full paper submitted to IEEE Automatic Control

Characterizing 2^{nd} Order Statistics of Vector Random Processes and Other Significant Results and Applications of Matrix Spectral Factorization and Realization Theory

T. H. Kerr
Lincoln Laboratory of Massachusetts Institute of Technology,
244 Wood St.
Lexington, MA 02173-9108

Abstract

The topics of Matrix Spectral Factorization in conjunction with results from Realization Theory are applied here to simulating a stationary multi-input/multi-output (MIMO) linear system from a specified power spectral density matrix. Matrix Spectral Factorization provides the appropriate transfer function matrix and Realization Theory specifies the corresponding parameters of a linear system having this transfer function. This approach can be used to correctly capture the cross-correlations that exist in a multichannel vector random process. Aspects of the solution to this problem are illustrated using an original representative example problem with a closed-form answer. Existing software programs for accomplishing Matrix Spectral Factorization are identified and one has been successfully validated using the known closed-form solution mentioned above. This computer program along with a mechanization of the novel simplified realization algorithm (offered for the first time here as an original theoretical contribution) can now be applied to application data (but the resulting real world answers will, in general, consist of long unwieldy representations rather than the nice closed-form expressions offered here in a simple example that was expressly chosen to ease the documentation effort and to serve as a convenient test case). Besides multichannel spectral estimation, other important apparently disparate applications (some old, some new) of the same solution methodology are summarized here including applicability to the complex case (to handle polarization issues related to coherent phase processing), multiport network synthesis, MIMO Wiener filtering, MIMO matched filtering, exact (rather than approximate) estimators for situations involving presence of multiplicative noise, robust control and estimation (i.e., H^{∞}), and in re-expressing second order statistics of an ARMA process as those of a simpler but mathematically equivalent AR process of slightly higher dimensions.

1 Introduction

A high level presentation is pursued here to boil down and reduce the results to a form that is understandable to the non-specialist. The topic of Matrix Spectral Factorization is particularly amenable to this type of broad brush handling now since the original breakthrough papers of [1], [2] as well as all the later papers on this topic are concerned with actually constructing a solution rather than with merely proving that such a solution exists. A similar situation exists with Realization Theory and its results. The illustrative numerical examples provided herein are original and of low-dimension and of closed-form so as to not obfuscate the theoretical results being conveyed here yet these examples are nondegenerate and nontrivial so that they are still representative of the computational effort that must be expended in going from tractable single channel spectral factorization [10] to the more formidable Matrix Spectral Factorization handled here as it arises for the multichannel case. These same numerical examples have been used to serve as a reasonable basis

*This work was sponsored by the United States Department of the Air Force.

for relative comparisons between intermediate options as well as facilitating software verification as a test case of known analytic closed-form solution (used in the same vein as occurs in [40], [41]). Of particular interest is a new theoretical result given in Section 3 and cast here in the role of a supporting algorithm. Part of the contribution of Section 3.2 is in providing a new result that allows a multichannel vector Auto-Regressive Moving Average (ARMA) process to be recast as a simpler more tractable AR process, but of slightly higher vector dimensions.

2 The Thrust of the Investigation

The techniques to be investigated here are applicable to situations where second order statistics of a random process are available in either the time-domain or in the frequency domain, respectively, as the correlation matrix or the power spectral density matrix already having been specified. The techniques investigated here will be applied to a particular low dimensional example for concreteness. Handling the particular 2×2 example (introduced below) is representative of what must be done in the general multi-input/multi-output (i.e., multichannel) case. The multichannel case is orders of magnitude more challenging than the single channel case treated in [10], which is one of the reasons why a concrete example is being used to illustrate the necessary computations. The 2×2 example retains the sufficient generality of having non-vacuous cross-correlation in order that it not be trivial or degenerate as merely the conjunction of two independent single channel examples. For perspective, if simple closed-form results were not available, the simplified discussion pursued here would not be feasible because the computational aspects for the multichannel case would be overwhelming and the general expressions would be long and unwieldy. This is the reason why results from actual data are not used in this illustrative example. However, several computer programs for accomplishing Matrix Spectral Factorization are identified in [42, Chapter 3], and these programs can handle the complexities of real world application data once they are validated (perhaps through using the examples of known solution, as successfully done in [42, Chapter 3] for the software corresponding to [17]).

Given a specified rational power spectral density matrix (i.e., the elements being ratios of polynomials), it is desirable to implement a computer simulation or emulation of a random process that has this same specified rational power spectral density matrix. Suppose that a typical representative autocorrelation function matrix example of immediate interest is

$$R_{yy}(\tau) = \begin{bmatrix} \frac{1}{6}e^{-2|\tau|} + \frac{1}{6}e^{-|\tau|} & \vdots & \frac{1}{4}e^{-2|\tau|} \\ \cdots & \cdots & \cdots \\ \frac{1}{4}e^{-2|\tau|} & \vdots & \frac{1}{2}e^{-|\tau|} \end{bmatrix}, \tag{1}$$

then the associated power spectral density matrix (obtained as the Bilateral Laplace [1] transform of Eq. 1) is

$$\Phi_{yy}(s) = \begin{bmatrix} \frac{2-s^2}{(4-s^2)(1-s^2)} & \frac{1}{4-s^2} \\ \frac{1}{4-s^2} & \frac{1}{1-s^2} \end{bmatrix}, \tag{2}$$

where the "abscissa of convergence" in Eq. 2 is for $|Re(s)| < 1$. Further, suppose that it is desired to model the above continuous-time power spectral matrix using a continuous-time multi-input/multi-output linear shaping filter driven by white noise. The proper shaping filter to be used to accomplish

[1] Throughout this discussion, s is the complex bilateral Laplace transform variable as in the Laplace transform kernel $\exp(-st)$. The Bilateral Laplace transform is closely related to the Fourier transform (i.e., $s = j\omega$) but is more convenient for the manipulations to be performed here of spectral factorization.

this can be obtained by applying standard Matrix Spectral Factorization algorithms [1], [2], to yield

$$\Phi_{yy}(s) = W^T(-s)\, W(s). \tag{3}$$

A right hand superscript T on a matrix denotes the transpose of the matrix. Since the relationship between the input spectral matrix, $\Phi_{xx}(s)$, and the output spectral matrix, $\Phi_{yy}(s)$, for a strictly stable linear time-invariant system is

$$\Phi_{yy}(s) = H(-s)\, \Phi_{xx}(s)\, H^T(s) \tag{4}$$

where

$H(s)$ — a not necessarily square transfer function matrix of the linear system;

it is convenient to treat the modeling as involving a zero mean, uncorrelated Gaussian white noise input (having spectral matrix being the identity matrix)

$$\Phi_{xx}(s) = I_n \tag{5}$$

as driving the linear system to yield the following degenerate simplification of Eq. 4 as representing the output spectral matrix by

$$\Phi_{yy}(s) = H(-s)\, H^T(s)\,. \tag{6}$$

By an obvious association between the factors of Eq. 3 and the representation of Eq. 6, it is seen that an appropriate transfer function to model the power spectral matrix $\Phi_{yy}(s)$ is

$$H(s) \equiv W^T(s)\,. \tag{7}$$

That each approach to Matrix Spectral Factorization or that the final result of computations does in fact yield a valid solution can easily be verified or confirmed merely by multiplying the asserted solution matrix by its conjugate transpose to again obtain the original power spectral density matrix of Eq. 2 as a check. Additionally, one should then proceed by checking both the solution matrix and its inverse to be certain that no poles occur in the right half s-plane or on the $j\omega$-axis, otherwise the result does not correspond to a stationary system.

Once the transfer function matrix $H(s)$ is thus completely specified (as illustrated for the example of Eq. 2 in Appendices B, C, and D of [42] using alternative factorization approaches), it is frequently of interest to then go further to specify the constant parameter system matrices C, F, G of an associated linear system:

$$\frac{d}{dt}z(t) = F\, z(t) + G\, x(t) \tag{8}$$

$$y(t) = C\, z(t) \tag{9}$$

that can subsequently be used to realize or synthesize (either digitally or as an analog circuit) the transfer function matrix via the well-known relationship [7, p. 389]

$$H(s) = C\,(sI - F)^{-1}\, G\,. \tag{10}$$

The accepted standard approach to obtaining the requisite matrices C, F, G of minimum degree is via a generalized Hankel Matrix using the method of Kalman and Ho [6], [7, pp. 390-392] as discussed in Appendix A or as an update [22], [32]; however, approaches based on the Hankel matrix are notoriously numerically ill-behaved for computer computation of the answer. However, an original algorithm for cutting through and simply specifying the parameters C, F, and G explicitly is illustrated here in Section 3 as applied to the *completion* of the modeling of the example of Eq. 2. After providing this plausible new candidate solution approach, it is made rigorous through recourse

to guarantees provided by the theoretical conclusions of the existing Kalman-Ho technique (without the need to actually exercise the Kalman-Ho technique itself), which is less computationally tractable than the one we demonstrate here.

Once the linear system parameters (C, F, G) of the underlying dynamical system are available by the above described techniques, it is then a routine matter to use the historically well-known Monte-Carlo simulation techniques to emulate a Gaussian vector random process that possesses the desired matrix autocorrelation function or, equivalently, the desired power spectral density matrix. The Gaussianess that is also the associated goal for the generated output process may be obtained by any one of four standard approaches [30, Section 26.8.6a] within the random number generator used as the input driver. It is well known that random processes with identical power spectra can have sample function realizations of radically different character in the time domain. The main goal, however, is to faithfully reproduce the second order statistics for applications as in, say, serving as a credible decoy target in Electronic Warfare.

2.1 Necessary and Sufficient Conditions to Perform Matrix Spectral Factorization

Well-known necessary and sufficient conditions for being able to perform a factorization of an $n \times n$ modified Lambda-matrix $\Phi(s)$ are [1] that the following four conditions be satisfied: (1) the elements of $\Phi(s)$ are ratios of polynomials in the complex variable s (i.e., elements of $\Phi(s)$ are rational functions); (2) the matrix must be such that $\overline{\Phi(s)} = \Phi(\overline{s})$, where the vinculum denotes taking the complex conjugate; (3) the matrix must be such that $\Phi^T(-s) = \Phi(s)$ (denoted as the property of being an Hermitian matrix or of being paraconjugate-Hermitian); (4) the matrix must be such that the following property of being *positive real* holds where $\Phi(s)|_{s=j\omega} = \Phi(j\omega)$ must be positive semidefinite (denoted as $\Phi(j\omega) \geq 0$) for all real scalar ω. The above mathematical conditions guarantee that the act of matrix spectral factorization can be performed successfully for a matrix satisfying these conditions. Conditions 2 through 4 in the above list are satisfied by every power spectral density matrix and any arbitrary general power spectral density matrix can be approximated arbitrarily closely by a matrix that satisfies condition 1 in addition to satisfying conditions 2 through 4.

As pointed out in [4, Eq. 6, p. 283], a necessary and sufficient condition for a matrix $\Phi_{yy}(s)$ to be the power spectral density matrix of some wide sense stationary vector random process $y(t)$ is that $\underline{c}^T \Phi_{yy}(j\omega)\underline{c} \geq 0$ for every complex n-vector $\underline{c} \neq 0$ for all real scalar ω. This condition is the most severe of several conditions that must be satisfied in order that a modified Lambda-matrix be classified as being positive real (as defined in [3, p. 117] and [5, p. 171]). However, the fact that a matrix $\Phi_{yy}(s)$ satisfies this last condition is not equivalent to the matrix being "positive real" since in order to be so, several other conditions must also hold such as $\Phi_{yy}(s)$ being analytic in $\mathrm{Re}(s) > 0$, which is *not* satisfied by arbitrarily general power spectral density matrices.

However, an arbitrary power spectral density matrix satisfies the above conditions 2 through 4 and it may be uniformly approximated arbitrarily closely (via Wierstrass' approximation theorem) by rational functions to also satisfy condition 1 in addition to conditions 2 through 4. This objective of approximating stationary spectral densities in the frequency domain by rational functions can be practically accomplished using any of the following techniques: (1) interpolation; (2) approximation based on a Fourier series expansion; (3) approximation based on a Laguerre series expansion; (4) approximation of the logarithmic curve of the power spectral density by intersecting straight line segments (each explained in [9, Chapt. V]).

