Available online at www.sciencedirect.com
SCIENCE DIRECT®
automatica
Automatica 41 (2005) 863–872
www.elsevier.com/locate/automatica

Brief Paper

Subspace identification of closed loop systems by the orthogonal decomposition method☆

Tohru Katayama[a,*], Hidetoshi Kawauchi[a], Giorgio Picci[b]

[a] *Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan*
[b] *Department of Information Engineering, University of Padova, 35131 Padova, Italy*

Received 3 May 2002; received in revised form 9 November 2004; accepted 25 November 2004

Abstract

In this paper, we consider a problem of identifying the deterministic part of a closed loop system by applying the stochastic realization technique of (Signal Process. 52 (2) (1996) 145) in the framework of the joint input–output approach. Using a preliminary orthogonal decomposition, the problem is reduced to that of identifying the plant and controller based on the deterministic component of the joint input–output process. We discuss the role of input signals in closed loop identification and the realization method based on a finite data, and then sketch a subspace method for identifying state space models of the plant and controller. Since the obtained models are of higher order, a model reduction procedure should be applied for deriving lower order models. Some numerical results are included to show the applicability of the present technique.

Keywords: Closed loop systems; Subspace identification; Stochastic realization; Orthogonal decomposition

1. Introduction

The identification problem for linear systems operating in closed loop has received much attention in the literature (Söderström & Stoica, 1989; Van den Hof, 1997; Forssell & Ljung, 1999). Also, the identification of multivariable systems operating in closed loop by subspace methods has been object of active research in the past decade; among early references using the joint input–output approach, we quote papers (Van der Klauw, Verhaegen, & Van den Bosch, 1991; Verhaegen, 1993).

Since the joint input–output approach requires assumptions of linearity of the feedback channel and is generally computationally demanding, several attempts have been made to identify directly the plant dynamics without estimating the feedback channel. For example modifying the N4SID method (Van Overschee & De Moor, 1996), a closed loop subspace identification method has been derived in Van Overschee and De Moor (1997). However this approach requires that a finite number of Markov parameters of the controller are known. Several other subspace-based closed loop identification have been proposed, e.g. (Chou & Verhaegen, 1999; Ljung & McKelvey, 1996). In general, these methods need some side information on the system and their statistical properties (e.g. consistency) have not been assessed; see e.g. Bauer (2004) and Chiuso and Picci (2004) for an up-to-date discussion of these aspects.

In this paper, we study a joint input–output subspace method for identifying the plant and controller operating in closed loop, by extending the orthogonal decomposition based technique in Katayama, Kawauchi, and Picci (2001, 2002). It is not assumed that all the input signals are purely non-deterministic (and hence persistently exciting of arbitrary order) as in the references above, but one of

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor T. Söderström.

* Corresponding author.
E-mail addresses: katayama@amp.i.kyoto-u.ac.jp (T. Katayama), picci@dei.unipd.it (G. Picci).

0005-1098/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2004.11.026

the exogenous inputs is allowed to be a purely deterministic (p.d.) (or "linearly singular" (Rozanov, 1967)) signal. First we compute the deterministic component of the joint input–output process, that is linearly related to the exogenous inputs, by means of an orthogonal decomposition as in Picci and Katayama (1996). Thereby, the identification of the closed loop system is reduced to that of obtaining the plant and controller based on the deterministic component of the joint input–output process. Based on the realization of the deterministic component, a subspace method for identifying the plant and controller is then derived by adapting some standard subspace method from the literature. In general the obtained models will be of higher order, so that a model reduction procedure should be applied for deriving lower order models.

The organization of the paper is as follows. In Section 2, the problem is stated along with the underlying assumptions. In Section 3, we formulate a joint input–output approach to the closed loop identification problem and derive the deterministic component of the joint input–output process by the preliminary orthogonal decomposition. Section 4 considers the state space realization of the deterministic component, with special emphasis on the role of input signals in closed loop identification and on the realization from a finite data. Section 5 derives some formulas for computing the plant and controller from overall transfer matrices. In Section 6, a subspace method of identifying the plant and controller is briefly sketched, together with a model reduction procedure. Section 7 includes some numerical results and Section 8 concludes the paper.

2. Problem statement

We consider the problem of identifying a closed loop system shown in Fig. 1, where $y \in \mathbb{R}^p$ is the output vector of the plant, and $u \in \mathbb{R}^m$ the input vector. The effect of stochastic unmeasurable disturbances, modeling errors, etc. is described by stationary error (or disturbance) processes $H(z)\xi$ and $F(z)\eta$, acting additively on the outputs of the plant and controller, respectively. The transfer matrices of noise filters $H(z)$ and $F(z)$ will be assumed square rational, minimum phase with $H(\infty) = I_p$ and $F(\infty) = I_m$. The (unmeasurable) inputs $\xi \in \mathbb{R}^p$ and $\eta \in \mathbb{R}^m$ are white noises with mean zero and positive definite covariance matrices. The measurable input signals $r \in \mathbb{R}^p$ and $d \in \mathbb{R}^m$ may be interpreted as the exogenous reference signal and a probing input (dither) or a measurable disturbance.

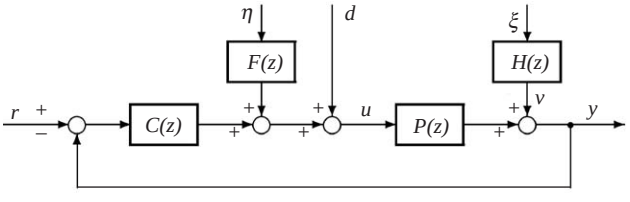

Fig. 1. Closed-loop system.

Let the plant be a finite dimensional, linear, time-invariant (FDLTI) system described by

$$y(t) = P(z)u(t) + H(z)\xi(t), \tag{1}$$

where $P(z)$ and $H(z)$ are the $(p \times m)$- and $(p \times p)$-dimensional transfer matrices of the plant and the noise filter, respectively. The control input is generated by

$$u(t) = d(t) + C(z)[r(t) - y(t)] + F(z)\eta(t), \tag{2}$$

where $C(z)$ and $F(z)$ denote the $(m \times p)$- and $(m \times m)$-dimensional transfer matrices of the FDLTI controller and of the measurement noise filter. The error term $F(z)\eta$ is added for increased generality, so the present setup is the same as in Verhaegen (1993).

The following assumptions are made on the closed-loop system, exogenous inputs, and noises.

A1. The closed loop system is well-posed in the sense that (u, y) are uniquely determined by the states of the plant and controller and by the exogenous inputs and noises. This generic condition is satisfied if $I_p + P(\infty)C(\infty)$ and $I_m + C(\infty)P(\infty)$ are non-singular. For the sake of simplicity, it is assumed throughout the paper that the plant is strictly proper, i.e. $P(\infty) = 0$.

A2. The closed loop system is internally stable.

A3: The random processes d, r, η, ξ are wide-sense stationary, zero-mean[1] and *mutually uncorrelated*, i.e. $d(t)$, $r(s)$, $\eta(\tau)$, $\xi(\sigma)$ are uncorrelated for any $t, s, \tau, \sigma \in \mathbb{Z}$.

