$, $(\gamma_0, \gamma_1, \cdots, \gamma_{k-1})$ are the Schur parameters of

$(1, c_1, c_2, \cdots, c_k)$. It is a classical result of Schur [50] that these parameters satisfy the condition

$$|\gamma_i| < 1, \qquad i = 0, 1, \cdots, n-1 \tag{151}$$

if and only if the Toeplitz matrices

$$T_i = \begin{bmatrix} 1 & c_1 & \cdots & c_i \\ c_1 & 1 & \cdots & c_{i-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_i & c_{i-1} & \cdots & 1 \end{bmatrix}, \qquad i = 1, 2, \cdots, n \tag{152}$$

are positive definite. Moreover, there is a bijection [50] between the class of strictly positive real (meromorphic) functions $v(z)$ and the class of sequences $(\gamma_0, \gamma_1, \gamma_2, \cdots)$ satisfying

$$|\gamma_i| < 1, \qquad \text{for } i = 0, 1, 2, \cdots. \tag{153}$$

Now returning to the covariance extension problem, we note that if the partial covariance sequence $(1, c_1, c_2, \cdots, c_n)$ is given, then the first n Schur parameters $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ are determined, and they satisfy (151). Moreover, there is a one-to-one correspondence between positive extensions $c_{n+1}, c_{n+2}, c_{n+3}, \cdots$ and extensions

$$\gamma_n, \gamma_{n+1}, \gamma_{n+2}, \cdots, \qquad |\gamma_i| < 1. \tag{154}$$

Consequently, (154) is a complete parameterization of all strictly positive real meromorphic functions interpolating $1, c_1, c_2, \cdots, c_n$.

However, the basic question of which extensions (154) are rational remains open. Partial results in this direction are provided in [23] in terms of asymptotic properties of the Schur parameter sequence. For example, it is noted that for rational modeling filters the Schur sequence is square summable and asymptotically rational. As it turns out, these properties are a consequence of stable manifold theory for a certain dynamical system. Indeed, in [10] we derived lower and upper bounds on the decay rates of the Schur sequence by using an interpretation of the fast filtering algorithm [38], [39] as a nonlinear dynamical system in (α, γ) space [13]. In fact, the vector sequence $\gamma(t) = (\gamma_t, \gamma_{t+1}, \cdots, \gamma_{t+n-1})'$ is generated by the recursion

$$\alpha(t+1) = A(\gamma(t))\alpha(t), \qquad \alpha(0) = \alpha \tag{155a}$$

$$\gamma(t+1) = G(\alpha(t+1))\gamma(t), \qquad \gamma(0) = \gamma \tag{155b}$$

where the matrix functions $A, G\colon \mathbb{R}^n \to \mathbb{R}^{n\times n}$ are defined as in (156) and (157), shown at the bottom of the page. In particular, this means that the Schur parameters are updated according to the recursion

$$\begin{aligned}\gamma_{t+n} = &-\alpha_1(t+1)\gamma_{t+n-1} - \alpha_2(t+1)\gamma_{t+n-2} \\ &- \cdots - \alpha_n(t+1)\gamma_t.\end{aligned} \tag{158}$$

It can be shown that if $(\alpha, \gamma) \in \mathcal{P}_n$, i.e., if (α, γ) corresponds to a strictly positive real $v(z)$, then so does $(\alpha(t), \gamma(t))$. Moreover, if $a_t(z)$ and $b_t(z)$ are the $a(z)$ and $b(z)$ polynomials corresponding to $(\alpha(t), \gamma(t))$, the pseudo-polynomial

$$d(z, z^{-1}) = r_t[a_t(z)b_t(z^{-1}) + a_t(z^{-1})b_t(z)] \tag{159}$$

is invariant along the trajectory of (155).

Recall that as pointed out above, rationality requires that $\gamma_t \to 0$ as $t \to \infty$, and therefore $\varphi_k(z) \to z^k$ and $\psi_k(z) \to z^k$, so that $a_t(z)$ and $b_t(z)$ tend to the same limit $\alpha_\infty(z)$. But then we must have

$$a_t(z)b_t(z^{-1}) + a_t(z^{-1})b_t(z) = 2r_\infty\alpha_\infty(z)\alpha_\infty(z^{-1}) \tag{160}$$

and consequently $\alpha_\infty(z) = \sigma(z)$ and $r_\infty = \rho^2$. Therefore, if (α, γ) corresponds to a positive real $v(z)$, the trajectory of the dynamical system (155) tends to $(\sigma, 0)$. In particular, the maximum entropy solution, corresponding to $\alpha = 0$, will converge in n steps to (0, 0) so that $\sigma(z) = z^n$ and $\rho = \sqrt{r_n}$, in harmony with the result reported above.