The requisite computed matrix power spectrum can be obtained from sample averages, then converted to the frequency domain via FFT's. An alternative well-known shortcut [8, Section 13-3] is to routinely FFT and record the accessible evolving sample functions, then shift and average the resulting FFT's of these sample functions directly in the frequency domain to obtain the corresponding power spectra in one fell swoop, thus circumventing the intermediate step of needing to explicitly

calculate the autocorrelation function in the time-domain. For numerically derived data (not yet approximated or fitted by the ratio of polynomials), a computational test of positive definiteness, rather than the analytical closed-form test utilized above, is available in [12].

2.2 Results of Applying Several Alternate Approaches to Matrix Spectral Factorization

The results of three different standard continuous-time analytical approaches in [1], [2] to Matrix Spectral Factorization are considered here as applied to the typical 2×2 example test problem of Eq. 2. This is done for the purpose of comparing the form of the various solutions (and to be aware of the relative computational complexity or computational burden to be incurred with each different approach for this same example).

For the specific example of Eq. 2, Matrix Spectral Factorization using Youla's first approach (detailed in [42, App. B]) yields the following factor as an answer

$$W(s) = \begin{bmatrix} \frac{-s-(\sqrt{7}/2)}{(2+s)(1+s)} & \frac{-s-(\sqrt{7}/2)}{(2+s)(1+s)} \\ \\ \frac{-1/2}{(2+s)(1+s)} & \frac{3/2}{(2+s)(1+s)} \end{bmatrix}, \tag{11}$$

which checks since for the factor in Eq. 11, the product $W^T(-s)\, W(s)$ is again the specific power spectral matrix of Eq. 2.

Again for the specific example of Eq. 2, Matrix Spectral Factorization using Youla's second approach (detailed in [42, App. C]) yields the following factor as an answer

$$W(s) = \begin{bmatrix} \frac{\sqrt{2}-s}{(2+s)(1+s)} & \frac{1-s}{(\sqrt{2}+s)(2+s)} \\ \\ 0 & \frac{\sqrt{7}-2s}{(\sqrt{2}+s)(2+s)(1+s)} \end{bmatrix}. \tag{12}$$

This result also checks since $W^T(-s)\, W(s)$ equals the original $\Phi_{yy}(s)$ of Eq. 2.

Again for the specific example of Eq. 2, Matrix Spectral Factorization using Davis' approach (detailed in [42, App. D]) yields the following factor as an answer

$$W(s) = \frac{\sqrt{2}}{(2+s)(1+s)} \begin{bmatrix} 1+(\sqrt{7}/4)s & (1/2)+(\sqrt{7}/4)s \\ s/4 & (\sqrt{7}/2)+(1/4)s \end{bmatrix}. \tag{13}$$

This answer checks as being a valid factorization since $W^T(-s)\, W(s)$ again equals the $\Phi_{yy}(s)$ in Eq. 2. For perspective, unabashedly performing Matrix Spectral Factorization by machine, even for simple problems, usually results in long unwieldy expressions for the answers. This original example handled here collapses to concise expressions not just for one but for three different solution approaches and is therefore a viable candidate to use in validating software implementation of Matrix Spectral Factorization for machine calculation, as further addressed in Section 2.6.

By observation, each of the above three alternate factorizations yield a 2×2 matrix result which, together with its inverse, is analytic in $\mathrm{Re}(s) > 0$. Thus each of the three factorizations above has the requisite properties and is a valid solution with which to proceed to realize or synthesize using the methods of Realization Theory as addressed in Section 3.

2.3 Relationship Existing Between Nonunique Alternative Factorizations

Matrix Spectral Factorizations are not unique and may differ by a "paraconjugate unitary" matrix pre-multiplying factor. A paraconjugate unitary matrix, $V(s)$, is a matrix having the following property that

$$V^T(-s)\, V(s) = I_n \tag{14}$$

where I_n is the $n \times n$ identity matrix.

Notice that if one factorization of $\Phi_{yy}(s)$ of Eq. 2 involves $W_1(s)$ as

$$\Phi_{yy}(s) = W_1^T(-s)\, W_1(s)\,, \tag{15}$$

then for any other $W_2(s)$ such that $W_2(s)$ is related to $W_1(s)$ as

$$W_2(s) = V_1(s)\, W_1(s)\,, \tag{16}$$

with $V_1(s)$ satisfying the condition of Eq. 14 as a paraconjugate unitary matrix, it follows that

$$\begin{aligned} W_2^T(-s)\, W_2(s) &= W_1^T(-s)\, V_1^T(-s)\, V_1(s)\, W_1(s) && (17)\\ &= W_1^T(-s)\, I_n\, W_1(s) && (18)\\ &= W_1^T(-s)\, W_1(s) && (19)\\ &= \Phi_{yy}(s) && (20) \end{aligned}$$

where, in the above, Eq. 14 was used in going between Eqs. 17 and 18 and via Eq. 15 in going between Eqs. 19 and 20. Therefore, by the simplified derivation offered above, $W_2(s)$ would be another valid factorization [2] of $\Phi_{yy}(s)$.

That the set of paraconjugate unitary matrices, $V(s)$, possessing the property or characteristic of Eq. 14 is neither vacuous nor trivial (as say being something as simple as the identity matrix) is illustrated next for the specific example of factoring the matrix portrayed in Eq. 2.

Via Eq. 16, the $W(s)$ that was obtained by Youla's first approach as Eq. 11, in the role of $W_1(s)$, to factorization and the $W(s)$ that was obtained here by Davis' factorization procedure as Eq. 12, in the role of $W_2(s)$, are related by

$$V(s)\frac{1}{(2+s)(1+s)}\begin{bmatrix} -s-(\sqrt{7}/2) & -s-(\sqrt{7}/2) \\ -1/2 & 3/2 \end{bmatrix} = \frac{1}{(2+s)(1+s)}\begin{bmatrix} \sqrt{2}+(\sqrt{14}/4)s & (\sqrt{2}/2)+(\sqrt{14}/4)s \\ \sqrt{2}s/4 & (\sqrt{14}/2)+(\sqrt{2}/4)s \end{bmatrix}.$$

The matrix $V(s)$ can be solved from the above equation via a matrix inversion to yield the following constant matrix

$$V(s) = V = \begin{bmatrix} -\frac{\sqrt{14}}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{4} & +\frac{7\sqrt{2}}{4} \end{bmatrix}.$$

However, the claim following Eq. 5 of [11] asserting that the $V(s)$ must be a constant matrix is *not* true. Even though Youla does treat some situations where $V(s)$ is a constant matrix (such as following Eq. 27 on p. 176 of [1] and at the end of the proof of Theorem 3 on p. 182 of [1]), in the general case $V(s)$ is *not* constant, a situation completely characterized in [1, Eqs. 164-164b] by Youla for the general case. Since [11] was very influential in its assertion that $V(s)$ must be a constant matrix and subsequently has been technically propagated and frequently parroted by all others that followed, it is important that this mis-assertion be refuted. A $V(s)$ that is *not* constant is offered here now.

[2] While factorizations are *usually* square and of the same dimension as the original $\Phi_{yy}(s)$, it is not mandatory that matrix factors always be square.

The $V(s)$ of Eq. 16 that relates the two factors of Eq. 12, in the role of $W_2(s)$, and Eq. 13, in the role of $W_1(s)$, is definitely not a constant matrix and serves as an exact matrix counterexample to the more recent assertions of [11, following Eq. 5] that $V(s)$ must always be a constant matrix. The $V(s)$ in this case is

$$V(s) = \frac{1}{(\sqrt{2}+s)(2s+\sqrt{7})} \begin{bmatrix} 4s^2-7 & -\sqrt{7}-2s \\ 2s-\sqrt{7} & 7-4s^2 \end{bmatrix} . \tag{21}$$

2.4 Attempting to Exploit a Relationship Between Matrix Spectral Factorization and Matrix Riccati Equation Solution

Rather than deal with the complexities of symbol manipulation as encountered in Matrix Spectral Factorization, many researchers realized by the beginning of the 1970's that there is a link between this problem and that of matrix Riccati equation solution [10, pp. 177-182], [11], [12]. However, Riccati equation solution, involving only numerical integration or application of a Newton-Raphson iterative procedure [7, pp. 358-363] (as a variant of the predecessor Kleinman's algorithm for the same purpose), is a much more straightforward problem since the matrices involved have only real (or complex) numbers as entries rather than ratios of polynomials of a complex variable, as encountered in Matrix Spectral Factorization. Yet [10], [12], and [45] only handle scalar one-dimensional spectral factorization problems via recourse to multidimensional matrix Riccati equations. Reference [11] ostensibly deals with the multidimensional Matrix Spectral Factorization problem via translation to a Riccati equation setting but encounters a barrier of a matrix inequality that needs to be satisfied and [11] then acknowledges that it is difficult to come up with a mechanism to force compliance [11, Section 3] even though some techniques were hypothesized in the conclusion section in [11, p. 414] but never demonstrated. [3] Thus the identification of an appropriate computer algorithm mechanization or program for accomplishing Matrix Spectral Factorization will be deferred to Section 2.6.

2.5 Update on Other Theoretical Developments Along the Principal Path of Automated Machine Calculation of Spectral Factors

Efficient greatest common denominator (GCD) extraction is central to most approaches to Matrix Spectral Factorization and is utilized in [15] (which serves as the basis for a modern implementation of Matrix Spectral Factorization being undertaken by Systems Engineering, Inc. [21]). An issue related to GCD extraction is "polynomial decomposition" such as that of [19]. Recent variations offer an algorithm for accomplishing this decomposition into proper polynomial factors or subfunctions in "polynomial time" rather than in "exponential time" as previously claimed for the algorithm of [19]. Perspectives on computational developments for determining when two polynomials are relatively prime beyond just using the fairly well-known "Euclid's algorithm" involve use of the "Cayley-Hamilton" theorem on "a square matrix satisfying its characteristic equation" as contorted into a slightly modified form, as discussed in [33].

Matrix Spectral Factorization, as discussed so far, is to be performed on either Lambda-matrices (polynomial matrices) or on *modified* Lambda-matrices whose entries are, respectively, either polynomials in a complex variable s with real rational coefficients or ratios of such polynomials. Factors interfering with or prohibiting our attaining otherwise straightforward operations counts for the Lambda-matrix operations arising in Matrix Spectral Factorization are: (1) aggregate number of polynomial entries to be dealt with (determined by dimension of the matrix); (2) need to computationally assess the degree of each polynomial entry; (3) complicating factors in manipulating several polynomials in performing elementary operations of addition, multiplication, subtraction, and division (as also encountered in MACSYMA) such as in determining when a resulting coefficient of a

[3] Additional perspective on this challenging problem is offered in [18].

specified fixed exponent of s is actually zero as contrasted to *appearing* to be zero due to the effects of round-off and truncation resulting from numerous operations as they affect the representation of numbers for the particular machine; (4) removing common factors that may occur in numerator and denominator to express results in simplest terms; (5) finding the least common denominator (LCD) of all the entries in a modified Lambda-matrix so that it may be multiplied throughout by this and therefore be converted to a more conveniently manipulated, exclusively polynomial Lambda-matrix, much more amenable to reduction to Smith Normal Form. Such are some of the issues along with others one is faced with in seeking to quantify the computational burden of Matrix Spectral Factorization that don't usually arise in other more conventional algorithms involving just numbers as entries in a matrix instead of symbol manipulation.

2.6 Computer Programs to Alleviate the Tedium and Minimize Human Error

Modern approaches to numerical computation of matrix factorizations were originally based on the observations of [11] that relate the Matrix Spectral Factorization problem to the solution of a matrix Riccati equation along with being required to satisfy a matrix inequality that was historically more difficult to enforce compliance with, as discussed in Section 2.4. Reference [20] seeks a more efficient computational approach for obtaining the greatest common denominator or divisor (GCD) that is needed for pre-multiplication in order to convert a modified Lambda-matrix (having entries consisting of ratios of polynomials) into Lambda-matrices (consisting of only polynomial entries) and thus easier to work with and/or easier to reduce to Smith Normal Form.