We shall realistically model the exogenous reference input r as a p.d. process, which may not necessarily be "exciting" enough to allow consistent identification. Instead d will be assumed to be a purely non-deterministic process with a full rank spectral density matrix. We shall comment on these conditions later on; for the time being, we just remark that a process of this kind is persistently exciting (PE) of any order. Under these assumptions, (u, y) are also jointly stationary second-order processes with zero mean.

The objective of this paper is to obtain state space models of the plant $P(z)$ and the controller $C(z)$ based on finite measurement data $\{d(t), r(t), u(t), y(t)\}$ by using a subspace identification method. The present problem is virtually the same as the one treated in Verhaegen (1993), but the approach is quite different. For, as shown below, we employ the approach based on the orthogonal decomposition of the joint input–output process into the deterministic and stochastic components (Picci & Katayama, 1996); the realization technique is then applied to the deterministic component to obtain state space realizations of the plant and

[1] The assumption that r is zero-mean ($E\{r(t)\} = 0$) does not preclude the sample trajectories of r to exhibit a non-zero constant component. It is made for notational convenience only and could be dispensed with by introducing some extra notations and using lengthier formulas; all results of the paper hold true if $E\{r(t)\} \neq 0$ as well.

controller, which are easily adapted to finite input–output data by using subspace methods.

As we have already pointed out, given the current state of knowledge of subspace identification in the presence of feedback, there seem to be no reliable closed loop subspace methods for identifying directly the transfer function $P(z)$ (and/or $C(z)$) from the measured signals, without assuming some a priori knowledge of certain parameters of the system. This is due either to the correlation of the noise and future inputs due to feedback, and/or to the fact that both $P(z)$ and $C(z)$ may very well have poles on or outside of the closed unit disk, which may lead to very unreliable subspace estimates. Some of these difficulties are analyzed in Chiuso and Picci (2003, 2004). For this reason, if we are to use subspace methods in the presence of feedback, joint identification seems, so far, to be a mandatory choice. Clearly, using the relations (1) and (2), the last of which rewritten as $u(t)-d(t)=C(z)(r(t)-y(t))+F(z)\eta(t)$ it would be possible to do separate identification by prediction error methods. This however is a different setting than the one we want to consider.

3. The joint input–output approach

In this section, we shall be concerned with the case where infinite data are available.

3.1. Joint input–output processes

Define the joint augmented input and output processes as

$$w := \begin{bmatrix} u \\ y \end{bmatrix} \in \mathbb{R}^l, \quad v := \begin{bmatrix} d \\ r \end{bmatrix} \in \mathbb{R}^l \quad \text{and} \quad \chi := \begin{bmatrix} \eta \\ \xi \end{bmatrix} \in \mathbb{R}^l,$$

where $l=m+p$. It then follows from Fig. 1 that these signals are related by the closed loop system transfer functions as

$$w(t) = T_{wv}(z)v(t) + T_{w\chi}(z)\chi(t), \tag{3}$$

where $T_{wv}(z)$ is the closed loop transfer matrix defined by

$$\begin{aligned} T_{wv}(z) &:= \begin{bmatrix} T_{ud}(z) & T_{ur}(z) \\ T_{yd}(z) & T_{yr}(z) \end{bmatrix} \\ &= \begin{bmatrix} S_i(z) & S_i(z)C(z) \\ P(z)S_i(z) & P(z)S_i(z)C(z) \end{bmatrix} \end{aligned} \tag{4}$$

and where $S_i(z)=(I_m+C(z)P(z))^{-1}$ is the input sensitivity matrix. Also, $T_{w\chi}(z)$ is the noise model in the closed loop system matrix; but the identification of the noise model is not treated in this paper.

Recall that the feedback system is internally stable if and only if the four transfer matrices in (4) are stable. Since in (3) v and χ are uncorrelated, there is no feedback from w to v, and hence we can employ an open loop identification method to get estimates of the transfer matrix $T_{wv}(z) := [T_{wd}(z)\ T_{wr}(z)]$ using the measurement of the input and output (v, w).

In order to deal with a well-posed estimation problem of $P(z)$ and $C(z)$, these transfer functions should be uniquely obtainable from the overall transfer function $T_{wv}(z)$ (i.e. at least in the ideal noise-free case when $\chi=0$). This is sometimes called "a priori" *identifiability* in the literature. That $P(z)$ and $C(z)$ are identifiable follows from the formula (4), which yields $P(z)=T_{yd}(z)T_{ud}^{-1}(z)$ and $C(z)=T_{ud}^{-1}(z)T_{ur}(z)$, where the inverse exists because of the invertibility of $S_i(z)$. Note that in order that both $P(z)$ and $C(z)$ be uniquely reconstructible from the data, in general we need to have *both* signals r and d acting on the system. If one of the two signals is absent, we can still solve for $P(z)$ and $C(z)$ from either $T_{wd}(z)$ or $T_{wr}(z)$. But in this case, to compute *both* $P(z)$ and $C(z)$ we need to use a left- or a right-inverse and we do not have uniqueness of the solution (i.e. identifiability) anymore. This is so unless we make special assumptions on $P(z)$ and $C(z)$ (like both of them being square invertible). For example, if $r=0$, so that we are allowed to use only $T_{wd}(z)$, then we can uniquely obtain $P(z)=T_{yd}(z)T_{ud}^{-1}(z)$. But, $C(z)$ would have to be obtained by solving $C(z)T_{yd}(z)=I-T_{ud}(z)$, which in general yields a non-unique solution $C(z)=(I-T_{ud}(z))T_{yd}^{\dagger}(z)$ depending on the choice of the right-inverse, so that $C(z)$ will not be identifiable.

Similarly, when $d=0$, we can write $P(z)=T_{yr}(z)T_{ur}^{\dagger}(z)$ and, unless the generalized inverse is a true inverse, the solution will not be unique. In this case $P(z)$ will not be identifiable. The controller transfer matrix $C(z)$ could instead be recovered uniquely using the formula $C(z)=T_{ur}(z)(I-T_{yr}(z))^{-1}$, where the inverse of the (square) matrix on the right now exists in virtue of the assumptions we have made.

3.2. Decomposition of the joint input–output process

Following Picci and Katayama (1996), we introduce the Hilbert spaces generated by second-order random variables of the exogenous inputs and of the joint input–output signals, which are respectively denoted by $\mathscr{D}=\overline{\text{span}}\{d(t) \mid t\in\mathbb{Z}\}$, $\mathscr{R}=\overline{\text{span}}\{r(t) \mid t\in\mathbb{Z}\}$, and $\mathscr{W}=\overline{\text{span}}\{w(t) \mid t\in\mathbb{Z}\}$. The joint input space $\mathscr{V}$ spanned by v is then the orthogonal direct sum $\mathscr{V}=\mathscr{D}\oplus\mathscr{R}$. We also define the Hilbert subspaces spanned by the infinite past and future of $d(t)$ at the present time t as

$$\mathscr{D}_t^- = \overline{\text{span}}\{d(\tau) \mid \tau<t\}, \quad \mathscr{D}_t^+ = \overline{\text{span}}\{d(\tau) \mid \tau\geqslant t\}.$$

Other subspaces $\mathscr{R}_t^-$, $\mathscr{R}_t^+$, $\mathscr{V}_t^-$, $\mathscr{V}_t^+$ are defined similarly. These are all subspaces of the ambient Hilbert space $\mathscr{H} := \mathscr{V}\vee\mathscr{W}$ spanned by the observable augmented input and output processes (v, w) (assumed for the time being, observable on an infinitely long time interval).

