RECENT PROGRESS ON THE PARTIAL STOCHASTIC REALIZATION PROBLEM

ANDERS LINDQUIST*

To Paul Fuhrmann on the occasion of his 60th birthday

In view of Paul Fuhrmann's many important contributions to realization theory, it seems quite appropriate to devote this lecture to the stochastic partial realization problem, when today we are honoring him on his 60th birthday. Some ten years ago Christopher I. Byrnes and I launched a joint research program on this topic, and by now we have some results which I think might interest this audience [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. Some of these results we have been obtain in collaboration with S. V. Gusev in particular, but also A. S. Matveev and H. J Landau. This short write-up is not a paper in itself but is merely intended to interest the audience in reading the papers [8, 7, 10, 11, 12] and also [32].

The stochastic partial realization problem has important applications in speech synthesis [17], spectral estimation [22, 33], stochastic systems theory [23], systems identification [32], and several other areas of systems and control.

1. The deterministic partial realization problem

Before turning to the main topic, let us consider for a moment the simpler deterministic partial realization problem; see, e.g., [24, 25, 21, 18]: Given a sequence of real numbers

$$c_0, c_1, c_2, \dots, c_n \tag{1.1}$$

find a *partial realization* of (1.1), i.e., a matrix F, a column vector g and a row vector h such that

$$c_k = hF^{k-1}g \quad \text{for } k = 1, 2, \dots, n. \tag{1.2}$$

This then defines an infinite extension

$$c_k = hF^{k-1}g \quad \text{for } k = n+1, n+2, \dots \tag{1.3}$$

of (1.1) such that the rational function

$$v(z) = h(zI - F)^{-1}g + \tfrac{1}{2}c_0 \tag{1.4}$$

has the Laurent expansion

$$v(z) = \tfrac{1}{2}c_0 + c_1 z^{-1} + c_2 z^{-2} + \dots$$

in the neighborhood of infinity. If the dimension of the matrix is a small as possible, we say that the partial realization is *minimal.*

* Division of Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, 100 44 Stockholm, Sweden; e-mail: alq@ math.kth.se

This problem can be motivated in the following way. Let us assume that we have a linear system

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

and that we would like to find matrices A, B, C, D so that

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ \quad\;\; y(t) = Cx(t) + Du(t) \end{cases} \tag{1.5}$$

models its input/output behavior. If we apply a impulse signal

$$u(t) = \begin{cases} 1 & \text{for } t = 0 \\ 0 & \text{for } t = 1, 2, 3, \dots \end{cases}$$

to the input, the output is

$$y(t) = \begin{cases} \frac{1}{2}c_0 & \text{for } t = 0 \\ CA^{t-1}B & \text{for } t = 1, 2, 3, \dots. \end{cases}$$

Thus, given the finite impulse response

$$y(0), y(1), \dots, y(n),$$

determining the matrices A, B, C, D obviously amounts to solving the partial realization problem.

We shall call the dimension of a minimal partial realization the *algebraic degree* of (1.1). The algebraic degree of (1.1) could be anything between zero and n, but it has a generic value $\left[\frac{n}{2}\right]$, and all other values are rare.

The deterministic partial realization problem is equivalent to Padé approximation and, as shown in [21], the solution can be determined recursively via Lanczos' algorithm [27]. There is no guarantee that F is Schur stable (all eigenvalues in the open unit disc). In fact, the property that F is Schur stable is not even generic [2].

From (1.2) it should be clear that a minimal partial realization (1.4) satisfies

$$H_{ij} := \begin{bmatrix} c_1 & c_2 & \cdots & c_j \\ c_2 & c_3 & \cdots & c_{j+1} \\ \vdots & \ddots & \vdots & \\ c_i & c_{i+1} & \cdots & c_{i+j} \end{bmatrix} = \begin{bmatrix} h \\ hF \\ \vdots \\ hF^{i-1} \end{bmatrix} \begin{bmatrix} g & Fg & \cdots & F^{j-1}g \end{bmatrix} \tag{1.6}$$

for each i, j such that $i + j = n$. Hence the maximum rank of these Hankel matrices is a lower bound for the algebraic degree of (1.1).

