MINIMAL CONTINUOUS STATE-SPACE PARAMETRIZATIONS

Alle-Jan van der Veen[*] *and Mats Viberg*[†]

[*]Delft University of Technology, Dept. Electrical Eng./DIMES, Delft, The Netherlands
[†]Chalmers University of Technology, Dept. Applied Electr., S-412 96 Göteborg, Sweden

We present a minimal continuous parametrization of all multivariate rational contractive transfer functions. In contrast to traditional minimal parametrizations, this parametrization does not contain any structural indices, which makes it very suitable for identification algorithms that use nonlinear optimization to estimate the parameters.

1. INTRODUCTION

State-space model identification is concerned with the fitting of state-space models to measured input-output data corresponding to an unknown dynamical system. An optimal approach to this problem requires solving a non-linear optimization problem. A critical issue is then to have a minimal number of unknown parameters to be estimated. Thus, there has been an active search for *canonical* system representations. For multi-input multi-output systems, a number of canonical forms are known, based *e.g.*, on the observer or controller canonical forms or on balanced realizations [1, 2].

For the purpose of identification, an important deficiency in all canonical representations known to date is that they require both continuously varying parameters (in $\mathbb{R}$, say), and discrete parameters (in a subset of $\mathbb{N}$). The latter parameters are extra parameters that specify the structure of the system, such as the Kronecker indices. Since these "structural" parameters often have little physical meaning, the only way to solve the optimization problem is to enumerate over a sufficient range of structural parameter values to cover all systems of a given order, and to perform a non-linear search over the continuous parameters for each such choice.

The purpose of this paper is to show that the class of contractive (alias passive, or bounded real) asymptotically stable systems is covered by a minimal representation *without* any structural parameters. The representation is not unique, but for each system $T(z)$ there is only a finite number of equivalent descriptions (unless the system is overparametrized). Because the solutions are isolated, this does not pose a problem for numerical optimization techniques, especially since we start with reasonably accurate initial points. The parametrization guarantees stability of the predictor, which is an obvious advantage. Although the contractive assumption may not be met by the "true" system to be identified, this can always be assured by appropriately scaling the input-output data.

The main ideas behind this parametrization are well known to insiders; the new aspects are in the details. Assuming $T(z)$ is rational and contractive, it is known that there is a (non-unique) embedding of $T(z)$ into an allpass (alias lossless, inner or para-unitary) system by adding extra inputs and outputs to it. Minimal parametrizations of allpass systems are known and are based on a factorization of the realization matrix into elementary Givens rotations. The rotation angles are the parameters of the representation. Although such embedding/factorization techniques have been known for a long time (*e.g.*, [3–7]) and go back to classical Darlington network synthesis theory, existing literature has overlooked certain minor details and either ended up with too many continuous parameters, or with an additional set of discrete (structural) parameters with values ± 1. We show in this paper that the freedom in the orthogonal embedding step is sufficient to get rid of all extra parameters.

Eusipco'96, Trieste (Italy), September 1996

2. STATE-SPACE PREDICTION-ERROR METHODS

Consider a linear time-invariant system with m inputs $\mathbf{u}_k$ and p outputs $\mathbf{y}_k$ (column vectors), k denoting discrete time. The objective is to find an n-th order state-space model that relates $\mathbf{y}_k$ and $\mathbf{u}_k$, based on a batch of input-output measurements.