Since there is no feedback from w to v, the future of v is conditionally uncorrelated with the past of w given the past

of v (Picci & Katayama, 1996). It is then easy to prove that

$$w_s(t) := w(t) - \hat{E}\{w(t) \mid \mathscr{V}_{t+1}^-\} = w(t) - \hat{E}\{w(t) \mid \mathscr{V}\} = \hat{E}\{w(t) \mid \mathscr{V}^\perp\},$$

where $\hat{E}\{\cdot \mid \cdot\}$ denotes the orthogonal projection, and $(\cdot)^\perp$ the orthogonal complement. This implies that $w_s(t) \perp \mathscr{V}$ for all $t \in \mathbb{Z}$, and hence $w_s(t)$ coincides with the augmented output of the feedback system which could be measured when the reference and dither signals are set identically equal to zero. The process w_s defined above is hence called the *stochastic component of w*. In the same way, the processes defined by the orthogonal projections

$$w_d(t) := \hat{E}\{w(t) \mid \mathscr{D}\}, \quad w_r(t) := \hat{E}\{w(t) \mid \mathscr{R}\} \tag{5}$$

are called the *d- and r-driven deterministic components*, meaning that they are the parts of w linearly related to the measurable exogenous inputs only.

Thus, we get an orthogonal decomposition of the output process $w = w_d + w_r + w_s$, or

$$\begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} u_d(t) \\ y_d(t) \end{bmatrix} + \begin{bmatrix} u_r(t) \\ y_r(t) \end{bmatrix} + \begin{bmatrix} u_s(t) \\ y_s(t) \end{bmatrix}, \tag{6}$$

where the above three components are uncorrelated at all times; in particular we have $E\{w_s(t)w_d^\top(\tau)\} = 0$ and $E\{w_s(t)w_r^\top(\tau)\} = 0$ for all $t, \tau \in \mathbb{Z}$. From this orthogonal decomposition, it easily follows that the deterministic components of w satisfy the following decoupled equations:

$$y_d(t) = P(z)u_d(t), \quad u_d(t) = d(t) - C(z)y_d(t),$$
$$y_r(t) = P(z)u_r(t), \quad u_r(t) = C(z)[r(t) - y_r(t)]$$

and from these relations we can easily see that the transfer matrices in (4) are determined by the deterministic components of the measurable signals.

Let $\hat{w} := w_d + w_r$, the component of w generated by the observable inputs (d, r). Suppose now that $\hat{w}$ has a minimal state space realization of dimension n of the form[2]

$$x(t+1) = Ax(t) + B_1 d(t) + B_2 r(t), \tag{7a}$$

$$\hat{w}(t) = Cx(t) + D_1 d(t) + D_2 r(t), \tag{7b}$$

where $A \in \mathbb{R}^{n\times n}$ must be a stability matrix, and $C := \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$, $D_i := \begin{bmatrix} D_{1i} \\ 0 \end{bmatrix}$, $i = 1, 2$. It is obvious that the d- and r-driven subsystems of (7), which are formally obtained by projecting the state equations onto $\mathscr{D}$ and $\mathscr{R}$ respectively, can be written in the form

$$x_d(t+1) = Ax_d(t) + B_1 d(t), \tag{8a}$$

$$w_d(t) = Cx_d(t) + D_1 d(t) \tag{8b}$$

and

$$x_r(t+1) = Ax_r(t) + B_2 r(t), \tag{9a}$$

$$w_r(t) = Cx_r(t) + D_2 r(t). \tag{9b}$$

We shall assume that both representations are minimal (i.e. observable and reachable) and hence both are of the same dimension n of the system (7).

Our program now is to construct a (minimal) state space model of the type (7), starting from "data" consisting of the input $v = (d, r)$ and output $\hat{w} = (\hat{u}, \hat{y})$ processes. This is an abstract prototype of the identification problem we have posed at the beginning of this paper.

4. Realization of closed-loop system

In this section, we shall construct a state space realization for the joint "deterministic" input–output process described in the previous section. The technique will be based on stochastic realization ideas described in Picci and Katayama (1996), Katayama and Picci (1999) and Verhaegen (1994).

4.1. Input signals for closed-loop identification

We have allowed the reference input signal r to be p.d. It is well known that the infinite past and infinite future spaces of such a process contain the same information, i.e. they satisfy the relation $\mathscr{R}_t^- = \mathscr{R}_t^+ = \mathscr{R}$ so that $\mathscr{R}_t^- \cap \mathscr{R}_t^+ = \mathscr{R}$ (see e.g. Rozanov (1967)). This fact forbids a straightforward application of the stationary stochastic realization procedure of Picci and Katayama (1996) and Katayama and Picci (1999). However, for most p.d. process which are in practice used to model deterministic input signals, suitably short *finite* past and future histories are in general different. We shall henceforth assume that there is a natural number $h > 0$ such that the finite past subspace $\mathscr{R}_{[t-h,t)} := \overline{\text{span}}\{r(t-j) \mid j = 1, \ldots, h\}$ and the finite future subspace $\mathscr{R}_{[t,t+h]} := \overline{\text{span}}\{r(t+j) \mid j = 0, \ldots, h\}$ are linearly independent, i.e. $\mathscr{R}_{[t-h,t)} \cap \mathscr{R}_{[t,t+h]} = \{0\}$ and

$$\mathscr{R}_{[t-h,t)} + \mathscr{R}_{[t,t+h]} = \mathscr{R}. \tag{10}$$

In fact, we shall denote by h the *smallest* integer for which this condition holds. The dimension of $\mathscr{R}$ apparently corresponds to the so-called *order of excitation* of the signal as defined in the literature (Ljung, 1999). For example, if r is a sum of h sinusoidal oscillations of different frequencies

$$r(t) = \sum_{k=-h}^{+h} r_k \, e^{j\omega_k t}, \quad \omega_{-k} = -\omega_k, \quad \omega_k \neq \omega_j, \tag{11}$$

where the r_k's are pairwise uncorrelated zero-mean random vectors satisfying $\text{cov}\{r_{-k}\} = \text{cov}\{r_k\} = \Sigma > 0$, and where $\text{cov}\{\cdot\}$ denoting the covariance matrix, then r satisfies condition (10). Note that r is just a zero-mean stationary process

[2] In the D matrix we have $D_{21} = 0$, $D_{22} = 0$, due to the fact that the plant $P(z)$ is strictly proper. Furthermore we have $D_{11} = S_i(\infty) = I_m$ and $D_{12} = D_c$, the feedthrough matrix of the controller.

with a pure point spectrum. The process r is of *full rank* if the $(2h+1)p \times (2h+1)p$ covariance matrix $\Sigma_{\mathbf{r}}$ of the stacked random vector

$$\mathbf{r} := [r^\top(-h)\ r^\top(-h+1)\ \cdots\ r^\top(h-1)\ r^\top(h)]^\top$$

is nonsingular. In this case it is easy to show that the dimension of $\mathscr{R}$ is equal to $(2h+1)p$. Sometimes for brevity we shall refer to h itself as the order of excitation of the signal r.