2. The stochastic partial realization problem

Now suppose A is a stability matrix. Then, passing (normalized) white noise $\{u(t)\}$ through the filter

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

and letting it come to statistical steady state, we obtain a stationary output process $\{y(t)\}$ with spectral density

$$\Phi(z) = w(z)w(z^{-1}) = \sum_{k=-\infty}^{\infty} c_k z^{-k}, \tag{2.1}$$

where

$$c_k = \mathrm{E}\{y(t+k)y(t)\}. \tag{2.2}$$

In view of the ergodic property of the process $\{y(t)\}$, the covariance lags (2.2) may also be represented as

$$c_k = \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} y_{t+k} y_t \tag{2.3}$$

almost surely, where

$$y_0, y_1, y_2, y_3, \dots$$

is a realization of the process $\{y(t)\}$. In practice, however, we only have access to finite string of output data

$$y_0, y_1, y_2, \dots, y_N. \tag{2.4}$$

In general N will be large, and therefore

$$c_k = \frac{1}{N+1-k} \sum_{t=0}^{N-k} y_{t+k} y_t \tag{2.5}$$

will be a reasonably good estimate. However, we can only estimate a finite number of covariance lags

$$c_0, c_1, c_2, \dots, c_n, \tag{2.6}$$

where $n << N$. For simplicity we assume that (2.6) is a *bona fide* partial covariance sequence in the sense that

$$T_n = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_n \\ c_1 & c_0 & c_1 & \cdots & c_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix}.$$

Of course, without loss of generality, we may normalize and take $c_0 = 1$. From now on we shall do this.

Now, reconstructing the filter $w(z)$ from the partial covariance sequence (2.6) is precisely the stochastic partial realization problem, first formulated in this context by Kalman [23]. Just as in the deterministic partial realization problem, we are faced with the problem of finding a triplet (F, g, h) such that

$$hF^{k-1}g = c_k \quad \text{for } k = 1, 2, \dots, n. \tag{2.7}$$

Then defining the infinite extension

$$c_k = hF^{k-1}g \quad \text{for } k = n+1, n+2, \dots, \tag{2.8}$$

the function

$$v(z) = \frac{1}{2} + c_1 z^{-1} + c_2 z^{-2} + \dots$$

is a rational function

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

where $a(z)$ and $b(z)$ are monic polynomials of degree n. However, there is now an additional requirement, namely that $v(z)$ is *strictly positive real*, i.e., $v(z)$ is analytic on and outside the unit circle and

$$\Phi(z) = v(z) + v(z^{-1}) > 0 \quad \text{on the unit circle.} \tag{2.9}$$

Then $\Phi(z)$ is a *bona fide* coercive spectral density, and therefore the spectral factorization problem

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

can be solved for the required filter, i.e., the minimum-phase, stable spectral factor

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

where we may choose coordinates so that $A = F$ and $C = h$.

Consequently we have a partial realization problem to determine a triplet (F, g, h) satisfying the interpolation condition (2.7), but with the additional constraint that $v(z)$ is strictly positive real. In particular this requires that F is Schur stable, but this is not enough. In fact, we must also have

$$a(z)b(z^{-1}) + b(z)a(z^{-1}) > 0 \quad \text{on the unit circle.}$$

These constraints make the stochastic partial realization problem considerably more complicated than the deterministic one. In particular, the infinite extension (2.8) is such that the Toeplitz matrix $T_k > 0$ for all $k \geq 0$.

Among all strictly positive real rational functions $v(z)$ interpolating the sequence $1, c_1, c_2, \dots, c_n$ in the sense that

$$v(z) = \tfrac{1}{2} + \hat{c}_1 z^{-1} + \hat{c}_2 z^{-2} + \dots \quad \text{with } \hat{c}_k = c_k \text{ for } k = 1, 2, \dots, n,$$

find one with minimum degree. This a minimum stochastic partial realization, and its degree p is called the *positive degree* of the partial covariance sequence $1, c_1, c_2, \dots, c_n$ and is greater or equal to the algebraic degree r of the same partial sequence. However, while r has a generic value, p does not. More precisely, whereas the generic value of the algebraic degree of (2.6) is $[\frac{n+1}{2}]$, it was proved in [7] that, for each $\nu = [\frac{n+1}{2}], [\frac{n+1}{2}] + 1, \dots, n$ there is a nonempty open set of covariance data in $\mathbb{R}^n$ for which the positive degree p is precisely ν. Thus the possibility that $p > r$ is nonrare. The positive degree can be characterized [7], but there is no easy way to compute it.