One-step ahead prediction-error methods (PEM) are based on the optimal (in the mean square sense) prediction of $\mathbf{y}_k$ by $\hat{\mathbf{y}}_k(\theta)$ from past measurements $\{\mathbf{y}_{k-s}\}_{s=1}^{\infty}, \{\mathbf{u}_{k-s}\}_{s=0}^{\infty}$, where the parameter vector θ symbolizes a given candidate model. If the model is correct, then the prediction error $\varepsilon_k(\theta) := \mathbf{y}_k - \hat{\mathbf{y}}_k(\theta)$ is a white noise sequence, and, assuming a Gaussian noise model, one can show that the ML estimate of θ is the minimizing argument of the concentrated criterion function [8]

$$V(\theta) = \log \det \sum_{k=1}^{N} \varepsilon_k(\theta)\varepsilon_k^T(\theta) . \tag{1}$$

Two major types of state-space models are widely used:
1. *Output error models (OE)*, which postulate a purely deterministic input-output relation corrupted by additive white measurement noise $\mathbf{e}_k$. Thus $\mathbf{y}_k = \hat{\mathbf{y}}_k(\theta) + \mathbf{e}_k$, $\varepsilon_k(\theta) = \mathbf{e}_k$, and the predictor model is simply

$$\begin{array}{rcrcl} \mathbf{x}_{k+1} & = & A\mathbf{x}_k & + & B\mathbf{u}_k \\ \hat{\mathbf{y}}_k(\theta) & = & C\mathbf{x}_k & + & D\mathbf{u}_k \,, \end{array} \tag{2}$$

where θ represents the unknown parameters in the system matrices A, B, C and D.
2. *Kalman predictor models (KP)*, which allow output measurements that are corrupted by several noise and disturbance

sources, some of which may have passed through parts of the system dynamics. A suitable state-space model is

$$\begin{array}{rcccccc} \mathbf{x}_{k+1} & = & A\mathbf{x}_k & + & B\mathbf{u}_k & + & K\mathbf{e}_k \\ \mathbf{y}_k & = & C\mathbf{x}_k & + & D\mathbf{u}_k & + & \mathbf{e}_k\,, \end{array} \tag{3}$$

where $\mathbf{e}_k$ is the innovation process and K is the (stationary) Kalman gain, to be estimated along with (A,B,C,D). If the model is correct, the innovations coincide with the one-step ahead prediction errors, *i.e.*, $\mathbf{y}_k = \hat{\mathbf{y}}_k(\theta)+\mathbf{e}_k$, $\varepsilon_k(\theta) = \mathbf{e}_k$. Substitution in (3) leads to the predictor model

$$\begin{array}{rcccccc} \mathbf{x}_{k+1} & = & \tilde{A}\mathbf{x}_k & + & \tilde{B}\mathbf{u}_k & + & K\mathbf{y}_k \\ \hat{\mathbf{y}}_k(\theta) & = & C\mathbf{x}_k & + & D\mathbf{u}_k & & \end{array} \tag{4}$$

where $\tilde{A} = A - KC$, $\tilde{B} = B - KD$, and θ represents the unknown parameters of $\tilde{A}$, $\tilde{B}$, C, D and K.

For both models, an implementation of the PEM (1) requires a non-linear multidimensional optimization. Reasonably accurate initial estimates can efficiently be obtained by subspace model identification techniques such as N4SID [9].

3. OUTPUT-ERROR MODEL PARAMETRIZATION

Parametrization of an OE structure (2) is equivalent to parametrizing a linear filter without any special constraints on the involved system matrices. Our proposed technique ensures stability as well, but it also imposes the additional property that the predictor is *contractive* (or *passive*). A discrete-time system $T(z)$ is called contractive if $\|T\|_\infty \le 1$, where $\|T\|_\infty = \sup_{\mathbf{u}\in\ell_2} \|T*\mathbf{u}\|/\|\mathbf{u}\|$. $T(z)$ is *lossless* if $T(z)T^*(\bar{z}^{-1}) = I$, $T^*(\bar{z}^{-1})T(z) = I$. In essence, the passivity property puts a bound on the gain of the predictor. This is sometimes an undesirable restriction and may require a scaling of the output vector so that the resulting scaled system is passive.