The 1971 approach of [7, Section 15.3, pp. 360-363] uses a Newton-Raphson iteration to solve the factorization problem (which had previously been observed in [11] to be equivalent to a steady-state matrix Riccati equation solution). The 1974 recursive Matrix Riccati equation solution technique of [27] does not require positive definiteness in the matrix to be factored. (This is a situation that frequently arises in network synthesis applications of Matrix Spectral Factorization but never arises in power spectral estimation applications.)

Despite the many new computer programs currently under development for accomplishing Matrix Spectral Factorization, most are either just getting started [45] [4] or are in a state of flux that will likely take considerable time for the transients to die down sufficiently to avail a usable trustworthy product. For the sake of conservatism and so that software to accomplish the objectives of Matrix Spectral Factorization with a demonstrated satisfactory track record and adequate documentation [17] at both the theoretical level and at the user level will be immediately available, software implementing the approach of [17] was procured in 1987 and successfully used, as exhibited in [42, Chap 3].

That similar considerations arise for both continuous-time and discrete-time formulations of the Matrix Spectral Factorization problem is evident from explanations including and following [16, Eq. 1] and from [17]. The discrete-time case involving the Z-transform complex variable z can either be handled directly via [14], [17] or via an easy conversion to the same form as would be encountered in the continuous-time case via a bilinear transformation (also known as a linear fractional transformation):

$$z = \frac{s+1}{1-s} \tag{22}$$

that re-expresses results in terms of the complex transform variable s. The inverse transformation of Eq. 22 (of similar form with signs on the 1's reversed) maps the inside of the unit circle in the z-plane into the left half s-plane. Concern about poles being present on the $j\omega$-axis in the s-plane

[4] Please notice that the only example of spectral factorization depicted in [45] is scalar (SISO) although the matrix (MIMO) case is ostensibly being pursued there.

being indicative of nonstationary systems now translate into worries in the discrete-time case about the presence of poles on the unit circle of the z-plane. Once the answers are obtained from the software, the transformation must be reversed, again a tedious burden and likely error source in hand calculations. However, the software of [17] performs these nasty little conversions for the user.

Tuel's Matrix Spectral Factorization program entitled "Wiener", as part of [17], can factor polynomial matrices of dimension 6×6 or less with order or degree not exceeding 20 for each polynomial entry or matrix element. The total software package requires approximately 10,000 computer words of storage on an IBM mainframe as host. Thus, Tuel's particular implementation of Matrix Spectral Factorization is bounded a posteriori or "after the fact" despite the difficulty encountered in predicting the associated computer burden beforehand, as discussed in Section 2.5.

3 Using Realization Theory to Convert Specified Transfer Function Matrices into State Variable Models and Associated Parameter Quantifications

Given a specified rational power spectral density matrix (i.e., the elements being ratios of polynomials), using Gaussian white noise as the input, it is desirable to implement a computer simulation or emulation of a random process that has this same specified rational power spectral density matrix. The next step is to go from the transfer function to the specification of the parameters of the underlying linear system that has this same transfer function. Realization Theory is the tool that accomplishes this. It is highly desirable that the resulting linear system be of minimal degree for the reasons offered in the first paragraph of Section 3.1. An apparently straightforward approach to how this linear system parameter specification can be accomplished will now be discussed in Sections 3.1.1 and 3.1.2 and made rigorous in Sections 3.1.3 and 3.2 and the results will be applied to the problem of simulating the multidimensional power spectrum of Eq. 2. An alternate conventional Realization Theory-based solution approach is also offered in Appendix A but it is less direct and more complicated to apply computationally than what is offered in Sections 3.1 for the application considered in this article. However, recourse is made to an interrelationship between the rigor available with the approach of Appendix A to ultimately also grant full rigor to the novel original approach of Sections 3.1.1 and 3.1.2

3.1 Taking an (Initially) Somewhat Heuristic but Direct Approach

As pointed out in [23, p. 324], [24], it is *impossible* to tell the order of the system simply from a cursory inspection of a specified transfer function block diagram since it is well known that a block diagram frequently exhibits more than the minimum number of integrators required to simulate the system. A design incorporating more integrators than are absolutely necessary is *extremely undesirable* not only from the obvious economic point of view (relating to the greater expense incurred in implementing additional integrators) but also from a stability viewpoint as explained next. Transfer functions represent the *observable and controllable* portion (in the sense of Kalman) and any unobservable integrators could possibly be going unstable without any obvious clues being externally visible until it is too late and something blows up internally and components are damaged.

A somewhat straightforward procedure to use in seeking an efficient simulation would be to obtain the per channel underlying differential equations describing the system directly from the transfer function and then proceed to obtain the desired final analog computer flow diagram directly from these intermediate differential equations. Even use of this method may initially indicate that differentiation of the input is needed, which is a situation that calls for use of more than the minimum number of integrators. "Differentiation of the input" (corresponding to having encountered zeroes

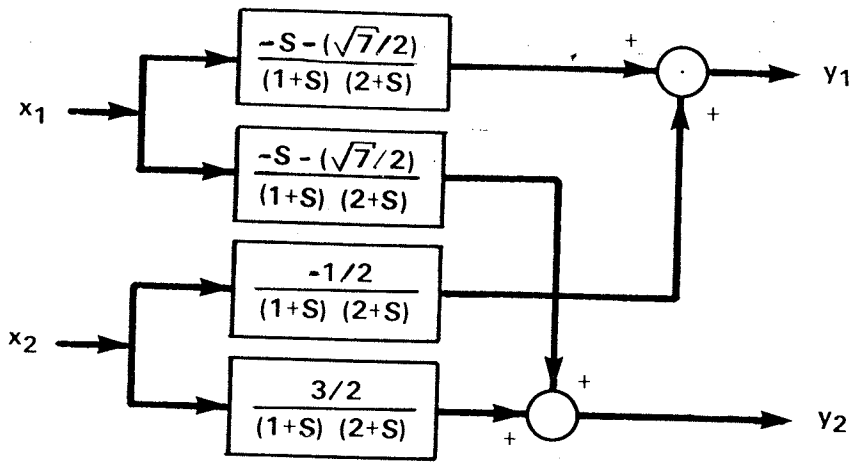

Figure 1: Two-Input/Two-Output Transfer Function Implementation of the Spectral Factor of Eq. 11.

of the transfer function) can be eliminated by applying the techniques found in [23, pp. 334-335] (and generalized here in Section 3.2 to handle multi-inputs). After applying this indicated technique to each individual channel, the inputs of the system-describing differential equations will no longer be differentiated and the resulting analog computer flow diagram so constructed from the reformulated or transformed differential equations that avoid "differentiation of the input" is less likely to have excess integrators. The per channel results are then aggregated as $\dot{x} = F_{TOTAL}\, x + G_{TOTAL}\, w\ ; y = H_{TOTAL}\, x$, where $F_{TOTAL} = \text{diag}\,[F_1, F_2, \ldots, F_m]$, $G_{TOTAL} = [G_1, \ldots, G_m]^T$, $H_{TOTAL} = \text{diag}\,[H_1, \ldots, H_m]$, and x is the augmented state vector of dimension $n = \sum_{i=1}^{m} n_i$. This is a *new* realization result different from the earlier results of Gilbert [28], of Kalman ([25], [26]), and of Kalman and Ho [6] that essentially resorts to block diagram algebra and honed intuition so familiar to classical control engineers (and easily implemented in a rule-based expert system using just 12 main rules [51, Section 7.5]) instead of the more mathematically esoteric (as well as less numerically stable) Hankel matrix-based approach, simply summarized in Appendix A. The obvious check is that the original transfer function matrix should satisfy the following as an identity: $\mathcal{H}(s) = H_{TOTAL}\ (sI - F_{TOTAL})^{-1}\ G_{TOTAL}$. For concreteness, the above mentioned technique is demonstrated next as applied to two different realizations or simulation implementations of systems, both having Eq. 2 as the output power spectral density matrix. Although initially advertised in Sections 3.1.1 and 3.1.2 as being merely heuristic, the results discussed in the second paragraph of Appendix A are used in Section 3.1.3 to make this original approach fully rigorous.

3.1.1 Heuristic Approach: Example 1

By the Matrix Spectral Factorization method used to obtain the result of Eq. 11, the resulting transfer function matrix is its transpose. The matrix transfer function block diagram, corresponding to the associated spectral factor that is to be the matrix transfer function via the standard association of Eq. 7, is depicted in Fig. 1. From the transfer function block diagram, the system differential equations are obtained which, unfortunately, also happen to contain differentiation of the inputs as the more general situation that is most likely to be encountered in practice. For this example, these differential equations are:

$$\ddot{y}_1 + 3\,\dot{y}_1 + 2\,y_1 = -\dot{x}_1 - \sqrt{7}/2\,x_1 - (1/2)\,x_2, \tag{23}$$

$$\ddot{y}_2 + 3\,\dot{y}_2 + 2\,y_2 = -\dot{x}_1 - \sqrt{7}/2\,x_1 + (3/2)\,x_2\,, \tag{24}$$

where in the above the use of $\cdot$ above a variable denotes differentiation with respect to time as $\frac{d}{dt}$.

In the next few steps, the procedure alluded to above and outlined in [23, pp. 334-335] for removing the differentiated input occurring in the differential equations is now applied here. Let

$$w_1 = y_1 \text{ and } w_2 = \dot{y}_1 - k_1\, x_1; \tag{25}$$

then

$$\dot{w}_1 = w_2 + k_1\, x_1 \tag{26}$$

and substituting the above in Eq. 23 and rearranging, yields

$$\dot{w}_2 = (-1 - k_1)\, \dot{x}_1 - (\sqrt{7}/2 + 3\, k_1)\, x_1 - (1/2)\, x_2 - 2\, w_1 - 3\, w_2\ ; \tag{27}$$

which upon taking

$$k_1 = -1, \tag{28}$$

yields

$$\dot{w}_2 = -2\, w_1 - 3\, w_2 + (6 - \frac{\sqrt{7}}{2})\, x_1 - (1/2)\, x_2. \tag{29}$$

Similarly, let

$$r_1 = y_2 \text{ and } r_2 = \dot{y}_2 - k_2\, x_1; \tag{30}$$

then

$$\dot{r}_1 = r_2 + k_2\, x_1\ , \tag{31}$$

and substituting the above in Eq. 24 and rearranging with

$$k_2 = -1, \tag{32}$$

yields

$$\dot{r}_2 = -2\, r_1 - 3\, r_2 + (6 - \frac{\sqrt{7}}{2})\, x_1 + (3/2)\, x_2. \tag{33}$$

The analog computer block diagram which summarizes this procedure is offered in Fig. 2, where x_1 and x_2 are independent Gaussian white noises with unit variance. Notice that only four integrators are required for this simulation!