3. What are the solutions of the stochastic partial realization problem?

If we remove the requirement that $v(z)$ be rational of at most degree n and allow it to be only meromorphic, the stochastic partial realization problem is reduced to a classical interpolation problem going back to Carathéodory [14, 15], Toeplitz [35] and Schur [34]. Schur developed a complete parameterization of the class of such meromorphic and strictly positive real interpolants in terms of what we now know as the Schur parameters $\gamma_0, \gamma_1, \gamma_2, \dots$. In fact, to say that $v(z)$ is positive real is to say that

$$|\gamma_k| < 1 \quad \text{for } k = 0, 1, 2, \dots. \tag{3.1}$$

Moreover, $\gamma_0, \gamma_1, \dots, \gamma_{n-1}$ are uniquely determined (via, for example, the Levinson algorithm) by $c_1, c_2, \dots, c_n$. Therefore, all meromorphic and strictly positive real interpolants are completely parameterized by the free Schur parameters $\gamma_n, \gamma_{n-1}, \dots$ satisfying (3.1).

Making the very natural choice that $\gamma_k = 0$ for $k = n, n+1, \dots$, we do obtain a rational $v(z)$ of degree n, which is a solution of precisely the type we are looking for. This is the *maximum entropy solution,* and the corresponding filter $w(z)$ is

$$w(z) = \sqrt{r_n}\frac{z^n}{\varphi_n(z)}, \tag{3.2}$$

where $\varphi_n(z)$ is the n:th Szegö polynomial and $\sqrt{r_n}$ is the corresponding normalization factor, which both can be determined by solving the normal equations, a linear system of equations.

Obviously the maximum entropy solution has no nontrivial zeros, since the ones that are cancel when the spectral density $\Phi(z)$ is formed. A natural question, therefore, is as follows: Given any Schur stable monic polynomial $\sigma(z)$ of degree n, is it possible to find a Schur stable polynomial $a(z)$ so that

$$w(z) = \frac{\sigma(z)}{a(z)} \tag{3.3}$$

is a solution to the stochastic partial realization problem? In [20, 19] Georgiou proved that this is possible and conjectured that this choice would be unique.

In [8] we proved an amplified version of this longstanding conjecture, showing not only that this parameterization is complete but also that the problem is well-posed in a the strong sense that the bijection is a diffeomorphism. This is proved as corollary of a more general theorem on the geometric duality between filtering and interpolation. In fact, we prove that filtering and interpolation induce complementary decompositions (foliations) of the space of positive real functions of degree less or equal to n. The leaves of the interpolation foliation are indexed by the partial covariance sequences and the leaves of filtering foliation by the zero polynomials $\sigma(z)$. The unique pole polynomial $a(z)$ is determined via the intersection of the appropriate leaves. (For details, also see [10].) In passing we mention that the leaves of the filtering foliation are precisely the stable manifolds of the fast filtering algorithm [28, 29], properly reformulated as in [13].

In [10] and [11] we provide alternative, but simpler, proofs of the somewhat weaker statement that the stochastic partial realization problem is well-posed in the sense of Hadamard.

However, all these proofs, including the proof of existence due to Georgiou [20], are non-constructive and thus do not offer any computational procedure. It should be noted that, although the computation of the maximum entropy solution is a linear problem, obtaining the solution corresponding to an arbitrary zero polynomial $\sigma(z)$ is a *nonlinear problem*, which explains why it is much more difficult.