For $-1 \le s \le 1$, $c = \sqrt{1-s^2}$, and integers n, m, p, define the plane rotations

$$Q_{ij}(s) = \begin{array}{c} \begin{array}{ccccc} & i & & n+j & \end{array} \\ \begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix} \end{array} \in \mathbb{R}^{(n+m+p)\times(n+m+p)} \tag{5}$$

$$Z_{ij}(s) = \begin{array}{c} \begin{array}{ccccc} & n+i & & n+p+j & \end{array} \\ \begin{bmatrix} I_n & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix} \end{array} \in \mathbb{R}^{(n+p+m)\times(n+p+m)} \tag{6}$$

Also define permutations $\Pi_{1,n+1}$ and Π_D by

$$\Pi_{1,n+1} \begin{bmatrix} x_1 \\ \vdots \\ x_{n+1} \\ x_{n+2} \\ \vdots \end{bmatrix} = \begin{bmatrix} x_2 \\ \vdots \\ x_{n+1} \\ x_1 \\ x_{n+2} \\ \vdots \end{bmatrix}, \quad \Pi_D = \begin{bmatrix} I_n & & \\ & 0 & I_p \\ & I_m & 0 \end{bmatrix}. \tag{7}$$

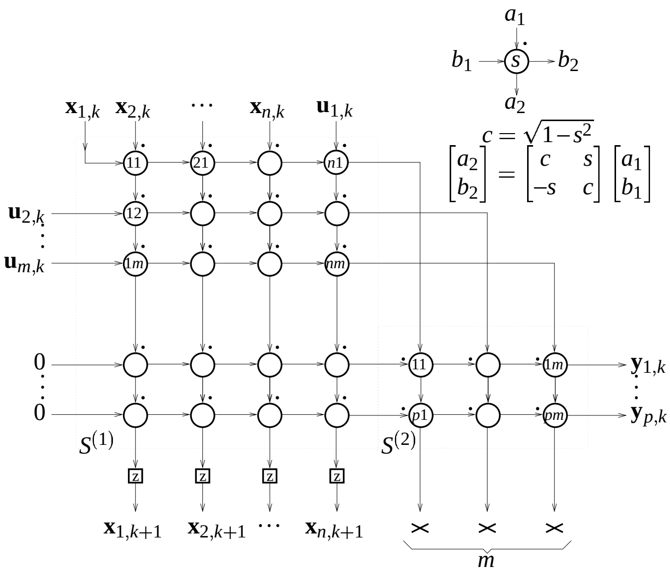


Figure 1. Cascade parametrization of a rational stable contractive MIMO system.

Theorem 1. *There is a minimal continuous parametrization with $n(m+p)+pm$ bounded coefficients which covers the set of all rational stable contractive systems with m inputs, p outputs and n states.*

In particular, every such system may be specified in terms of two matrices $S^{(1)} : n\times(m+p)$, $S^{(2)} : p\times m$ with entries $-1 \le s_{ij}^{(\cdot)} \le 1$ as $T(z) = D + C(I-Az)^{-1}zB$ where

$$\begin{array}{cc} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ p \end{array} & \begin{bmatrix} A & B \\ C & D \end{bmatrix} \end{array} = [I_{n+p} \quad 0_{(n+p)\times m}]Z_{p,m}\cdots Z_{12}Z_{11} \tag{8}$$
$$\cdot \Pi_D \cdot Q_{n,m+p}\cdots Q_{12}Q_{11} \cdot \Pi_{1,n+1} \begin{bmatrix} I_{n+m} \\ 0_{p\times(n+m)} \end{bmatrix}$$

for $Q_{ij} := Q_{ij}(s_{ij}^{(1)}), Z_{ij} := Z_{ij}(s_{ij}^{(2)})$. The parametrization is not unique, but for strictly contractive systems which are controllable via the first input, only a finite (discrete) set of parameter matrices lead to the same $T(z)$.

The structure of this parametrization is perhaps better understood from figure 1, which shows the state space mapping

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix}$$

in terms of the factorization (8). The proof is in three steps.