The resulting augmented system, using the technique mentioned near the end of 2^{nd} paragraph of Section 3.1, can be reexpressed in vector/matrix form as

$$\frac{d}{dt}\begin{bmatrix} w_1 \\ w_2 \\ r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -3 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ r_1 \\ r_2 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ (6-\sqrt{7})/2 & -1/2 \\ -1 & 0 \\ (6-\sqrt{7})/2 & 3/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{34}$$

with

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ r_1 \\ r_2 \end{bmatrix}. \tag{35}$$

The stability of this augmented system can easily be checked by applying the Routh-Hurwitz criterion to the characteristic equation for the system matrix F_1 in Eq. 34. The characteristic equation is

$$\begin{aligned} 0 &= \det[\lambda I_4 - F_1] = \det\begin{bmatrix} \lambda & -1 & 0 & 0 \\ 2 & (\lambda+3) & 0 & 0 \\ 0 & 0 & \lambda & -1 \\ 0 & 0 & 2 & (\lambda+3) \end{bmatrix} \\ &= \lambda^4 + 6\lambda^3 + 13\lambda^2 + 12\lambda + 4. \end{aligned} \tag{36}$$

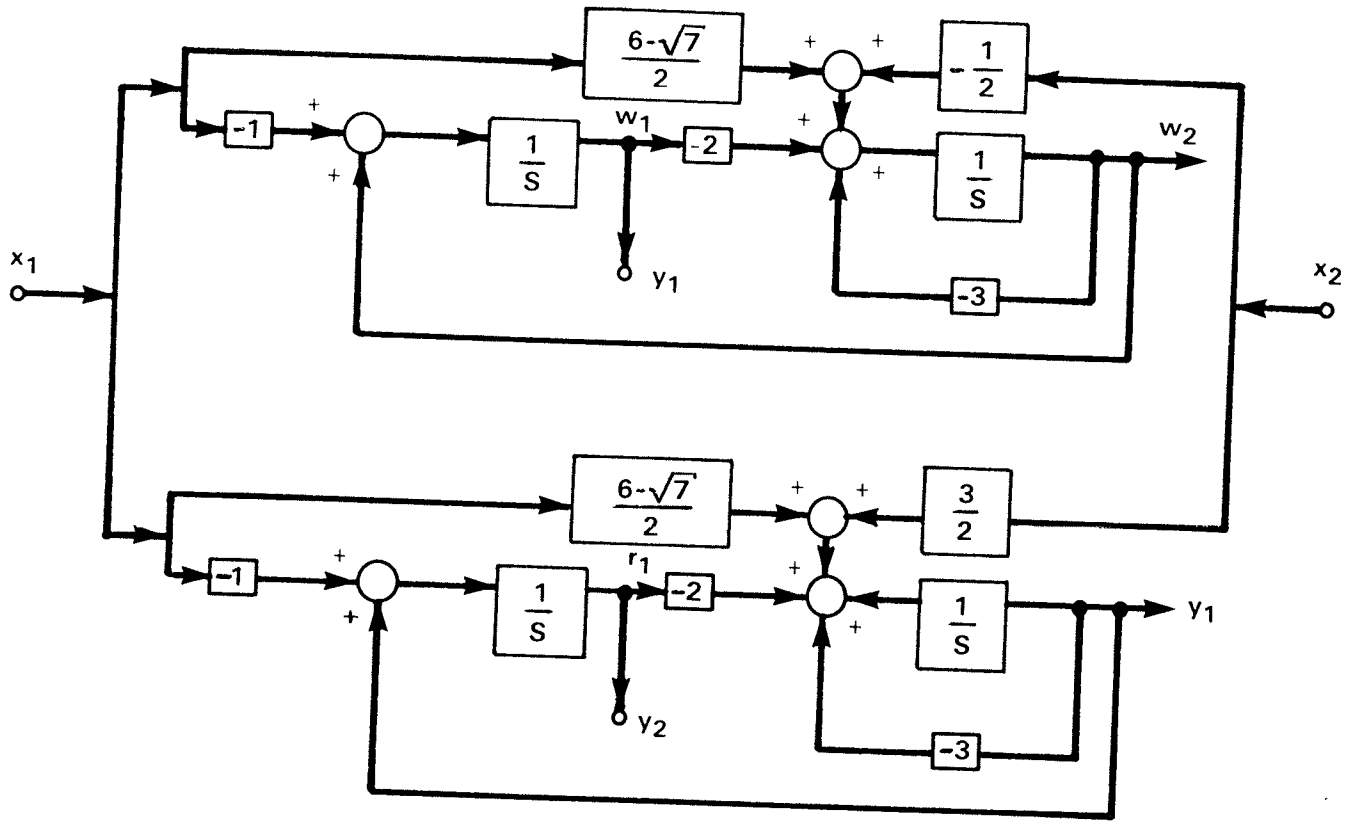

Figure 2: Implementation Detail of Transfer Function Realization of Matrix Factor Solution of Eq. 11.

The corresponding Routh-Hurowitz array, from which stability may be inferred, is

$$\begin{array}{c|ccc} \lambda^4 & 1 & 13 & 4 \\ \lambda^3 & 6 & 12 & \\ \lambda^2 & 11 & 4 & \\ \lambda^1 & \frac{108}{11} & & \\ \lambda^0 & 4. & & \end{array} \tag{37}$$

Since there are no changes in sign along the first column to the right of the vertical line in the above Routh-Hurowitz array, the interpretation from its use is that there are no zeroes of the characteristic equation that have real parts greater than or equal to zero. Therefore, this particular augmented system of Eqs. 34 and 35 is stable (and, as a consequence, is a stationary random process).

By the Kalman rank test on the controllability Grammian, for the augmented system (F_1, G_1) to be a controllable pair, it must be that

$$\text{rank}[G_1 \vdots F_1 G_1 \vdots F_1^2 G_1 \vdots F_1^3 G_1] = 4. \tag{38}$$

It is sufficient to just test the 4×4 matrix that results from adjoining the first two indicated components as:

$$\det\left[G_1 \vdots F_1 G_1\right] = \det\begin{bmatrix} -1 & 0 & (6-\sqrt{7})/2 & -1/2 \\ (6-\sqrt{7})/2 & -1/2 & (3\sqrt{7}/2)-7 & 3/2 \\ -1 & 0 & (6-\sqrt{7})/2 & 3/2 \\ (6-\sqrt{7})/2 & 3/2 & (3\sqrt{7}/2)-7 & -9/2 \end{bmatrix} = 5\sqrt{7} - 15 \neq 0. \tag{39}$$

Since this determinant is nonsingular, we have that Eq. 38 holds. This in turn assures that (F_1, G_1) is a controllable pair so the system of Eq. 34 is controllable.

By a similar Kalman rank test on the observability Grammian, for (C_1, F_1) to be an observable pair, it must be that

$$\text{rank}\left[C_1^T \vdots F_1^T C_1^T \vdots (F_1^T)^2 C_1^T \vdots (F_1^T)^3 C_1^T\right] = 4. \tag{40}$$

and since

$$\left[C_1^T \vdots F_1^T C_1^T\right] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{41}$$

has four independent columns, it follows that Eq. 40 does in fact hold. Therefore, the augmented system (F_1, C_1) is an observable pair so the system of Eqs. 34 and 35 is observable.

Further,

$$\begin{aligned} (sI - F_1)^{-1} &= \frac{\text{adj}[sI-F_1]}{\det[sI-F_1]} = \frac{\text{adj}[sI-F_1]}{(s^2+3s+2)^2} \\ &= \frac{1}{(s^2+3s+2)^2} \cdot \\ &\quad \begin{bmatrix} (s+3)(s^2+3s+2) & -2(s^2+3s+2) & 0 & 0 \\ (s^2+3s+2) & s(s^2+3s+2) & 0 & 0 \\ 0 & 0 & (s+3)(s^2+3s+2) & -2(s^2+3s+2) \\ 0 & 0 & (s^2+3s+2) & s(s^2+3s+2) \end{bmatrix} \\ &= \frac{1}{s^2+3s+2} \begin{bmatrix} s+3 & -2 & 0 & 0 \\ 1 & s & 0 & 0 \\ 0 & 0 & s+3 & -2 \\ 0 & 0 & 1 & s \end{bmatrix} . \end{aligned} \tag{42}$$

Further pre- and post-multiplying the result of Eq. 42 by the observation matrix of Eq. 35 and the noise gain matrix of Eq. 34, respectively, again yields the transfer function of Eq. 11, transposed, as a type of check on the correctness of these calculations.

3.1.2 Heuristic Approach: Example 2

By the Matrix Spectral Factorization method used to obtain the result of Eq. 12, the resulting transfer function matrix is its transpose. From the theory of the stochastic identification problem already illustrated in Section 3.1.1, we must similarly find the C_2, F_2, and G_2 such that

$$\begin{bmatrix} \frac{\sqrt{2}-s}{(2+s)(1+s)} & 0 \\ \frac{1-s}{(\sqrt{2}+s)(2+s)} & \frac{\sqrt{7}-2s}{(\sqrt{2}+s)(2+s)(1+s)} \end{bmatrix} = C_2(sI - F_2)^{-1}G_2. \tag{43}$$

By exactly the same method as used on the first factorization solution matrix (interpreted as a transfer function matrix specification), the differential equations for the y_i (i=1,2) are found to be:

$$\ddot{y}_1 + 3\dot{y}_1 + 2y_1 = -\dot{x}_1 + \sqrt{2}x_1, \tag{44}$$

$$\dddot{y}_2 + \left[3+\sqrt{2}\right]\ddot{y}_2 + \left[2+3\sqrt{2}\right]\dot{y}_2 + 2\sqrt{2}y_2 = -\ddot{x}_1 + \dot{x}_1 - 2\dot{x}_2 + \sqrt{7}x_2. \tag{45}$$

By again applying the method outlined in [23, pp. 334-335], as already applied to the first differential equation in Section 3.1.1, removes the differentiated input in Eq. 45 by use of the following

substitutions:

$$\begin{aligned} w_1 &= y_2, \\ w_2 &= \dot{y}_2 - k_1 x_1, \\ w_3 &= \ddot{y}_2 - k_1 \dot{x}_1 - k_2 x_1 - k_3 x_2, \end{aligned} \tag{46}$$

with

$$\begin{aligned} k_1 &= -1, \\ k_2 &= (3+\sqrt{2}), \\ k_3 &= -2 \end{aligned} \tag{47}$$

(the k_i's being chosen to eliminate the derivative terms of the input). The resulting state variable representation is:

$$\begin{bmatrix} \dot{w}_1 \\ \dot{w}_2 \\ \dot{w}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2\sqrt{2} & -(2+3\sqrt{2}) & -(3+\sqrt{2}) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 3+\sqrt{2} & -2 \\ (-8-3\sqrt{2}) & (6+2\sqrt{2}+\sqrt{7}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{48}$$

The method outlined in [23, pp. 334-335], when applied to the differential equation in Eq. 44, removes the differentiated input by use of the following substitutions:

$$\begin{aligned} r_1 &= y_1, \\ r_2 &= \dot{y}_1 - k_1 x_1, \end{aligned} \tag{49}$$

with

$$k_1 = -1. \tag{50}$$

The resulting state variable representation is

$$\begin{bmatrix} \dot{r}_1 \\ \dot{r}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 3+\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{51}$$

The full augmented system formed from these uncoupled state variable representations, using the technique near the end of the 2^{nd} paragraph of Section 3.1, is:

$$\begin{bmatrix} \dot{r}_1 \\ \dot{r}_2 \\ \dot{w}_1 \\ \dot{w}_2 \\ \dot{w}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -2 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -2\sqrt{2} & -(2+3\sqrt{2}) & -(3+\sqrt{2}) \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 3+\sqrt{2} & 0 \\ -1 & 0 \\ 3+\sqrt{2} & -2 \\ -8+\sqrt{2} & (6+2\sqrt{2}+\sqrt{7} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{52}$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} . \tag{53}$$

In the general theory for obtaining (C, F, G) from the factor $W^T(s)$ (which this second example closely follows), it is demonstrated that the augmented system is stable if and only if all of the individual subsystems are stable. Since the eigenvalues of the individual subsystems are obtained from the denominators of the elements of $W^T(s)$ and $W^T(s)$ is analytic in $\mathrm{Re}(s) \geq 0$ (as one of the properties inherited by valid Matrix Spectral Factorization solutions), it is known that all the subsystems are stable. Therefore, the augmented system F_2 is stable. Also, from the general theory for obtaining (C, F, G) from $W^T(s)$ by this method, it is demonstrated that (C, F) is an observable pair and that (F, G) is a controllable pair. Therefore, it is unnecessary to check (H_2, F_2) for observability or to check (F_2, G_2) for controllability since the general theory applies here to assure that they are in fact already observable and controllable pairs, respectively, by virtue of how they were constructed.