Appendix B
Resultants and Resultant Varieties

The third major ingredient of the partial stochastic realization problem involves the notion of degree of a rational function, i.e., understanding when there exists a solution to the rational covariance extension problem having degree less than n. This occurs of course when the numerator b and

$$A(\gamma) = \begin{bmatrix} \dfrac{1}{1-\gamma_{n-1}^2} & \dfrac{\gamma_{n-1}\gamma_{n-2}}{(1-\gamma_{n-1}^2)(1-\gamma_{n-2}^2)} & \cdots & \dfrac{\gamma_{n-1}\gamma_0}{(1-\gamma_{n-1}^2)\cdots(1-\gamma_0^2)} \\ 0 & \dfrac{1}{1-\gamma_{n-2}^2} & \cdots & \dfrac{\gamma_{n-2}\gamma_0}{(1-\gamma_{n-2}^2)\cdots(1-\gamma_0^2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{1-\gamma_0^2} \end{bmatrix} \tag{156}$$

$$G(\alpha) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_n & -\alpha_{n-1} & -\alpha_{n-2} & \cdots & -\alpha_1 \end{bmatrix} \tag{157}$$

the denominator a of the rational interpolant have roots in common.

There are many classical approaches to determining whether a pair of polynomials have a root in common. Each of these gives rise to a particular polynomial test, typically that of determining whether the resultant of a and b vanishes. Indeed, it is because of the variety of possible constructions of the resultant that the following result (see, e.g., [36]) is so important.

Theorem B.1: The resultant of two polynomials a, b of a single complex variable z is, up to a nonzero multiplicative constant, the unique irreducible polynomial in the coefficients of a and b which vanishes if and only if the polynomials a and b have a root in common.

We denote by $R(a, b)$ the irreducible polynomial constructed as follows. Suppose the maximum of the degrees of a and b is n. Denote by V_{n-1} the n-dimensional vector space of polynomials of degree less than or equal to $n-1$. Consider the linear map $M_{(a,b)}$: $V_{n-1} \times V_{n-1} \to V_{2n-1}$ defined by

$$M_{(a,b)}(u, v) = au + bv. \tag{161}$$

Then the *resultant* of a and b is given by

$$R_{(a,b)} = \det M_{(a,b)}. \tag{162}$$

We observe that if (u, v) belongs to the nullspace, $\ker M_{(a,b)}$ of $M_{(a,b)}$, then

$$\frac{u}{v} = -\frac{b}{a}.$$

Since u and v are polynomials of degree at most $n-1$, this implies that a and b must have a nontrivial common factor $\theta = (a, b)$.

Remark B.2: More generally, the range space of $M_{(a,b)}$ consists of all polynomials $p \in V_{2n-1}$ having $\theta = (a, b)$ as a common factor, i.e.,

$$\operatorname{Im} M_{(a,b)} = \{p \in V_{2n-1} \mid p = q\theta\}. \tag{163}$$

We also remark that the matrix representation of $M_{(a,b)}$, with respect to the bases consisting of the standard monomials, is the matrix arising in the classical determinant expressions for the resultant (see, e.g., [36]). In particular, $R(a, b)$ is an irreducible polynomial. Alternatively, Kronecker constructed the resultant as the determinant of the Hankel matrix. We first note that this polynomial, which of course vanishes precisely when a and b have a root in common, is irreducible. Indeed, over the complex field, the zero locus, Z_n, consists of those complex Hankel matrices having rank less than or equal to $n-1$ and is well known to be an irreducible algebraic set. In particular, the determinant of the Hankel matrix is a nonzero real multiple of $R(a, b)$.