Step 1: Lossless embedding

Assume that $T(z)$ is specified in terms of a minimal realization (A,B,C,D) with $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, $C \in \mathbb{R}^{p,n}$, $D \in \mathbb{R}^{p,m}$. Step 1 is to find an invertible state transformation R, and state matrices B_2, D_{12} such that

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} R^{-1} & \\ & I \end{bmatrix} \begin{array}{cc} & \begin{array}{ccc} n & m & p \end{array} \\ \begin{array}{c} n \\ p \end{array} & \begin{bmatrix} A & B & B_2 \\ C & D & D_{12} \end{bmatrix} \end{array} \begin{bmatrix} R & & \\ & I & \\ & & I \end{bmatrix} \tag{9}$$

is isometric: $\mathbf{\Sigma}_1\mathbf{\Sigma}_1^* = I$, so that the corresponding transfer function $\Sigma_1(z)$ is (partially) lossless and embeds $T(z)$. Defining $M = RR^*$, this is equivalent to solving

$$\begin{array}{ccccccc} AMA^* & + & BB^* & + & B_2B_2^* & = & M \\ AMC^* & + & BD^* & + & B_2D_{12}^* & = & 0 \\ CMC^* & + & DD^* & + & D_{12}D_{12}^* & = & I\,. \end{array} \tag{10}$$

Under the conditions of theorem 1, the bounded real lemma (see [6, 10]) claims that solutions $M > 0$ exist, and that for each solution $I - DD^* - CMC^* \geq 0$ (> 0 holds if T is strictly contractive). M is found as the solution of a discrete algebraic Riccati equation. It is not unique but there are at most 2^{n-1} isolated solutions. The non-uniqueness has to do with the choice of the spectral factor of $I - T(z)T^*(\bar{z}^{-1})$.

Take any solution M. Then D_{12} and B_2 follow from

$$\begin{array}{rcl} D_{12}D_{12}^* & = & I - DD^* - CMC^* \\ B_2 & = & -(AMC^* + BD^*)D_{12}^\dagger \end{array} \tag{11}$$

where $(\cdot)^\dagger$ denotes the pseudo-inverse. D_{12} is a square root of a positive semidefinite matrix. We choose D_{12} to be *lower triangular* with $\mathrm{diag}(D_{12}) \geq 0$. (This aspect is new in comparison with [7].) If T is strictly contractive, then for the chosen M this D_{12} is unique and $\mathrm{diag}(D_{12}) > 0$, otherwise $D_{12}D_{12}^*$ might be singular with a continuum of suitable factors.

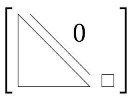
Step 2: Transformation into Hessenberg form

Suppose at this point that we have

$$\mathbf{\Sigma}_1 =: \begin{array}{c} \\ n \\ p \end{array} \begin{array}{c} \begin{array}{ccc} n & m & p \end{array} \\ \begin{bmatrix} A & B & B_2 \\ C & D & D_{12} \end{bmatrix} \end{array}$$

where $\mathbf{\Sigma}_1\mathbf{\Sigma}_1^* = I$, D_{12} is lower triangular and $\mathrm{diag}(D_{12}) \geq 0$. Denote by $\mathbf{b}_1$ the first column of B. Step 2 is to find a unitary state transformation Q such that

$$[A' \quad \mathbf{b}_1'] := Q[A \;\; \mathbf{b}_1] \begin{bmatrix} Q^* & \\ & 1 \end{bmatrix} =$$



is in lower Hessenberg form. We omit a description of this (standard) procedure. Some freedom is left; we can use it to guarantee that all entries on the super-diagonal are nonnegative (a "positive lower Hessenberg" form). It can be shown that the entries of the superdiagonal of $[A' \;\; \mathbf{b}_1']$ are strictly positive and Q is unique if and only if the system is controllable via its first input. Without this condition, there might exist a continuum of suitable Q.