Regarding the purely non-deterministic component d, we shall require that it satisfies the so-called *richness condition*, $\mathscr{D}_t^+ \cap \mathscr{D}_t^- = \{0\}$, so that, for each t we have the direct sum decomposition

$$\mathscr{D}_t^- + \mathscr{D}_t^+ = \mathscr{D}. \tag{12}$$

This is equivalent to the canonical angles between the past and future subspaces of d being strictly positive and is in turn equivalent to the spectral density matrix of d being strictly positive definite on the unit circle, i.e. $\Phi_d(\mathrm{e}^{\mathrm{j}\omega}) > cI_l$, $c > 0$ (Hannan & Poskitt, 1988).

By construction, the n-dimensional state vector of (7) can be decomposed as $x(t) = x_d(t) + x_r(t)$, where $x_d(t)$ and $x_r(t)$ are (componentwise) orthogonal and span the state spaces $\mathscr{X}_t^d$ and $\mathscr{X}_t^r$ of the models (8) and (9), respectively. The following observation will be needed later.

Lemma 1. *Let $T \geqslant t$. For a stationary minimal model of $(\hat{w}, v)$ of the type* (7)*, we have the orthogonal decomposition*

$$\tilde{x}(t) := x(t) - \hat{E}\{x(t) \mid \mathscr{V}_{[t,T]}\} = \tilde{x}_d(t) + \tilde{x}_r(t), \tag{13}$$

where $\tilde{x}_d(t) := x_d(t) - \hat{E}\{x_d(t) \mid \mathscr{D}_{[t,T]}\}$ and $\tilde{x}_r(t) := x_r(t) - \hat{E}\{x_r(t) \mid \mathscr{R}_{[t,T]}\}$. Whenever a non-zero dither signal is acting on the system ($d \neq 0$), the covariance matrices of $\tilde{x}_d(t)$ and of $\tilde{x}(t)$ are positive definite.

Proof. The first statement follows from the orthogonal decomposition

$$\hat{E}\{x(t) \mid \mathscr{V}_{[t,T]}\} = \hat{E}\{x(t) \mid \mathscr{D}_{[t,T]}\} + \hat{E}\{x(t) \mid \mathscr{R}_{[t,T]}\},$$

where in the two terms at the right-hand side we can substitute $x(t) = x_d(t) + x_r(t)$. The second statement follows from the richness condition (12) which implies that the covariance matrix of $\tilde{x}_d(t)$ must be nonsingular (for the components of $x_d(t)$ belong to $\mathscr{D}_t^-$ and no nontrivial linear combination of them can be estimated without error based on the future $\mathscr{D}_{[t,T]}$). Then just recall that $\tilde{x}_d(t) + \tilde{x}_r(t)$ is an orthogonal sum. □

The case $d = 0$ will be examined separately.

Lemma 2. *Consider the r-driven system* (9) *and assume that the input process is a p.d. harmonic process of full rank of the form* (11)*. Then $\tilde{x}_r(t)$ defined in Lemma* 1 *has a positive definite covariance matrix if $T - t < 2h + 1 - n$. Conversely, if $T - t \geqslant 2h + 1 - n$, there is a generic set of systems* [3] (9) *for which the covariance matrix of $\tilde{x}_r(t)$ is singular.*

Proof. See appendix. □

4.2. Realization with finite data

In subspace identification we have finite data and for this reason we are allowed to work with (sample) covariance matrices involving only a finite number of lags. In our setting this is equivalent to having access only to observed random variables $\{v(t)\}$ and $\{\hat{w}(t)\}$ with t belonging to a certain finite interval $[t_0, T]$. Our first concern will be to exhibit a state space realization in which all random quantities are functions of the available random data $\mathscr{V}_{[t_0,T]}$. This realization will involve the same parameters of the stationary model (7). Let $\hat{w}_v(t) := \hat{E}\{\hat{w}(t) \mid \mathscr{V}_{[t_0,T]}\}$ be the projected deterministic output onto the finite available data space. The following result is taken from Picci and Katayama (1996).

Lemma 3. *The n-dimensional projected state vector $\hat{x}(t) := \hat{E}\{x(t) \mid \mathscr{V}_{[t_0,T]}\}$ satisfies the same state space equations of* (7)*; in fact,*

$$\hat{x}(t+1) = A\hat{x}(t) + B_1 d(t) + B_2 r(t), \tag{14a}$$

$$\hat{w}_v(t) = C\hat{x}(t) + D_1 d(t) + D_2 r(t), \tag{14b}$$

$$\hat{x}(t_0) = \hat{E}\{x(t_0) \mid \mathscr{V}_{[t_0,T]}\} \tag{14c}$$

is a state space realization of the projected deterministic output $\hat{w}_v$ in terms of random variables of $\mathscr{V}_{[t_0,T]}$.

If we could construct the state vector $\hat{x}(t)$ as a function of the data $\mathscr{V}_{[t_0,T]}$, we could estimate the parameters (A, B, C, D) of the model by solving a linear regression problem. In practice, the integer $k := t - t_0$, which we shall name the *regression horizon* of the algorithm, is a user-defined parameter which should be chosen large enough so as to satisfy certain conditions which will be discussed below.

Obviously for $k \to \infty$, we have $\hat{x}(t) \to x(t)$ and the model (14) reduces to the stationary realization (7), for which we know how to construct the state space. Unfortunately, for data in a finite interval, the construction of the state space is not entirely straightforward. This difficulty can be fixed in several ways. We shall discuss below a procedure due to Verhaegen (1994), which uses an orthogonal projection onto the orthogonal complement, denoted $\mathscr{V}_{[t,T]}^{\perp}$, of the future inputs $\mathscr{V}_{[t,T]} \subset \mathscr{V}_{[t_0,T]}$. This procedure has the merit of avoiding the computation of oblique projections, which may lead to numerically ill-conditioned problems.

[3] This means that there exists an open, dense set of pairs (A, B_2).

Consider the stacked vectors

$$\hat{w}_t^+ := \begin{bmatrix} \hat{w}_v(t) \\ \hat{w}_v(t+1) \\ \vdots \\ \hat{w}_v(T) \end{bmatrix}, \quad d_t^+ := \begin{bmatrix} d(t) \\ d(t+1) \\ \vdots \\ d(T) \end{bmatrix}$$

and r_t^+ is defined similarly. It follows from (14) that

$$\hat{w}_t^+ = \Gamma_\tau \hat{x}(t) + \Psi_\tau^1 d_t^+ + \Psi_\tau^2 r_t^+, \tag{15}$$

where $\tau := T - t + 1$, and where

$$\Gamma_\tau := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\tau-1} \end{bmatrix} \in \mathbb{R}^{l\tau \times n}$$

is the (extended) observability matrix, and

$$\Psi_\tau^i = \begin{bmatrix} D_i & & & 0 \\ CB_i & D_i & & \\ \vdots & \ddots & \ddots & \\ CA^{\tau-2}B_i & \cdots & CB_i & D_i \end{bmatrix}, \quad i = 1, 2 \tag{16}$$

are the Toeplitz matrices of the first $\tau-1$ Markov parameters of the system.