The hypersurface Z_n is of course singular; for example, the determinant and all of its partial derivatives vanish at the zero matrix. We are most interested in the set of regular points of Z_n, especially over $\mathbb{R}$, although our analysis extends *mutatis mutandis* to $\mathbb{C}$. Slightly modifying the notation in Brockett [5], we introduce the notation $\operatorname{Hank}(n, r)$ for the space of sequences $(c_1, c_2, \cdots, c_{2n})$ for which the corresponding Hankel matrix has rank r. In this notation

$$Z_n = \bigcup_{r \le n-1} \operatorname{Hank}(n, r). \tag{164}$$

$\operatorname{Hank}(n, n)$ is of course an open, dense subset of $\mathbb{R}^{2n}$ with Z_n as its complement. According to [5, Th. 1], for each $r < n$ the subset $\operatorname{Hank}(n, r)$ is a smooth manifold of dimension $n + r$. In particular, the set of regular points of Z_n coincides with $\operatorname{Hank}(n, n-1)$, i.e., with the pairs (a, b) having a simple common root.

We shall now describe the tangent space, $T_{(a,b)}(Z_n)$, to Z_n at a pair (a, b) in $\operatorname{Hank}(n, n-1)$. Tangent vectors to any point (a, b) in $\mathbb{R}^{2n}$ can be represented as pairs of arbitrary polynomials (u, v), where $\deg u \le n-1$ and $\deg v \le n-1$.

Proposition B.3: Suppose now that $(a, b) = \theta$, where $\deg \theta = 1$, and that $a = a_0\theta$ and $b = b_0\theta$. Then

$$T_{(a,b)}(Z_n) = \{(u_0\theta, v_0\theta) + \rho(a_0, b_0) \mid \deg u_0 \le n-2, \deg u_0 \le n-2, \rho \in \mathbb{R}\}. \tag{165}$$

We may of course take $R(a, b)$ as the defining equation for Z_n. It is then clear from calculating the Newton quotient for $R(a, b)$ that each vector in the right-hand side of (165) is contained in $T_{(a,b)}(Z_n)$. On the other hand, both sides of (165) are vector spaces of dimension $2n-1$, so that equality in (165) is obtained.

REFERENCES

[1] N. I. Akhiezer, *The Classical Moment Problem.* New York: Hafner, 1965.

[2] B. D. O. Anderson, "The inverse problem of stationary covariance generation," *J. Statistical Phys.*, vol. 1, pp. 133–147, 1969.

[3] M. Aoki, *State Space Modeling of Time Series.* New York: Springer-Verlag, 1987.

[4] K. J. Åström, *Introduction to Stochastic Realization Theory.* New York: Academic, 1970.

[5] R. W. Brockett, "The geometry of the partial realization problem," in *Proc. 1978 IEEE Decision Contr. Conf.*, pp. 1048–1052.

[6] C. I. Byrnes and A. Lindquist, "The stability and instability of partial realizations," *Syst. Contr. Lett.*, vol. 2, pp. 2301–2312, 1982.

[7] ———, "On the geometry of the Kimura–Georgiou parameterization of modeling filter," *Int. J. Contr.*, vol. 50, pp. 99–105, 1989.

[8] ———, "Toward a solution of the minimal partial stochastic realization problem," *Comptes Rendus Acad. Sci. Paris*, t. 319, Series I, pp. 1231–1236, 1994.

[9] ———, "Some recent advances on the rational covariance extension problem," in *Proc. IEEE European Wkshp. Computer-Intensive Methods Contr. Signal Processing*, Sept. 1994.

[10] C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Mateev, "A complete parametrization of all positive rational extensions of a covariance sequence," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1841–1857, 1995.

[11] C. I. Byrnes, A. Lindquist, and T. McGregor, "Predictability and unpredictability in Kalman filtering," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 563–579, 1991.

[12] C. I. Byrnes, A. Lindquist, and Y. Zhou, "Stable, unstable and center manifolds for fast filtering algorithms," in *Modeling, Estimation and Control of Systems with Uncertainty*, G. B. Di Masi, A. Gombani, and A. Kurzhanski, Eds. Boston: Birkhauser, 1991.

[13] ———, "On the nonlinear dynamics of fast filtering algorithms," *SIAM J. Contr. Optimization*, vol. 32, pp. 744–789, 1994.

[14] J. A. Cadzow, "Spectral estimation: An overdetermined rational model equation approach," *Proc. IEEE*, vol. 70, pp. 907–939, 1982.

[15] C. Carathéodory, "Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen," *Math. Ann.*, vol. 64, pp. 95–115, 1907.