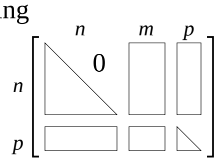
Step 3: Factorization of $\mathbf{\Sigma}_1$

Suppose at this point that we have an embedding $\mathbf{\Sigma}_1$, isometric, in the required positive Hessenberg form, and with D_{12} lower triangular with nonnegative main diagonal. The final step is to factor $\mathbf{\Sigma}_1$ into elementary Givens rotations, producing the actual parameters of the state space model. It suffices for our purposes to consider rotations of the form

$$q(s) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \qquad c = \sqrt{1-s^2}, \qquad -1 \leq s \leq 1\,.$$

For ease of description, we move column 1 of $\mathbf{\Sigma}_1$ behind column $n+1$, giving

$$\Phi = \mathbf{\Sigma}_1\Pi_{1,n+1}^* =$$



$$=: \begin{array}{c} \\ n \\ p \end{array} \begin{array}{c} \begin{array}{ccc} n & m & p \end{array} \\ \begin{bmatrix} A & B & B_2 \\ C & D & D_{12} \end{bmatrix} \end{array}$$

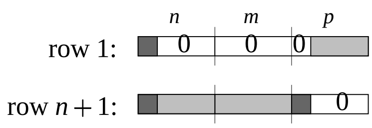
where the permutation $\Pi_{1,n+1}$ is defined in (7). (Note that we redefined $A, B, \cdots$ for ease of notation.) Subsequently, we apply a sequence of rotations to the columns of Φ to reduce it to a submatrix of the identity matrix, taking care that A and D_{12} remain lower triangular with nonnegative diagonal entries throughout the transformations.

- Apply a Givens rotation $q_{11}^* = \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]$ to columns 1 and $n+1$ of Φ, to cancel b_{11} against a_{11}, *i.e.*, $[a_{11} \quad b_{11}]q_{11}^* = [a_{11}' \quad 0]$. This rotation is specified by

$$s = b_{11}\,(a_{11}^2 + b_{11}^2)^{-1/2}\,, \quad c = \sqrt{1-s^2}\,.$$

(If both $a_{11} = 0$ and $b_{11} = 0$, then we may select any s in the range $[-1, 1]$.) Because $a_{11} \geq 0$, $c \geq 0$ and $\mathrm{sign}(s) = \mathrm{sign}(b_{11})$, we have $a_{11}' \geq 0$, so that the positivity property of the main diagonal of A is invariant.

- In the same way, use the transformed a_{11} to zero all entries of the top row of $[B \;\; B_2]$. This defines a sequence of Givens rotations $q_{12}^*, \cdots, q_{1,m+p}^*$ which are applied in turn to Φ. Because Φ is isometric, the norm of each row is 1. This property is retained by the rotations, so that after the transformations we must have $a_{11} = +1$.
- It is clear that A remains lower triangular during the rotations. We have to show that D_{12} also remains lower triangular, with nonnegative main diagonal, and that the first column of C is zero. This nontrivial fact follows from the orthonormality of the rows of Φ, which is invariant under the transformations. Indeed, after the first row of B has been zeroed, $a_{11} > 0$ because the realization is controllable. After $(B_2)_{11}$ has been zeroed, we have for the transformed Φ,



(12)

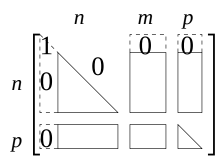
Since the rows are orthonormal, the first entry of the $n+1$-st row, the transformed c_{11}, must be zero at this point. Hence, subsequent rotations of the first column and columns 2 to p of $\left[\begin{smallmatrix} B_2 \\ D_{12} \end{smallmatrix}\right]$ do not destroy the zeros on the $n+1$-st row. The same holds for rows $n+2, \cdots$, so that D_{12} stays lower triangular while B_2 is made zero. The fact that $(D_{12})_{11} \geq 0$ after rotation $q_{1,m+1}$ follows directly from the following lemma:

Suppose $[a \;\; b]\left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right] = [0 \;\; r]$ *where* $b \geq 0$, $c = \sqrt{1-s^2} \geq 0$. *Then* $r \geq 0$.