The first step toward finding a computational method for solving the stochastic partial realization problem was presented in [7], where the problem was reduced to solving a Riccati-type covariance extension equation. Although there is yet no general method of solving this equation, it provides some additional insight into the issue of minimality, relating the degree of the modeling filter to the rank of the unique positive semidefinite solution of the covariance extension equation.

In a recent paper [12], we present a convex optimization problem for solving the rational covariance extension problem. Given a partial covariance sequence and the desired zeros of the modeling filter (3.3), the poles are uniquely determined from the unique minimum of the corresponding optimization problem. In this way we obtain an algorithm for solving the covariance extension problem, as well as a constructive proof of Georgiou's existence result and his conjecture.

4. Why are we interested in this problem?

There are several interesting consequences and applications to what has been discussed above. We mention just two here and refer the audience to the reference list for further reading.

The maximum entropy solution is regularly used in speech coding in designing the so called LPC filter. The speech is typically broken into segments of 20 ms. An unvoiced segment can be regarded as a realization of a stationary stochastic process obtained by passing white noise through a modeling filter $w(z)$. (The voiced segments are modeled in a similar way exchanging the white noise for a periodic "pulse train".) Now, the disadvantage with the LPC filter is that the maximum entropy filter produces a rather "flat" speech, which does not represent nasals and fricative very well. For this we need to model the notches in the spectrum by placing zeros close to the unit circle. Hence a nontrivial zero polynomial $\sigma(z)$ would be desired. The more general solutions of the stochastic partial realization problem gives us the freedom to choose the zeros arbitrarily.

Recently there has been quite some interest in a new type of stochastic identification procedure known as subspace identification [1, 36]. In [32] it was pointed out that there is no guarantee that these subspace identification algorithms will actually work for generic data. This is for the reasons mention above. In fact, at least indirectly these algorithms are based on the assumption that the algebraic and positive degrees of a partial covariance sequence coincide. Also the geometry of [36] is equivalent to the splitting geometry of stationary stochastic systems [30, 31], developed for infinite strings of data, and which strictly speaking requires stronger conditions to hold here. In the context of this talk, the procedure of [36] is to perform minimal factorization of a square Hankel matrix H_{ii}, where $i = [n/2]$; see (1.6). But as pointed out above, the rank of H_{ii} only provides a lower bound of the algebraic degree, which in turn may be smaller than the positive degree, in which case the identification method will fail. It is easy to generate data for which such failure will occur even for large n; see [16].