[16] ———, "Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Functionen," *Rend. di Palermo*, vol. 32, pp. 193–217, 1911.

[17] J. B. Conway, *Functions of One Complex Variable.* New York: Springer-Verlag, 1993.
[18] Ph. Delsarte, Y. Genin, Y. Kamp, and P. van Dooren, "Speech modeling and the trigonometric moment problem," *Philips J. Res.*, vol. 37, pp. 277–292, 1982.
[19] C. J. Demeure and C. T. Mullis, "The Euclid algorithm and the fast computation of cross-covariance and autocovariance sequences," *IEEE Trans. Acoust., Speech Signal Processing*, vol. 37, pp. 545–552, 1989.
[20] L. Euler, "De Fractionibus continuis dissertatio," in *Proc. Nat. Acad. St. Petersburg*, Opera Omnia I.14, 1744 (in Latin); English translation by M. F. Wyman and B. F. Wyman in *Mathematical Systems Theory*, vol. 18, pp. 295–328, 1985.
[21] P. Faurre, M. Clerget, and F. Germain, *Opérateurs Rationnels Positifs.* Dunod, 1979.
[22] T. T. Georgiou, "Partial realization of covariance sequences," Ph.D. dissertation, Univ. Florida, Gainesville, 1983.
[23] ———, "Realization of power spectra from partial covariance sequences," *IEEE Trans. Acoust., Speech Signal Processing*, vol. ASSP-35, pp. 438–449, 1987.
[24] K. Glover, "All optimal Hankel norm approximations of linear multivariable systems and their L^∞ error bounds," *Int. J. Contr.*, vol. 39, pp. 1115–1193, 1984.
[25] W. B. Gragg, "Matrix interpretations and applications of the continued fraction algorithm," *Rocky Mountain J. Math.*, vol. 4, pp. 213–225, 1974.
[26] W. B. Gragg and A. Lindquist, "On the partial realization problem," *Linear Algebra its Appl.*, vol. 50, pp. 277–319, 1983.
[27] U. Grenander and G. Szegö, *Nonlinear Methods of Spectral Analysis.* Univ. California Press, 1958.
[28] S. Haykin, *Toeplitz Forms and Their Applications.* New York: Springer-Verlag, 1979.
[29] N. Jacobson, *Lectures in Abstract Algebra*, vol. III. New York: Van Nostrand, 1964.
[30] R. E. Kalman, "Realization of covariance sequences," in *Proc. Toeplitz Memorial Conf.*, Tel Aviv, Israel, 1981.
[31] ———, "On minimal partial realizations of a linear input/output map," in *Aspects of Network and System Theory*, R. E. Kalman and N. de Claris, Eds. New York: Holt, Reinhart and Winston, 1971, pp. 385–408.
[32] ———, "On partial realizations, transfer functions and canonical forms," *Acta Polytech. Scand.*, vol. MA31, pp. 9–39, 1979.
[33] S. M. Kay and S. L. Marple, Jr., "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, pp. 1380–1419, 1981.
[34] H. Kimura, "Positive partial realization of covariance sequences," *Modeling, Identification and Robust Control*, C. I. Byrnes and A. Lindquist, Eds. North-Holland, 1987, pp. 499–513.
[35] L. Kronecker, *Zur Teorie der Elimination einer Variabeln aus zwei algebraischen Gleichnungen.* Berlin: Monatsber, König. Preuss. Akad. Wiss., 1881.
[36] S. Lang, *Algebra.* New York: Addison-Wesley, 1970.
[37] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, vol. 45, pp. 255–282, 1950.
[38] A. Lindquist, "A new algorithm for optimal filtering of discrete-time stationary processes," *SIAM J. Contr.*, vol. 12, pp. 736–746, 1974.
[39] ———, "Some reduced-order non-Riccati equations for linear least-squares estimation: The stationary, single-output case," *Int. J. Contr.*, vol. 24, pp. 821–842, 1976.
[40] ———, "On Fredholm integral equations, Toeplitz equations and Kalman–Bucy filtering," *Appl. Math. Optimization*, vol. 1, pp. 355–373, 1975.
[41] A. Lindquist and G. Picci, "On the stochastic realization problem," *SIAM J. Contr. Optim.*, vol. 17, pp. 365–389, 1979.
[42] ———, "On 'subspace method identification' and stochastic model reduction," in *Proc. 10th IFAC Symp. Syst. Identification*, Copenhagen, June 1994, pp. 397–403.
[43] ———, *Canonical Correlation Analysis, Approximate Covariance Extension, and Identification of Stationary Time Series*, to be published.
[44] A. Magnus, "Certain continued fractions associated with the Padé table," *Math. Z.*, vol. 78, pp. 361–374, 1962.
[45] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, 1975.
[46] S. L. Marple, Jr., *Digital Spectral Analysis and Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1987.
[47] J. W. Milnor, *Singular Points on a Complex Hypersurface.* Princeton, NJ: Princeton Univ. Press.
[48] B. P. Molinari, "The time-invariant linear-quadratic optimal-control problem," *Automatica*, vol. 13, pp. 347–357, 1977.
[49] J. Rissanen, "Recursive identification of linear systems," *SIAM J. Contr.*, vol. 9, pp. 420–430, 1971.
[50] I. Schur, "On power series which are bounded in the interior of the unit circle I and II," *J. fur die reine und angewandte Math.*, vol. 148, pp. 122–145, 1918.
[51] P. van Overschee and B. De Moor, "Subspace algorithms for stochastic identification problem," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382–387, 1982.
[52] A. S. Willsky, *Digital Signal Processing and Control and Estimation Theory.* Cambridge, MA: MIT Press, 1979.