Thus, $\mathrm{diag}(D_{12}) \geq 0$ is invariant under the transformations.

- At this point, we have obtained

$$\mathbf{\Sigma}_1\Pi_{1,n+1}^*Q_{11}^*\cdots Q_{1,m+p}^* =$$



where each Q_{ij} is an embedding of q_{ij}, as defined in (5). The zeroing of the second through the n-th row of $[B\ B_2]$ proceeds similarly. This gives

$$\Phi' = \mathbf{\Sigma}_1 \Pi^*_{1,n+1} Q^*_{11} Q^*_{12} \cdots Q^*_{n,m+p} = \begin{array}{c} \\ n \\ p \end{array} \begin{array}{c} \begin{array}{ccc} n & m & p \end{array} \\ \begin{bmatrix} I & 0 & 0 \\ 0 & D' & D'_{12} \end{bmatrix} \end{array} \quad (13)$$

where D'_{12} is lower triangular with nonnegative main diagonal. In similar ways, we now use the main diagonal entries of D'_{12} to zero the entries of D'. As before, this can be done by Givens rotations $z^*_{11}, \cdots$, maintaining the positivity of these diagonal entries at all times. In the end, we obtain

$$\Phi' \Pi^*_D Z^*_{11} Z^*_{12} \cdots Z^*_{pm} = [I_{n+p} \quad 0]$$

where each Z_{ij} is an embedding of z_{ij} as defined in (6). Conversely, after substituting (13) and inverting all rotations, we have a factorization for $\mathbf{\Sigma}_1$. Since $T(z)$ is specified by the first $n+m$ columns of $\mathbf{\Sigma}_1$, it follows that equation (8) holds. $\square$

4. KALMAN PREDICTOR PARAMETRIZATION

Consider the KP model (4). The model takes $\mathbf{u}_k$ and $\mathbf{y}_k$ as inputs and delivers the optimal one-step ahead predictor $\hat{\mathbf{y}}_k$ as output. However, $\mathbf{y}_k$ does not appear in the measurement equation of (4). Thus, a slightly modified parametrization strategy is necessary for this case. The system matrix describing the KP model takes the form

$$\mathbf{T} = \begin{array}{c} \\ n \\ p \end{array} \begin{array}{c} \begin{array}{ccc} n & m & p \end{array} \\ \begin{bmatrix} \tilde{A} & \tilde{B} & K \\ C & D & 0 \end{bmatrix} \end{array} . \quad (14)$$

If we scale the input-output data such that $\mathbf{T}$ is (strictly) contractive, then a parametrization of $\mathbf{T}$ may be obtained as in theorem 1 (assuming $\tilde{A}$ is stable). However, that parametrization is not minimal, since we already know that the (2,3) block of $\mathbf{T}$ should be equal to zero (in fact the problem is ill-defined without this constraint). The following parametrization is minimal and ensures that the (2,3) block is zero.

Define Q_{ij} and Z_{ij} as in (5), except that the Givens rotations are embedded in matrices of size $n+m+2p$. Also define

$$\hat{\Pi}_D = \begin{bmatrix} I_{n+m} & & \\ & 0 & I_p \\ & I_p & 0 \end{bmatrix}, \quad \tilde{\Pi}_D = \begin{bmatrix} I_n & & & \\ & 0 & I_p & \\ & I_m & 0 & \\ & & & I_p \end{bmatrix}$$

Theorem 2. *There is a minimal continuous parametrization with $n(m+2p)+pm$ bounded coefficients, which covers the set of all frational stable contractive Kalman predictors (4) with m inputs, p outputs and n states.*