We have that

$$[sI - F_2]^{-1} = \begin{bmatrix} s+3 & 1 & 0 & 0 & 0 \\ -2 & s & 0 & 0 & 0 \\ 0 & 0 & \frac{s^2+(3+\sqrt{2})s+(2+3\sqrt{2})}{(s+\sqrt{2})} & \frac{s+3+\sqrt{2}}{s+\sqrt{2}} & \frac{1}{s+\sqrt{2}} \\ 0 & 0 & \frac{-2\sqrt{2}}{(s+\sqrt{2})} & \frac{s^2+(3+\sqrt{2})s}{s+\sqrt{2}} & \frac{s}{s+\sqrt{2}} \\ 0 & 0 & \frac{-2\sqrt{2}s}{s+\sqrt{2}} & \frac{-(2+3\sqrt{2})s-2\sqrt{2}}{s+\sqrt{2}} & \frac{s^2}{s+\sqrt{2}} \end{bmatrix} . \tag{54}$$

and the final check on (C_2, F_2, G_2) is that

$$\begin{aligned} C_2 [sI - F_2]^{-1} G_2 &= \frac{1}{(s^2+3s+2)} \cdot \begin{bmatrix} s+3 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{s^2+(3+\sqrt{2})s+(2+3\sqrt{2})}{s+\sqrt{2}} & \frac{s+3+\sqrt{2}}{s+\sqrt{2}} & \frac{1}{s+\sqrt{2}} \end{bmatrix} G_2 \\ &= \begin{bmatrix} \frac{\sqrt{2}-s}{(2+s)(1+s)} & 0 \\ \frac{1-s}{(\sqrt{2}+s)(2+s)} & \frac{\sqrt{7}-2s}{(\sqrt{2}+s)(2+s)(1+s)} \end{bmatrix} = W_2^T(s). \end{aligned} \tag{55}$$

Since from the above, $W_2^T(-s)W_2(s) = \Phi_{yy}(s)$, the triple (C_2, F_2, G_2) is also a valid implementation of a particular representation of a solution to Eq. 2.

3.1.3 Demonstrating These Realizations to be Minimal

Using a result of Gilbert [28], it is possible to easily test the above two realizations (obtained by the novel direct approach) for irreducibility. The supporting definitions and necessary theorems are presented below without proof (but a reference is supplied where the requisite proof may be found for each).

Theorem 1: Every stationary impulse-response matrix

$$h(t, \tau) = h(t - \tau), \tag{56}$$

having the property that

$$h(t - \tau) = \tilde{P}(t)\tilde{Q}(t) \text{ for all t}, \tau \tag{57}$$

has constant irreducible realizations. [5] All such realizations are strictly equivalent [26, p. 160].

For the restricted class of stationary impulse-response matrices, we have the following:

Definition: For a realization (C, F, G) of $W^T(s)$, the transfer function matrix $H(s)$ corresponding to the matrix $h(t-\tau)$, is *reducible* if and only if it has the following property: there exists a nonsingular matrix T such that proper submatrices (obtained by deleting a set of rows and the correspondingly numbered set of columns from $T\ F\ T^{-1}$; then deleting the same set of rows from $T\ G$ and the same set of columns from $C\ T^{-1}$) of $T\ F\ T^{-1}$, respectively, also constitute a realization of $W^T(s)$. Otherwise the realization is *irreducible* [25, p. 523].

The following theorem shows that any two irreducible realizations of $W^T(s)$ are strictly equivalent.

Theorem 2: Any two irreducible realizations (C, F, G) and $(\hat{C}, \hat{F}, \hat{G})$ of a proper rational matrix $W^T(s)$ are related as

$$\hat{F} = TFT^{-1}, \tag{58}$$

$$\hat{G} = TG, \tag{59}$$

$$\hat{C} = CT^{-1}, \tag{60}$$

where T is a nonsingular matrix.

For any two realizations related by a nonsingular constant matrix, as in the above Theorem 2, we have that

$$\begin{aligned} \hat{W}^T(s) &= \hat{C}(sI - \hat{F})^{-1}\hat{G} &&= CT^{-1}(sI - TFT^{-1})^{-1}TG \\ &= CT^{-1}\left[T(sI - F)T^{-1}\right]^{-1}TG &&= C(sI - F)^{-1}G \\ &= W^T(s). \end{aligned} \tag{61}$$

In other words, $W^T(s)$ is invariant under a change of the underlying coordinates describing the state vector x in the state space $\mathcal{R}^n$ (i.e., $W^T(s)$ is a coordinate-free description). Hence, we may select coordinates with respect to which F has a suitable canonical form [25, p. 526]. This is exactly what is being suggested by the technique that is being applied to the examples of Sections 3.1.1 and 3.1.2 (and at the beginning of Section 3.2) so that a suitable canonical form is obtained and demonstrably shown to be observable, controllable, and stable. The following theorem assures that the realizations described in Sections 3.1.1 and 3.1.2 are irreducible.

Theorem 3: A linear dynamical system is an irreducible realization of an impulse-response matrix if and only if the system is completely controllable and completely observable [26, p. 153].

In what follows, for the examples of stochastic or random process modeling previously appearing in Sections 3.1.1 and 3.1.2, each realization of both alternative factorizations will be shown to be irreducible by a relatively easy test proposed by Gilbert [28, p. 141] based on matrix partial fraction expansion [35]. Gilbert's test is also concisely summarized by Kalman [26, p. 180, Theorem 11] as the following theorem.

[5] Irreducible realizations are ones that are of minimum or least degree (i.e., requiring use of the smallest number of state variables or, equivalently, integrators).

Theorem 4: *Assumptions*: No element of the transfer matrix $W(s)$ has multiple poles. $W^T(s)$ has a total of q distinct poles $s_1, s_2, \ldots, s_q$, with corresponding matrices $R(1), R(2), \ldots, R(q)$, where

$$R(k) \triangleq \lim_{s \to s_k} (s - s_k) W^T(s). \tag{62}$$

Conclusions:

i The dimension of irreducible realizations of $W(s)$ is

$$n = \sum_{k=1}^{q} r_k \ , \tag{63}$$

where

$$r_k = \text{rank}\,[R(k)]\,. \tag{64}$$

ii Each matrix $R(k)$ may be factored (for example, by using an SVD algorithmic implementation) into

$$R(k) = H(k)G(k) \ \ (k = 1, 2, \ldots, q), \tag{65}$$

where $H(k)$ is an $(s \times r_k)$ matrix and $G(k)$ is an $(r_k \times m)$ matrix, both of rank r_k.

iii Finally, $W^T(s)$ has the associated irreducible realization being:

$$F = \text{diag}\,[s_1 I_{r_1}, \ldots, s_q I_{r_q}]\,, \tag{66}$$

$$G = [G(1), \ldots, G(q)]^T \ , \tag{67}$$

and

$$H = [H(1), \ldots, H(q)]\,, \tag{68}$$

where in the above, $I_r = (r \times r)$ identity matrix.

Returning now to apply Gilbert's test from the above Theorem 4 to the following numerical example of Section 3.1.1 having the matrix factor

$$W_1^T(s) = \frac{1}{(2+s)(1+s)} \begin{bmatrix} -s - \sqrt{7}/2 & -1/2 \\ -s - \sqrt{7}/2 & 3/2 \end{bmatrix} \tag{69}$$

yields

$$R(1) = \lim_{s \to -2} (2+s) W_1^T(s) = \begin{bmatrix} (\sqrt{7}/2) - 2 & 1/2 \\ (\sqrt{7}/2) - 2 & -3/2 \end{bmatrix}\,, \tag{70}$$

$$R(2) = \lim_{s \to -1} (1+s) W_1^T(s) = \begin{bmatrix} 1 - (\sqrt{7}/2) & -1/2 \\ 1 - (\sqrt{7}/2) & 3/2 \end{bmatrix}\,, \tag{71}$$

with

$$r_1 = \text{rank}\,[R(1)]\,, \tag{72}$$

$$r_2 = \text{rank}\,[R(2)]\,, \tag{73}$$

and for this example

$$\sum_{i=1}^{2} r_i = 2 + 2 = 4. \tag{74}$$

The conclusion of Gilbert's test states that $W_1^T(s)$ cannot be realized using a square F of dimension less than $n = \sum_{i=1}^{2} r_i = 4$. Since F_1 is 4×4, this realization is minimal.

Somewhat surprisingly, applying Gilbert's test to the example of Section 3.1.2 yields

$$R(1) = \lim_{s \to -\sqrt{2}} (\sqrt{2}+s) W_2^T(s) = \frac{1}{(2-\sqrt{2})(1-\sqrt{2})} \begin{bmatrix} 0 & 0 \\ -1 & \sqrt{7}+2\sqrt{2} \end{bmatrix}, \tag{75}$$

$$R(2) = \lim_{s \to -1} (1+s) W_2^T(s) = \frac{1}{(\sqrt{2}-1)} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{7}+2 \end{bmatrix}, \tag{76}$$

$$R(3) = \lim_{s \to -2} (2+s) W_2^T(s) = \frac{1}{(2-\sqrt{2})} \begin{bmatrix} -2 & 0 \\ -1 & \sqrt{7}+4 \end{bmatrix}, \tag{77}$$

with

$$r_1 = \text{rank}\,[R(1)] = 1, \tag{78}$$

$$r_2 = \text{rank}\,[R(2)] = 2, \tag{79}$$

$$r_3 = \text{rank}\,[R(3)] = 2, \tag{80}$$

and

$$\sum_{i=1}^{3} r_i = 1 + 2 + 2 = 5. \tag{81}$$

Since $W_2^T(s)$ cannot be realized using any F of dimension less than $n = \sum_{i=1}^{3} r_i = 5$, the F_2 exhibited is evidently also a realization that is a minimal realization for this particular factor. Although both factors used in Sections 3.1.1 and 3.1.2 correspond to the same underlying specified output power spectral density matrix, the two final models are of different dimensions.

The observation that once *any* realization of an impulse-response matrix (or, equivalently, a transfer function matrix) is specified or given (as done above as the preliminary step), then [32, Theorems 6.5, 6.6] claims that it can be used to reduce a nonminimal realization to one that is minimal. Our experience is that the result of our preliminary step have always already been minimal so we haven't had to further invoke using the Kalman-Ho procedure. This is the crux that makes the technique offered in Section 3 rigorous. Now, according to [32, p. 117], since both of the above considered realizations are minimal or irreducible realizations for their particular transfer function matrix instantiations or factorizations of the same underlying power spectrum of Eq. 2, one should then apply Occam's razor ruthlessly and choose the simpler of the two realizations (being the representation of Eqs. 34, 35 having the lesser dimension of 4 rather than 5) to actually implement for emulating the power spectrum of Eq. 2.

3.2 A Final Algorithm that *is* Rigorous

Insights gained from the results of the previous two modeling examples of Section 3.1 are useful in establishing general properties of the algorithm that was invoked both in Sections 3.1.1 and 3.1.2 for getting rid of differentiated input (i.e., for getting rid of the MA portion of an ARMA process to result in just a pure AR process of slightly greater dimension). This approach is formally generalized here and demonstrated to be a totally rigorous approach from which to validly obtain a state variable model which is a minimal realization (by virtue of convenient supporting results now available and presented in Appendix A). The main innovative theoretical contribution of this article is summarized below.

Consider the scalar constant coefficient linear differential equation of order n_i represented by the following:

$$y^{(n_i)} + \sum_{\ell=0}^{(n_i-1)} a_\ell y^{(\ell)} = \sum_{s=0}^{n_i} \sum_{j=1}^{r} b_{j_s} x_j^{(s)}, \tag{82}$$

where the superscript in parentheses denotes the order of the derivative with respect to time, t. Define auxiliary variables z_i $(i = 1, \ldots, n_i)$ such that:

$$z_1 = y - \sum_{j=1}^{r} k_{1j} x_j, \tag{83}$$

$$z_2 = y^{(1)} - \sum_{j=1}^{r} k_{1j} x_j^{(1)} - \sum_{j=1}^{r} k_{2j} x_j = y^{(1)} - \sum_{m=1}^{2} \sum_{j=1}^{r} k_{mj} x_j^{(2-m)}, \tag{84}$$

$$\vdots$$

$$z_\ell = y^{(\ell-1)} - \sum_{m=1}^{\ell} \sum_{j=1}^{r} k_{mj} x_j^{(\ell-m)}, \tag{85}$$

$$z_{\ell+1} = y^{(\ell)} - \sum_{m=1}^{\ell+1} \sum_{j=1}^{r} k_{mj} x_j^{(\ell+1-m)}, \tag{86}$$

$$\vdots$$

$$z_{n_i} = y^{(n_i-1)} - \sum_{m=1}^{n_i} \sum_{j=1}^{r} k_{mj} x_j^{(n_i-m)}, \tag{87}$$

where, for each j, the sequence $\{k_{mj}\}_{m=1}^{n_i}$ must yet be specified. Determining just what the sequence of constants $\{k_{mj}\}_{m=1}^{n_i}$ must be in order that all the differentiated input terms of the x_i's be removed so that the resulting differential equation in z_i's may be represented in standard state variable form is the goal of the analysis of this section.