Projecting (15) onto $\mathscr{V}_{[t,T]}^\perp$, we get

$$\hat{E}\{\hat{w}_t^+ \mid \mathscr{V}_{[t,T]}^\perp\} = \Gamma_\tau \hat{x}^c(t), \tag{17}$$

where $\hat{x}^c(t)$ is the *complementary state*

$$\hat{x}^c(t) := \hat{E}\{\hat{x}(t) \mid \mathscr{V}_{[t,T]}^\perp\} = \hat{x}(t) - \hat{E}\{\hat{x}(t) \mid \mathscr{V}_{[t,T]}\}.$$

A technical condition which we need in this setting is the following "consistency condition"

$$\hat{\mathscr{X}}_t \cap \mathscr{V}_{[t,T]} = \{0\}, \tag{18}$$

where $\hat{\mathscr{X}}_t := \operatorname{span}\{\hat{x}(t)\}$. This condition is essentially the same consistency condition discussed in Jansson and Wahlberg (1998), and it is equivalent to $\hat{x}^c(t)$ *having a positive definite covariance matrix* (to show this one can use the same argument used in the proof of Lemma 1). Although $\hat{x}^c(t)$ does not satisfy state equations of the form (14), if $T - t \geqslant n$ and (18) holds, one can nevertheless identify Γ_τ, modulo a change of basis, from the orthogonal projection in (17) and thereby compute the matrices A, C as in Verhaegen (1994).

Now, provided $d \neq 0$, (18) holds trivially in the stationary case, where $t_0 = -\infty$ (or $k = \infty$) in force of Lemma 1. In fact, in this case we have $\hat{x}^c(t) = \tilde{x}(t)$. It is shown in the literature that for a finite regression horizon, the condition (18) can certainly be satisfied for k and $T - t$ large enough (Bauer & Jansson, 2000). The following statement summarizes the discussion above.

Proposition 1. *Assume $d \neq 0$. If k and $T - t$ are chosen larger than the system order n, the consistency condition* (18) *holds generically and the observability matrix Γ_τ can be obtained from the orthogonal projection in* (17)*, modulo a right multiplication by an $n \times n$ nonsingular matrix.*

Thus when the dither signal is non-zero, the consistency condition (18) can hold even if r is not exciting enough, in particular h could be smaller than n and still we could consistently identify the A, C parameters since Proposition 1 covers in particular the case where $r \equiv 0$. However, in case $d = 0$ the condition may fail when k and $T - t$ are chosen too large since, as we shall see, the two subspaces in (18) may then have a non-empty intersection, which will be discussed in the next subsection.

4.3. The case of no dither signal

We discuss what can happen when there is no dither signal injected in the loop, i.e. $d = 0$. Naturally, we assume that we are using all "excitation power" of the input signal so that $[t_0, T] \equiv [t - h, t + h]$. Taking a smaller interval would be equivalent to working with a fictitious input signal $\bar{r}$ with a smaller number of independent sinusoidal components, generating a smaller input space $\bar{\mathscr{R}} := \mathscr{R}_{[t-\bar{h},t+\bar{h}]} \subset \mathscr{R}$. This means that hereafter we shall have $k = h$ and $T - t = h$.

Proposition 2. *Let $d = 0$. Assume that the input signal r is a p.d. full rank harmonic signal persistently exciting of order h, and that both the regression horizon k and $T - t$ are chosen equal to h. Then the consistency condition* (18) *holds if and* (*generically*) *only if $h \geqslant n$.*

Proof. Since $\mathscr{V}_{[t_0,T]} = \mathscr{R}_{[t-h,t+h]} \equiv \mathscr{R}$, the state of the finite data model (14) (with $d = 0$) coincides with the stationary state $x_r(t)$. Therefore Lemma 2 applies and (18) holds in this case if and only if $n + T - t < 2h + 1$. Substituting $T - t = h$ gives the assertion. □

We see that under the conditions of Proposition 2, the covariance matrix of $\hat{x}^c(t)$ is still invertible and we can identify the A, C parameters of the r-driven model consistently. Note however that if we are to identify (B_2, D_{12}) in (7), the signal r should have an order of excitation large enough; there is a simple argument to get at least a sufficient upper bound. Observe that, once A and C have been determined in a minimal realization, the basis in which the system (7) is represented is fixed so that the Markov parameters of (9) determine B_2 and D_{12} uniquely. In fact, these unknown parameters can be obtained uniquely from the Toeplitz matrix Ψ_τ^2 of (16).

We show how to compute the Toeplitz matrix Ψ_τ^2. From (15), we can compute $Z := \hat{E}_{\parallel \mathscr{X}_t^r}\{\hat{w}_t^+ \mid \mathscr{R}_{[t,t+h]}\}$, the oblique projection of $\hat{w}_t^+$ onto $\mathscr{R}_{[t,t+h]}$ along $\mathscr{X}_t^r$, to obtain the linear relation $Z = \Psi_\tau^2 r_t^+$, which can be interpreted as a linear regression in the Markov parameters, so that it can be solved by using the least-squares method.

5. Extracting plant and controller models

Assume now that we have estimated the joint model (7). Then by simple manipulations of the joint state space equations (7), we can derive state space models for the plant and controller. The formulas are collected in the following proposition.

Proposition 3. *A* (*non-minimal*) *state space representations of the plant and controller are respectively given by*

$$P(z)=\left[\begin{array}{c|c} A-B_1C_1 & B_1 \\ \hline C_2 & 0 \end{array}\right] \tag{19}$$

and

$$C(z)=\left[\begin{array}{c|c} A-B_1C_1 & B_2-B_1D_{12} \\ \hline C_1 & D_{12} \end{array}\right]. \tag{20}$$

Proof. It follows from (7) that the closed loop transfer matrices are expressed as

$$\begin{bmatrix} T_{ud}(z) & T_{ur}(z) \\ T_{yd}(z) & T_{yr}(z) \end{bmatrix}=\left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & 0 & 0 \end{array}\right].$$

We see from (4) that the plant and controller are computed from $P(z)=T_{yd}(z)T_{ud}^{-1}(z)$ and $C(z)=T_{ud}^{-1}(z)T_{ur}(z)$, respectively. The formulas are easily derived by using standard operations on the state space models. □

Let $x_p(t)\in\mathbb{R}^{n_p}$ and $x_c(t)\in\mathbb{R}^{n_c}$ be state vectors of the plant and the controller, respectively and let the state space models of the plant and of the controller be respectively given by

$$x_p(t+1)=A_px_p(t)+B_pu(t), \tag{21a}$$

$$y(t)=C_px_p(t) \tag{21b}$$

and

$$x_c(t+1)=A_cx_c(t)+B_c\{r(t)-y(t)\}, \tag{22a}$$

$$u(t)=d(t)+C_cx_c(t)+D_c\{r(t)-y(t)\}. \tag{22b}$$

The following proposition shows that the models in Proposition 3 are not necessarily minimal.

Proposition 4. *Let* (21) *and* (22) *be minimal realizations of the plant and of the controller, and assume there are no pole-zero cancellations in forming the product* $C(z)P(z)$. *Then the realizations*

$$P(z)=\left[\begin{array}{cc|c} A_p & 0 & B_p \\ -B_cC_p & A_c & 0 \\ \hline C_p & 0 & 0 \end{array}\right] \tag{23}$$

and

$$C(z)=\left[\begin{array}{cc|c} A_p & 0 & 0 \\ -B_pC_p & A_c & B_c \\ \hline -D_cC_p & C_c & D_c \end{array}\right] \tag{24}$$

are similar to (19) *and* (20), *i.e. they can be obtained one from the other by a* (*nonsingular*) *change of basis in the state space. Hence* (19) *and* (20) *are non-minimal*.