References

1. M. Aoki, *State Space Modeling of Time Series*, Springer-Verlag, 1987.
2. C. I. Byrnes and A. Lindquist, *The stability and instability of partial realizations*, Systems and Control Letters **2** (1982), 99–105.
3. C. I. Byrnes and A. Lindquist, *An algebraic description of the rational solutions of the covariance extension problem*, in Linear Circuits, Systems and Signal Processing, C.I. Byrnes, C.F. Martin and R.E. Saeks (editors), Elsevier 1988, 9-17.
4. C. I. Byrnes and A. Lindquist, *On the geometry of the Kimura-Georgiou parameterization of modelling filter*, Inter. J. of Control **50** (1989), 2301–2312.
5. C. I. Byrnes and A. Lindquist, *Some recent advances on the rational covariance extension problem*, Proc. IEEE European Workshop on Computer-Intensive Methods in Control and Signal Processing, September 1994.
6. C. I. Byrnes and A. Lindquist, *Toward a solution of the minimal partial stochastic realization problem*, Comptes Rendus Acad. Sci. Paris, t. 319, Série I (1994), 1231–1236.
7. C. I. Byrnes and A. Lindquist, *On the partial stochastic realization problem*, IEEE Trans. Automatic Control AC-42 (1997), to be published.
8. C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev, *A complete parametrization of all positive rational extensions of a covariance sequence*, IEEE Trans. Automatic Control AC-40 (1995), 1841–1857.
9. C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev, *The geometry of positive real functions with applications to the rational covariance extension problem*, Proc. 33rd Conf. on Decision and Control, 3883–3888.
10. C. I. Byrnes and A. Lindquist, *On duality between filtering and interpolation*, in *Systems and Control in the Twenty-First Century*, C.I.Byrnes, B.N.Datta, D.S.Gilliam, and C.F.Martin (editors), pp. 101–136.
11. C. I. Byrnes, H. J. Landau, and A. Lindquist, *On the well-posedness of the rational covariance extension problem*, in *Current and Future Directions in Applied Mathematics*, M. Alber, B. Hu and J. Rosenthal (editors), pp. 81–108..
12. C. I. Byrnes, S. V. Gusev and A. Lindquist, *A convex optimization approach to the rational covariance extension problem*, submitted for publication.
13. C. I. Byrnes, A. Lindquist, and Y. Zhou, *On the nonlinear dynamics of fast filtering algorithms*, SIAM J. Control and Optimization, **32**(1994), 744–789.
14. C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen*, Math. Ann. **64** (1907), 95–115.
15. C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Functionen*, Rend. di Palermo **32** (1911), 193–217.
16. A. Dahlén, A. Lindquist, and J. Mari *Experimental evidence showing that stochastic subspace identification methods may fail*, to appear.
17. Ph. Delsarte, Y. Genin, Y. Kamp and P. van Dooren, *Speech modelling and the trigonometric moment problem*, Philips J. Res. **37** (1982), 277–292.
18. Paul A. Fuhrmann, *On the partial realizaion problem and the recursive inversion of Hankel and Toeplitz matrices*, Contemporary Mathematics **47** (1985), 249–161.
19. T. T. Georgiou, *Partial realization of covariance sequences*, CMST, Univ. Florida, Gainesville, 1983.
20. T. T. Georgiou, *Realization of power spectra from partial covariance sequences*, IEEE Transactions Acoustics, Speech and Signal Processing **ASSP-35** (1987), 438–449.
21. W. B. Gragg and A. Lindquist, *On the partial realization problem*, Linear Algebra and its Applications **50** (1983), 277–319.

22. S. Haykin, *Toeplitz forms and their applications*, Springer-Verlag, 1979.
23. R. E. Kalman, *Realization of covariance sequences*, Proc. Toeplitz Memorial Conference (1981), Tel Aviv, Israel, 1981.
24. R. E. Kalman, *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.), Holt, Reinhart and Winston, 1971, 385–408.
25. R. E. Kalman, *On partial realizations, transfer functions and canonical forms*, Acta Polytech. Scand. **MA31** (1979), 9–39.
26. H. Kimura, *Positive partial realization of covariance sequences*, Modelling, Identification and Robust Control (C. I. Byrnes and A. Lindquist, eds.), North-Holland, 1987, pp. 499–513.
27. C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Standards **45** (1950), 255–282.
28. A. Lindquist, *A new algorithm for optimal filtering of discrete-time stationary processes*, SIAM J. Control **12** (1974), 736–746.
29. A. Lindquist, *Some reduced-order non-Riccati equations for linear least-squares estimation: the stationary, single-output case*, Int. J. Control **24** (1976), 821–842.
30. A. Lindquist and G. Picci, *Realization theory for multivariate stationary Gaussian processes,* SIAM J. Control and Optimization **23** (1985), 809–857.
31. A. Lindquist and G. Picci, *A geometric approach to modelling and estimation of linear stochastic systems,* Journal of Mathematical Systems, Estimation and Control **1** (1991), 241–333.
32. A. Lindquist and G. Picci, *Canonical correlation analysis, aprroximate covariance extension, and identification of stationary time series*, Automatica **32** (1996), 709–733.
33. S. L.Marple, Jr., *Digital Spectral Analysis and Applications*, Prentice-Hall, 1987.
34. I. Schur, *On power series which are bounded in the interior of the unit circle I and II*, Journal fur die reine und angewandte Mathematik **148** (1918), 122–145.
35. O. Toeplitz, *Über die Fouriersche Entwicklung positiver Funktionen*, Rendiconti del Circolo Matematico di Palermo **32** (1911), 191–192.
36. P. van Overschee and B. De Moor, *Subspace algorithms for stochastic identification problem*, IEEE Trans. Automatic Control **AC-27** (1982), 382–387.