In particular, every such system may be specified (up to state equivalence) in terms of two matrices $S^{(1)} : n\times(m+2p)$, $S^{(2)} : p\times m$ with entries $-1 \le s^{(\cdot)}_{ij} \le 1$ as

$$\begin{array}{c} \\ n \\ p \end{array} \begin{array}{c} \begin{array}{ccc} n & m & p \end{array} \\ \begin{bmatrix} \tilde{A} & \tilde{B} & K \\ C & D & 0 \end{bmatrix} \end{array} = [I_{n+p} \quad 0_{(n+p)\times(m+p)}] Z_{p,m} \cdots Z_{12} Z_{11} \quad (15)$$

$$\cdot \tilde{\Pi}_D \cdot Q_{n,m+2p} \cdots Q_{12} Q_{11} \cdot \hat{\Pi}_D \Pi_{1,n+1} \begin{bmatrix} I_{n+m+p} \\ 0_{p\times(n+m+p)} \end{bmatrix}$$

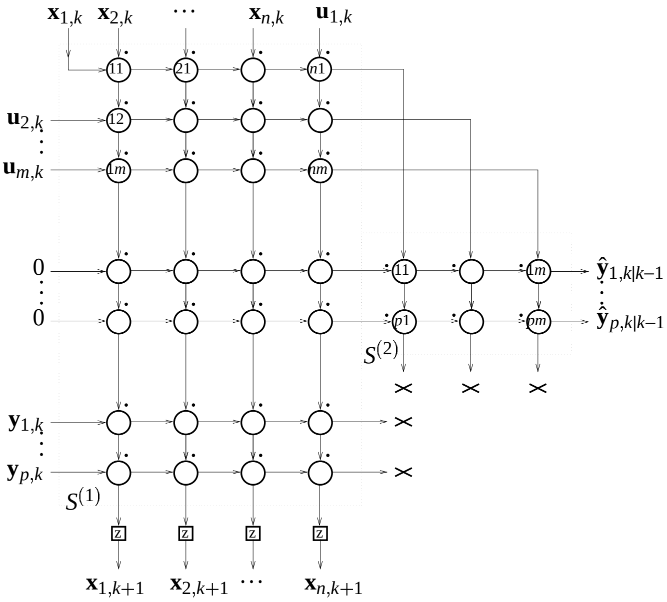


Figure 2. Kalman filter parametrization.

for $Q_{ij} := Q_{ij}(s^{(1)}_{ij}), Z_{ij} := Z_{ij}(s^{(2)}_{ij})$. The parametrization is not unique, but for strictly contractive systems which are controllable via the first input, only a finite (discrete) set of parameter matrices lead to state-equivalent systems.

The proof is mostly the same as the proof of theorem 1 and omitted for brevity. The structure of this parametrization is depicted in figure 2.

References

[1] R. Ober, "Balanced parametrization of classes of linear systems," *SIAM J. Contr. Optim.*, v. 29, pp. 1251–1287, 1991.

[2] C.T. Chou, *Geometry of Linear Systems and Identification.* PhD thesis, Trinity College, Cambridge (UK), March 1994.

[3] E. Deprettere and P. Dewilde, "Orthogonal cascade realization of real multiport digital filters," *Circuit Theory and Appl.*, v. 8, pp. 245–272, 1980.

[4] P. Van Dooren and P.M. Dewilde, "Minimal cascade factorization of real and complex rational transfer matrices," *IEEE Trans. Circuits Syst.*, v. 28, pp. 390–400, May 1981.

[5] P. Dewilde, "Advanced digital filters," in *Modern Signal Processing* (T. Kailath, ed.), pp. 169–209, Springer, 1985.

[6] P.P. Vaidyanathan, "The discrete-time Bounded-Real Lemma in digital filtering," *IEEE Tr. Circ. Syst.*, v. 32, pp. 918–924, 1985.

[7] U.B. Desai, "A state-space approach to orthogonal digital filters," *IEEE Tr. Circ. Syst.*, v. 38, pp. 160–169, Feb. 1991.

[8] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[9] P. VanOverschee and B. DeMoor, "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, v. 30, pp. 75–93, Jan. 1994.

[10] A.J. van der Veen and P.M. Dewilde, "Embedding of time-varying contractive systems in lossless realizations," *Math. Control Signals Systems*, v. 7, pp. 306–330, 1994.