Proof. Follows from the argument in Verhaegen (1993). □

Remark 1. Since the obtained transfer matrices of the plant and the controller have higher dimension than the true ones, we shall need to perform a model reduction step in order to recover lower dimensional models.

Remark 2. As noted earlier, a number of different expressions for $P(z)$ and $C(z)$ can be obtained from the transfer function $T_{wv}(z)$. One may wonder which of these expressions should provide the most reliable estimates of $P(z)$ and $C(z)$. Even if at this stage we do not know an answer to this question, it should be noted that when *both* input signals are used for the estimation, all of these expressions come out from formal manipulations of the *same* joint state space model, which is the one estimated using the data. For example we may insist in using only entries of the *r*-driven subsystem transfer matrix, say $P(z)=T_{yr}(z)T_{ur}^{\dagger}(z)$, etc. Since $T_{yr}(z)$ and T_{ur} represent subsystems of the model (7), which was estimated using both input signals, there seems to be no good reason why this estimate should in general be worse (or better) than the one computed in Proposition 3. What really seems to matter most is the accuracy in estimating the overall model (7). However further analysis is needed to clarify this question.

6. Closed-loop subspace identification algorithm

We briefly discuss a subspace identification method based on the data measured on a finite interval. Suppose that the input–output data[4] $\{d(t), r(t), u(t), y(t), t=0, 1, \ldots, N+2k-2\}$, with $k>n$ and N very large, be sample values from the jointly stationary "true" input–output processes, where n is the dimension of the "true system".

As usual, we fix the present time $t=k$ and define the $kl\times N$ block Hankel matrix generated by the past inputs as

$$V_{0|k-1}=\begin{bmatrix} v(0) & v(1) & \cdots & v(N-1) \\ v(1) & v(2) & \cdots & v(N) \\ \vdots & \vdots & \ddots & \vdots \\ v(k-1) & v(k) & \cdots & v(N+k-2) \end{bmatrix}.$$

The block Hankel matrices $W_{0|k-1}\in\mathbb{R}^{kl\times N}$ formed by the past outputs and $V_{k|2k-1}, W_{k|2k-1}\in\mathbb{R}^{kl\times N}$ formed by the

[4] In this section we shall denote measured sample values of a stochastic process by the same symbol of the corresponding random quantity. Since from now on we shall only work with measured data, this should cause no confusion.

future data are defined similarly. We also define the stacked block Hankel matrices $V_{0|2k-1}$ and $W_{0|2k-1}$. The subspace of $\mathbb{R}^N$ generated by the rows of the Hankel matrix $V_{0|2k-1}$ is denoted by $\mathscr{V}_{0|2k-1}$.

The first step of the subspace identification is the preliminary decomposition of w into the deterministic and stochastic components based on finite input–output data. This is performed by computing the LQ decomposition (Verhaegen, 1994; Van Overschee & De Moor, 1996)

$$\begin{bmatrix} V_{k|2k-1} \\ V_{0|k-1} \\ W_{0|k-1} \\ W_{k|2k-1} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ L_{31} & L_{32} & L_{33} & 0 \\ L_{41} & L_{42} & L_{43} & L_{44} \end{bmatrix} \begin{bmatrix} Q_1^\top \\ Q_2^\top \\ Q_3^\top \\ Q_4^\top \end{bmatrix}, \quad (25)$$

where $L_{11}, L_{22}, L_{33}, L_{44} \in \mathbb{R}^{kl\times kl}$ are the lower triangular matrices and where $Q_i^\top Q_j = I_{kl}\delta_{ij}$. Noting that the rows of $Q_1^\top, Q_2^\top$ form an orthonormal basis for the rowspace $\mathscr{V}_{0|2k-1}$, it follows that the (sample) deterministic component $\hat{W}^v_{0|2k-1} := \hat{E}\{W_{0|2k-1} \mid \mathscr{V}_{0|2k-1}\}$ is given by

$$\hat{W}^v_{0|2k-1} = \begin{bmatrix} L_{31} & L_{32} \\ L_{41} & L_{42} \end{bmatrix} \begin{bmatrix} Q_1^\top \\ Q_2^\top \end{bmatrix}. \quad (26)$$

Let $X_t \in \mathbb{R}^{n\times N}$ be the string of sample state vectors of the true model (7), and let $\hat{X}^v_t$ be the string X_t projected onto $\mathscr{V}_{0|2k-1}$. Clearly the tth block row $\hat{W}^v_t$ of $\hat{W}^v_{0|2k-1}$ satisfies the matrix state space equations

$$\hat{X}^v_{t+1} = A\hat{X}^v_t t + [B_1\ B_2]V_t, \quad \hat{W}^v_t = C\hat{X}^v_t + [D_1\ D_2]V_t.$$

It may be noted from Lemma 3 that $\hat{W}^v_t$ corresponds to (15), the same state-space equation satisfied by the optimal W^v_t of (7), but with different initial states. Hence, neglecting the difference in initial states, we have the matrix input–output equation (Verhaegen, 1994; Van Overschee & De Moor, 1996)

$$\hat{W}^v_{k|2k-1} = \Gamma_k \hat{X}^v_k + \Psi_k V_{k|2k-1}, \quad (27)$$

where $\Psi_k := [\Psi_k^1\ \Psi_k^2]$. Clearly, it follows from (25)–(27) that

$$\hat{W}^v_{k|2k-1} = \Gamma_k \hat{X}^v_k + \Psi_k L_{11} Q_1^\top = L_{41}Q_1^\top + L_{42}Q_2^\top.$$

Post-multiplying the above equation by Q_2 yields

$$\hat{W}^v_{k|2k-1}Q_2 = \Gamma_k \hat{X}^v_k Q_2 = L_{42}, \quad (28)$$

where the first equality of the above equation corresponds to (17). Assuming that $\hat{X}^v_k Q_2$ has full rank (Verhaegen, 1994), we see that $\mathrm{Im}(\Gamma_k) = \mathrm{Im}(L_{42})$. Let the SVD of (28) be given by $L_{42} = \hat{U}\hat{\Sigma}\hat{V}^\top$. Then, the estimate of the extended observability matrix is given by $\Gamma_k = \hat{U}\hat{\Sigma}^{1/2}$. The subsequent steps of the MOESP-like subspace identification algorithm for the B, D parameters are well known and will not be discussed here. The above subspace identification method is called the ORT method; see also Chiuso and Picci (2001).

All the identified state space models have higher dimensions than the true ones, so that we need a model reduction procedure to delete nearly unreachable and/or unobservable modes. Since the plant and/or the controller may possibly be unstable, we use a model reduction technique called the square root (SR) method (Varga, 2001), which can be applied to unstable transfer matrices.