Christopher I. Byrnes (M'81–SM'86–F'89) received the B.S. degree from Manhattan College, Bronx, NY, in 1971 and the M.S. and Ph.D. degrees from the University of Massachusetts, Amherst, in 1973 and 1975, respectively.

He joined Washington University, St. Louis, in 1989 as Chairman of and Professor in the Department of Systems Science and Mathematics. Since 1991, he has been Dean of the School of Engineering and Applied Science. Prior to that, he held positions at Arizona State University, Harvard, and the University of Utah and was also an Adjunct Professor of mathematical system theory at the Royal Institute of Technology (KTH), Stockholm, Sweden. He has also held visiting positions at Bremen, Groningen, Harvard, IIASA, Soviet Union, Austria, Germany, Kansas, KTH, Osaka, Paris-Dauphine, La Sapienza, Rome, Stanford, and the University of Tokyo. His research interests include adaptive control, algebraic system theory, distributed parameter systems, linear multivariable control, nonlinear control, and the applications of nonlinear dynamics in control and estimation. He is the editor of 11 research volumes and the author of 150 technical articles.

Dr. Byrnes is a member of AAAS, AMS, and SIAM and was named a Fellow of the Japan Society for the Promotion of Science in 1986 and the Graduate School Distinguished Research Professor at Arizona State University in 1988.

Anders Lindquist (M'77–SM'86–F'89) was born in Lund, Sweden, in 1942. He received the civiling, TeknL, and TeknD degrees from the Royal Institute of Technology, Stockholm, Sweden, and in 1972 he was appointed a Docent of Optimization and Systems Theory there.

From 1972 to 1974, he held visiting positions at the University of Florida, Gainesville, Brown University, Providence, RI, and State University of New York, Albany. From 1975 to 1976, he was a SIAM Visiting Lecturer. In 1974, he became an Associate Professor and in 1980 a (full) Professor of Mathematics at the University of Kentucky, Lexington, where he remained until 1983. He is presently a Professor at the Royal Institute of Technology, where in 1982 he was appointed to the Chair of Optimization and Systems Theory and an Affiliate Professor at Washington University, St. Louis. He has also held visiting positions at the University of Padova, Italy, University of Arizona, USSR Academy of Sciences, Moscow, East China Normal University, Shanghai, and Technion, Haifa, Israel. He is the author of many papers in the area of systems and control, especially stochastic control, filtering, stochastic systems theory, realization theory, and applications of nonlinear dynamics in estimation and control.

Dr. Lindquist is a member of the Royal Swedish Academy of Engineering Sciences and an honorary member of the Hungarian Operations Research Society. He is a Communicating Editor of the *Journal of Mathematical Systems, Estimation, and Control* and, until 1993, he was an Associate Editor of *Systems and Control Letters.* He also serves on the editorial boards of *Adaptive Control* and *Signal Processing* and of the book series *Systems and Control: Foundations and Applications* and *Applied and Computational Control, Signals, and Circuits.* Until 1993, he also served on the editorial board of the book series *Progress in Systems and Control.*