From the above, it can be seen that

$$\frac{d}{dt} z_1 = y^{(1)} - \sum_{j=1}^{r} k_{1j} x_j^{(1)} = z_2 + \sum_{j=1}^{r} k_{2j} x_j; \tag{88}$$

indeed, the general expression can be seen to be

$$\frac{d}{dt} z_\ell = y^{(\ell)} - \sum_{m=1}^{\ell} \sum_{j=1}^{r} k_{mj} x_j^{(\ell+1-m)} = z_{\ell+1} + \sum_{j=1}^{r} k_{(\ell+1)j} x_j \ , \tag{89}$$

or

$$\frac{d}{dt} z_\ell = z_{\ell+1} + \sum_{j=1}^{r} k_{(\ell+1)j} x_j \ \text{ for } \ 1 \leq \ell \leq (n_i - 1). \tag{90}$$

Obtaining expressions for the y_i's from the equations above, we have that

$$y = z_1 + \sum_{j=1}^{r} k_{1j} x_j, \tag{91}$$

$$y^{(1)} = z_2 + \sum_{m=1}^{2} \sum_{j=1}^{r} k_{mj} x_j^{(2-m)}, \tag{92}$$

$$\vdots$$

$$y^{(\ell)} = z_{\ell+1} + \sum_{m=1}^{\ell+1} \sum_{j=1}^{r} k_{mj} x_j^{(\ell+1-m)}, \tag{93}$$

$$\vdots$$

$$y^{(n_i-1)} = z_{n_i} + \sum_{m=1}^{n_i}\sum_{j=1}^{r} k_{mj} x_j^{(n_i-m)}. \tag{94}$$

Using all of the above expressions for the $y^{(\cdot)}$'s and substituting into the original differential equation of Eq. 82, we obtain an expression involving $\frac{d}{dt}z_{n_i}$; this expression is:

$$\frac{d}{dt}z_{n_i} + \sum_{m=1}^{n_i}\sum_{j=1}^{r} k_{mj} x_j^{(n_i+1-m)} = y^{(n_i)} = -\sum_{\ell=0}^{(n_i-1)} a_\ell y^{(\ell)} + \sum_{s=0}^{n_i}\sum_{j=1}^{r} b_{js} x_j^{(s)}. \tag{95}$$

Now by working with the above expression to eliminate all expressions involving y's and to maneuver it into a more manageable form by performing changes of the dummy index of summation and by performing interchanges in the order of summation, the resulting expression is:

$$\begin{aligned}
\frac{d}{dt}z_{n_i} &= -\sum_{\ell=0}^{(n_i-1)} a_\ell \left[z_{\ell+1} + \sum_{t=1}^{(\ell+1)}\sum_{j=1}^{r} k_{tj} x_j^{(\ell+1-t)} \right] \\
&\quad + \sum_{s=o}^{n_i}\sum_{j=1}^{r} b_{js} x_j^{(s)} - \sum_{m=1}^{n_i}\sum_{j=1}^{r} k_{mj} x_j^{(n_i+1-m)} \\
&\quad - \sum_{\ell=0}^{(n_i-1)} z_{\ell+1} + \\
&\quad \sum_{j=1}^{r} \left[- \sum_{\ell=0}^{(n_i-1)}\sum_{t=1}^{(\ell+1)} a_\ell k_{tj} x_j^{(\ell+1-t)} + \sum_{s=0}^{n_i} b_{js} x_j^{(s)} - \sum_{m=1}^{n_i} k_{mj} x_j^{(n_i+1-m)} \right].
\end{aligned} \tag{96}$$

Regrouping two of the terms of the above summation, we have that

$$\begin{aligned}
\frac{d}{dt}z_{n_i} &= -\sum_{\ell=0}^{(n_i-1)} a_\ell z_{\ell+1} \\
&\quad + \sum_{j=1}^{r} \left[- \sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell} a_\ell k_{tj} x_j^{(\ell+1-t)} - \sum_{\ell=0}^{(n_i-1)} a_\ell k_{(\ell+1)j} x_j + b_{j0} x_j + \sum_{s=1}^{n_i} b_{js} x_j^{(s)} - \sum_{m=1}^{n_i} k_{mj} x_j^{(n_i+1-m)} \right].
\end{aligned}$$

Making a change of the dummy index of summation and regrouping terms, we have that

$$\begin{aligned}
\frac{d}{dt}z_{n_i} &= -\sum_{\ell=0}^{(n_i-1)} a_\ell z_{\ell+1} + \sum_{j=1}^{r} \left[b_{j0} - \sum_{\ell=0}^{(n_i)} a_\ell k_{(\ell+1)j} \right] x_j \\
&\quad + \left\{ \sum_{j=1}^{r} \left[- \sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell} a_\ell k_{tj} x_j^{(\ell+1-t)} + \sum_{m=1}^{n_i} b_{j(n_i+1-m)} x_j^{(n_i+1-m)} + \sum_{m=1}^{n_i} k_{mj} x_j^{(n_i+1-m)} \right] \right\}.
\end{aligned}$$

However, we want to have the following hold

$$\frac{d}{dt}z_{n_i} = -\sum_{\ell=0}^{(n_i-1)} a_\ell z_{\ell+1} + \sum_{j=1}^{r} \left[b_{j0} - \sum_{\ell=0}^{(n_i-1)} a_\ell k_{(\ell+1)j} \right] x_j \ , \tag{97}$$

so we must require that the term in braces in the above unnumbered equation preceding Eq. 97 be absent (which occurs when it is equivalent to zero) as enforced when the following condition or constraint is satisfied:

$$\sum_{j=1}^{r}\left[-\sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell}a_\ell k_{tj}x_j^{(\ell+1-t)}+\sum_{m=1}^{n_i}b_{j(n_i+1-m)}x_j^{n_i+1-m)}+\sum_{m=1}^{n_i}k_{mj}x_j^{(n_i+1-m)}\right]\equiv 0\,. \tag{98}$$

Therefore, we must find or discover which values of the finite sequence or collection of constants $\{k_{mj}\}_{m=1}^{n_i}$ (whose specification is at the discretion of the user) are required in order that the condition of Eq. 98 holds. Furthermore, the above equation implies that for each fixed j, it must be the case that

$$\left[-\sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell}a_\ell k_{tj}x_j^{(\ell+1-t)}+\sum_{m=1}^{n_i}b_{j(n_i+1-m)}x_j^{(n_i+1-m)}+\sum_{m=1}^{n_i}k_{mj}x_j^{(n_i+1-m)}\right]\equiv 0\,. \tag{99}$$

Separating terms in $x_j^{n_i}$ in the above, we obtain

$$\left[-\sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell}a_\ell k_{tj}x_j^{(\ell+1-t)}+\sum_{m=2}^{n_i}b_{j(n_i+1-m)}x_j^{(n_i+1-m)}-\sum_{m=2}^{n_i}k_{mj}x_j^{(n_i+1-m)}+b_{jn_i}x_j^{(n_i)}-k_{1j}x_j^{(n_i)}\right]\equiv 0\,. \tag{100}$$

Since there are only two terms in $x_j^{(n_i)}$, it must be that

$$b_{jn_i}=k_{1j}\text{ for each }j. \tag{101}$$

Another change in the dummy index of summation results in

$$\left[-\sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell}a_\ell k_{tj}x_j^{(\ell+1-t)}+\sum_{m=1}^{(n_i-1)}b_{j(n_i-m)}x_j^{(n_i-m)}-\sum_{m=1}^{(n_i-1)}k_{(m+1)j}x_j^{(n_i-m)}\right]\equiv 0\,. \tag{102}$$

Letting

$$n_i-m=s \tag{103}$$

in the above, results in

$$\left[-\sum_{\ell=1}^{(n_i-1)}\sum_{t=1}^{\ell}a_\ell k_{tj}x_j^{(\ell+1-t)}+\sum_{s=1}^{(n_i-1)}\{b_{js}-k_{(n_i-s+1)j}\}x_j^{(s)}\right]\equiv 0\,. \tag{104}$$

Finally, when

$$t=\ell+1-s, \tag{105}$$

the result is

$$\left[-\sum_{\ell=1}^{(n_i-1)}\sum_{s=1}^{\ell}a_\ell k_{(\ell+1-s)j}x_j^{(s)}+\sum_{s=1}^{(n_i-1)}\{b_{js}-k_{(n_i-s+1)j}\}x_j^{(s)}\right]\equiv 0\,. \tag{106}$$

Permuting the order of summation in the first term yields

$$\sum_{s=1}^{(n_i-1)}\left[\sum_{\ell=s}^{(n_i-1)}a_\ell k_{(\ell+1-s)j}+b_{js}-k_{(n_i-s+1)j}\right]x_j^{(s)}\equiv 0\,. \tag{107}$$

This implies that for each fixed s, we have that

$$k_{(n_i-s+1)j} = b_{js} - \sum_{\ell=s}^{(n_i-1)} a_\ell k_{(\ell+1-s)j} \cdot \tag{108}$$

Thus each of the k_{mj}'s for $(1 < m \leq n_i)$ is now specified for each j via the finite recursion of Eq. 108, as initiallized with Eq. 101.

The above procedure, as generalized here for handling multiple inputs, allows us to recast or re-represent the general system-describing differential equations of Eqs. 82 (having differentiated inputs) in the equivalent state variable form devoid of differentiated inputs as follows

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ \vdots \\ z_{n_i} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & 0 & \ddots & 1 & \vdots \\ \vdots & \vdots & \vdots & & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & \cdots & -a_{n_i-1} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ \vdots \\ z_{n_i} \end{bmatrix}$$

$$+ \begin{bmatrix} k_{21} & k_{22} & \cdots & k_{2r} \\ k_{31} & k_{32} & \cdots & k_{3r} \\ k_{41} & k_{42} & \cdots & k_{4r} \\ \vdots & \vdots & & \vdots \\ k_{n_i 1} & k_{n_i 2} & \cdots & k_{n_i r} \\ \left[b_{10} - \sum_{\ell=0}^{(n_i-1)} a_\ell k_{(\ell+1)1}\right] & \left[b_{20} - \sum_{\ell=0}^{(n_i-1)} a_\ell k_{(\ell+1)2}\right] & \cdots & \left[b_{r0} - \sum_{\ell=0}^{(n_i-1)} a_\ell k_{(\ell+1)r}\right] \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ x_r \end{bmatrix}, \tag{109}$$

where the k_{mj}'s are found as recursively specified above in Eqs. 101, 108. Recall that this procedure will be repetitively used in conjunction with the procedure near the end of the 2^{nd} paragraph of Section 3.1 in achieving realizations.

Although it does not arise in the application of interest in our underlying application, the unmodified approach of this section cannot handle the case when the MA portion has a degree that exceeds that of the original AR portion in Eqs. 82. This degenerate case of an excess MA component can still be handled by decomposing the ARMA process into the sum of an ARMA process (where $p = q$) and a process that is exclusively or purely MA. The pure MA component or portion can be handled by the recent technique of [39].