7. Simulation results

Some simulation results are included to show the applicability of the present technique. Suppose that the plant, controller and two noise models are given by Van den Hof and Schrama (1993)

$$P(z) = \frac{z^{-1}}{1 - 1.6z^{-1} + 0.89z^{-2}}, \quad C(z) = 1 - 0.8z^{-1}$$

and

$$F(z) = 1, \quad H(z) = \frac{1 - 1.56z^{-1} + 1.045z^{-2} - 0.3338z^{-3}}{1 - 2.35z^{-1} + 2.09z^{-2} - 0.6675z^{-3}}.$$

The configuration of the feedback system is the same as the one shown in Fig. 1, where d, η and ξ are Gaussian white noises with variances $\sigma_d^2 = 0.2$, $\sigma_\eta^2 = 0.01$ and $\sigma_\xi^2 = \frac{1}{9}$, respectively. The reference input r is a sinusoidal signal generated by $r(t) = \rho\sum_{j=1}^{30} A_j \sin(\omega_j t + \phi_j)$, $t = 0, 1, \ldots, N + 2k - 2$, where ρ is a normalizing constant yielding $\sigma_r^2 = 1$, and A_j are Gaussian random numbers with $\mathcal{N}(0, 1)$, and ω_j, ϕ_j are uniformly distributed over $(0, \pi)$.

Assuming that the orders of plant and controller are known, third-order state space models are fitted to the input v and output w, because the order of the deterministic component is three. Then, the identified plants and controllers are reduced to two- and one-dimensional models, respectively. We take the number of data points $N = 2000$ and the number of block rows $k = 15$, and generated 30 data set, each with different samples for d, r, η and ξ. Fig. 2 shows the estimated poles of plant and controller, where + and ×

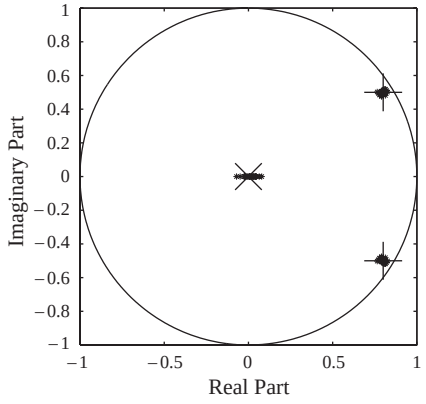

Fig. 2. Estimated poles of plant and controller over 30 runs, where + and ×, respectively denote the true poles of plant and controller.

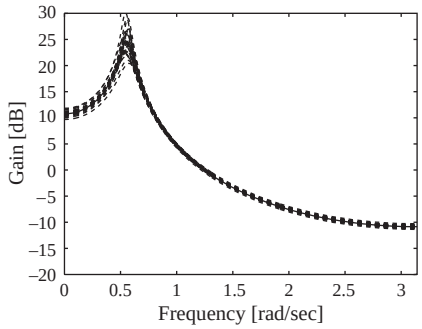

Fig. 3. Estimated bode plots of plant transfer function $P(z)$.

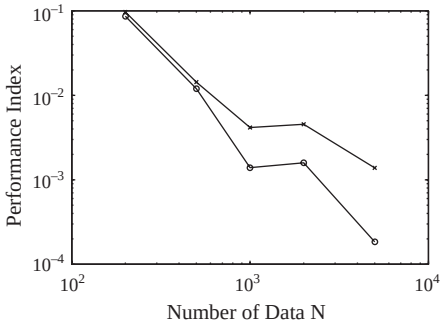

Fig. 4. Performance of identification by ORT (∘) and CCA (×).

denote the true poles of plant and controller, respectively. Fig. 3 shows Bode plots of the estimated plant. We see from these results that the identification result of this simulation is quite good.

We also consider the effect of the number of data on the closed loop identification. The performance is measured by the norm of the estimation error of the plant parameter vector $\theta := (-1.6\ 0.89\ 1\ 0) \in \mathbb{R}^4$, i.e.

$$I_N = \frac{1}{M}\sum_{i=1}^{M} \|\theta - \hat{\theta}(i, N)\|^2,$$

where $\hat{\theta}(i, N) \in \mathbb{R}^4$, $N = 200, 500, 1000, 2000, 5000$ denotes the estimate of θ at ith run with data N, and the number of runs is $M = 30$ in each case. Fig. 4 compares the performance of the identification of plant transfer function by the ORT method and CCA method (Katayama & Picci, 1999). This shows the advantage of the present ORT-based algorithm.

8. Conclusions

In this paper we have developed a subspace method for identifying the deterministic part, i.e. the plant and controller, of closed loop systems in the joint input–output framework. It is assumed that one of the exogenous inputs is purely deterministic and the other is purely non-deterministic. We have discussed the realization method based on a finite data and the role of input signals in closed loop system identification, and derived a subspace method to identify the plant and controller. Numerical results are included to show the applicability of the present approach.

Acknowledgements

The authors wish to thank H. Tanaka for his assistance in performing simulations.

Appendix A. Proof of Lemma 2

Because of stationarity it is enough to prove the statement for $t = 0$. It is easy to check that the covariance matrix of $\tilde{x}_r(0)$ is positive definite if and only if $\mathscr{X}_0^r \cap \mathscr{R}_{[0,T-t]} = \{0\}$, i.e. the two subspaces are linearly independent. Now introduce

$$A(z) := (zI - A)^{-1}B_2$$

and let $z_k := e^{j\omega_k}$, $k = 0, \pm 1, \ldots, \pm h$. Since $x_r(0) = \sum_{k=-h}^{h} A(z_k)r_k$, the space $\mathscr{X}_0^r$ is spanned by the components of the (n-dimensional) vector

$$[A(z_{-h})\ A(z_{-h+1})\ \cdots\ A(z_{h-1})\ A(z_h)]\mathbf{r} =: \mathfrak{A}\,\mathbf{r}$$

while $\mathscr{R}_{[0,T-t]}$ is spanned by the components of

$$\left[\begin{bmatrix} 1 & \cdots & 1 \\ z_{-h} & \cdots & z_{-h} \\ \cdots & \cdots & \cdots \\ z_{-h}^{T-t} & \cdots & z_{-h}^{T-t} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ z_{-h+1} & \cdots & z_{-h+1} \\ \cdots & \cdots & \cdots \\ z_{-h+1}^{T-t} & \cdots & z_{-h}^{T-t} \end{bmatrix} \cdots \begin{bmatrix} 1 & \cdots & 1 \\ z_{h-1} & \cdots & z_{h-1} \\ \cdots & \cdots & \cdots \\ z_{h-1}^{T-t} & \cdots & z_{h-1}^{T-t} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ z_h & \cdots & z_h \\ \cdots & \cdots & \cdots \\ z_h^{T-t} & \cdots & z_h^{T-t} \end{bmatrix}\right] \times \mathbf{r} =: \mathfrak{Z}^+\mathbf{r},$$

where each block matrix has p columns. Hence the random variables generating $\mathscr{X}_0^r$ and $\mathscr{R}_{[0,T-t]}$ are represented isometrically by the rows of the matrices $\mathfrak{A}$ and $\mathfrak{Z}^+$ in the space $\mathbb{C}^{(2h+1)p}$ with inner product weighted by the positive definite covariance $\Sigma_{\mathbf{r}}$. It follows that $\mathscr{X}_0^r \cap \mathscr{R}_{[0,T-t]} = \{0\}$, if and only if the rowspaces of $\mathfrak{A}$ and $\mathfrak{Z}^+$ have only the zero vector in common. Consider first the case where $p = 1$, and assume that there exist two vectors $c \in \mathbb{C}^n$ and $a \in \mathbb{R}^{T-t+1}$ such that $c^*\mathfrak{A} = a^*\mathfrak{Z}^+ \neq 0$, where $*$ denotes the conjugate transpose. Since the rows of $\mathfrak{Z}^+$ are linearly independent ($z_k \neq z_j$), c cannot be zero and hence there is a proper rational function