An alternate interpretation can be applied to our above generalization of a very old trick from differential equation manipulation, as offered next. Notice that within the realm of random processes (where the inputs to Eq. 82 are Gaussian white noises), the straightforward transformation offered here essentially converts a scalar ARMA process into an equivalent vector AR process. Typically, AR processes are usually much simpler to work with (especially with regard to spectral estimation). When this result is used in conjunction with the approach illustrated in Sections 3.1.1 and 3.1.2, then the above described method is applicable for converting multi-input/multi-output ARMA processes into AR processes. This is the difference between this new more general method offered for the first time here and more standard methods for getting rid of differentiated input terms such as that offered in 1987 in [36, Eqs. 1 and 2]. The approach of [36] is only applicable to single-input/single-output (SISO) transfer functions of the form

$$h_{1,1}(z) = \frac{1 + b_1 z^{-1} + \cdots + b_q z^{-q}}{1 + a_1 z^{-1} + \cdots + a_p z^{-p}}, \tag{110}$$

where the numerator and denominator polynomials in Eq. 110 are Hurwitz and additionally [36] assumes that $p = q$ to yield the realization result in terms of so-called "phase-variable" or "companion form" [37] that:

$$x_{n+1} = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & & & & -a_p \\ & \ddots & & & \vdots \\ 0 & & 1 & & -a_1 \end{bmatrix} x_n + \begin{bmatrix} b_{p+1} \\ \vdots \\ \vdots \\ b_1 \end{bmatrix} e_{n+1} \,, \qquad (111)$$

$$\text{with } y_n = [0, \cdots, 0, 1]\, x_n \,,$$

while for the application of interest here we seek approaches that can handle more general vector channels (such as the original technique developed within Section 3.1).

The approach of [36] above can be seen to be controllable because it results in an AR system matrix that is in "phase variable" canonical form and this representation is reasonably well-known to be controllable ([37]) and being controllable and also obviously observable, it is a minimal realization as established by the Kalman/Gilbert results that are reviewed in this context in Section 3.1.3. The approach of [44, Eqs. 3.9-13 to 3.9-16] offers an equivalent AR model for the (scalar single channel SISO) ARMA process and just "hopes for the best" without establishing that the resulting model is well-behaved (controllable, observable, stable) and satisfactory in this role. The structure to date has defied establishing controllability and observability for the equivalent AR representation of [44]. The beauty of the approach trailblazed herein for multi-input/multi-output (multichannel) ARMA process re-representation as an equivalent AR process is that the resulting AR system matrix is also demonstrably minimal, as explained in the last paragraph of Section 3.1.3. The significance of this result is that in pursuing model-based spectral estimation in a situation where the structure of the model is unknown *a priori*, we can now (without loss in generality) replace the more general ARMA structural assumption (otherwise requiring separate handling of AR and MA portions as a horribly nonlinear optimization) by a more benign and routinely tractable AR assumption (of slightly higher but finite vector dimension). This completes the novel contribution of this article.

4 Conclusions

Given a specified power spectral density matrix, the objective is to correctly emulate it in simulations with as little complexity as possible. The use of Matrix Spectral Factorization in conjunction with the results of Realization Theory have been investigated here and are shown to achieve this goal using explicit closed-form solutions to illustrate the main technique. A goal within the application of interest is to correctly emulate the cross-correlations of the original vector or multichannel random process and the approach presented here will handle this objective. The computer program identified in Section 2.6, along with a mechanization of the novel original realization algorithm (offered in Section 3), can now be applied off-line to actual application data to specify realistic model parameters as a prelude to actual implementation for emulating second order statistics of the random process (either in simulation or in hardware).

From a control and estimation theory viewpoint, an initially unexpected result was that the realization of Section 3.1.1 and the realization of Section 3.1.2 are *both minimal or irreducible* despite the differences in the degree of the underlying state variable models for the two cases being 4- and 5-states, respectively. If both realizations were from the same transfer function matrix, only the model of lesser state size could have been a minimal realization; however, the two realizations were an outcome of different matrix spectral factors so both realizations are valid minimal ones as demonstrated in Section 3.1.3.

Following the introduction in Section 1 to set the stage, the thrust of this investigation was laid out in Section 2. Section 2 also offered additional insights in scoping the problem of Matrix Spectral Factorization and this effort culminated with the validation activity for computer software computations (detailed in [42]). Section 3 completes the investigation by demonstrating how the state variable representation of the model can be rigorously obtained from the transfer function matrix (that is the transpose of any particular valid factor resulting from performing Matrix Spectral Factorization). All really detailed supporting analytic discussions are relegated to Appendices A to I of [42] in order that the engineering impact of this investigation be easily understandable.

The approach that was pursued with more fervor in [43, Sections 3.3, 3.4] (including results obtained from use of actual application data) uses multichannel extensions of high-resolution Maximum Entropy Method (MEM) techniques in conjunction with Levinson recursion, as is currently standard, as also reported in the application of the concurrent associated classified project report. Despite this departure from what we used in practice, the Matrix Spectral Factorization approach reported here was retained as a back-up alternative approach for possible analog implementation in case realizations become more practical in this domain because of recent breakthroughs in MMIC technology rivaling those already present within the domain of VLSI digital implementation. For the sake of completeness, another candidate random process emulation approach has emerged within the last two years [46], based on Akaike's Canonical Variate Analysis (CVA), and now implemented by Integrated Systems Inc. as an adjunct to $MATRIX_x$, using the numerically stable Singular Value Decomposition (SVD).

A Kalman-Ho Technique of Specifying Linear System Parameters (C, F, G) Using Realization Theory

The Kalman-Ho algorithm [6] for the construction of a minimal (irreducible) realization makes use of Hankel matrices (to be defined below) via the following structural observations. Starting from (almost) first principles, recall that the inverse Laplace transform of the resolvent matrix $(sI_n - F)^{-1}$ is the matrix exponential as in

$$\mathcal{L}^{-1}\{(sI_n - F)^{-1}\} = e^{Ft} = \sum_{i=0}^{\infty} F^i \frac{t^i}{i!} \,. \tag{112}$$

The definition of the matrix exponential (based on the Taylor series expansion) accounts for the term on the right hand side in Eq. 112. Upon taking the Laplace transform throughout Eq. 112, we have that

$$\begin{aligned} (sI_n - F)^{-1} &= \mathcal{L}\{e^{Ft}\} &&= \mathcal{L}\textstyle\sum_{i=0}^{\infty} F^i \frac{t^i}{i!} \\ &= \textstyle\sum_{i=0}^{\infty} F^i \mathcal{L}\{\frac{t^i}{i!}\} &&= \textstyle\sum_{i=0}^{\infty} F^i \frac{1}{s^i} \end{aligned} \tag{113}$$

or, simply

$$(sI_n - F)^{-1} = \sum_{i=0}^{\infty} F^i \frac{1}{s^i} \tag{114}$$

as a representation that is, perhaps, somewhat less familiar to the reader than that of Eq. 112. Upon pre- and post-multiplying Eq. 114 by C and G, respectively, we have that

$$C(sI_n - F)^{-1}G = \sum_{i=0}^{\infty} [\, CF^iG \,] \frac{1}{s^i} \tag{115}$$

and the left hand side of Eq. 115 can be recognized to be the expression for the transfer function matrix $H(s)$ in Eq. 10 for a physical interpretation of Eq. 115 in the frequency domain.

Analogously, we could have branched out from the familiar Eq. 112 by pre- and post-multiplying by C and G, respectively, to obtain a time-domain physical interpretation of the impulse-response matrix as

$$Ce^{Ft}G = \sum_{i=0}^{\infty} CF^iG\,\frac{t^i}{i!}\,. \tag{116}$$

The whole point of this discussion is to call attention to the presence of the product CF^iG in both Eqs. 115 and 116 as fundamental invariants that completely characterize the system. Now defining a sequence of matrices J_i as follows

$$\begin{aligned} J_0 &= CF^0G = CI_nG = CG && (117)\\ J_1 &= CF^1G && (118)\\ J_2 &= CF^2G && (119)\\ &\vdots \quad \vdots\\ j_i &= CF^iG\,, && (120) \end{aligned}$$

then the associated Hankel matrix $\mathcal{H}_{i,j}$, defined as follows

$$\mathcal{H}_{i,j} = \begin{bmatrix} J_0 & J_1 & \cdots & J_{j-1} \\ J_1 & J_2 & \cdots & J_j \\ \vdots & \vdots & \vdots & \vdots \\ J_{i-1} & J_i & \cdots & J_{i+j-2} \end{bmatrix}\,, \tag{121}$$

may be investigated to determine whether a system is realizable. [6] Details on how to do this (based, respectively, on the associated j-observability and j-controllability Grammians

$$\left[G \vdots FG \vdots F^2G \vdots \ldots \vdots F^{j-1}\right] \tag{122}$$

$$\begin{bmatrix} C \\ \cdots \\ CF \\ \cdots \\ \vdots \\ \cdots \\ CF^{j-1} \end{bmatrix}\,, \tag{123}$$

and their requisite projective subspaces, an invocation of Sylvester's theorem on how a matrix satisfies its own characteristic equation, thus allowing an matrix infinite series, such as arise in Eqs. 112 to 116, to be reexpressed as a finite series with an upper limit of the sum now being $n-1$ (where n is the dimension of the underlying system matrix F)) are relegated to [32, Chapter 6].

More germane to our particular problem is the construction of a realization, given a specified transfer function matrix $H(s)$ as the starting point. First expand it or represent it as

$$H(s) = \sum_{i=0}^{\infty} J_i \frac{1}{s^i} \tag{124}$$

[6] Where such investigations are nontrivial is when only $\mathcal{H}_{i,j}$ is provided (as obtained or derived by actual measurements yielding a string of deterministic data uncorrupted by noise) without prior access to the information of Eqs. 112 to 120.

(corresponding to Eq. 115), then use the J_i's that can be obtained from the associated Hankel matrix array of Eq. 121 that is subsequently constructed. Finally (via Theorems 6.7 and 6.9 of [32]), the explicit constructive procedure of Kalman and Ho is available directly from the Hankel matrix of Eq. 121. Again emphasizing that the critical step before the Kalman-Ho realization procedure [6] can commence is to first extract the J_i's (from Eq. 124) and use these to form the Hankel matrix of Eq. 121. The submatrix-based Kalman-Ho procedure is concisely summarized as an algorithm reported on two pages of [32, pp. 126-127].

B Other Applications of Matrix Spectral Factorization

Matrix Spectral Factorization continues to be of interest to many specialists for the following applications:

- Putting a problem involving serially time-correlated additive noises into standard Kalman filter form (which assumes only uncorrelated white additive noises) by modeling the second order statistical structure of the time-correlated Gauss-Markov noise via spectral factorization and then using state augmentation [44, pp. 133-135], [29, pp. 146-147].
- Pre-whitening for multichannel applications as a prelude to performing matched filtering (see Chapts. 1, 3, 6, 7, 9, 11, 14 of [31] for ramifications in radar applications) when the additive noise is not originally white but is serially correlated in time.
- In conjunction with Realization Theory [6], [28] as a basis for constructing reduced-order models of lesser computational burden which still adequately "capture the essence of the true system" for purposes of modeling/emulation, control, or estimation. Recent techniques used to accomplish the realizations include continued-fraction expansions and linear fractional transformations.
- Modeling and control for flexible structures, as is being pursued by NASA for the Space Station [21].
- In the design of optimal feedback controllers for linear discrete-time multivariable stochastic systems [34].
- Establishing stability of a system or of a digital filtering algorithmic mechanization.
- In multiport network synthesis [3, pp. 190-193] as applied to such practical problems as providing a balanced lossless reciprocal feed network between several transceivers and a common antenna [3, p. 235]. (An explicit step-by-step connection between the property of being "bounded real", as is necessary for passive MIMO circuits to satisfy in order to be realized, and the property of being "positive real", as valid power spectral density matrices need to satisfy, is carefully established in Section I.3 of [42].)
- As an *exact* solution to handling the effect of multiplicative noise in certain otherwise linear (or bilinear) applications [13] using Kronecker products rather than attempting to use Lie group techniques (apparently without sufficient supporting computational software yet being available for numerically realizing the Lie group approaches for practical problems of realistic dimensions).
- Complex random processes (as arise in handling PP and OP or VV and HH polarization components in radar) can also be routinely handled by the original solution method identified in [42, pp. 11, 12 and bottom of p. 25] by decomposing the associated power spectral density matix, consisting of ratios of polynomials now having complex rather than real coefficients as

in Sections 2 and 3 here, into pure real and pure imaginary parts (in the manner discussed in [42, Section 1.4] and [43, Section 3.3.1] using the technique of [38]), then applying the usual techniques already covered herein separately to each. The computational burden for handling complex processes in this way is merely double what it would be for comparable real processes. For example, in considering target cross-correlations across three adjoining radar range gates, the state size to be used for both PP and OP to capture all possible cross-correlations would need to be 6, which is within the capabilities of the software of Section 2.6.