$$\frac{n(z)}{d(z)} := c^*(zI - A)^{-1}B_2,$$

which (generically) is of degree n satisfying

$$\frac{n(z_k)}{d(z_k)} = a(z_k), \quad k = 0, \pm 1, \ldots, \pm h, \tag{A.1}$$

where $a(z) = a^*[1\ z\ \cdots\ z^{T-t}]$ is a polynomial with degree $T-t$. This means that the polynomial $a(z)d(z)-n(z)$ with degree $n+(T-t)$ must have $2h+1$ distinct zeros $\{z_k,\ k=0,\pm1,\ldots,\pm h\}$. Clearly, for this to be possible, $n+(T-t)\geqslant 2h+1$ must hold. On the other hand, since all polynomials $a(z)d(z)-n(z)$ can have at most degree $n+(T-t)$, if $n+(T-t)<2h+1$, the equality in (A.1) cannot hold, and hence $\mathscr{X}^r_0\cap\mathscr{R}_{[0,T-t]}=\{0\}$. The case where $p>1$ can be dealt with similarly.

References

Bauer, D. (2004). Asymptotic properties of subspace estimates. *Preprint*, Technical University, Wien.

Bauer, D., & Jansson, M. (2000). Analysis of the asymptotic properties of the MOESP type of subspace algorithms. *Automatica*, *36*(4), 497–509.

Chiuso, A., & Picci, G. (2001). Some algorithmic aspects of subspace identification with inputs. *International Journal of Applied Mathematics & Computer Science*, *11*(1), 55–75.

Chiuso, A., & Picci, G. (2003). Constructing the state of random processes with feedback. *Proceedings of the 13th IFAC Symposium on System Identification* (*SYSID*) (pp. 881–886). Rotterdam.

Chiuso, A., & Picci, G. (2004). Consistency analysis of some closed-loop subspace identification methods. *Proceedings of the 43rd IEEE Conference on Decision & Control*. Paradise Island, The Bahamas.

Chou, C. T., & Verhaegen, M. (1999). Closed-loop identification using canonical correlation analysis. *Proceedings of the 1999 European Control Conference*. Karlsruhe, F-162.

Forssell, U., & Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, *35*(7), 1215–1241.

Hannan, E. J., & Poskitt, D. S. (1988). Unit canonical correlations between future and past. *Annals of Statistics*, *6*, 784–790.

Jansson, M., & Wahlberg, B. (1998). On consistency of subspace methods for system identification. *Automatica*, *34*(12), 1507–1519.

Katayama, T., & Picci, G. (1999). Realization of stochastic systems with exogenous inputs and subspace identification methods. *Automatica*, *35*(10), 1635–1652.

Katayama, T., Kawauchi, H., & Picci, G. (2001). A subspace identification of the deterministic part of state space model operating in closed loop. *Proceedings of the 2001 European Control Conference* (pp. 2505–2510). Porto.

Katayama, T., Kawauchi, H., & Picci, G. (2002). Subspace identification of closed loop systems by stochastic realization. *Preprints 15th IFAC World Congress*. Barcelona, #T-Mo-M02-2.

Ljung, L. (1999). *System identification—theory for the user*. (2nd ed.), Englewood Cliffs, NJ: Prentice-Hall.

Ljung, L., & McKelvey, T. (1996). Subspace identification from closed loop data. *Signal Processing*, *52*(2), 209–216.

Picci, G., & Katayama, T. (1996). Stochastic realization with exogenous inputs and "subspace methods" identification. *Signal Processing*, *52*(2), 145–160.

Rozanov, Y. A. (1967). *Stationary random processes*. Holden-Day.

Söderström, T., & Stoica, P. (1989). *System identification*. Englewood Cliffs, NJ: Prentice-Hall.

Van den Hof, P. M. J. (1997). Closed-loop issues in system identification. *Proceedings of the 11th IFAC Symposium on System Identification* (*SYSID*) (pp. 1651–1664). Kitakyushu, Japan.

Van den Hof, P. M. J., & Schrama, R. J. P. (1993). An indirect method for transfer function estimation from closed loop data. *Automatica*, *29*(6), 1523–1527.

Van der Klauw, A. C., Verhaegen, M., & Van den Bosch, P. P. J. (1991). State space identification of closed loop systems. *Proceedings of the 30th IEEE Conference on Decision & Control* (pp. 1327–1332). Brighton.

Van Overschee, P., & De Moor, B. (1996). *Subspace identification for linear systems*. Dordrecht: Kluwer Academic Publications.

Van Overschee, P., & De Moor, B. (1997). Closed loop subspace systems identification. *Proceedings of the 36th IEEE Conference on Decision & Control* (pp. 1848–1853). San Diego, CA.

Varga, A. (2001). On balancing and order reduction of unstable periodic systems. *Preprints IFAC Workshop on Periodic Control Systems* (pp. 177–182). Como, Italy.

Verhaegen, M. (1993). Application of a subspace model identification technique to identify LTI systems operating in closed loop. *Automatica*, *29*(4), 1027–1040.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input–output data. *Automatica*, *30*(1), 61–74.

Tohru Katayama received the B.E., M.E. and Ph.D. degrees all in applied mathematics and physics from Kyoto University, Kyoto, in 1964, 1966 and 1969, respectively. Since 1986, he has been in the Department of Applied Mathematics and Physics, Kyoto University. He had visiting positions at UCLA from 1974–1975, and at University of Padova in 1997. He was an Associate Editor of IEEE Transactions on Automatic Control from 1996 to 1998, and is now a Subject Editor of International Journal of Robust and Nonlinear Control, and the Chair of IFAC Technical Committee of Stochastic Systems for 1999–2002, and is now the Chair of IFAC Coordinating Committee of Signals and Systems for 2002–2005. His research interests includes estimation theory, stochastic realization, subspace method of identification, blind identification, and control of industrial processes.

Hidetoshi Kawauchi received the B.E. and M.E. degrees both in applied mathematics and physics from Kyoto University, Kyoto, in 1999 and 2001, respectively. Since 2001, he has been an engineer of development at the Sony Corporation, Shinagawa, Japan. His research interests includes theory of subspace system identification of closed loop systems.

Giorgio Picci holds a full professorship with the University of Padova, Italy, Department of Information Engineering, since 1980. He has held several long-term visiting appointments with various American and European universities among which Brown University, M.I.T., the University of Kentucky, Arizona State University, the Center for Mathematics and Computer Sciences (C.W.I.) in Amsterdam, the Royal Institute of Technology, Stockholm, Sweden, Kyoto University and Washington University, St. Louis, MO. He has been contributing to Systems and Control theory mostly in the area of modeling, estimation and identification of stochastic systems and published over 100 papers and edited three books in this area. He has been involved in various joint research projects with industry and state agencies. He is currently general coordinator of the Commission of European Communities IST project RECSYS, of the fifth Framework Program. Giorgio Picci is a Fellow of the IEEE, past chairman of the IFAC Technical Committee on Stochastic Systems and a member of the EUCA council.