- In lieu of Linear Quadratic Gaussian (LQG) optimal control now encountering practical implementation difficulties when system parameters change or when unmodeled higher order frequency response characteristics cause the system response to unfortunately deviate from the design goals, the techniques of Robust Control (also known as H^∞ techniques) have emerged in an attempt to bridge the gap between optimal control theory and practice. There are several paths to establishing H^∞ solutions (such as via min-max game theory, or via use of two coupled non-standard Riccati equations, etc.) and oversimplified benchmark problems have been agreed upon at 1990 control conferences to serve as a common design challenge to narrow the field of entries and select the most expedient solution approach, as sought in [50]. (Since the benchmark problem is still a single channel SISO problem, it is recommended that the closed-form two channel example of Sections 2.2, 3.1.1, 3.1.2 would better serve as the basis for an easy MIMO benchmark problem.) Recently, it has been recognized that the Matrix Spectral Factorization approach also leads to H^∞ control solutions [47], [48] and, moreover, the frequency domain has historically offered an alternative approach to optimal control design (via either the 1976 approach of Youla, Bongiorno, and Jabr (1976) or via the approach of Rosenbrock). Recently, others [49] have now recognized the connection between the topic of Realization Theory and that of H^∞ control.

References

[1] Youla, D. C., "On the Factorization of Rational Matrices," *IRE Transactions on Information Theory,* Vol. IT-7, No. 4, pp. 172-189, July 1961.

[2] Davis, M. C., "Factoring the Spectral Matrix," *IEEE Transactions on Automatic Control,* Vol. AC-8, No. 5, pp. 296-305, October 1963.

[3] Newcomb, R. W., *Linear Multiport Synthesis,* McGraw-Hill, N.Y., 1966.

[4] Kras, G., and Potzl, H., "Limiting Conditions on the Correlation Properties of Random Signals," *IRE Transactions on Circuit Theory,* Vol. CT-3, No. 4, pp. 282-285, December 1956.

[5] Anderson, B. D. 0., "A System Theory Criterion for Positive Real Matrices," *J. S. I. A. M. Control,* Vol. 5, No. 2, pp. 171-182, May 1967.

[6] Ho, B. L., and Kalman, R. E., "Effective Construction of Linear State-Variable Models from Input/Output Data," *Proceedings of Third Annual Allerton Conference on Circuit and System Theory,* Ed. by M. E. Van Valkenburgh, Urbana, IL, pp. 449-459, 20-22 October 1965.

[7] Anderson, B. D. 0., and Moore, J. B., *Linear Optimal Control,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

[8] Papoulis, A., *Probability, Random Variables and Stochastic Processes,* McGraw-Hill Book Co., Inc., N.Y., 1965.

[9] Solodovnikov, V. V., *Introduction to the Statistical Dynamics of Automatic Control Systems,* (Russian Version in 1952), Translated by J. B. Thomas and L. A. Zadeh, Dover, N.Y., 1960.

[10] Brockett, R. W., *Finite Dimensional Linear Systems,* John Wiley and Sons Inc., N.Y., 1970.

[11] Anderson, B. D. 0., "An Algebraic Solution to the Spectral Factorization Problem," *IEEE Transactions on Automatic Control,* Vol. AC-12, No. 4, pp. 410-414, August 1967.

[12] Ho, B. L., and Kalman, R. E., "Spectral Factorization Using the Riccati Equation," *Proceedings of Fourth Annual Allerton Conference on Circuit and System Theory,* Ed. by W. R. Perkins and J. B. Cruz, pp. 388-399, 5-7 October 1966.

[13] Loo, S. G., "Statistical Design of Optimum Multivariate Systems Whose Inputs are Corrupted by Multiplicative Noise," *IEEE Transactions on Automatic Control,* Vol. AC-13, No. 3, pp. 300-301, June 1968.

[14] Denham, M. J., "On the Factorization of Discrete-Time Rational Spectral Density Matrices," *IEEE Transactions on Automatic Control,* Vol. AC-20, No. 4, pp. 535-537, August 1975.

[15] Van Dooren, P., "A Generalized Eigenvalue Approach for Solving Riccati Equations," *SIAM Journal of Scientific and Statistical Computations,* Vol. 2, No. 2, pp. 121-135, June 1981.

[16] Van Dooren, P., "The Generalized Eigenstructure Problem in Linear System Theory," *IEEE Transactions on Automatic Control,* Vol. AC-26, No. 1, pp. 111-129, February 1981.

[17] Tuel, W. G., "Computer Algorithm for Spectral Factorization of Rational Matrices," *IBM Journal of Research and Development,* Vol. 12, No. 2, pp. 163-170, March 1968.

[18] Laub, A., "A Schur Method for Solving Algebraic Riccati Equations," *IEEE Transactions on Automatic Control,* Vol. AC-24, No. 6, pp. 913-921, December 1979.

[19] Barton, D. R., and Zippel, R., "Polynomial Decomposition Algorithms," *Journal of Symbolic Computation,* Vol. 1, No. 2, pp. 159-168, 1985.

[20] Silverman, L. M., and Van Dooren, P., "A System Theoretic Interpretation for GCD Extraction," *IEEE Transactions on Automatic Control,* Vol. AC-26, No. 6, pp. 1273-1276, December 1981.

[21] Bennett, W. H., Blankenship, G. L., and Kwatny, H. G., "Spectral Factorization and Homogenization Methods for Modeling and Control of Flexible Structures," Technical Report SEI TR-86-14, AF OSR Contract No. F49620-84-C-0115, Systems Engineering, Inc., Greenbelt, MD, January 1987.

[22] Sveinsson, J. R., and Fairman, F. W., "Minimal Balanced Realization of Transfer Function Matrices Using Markov Parameters," *IEEE Transactions on Automatic Control,* Vol. AC-30, No. 10, pp. 1014-1016, October 1985.

[23] DeRusso, P. M., Roy, R. J., and Close, C. M., *State Variables for Engineers,* John Wiley and Sons, N.Y., 1967.

[24] Anderson, B. D. 0., and Newcomb, R. W., "Canonical Simulation of a Transfer Function Matrix," *IEEE Transactions on Automatic Control,* Vol. AC-11, No. 2, pp. 319-320, April 1966.

[25] Kalman, R. E., "Irreducible Realizations and the Degree of a Rational Matrix," *Journal of Society for Industrial and Applied Mathematics (SIAM),* Vol. 13, No. 2, June 1965.

[26] Kalman, R. E., "Mathematical Description of Linear Systems," *Journal of SIAM: Control,* Series A, Vol. 1, No. 2, 1963.

[27] Anderson, B. D. 0., Hitz, K. L., and Diem, N. D., "Recursive Algorithm for Spectral Factorization," *IEEE Transactions on Circuits and Systems,* Vol. CAS-21, No. 6, pp. 742-750, November 1974.

[28] Gilbert, E., "Controllability and Observability in Multivariable Control Systems," *SIAM Journal on Control,* Series A, Vol. 2, No. 1, pp. 128-151, 1963.

[29] Maybeck, P. S., *Stochastic Models, Estimation, and Control,* Vol. 1, Academic Press, N.Y., 1979.

[30] Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Tables,* National Bureau of Standards, AMS Series 55, 1966.

[31] Cook, C. E., and Bernfield, M., *Radar Signals: An Introduction to Theory and Application,* Academic Press, N.Y., 1967.

[32] Casti, J. L., *Linear Dynamical Systems,* Academic Press, N.Y., 1987.

[33] Mortensen, R. E., "Comments on 'A Method to Determine Whether Two Polynomials are Relatively Prime'," *IEEE Transactions on Automatic Control,* Vol. AC-16, No. 1, pp. 112-113, February 1971.

[34] Sebek, M., "Direct Polynomial Approach to Discrete-Time Stochastic Tracking," *Probl. Control and Inf. Theory (Hungary),* Vol. 12, No. 4, pp. 293-300, 1983.

[35] Ling, Y.-L., and Wang, B.-C., "Matrix Partial Fraction Expansions of Rational Matrices," *IEEE Transactions on Automatic Control,* Vol. AC-32, No. 4, pp. 330-333, April 1987.

[36] Alengrin, G., and Zerubia, J., "A Method to Estimate the Parameters of an ARMA Model," *IEEE Transactions on Automatic Control,* Vol. AC-32, No. 12, pp. 1113-1115, December 1987.

[37] Tuel, W. G., "On the Transformation to (Phase-Variable) Canonical Form," *IEEE Transactions on Automatic Control,* Vol. AC-11, No. 4, p. 607, July 1966.

[38] Jacovitti, G., Neri, A., Cusani, R., "Methods for Estimating the Autocorrelation Function of Complex Gaussian Stationary Processes," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* Vol. ASSP-35, No. 8, pp. 1126-1138, Aug. 1987.

[39] Ohsmann, M., "An Iterative Method for the Identification of MA Systems," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-36, No. 1, pp. 106-109, January 1988.

[40] Kerr, T. H., "An Analytic Example of a Schweppe Likelihood Ratio Detector," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-25, No. 4, pp. 545-558, July 1989.

[41] Kerr, T. H., "A Constructive Use of Idempotent Matrices to Validate Linear Systems Analysis Software," ~~accepted by *IEEE Transactions on Aerospace and Electronic Systems*~~ *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-26, No. 6, pp. 935-952, November 1990.

[42] Kerr, T. H., "Multichannel Shaping Filter Formulations for Vector Random Process Modeling Using Matrix Spectral Factorization," MIT Lincoln Laboratory Project Report No. PA-500, Lexington, MA, 27 March 1989.

[43] Kerr, T. H., "Rationale for Monte-Carlo Simulator Design to Support Multichannel Spectral Estimation and/or Kalman Filter Performance Testing and Software Validation/Verification Using Closed Form Test Cases," MIT Lincoln Laboratory Project Report No. PA-512, Lexington, MA, 22 December 1989.

[44] Gelb, A. (ed.), "Applied Optimal Estimation," MIT Press, Cambridge, MA, 1974.

[45] Wette, M., and Laub, A. J., "Numerical Algorithms and Software for Spectral Factorization," *Proceedings of the 1988 American Control Conference*, Vol. 3, 2503 ff., Atlanta, GA, 15-17 June 1988.

[46] Larrimore, W. E., "Canonical Variate Analysis in Identification, Filtering, and Adaptive Control," *Proceedings of the* 29^{th} *Conference on Decision and Control*, pp. 596-604, Honolulu, HI, December 1990.

[47] Green, M., Glover, K., Limebeer, D., and Doyle, J., "A J-Spectral Factorization Approach to H^∞ Control," *SIAM Journal of Control and Optimization*, Vol. 28, No. 6, pp. 1350-1371, November 1990.

[48] Sideris, A., "H^∞ Optimal Control as a Weighted Wiener-Hopf Problem," *IEEE Transactions on Automatic Control*, Vol. AC-35, No. 3, pp. 361-366, March 1990.

[49] Chang, B.-C., "A Stable State-Space Realization in the Formulation of H^∞ Norm Computation," *IEEE Transactions on Automatic Control*, Vol. AC-32, No. 9, pp. 811-815, September 1987.

[50] Williams, S. J. and Hyde, R. A., "A Comparison of Different H^∞ Methods in VSTOL Flight Control System Design," *Proceedings of the 1990 American Control Conference*, Vol. 3, 2503 ff., Atlanta, GA, 15-17 June 1988.

[51] DiStefano, J. J., Stubberud, A. R., and Williams, I. J., *Schaum's Outline of Theory and Problems of Feedback and Control Systems*, Schaum Publishing Co., N.Y., N. Y., 1967.