Geometry of oblique splitting subspaces, minimality and Hankel operators

To Anders (Professor Noci) as a token of esteem and friendship over many years

Alessandro Chiuso and Giorgio Picci

Department of Information Engineering, University of Padova, Padova, Italy; Email: {chiuso,picci}@dei.unipd.it

Key words and phrases. Stochastic Realization, Identification, Stochastic systems with exogenous inputs, Subspace-methods Identification.

CHAPTER 1

Geometry of oblique splitting subspaces, minimality and Hankel operators

Abstract

Stochastic realization theory provides a natural theoretical background for recent identification methods, called *subspace methods*, which have shown superior performance for multivariable state-space model-building. The basic steps of subspace algorithms are geometric operations on certain vector spaces generated by observed input-output time series which can be interpreted as "sample versions" of the abstract geometric operations of stochastic realization theory. The construction of the state space of a stochastic process is one such basic operation.

In the presence of exogenous inputs the state should be constructed starting from input-output data observed on a finite interval. This and other related questions still seems to be not completely understood, especially in presence of *feedback* from the output process to the input, a situation frequently encountered in applications. This is the basic motivation for undertaking a first-principle analysis of the stochastic realization problem with inputs, as presented in this paper. It turns out that stochastic realization with inputs is by no means a trivial extension of the well-established theory for stationary processes (time-series) and there are fundamentally new concepts involved, e.g. in the construction of the state space under possible presence of feedback from the output process to the input. All these new concepts lead to a richer theory which (although far from being complete) substantially generalizes and puts what was known for the time series setting in a more general perspective.

1. Introduction

In this paper we shall study the stochastic realization problem with inputs. Our aim will be to discuss procedures for constructing state space models for a stationary process $\boldsymbol{y}$ "driven" by an exogenous observable input signal $\boldsymbol{u}$, also modelled as a stationary process, of the form

$$\begin{cases} \boldsymbol{x}(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + G\boldsymbol{w}(t) \\ \mathbf{y}(t) = C\boldsymbol{x}(t) + D\boldsymbol{u}(t) + J\boldsymbol{w}(t) \end{cases} \tag{1.1}$$

where $\boldsymbol{w}$ is a normalized white noise process. We will especially be interested in coordinate-free (i.e. "geometric") procedures by which one could abstractly *construct* state space models of the form (1.1), starting from certain Hilbert spaces of random variables generated by the "data" of the problem, namely the processes $\boldsymbol{y}$ and $\boldsymbol{u}$. We shall also characterize some structural properties of state space models of this kind, like minimality absence of feedback etc.

Stochastic realization theory lies at the grounds of a recent identification methodology, called *subspace identification*, which has shown superior performance especially for multivariable state-space model-building, and has been intensively investigated in the past ten years [**11, 18, 19, 16, 27**]. The basic steps of subspace algorithms are geometric operations on certain vector spaces generated by observed input-output time series. These operations can be interpreted as "sample versions" of certain abstract geometric operations of stochastic realization theory [**16, 22, 24**]. In fact, it is by now well understood that stochastic realization theory provides a natural theoretical background for subspace identification of time-series (no inputs). The celebrated subspace algorithm of [**18**] uses the sample-version of a standard geometric construction of the state space (projection of the future onto the past) and computes the G, J parameters of the model by solving the Riccati equation of stochastic realization.

The situation is unfortunately not as clear for identification in the presence of exogenous inputs. Some basic conceptual issues underlying the algorithms remain unclear (see [**4**]). One such issue is how the state space of a stochastic process in the presence of exogenous inputs should be constructed starting from input-output data observed on a finite interval. This and other related questions are examined in the recent paper [**4**]. On the basis of the analysis of this paper, one may be led to conclude that all identification procedures with inputs appeared so far in the literature use *ad hoc* approximations of the basic step of state-space construction of the output process $\boldsymbol{y}$, and can only lead to suboptimal performance.

This state of affairs is the basic motivation for undertaking a first-principle analysis of the stochastic realization problem with inputs, as presented in this paper. We warn the reader that stochastic realization with inputs is *not* a trivial extension of the well-established theory for stationary processes (time-series) as there are fundamentally new concepts involved, relating to the construction of the state space, like the possible presence of *feedback* [**7, 6**] from the output process to the input, the diverse notion of minimality etc.. All these new concepts lead to a richer theory which substantially generalizes and puts what was known for the time series setting in a more general perspective.

In order to construct state-space descriptions of $\boldsymbol{y}$ driven by a non-white process $\boldsymbol{u}$ of the above form it is necessary to generalize the theory of stochastic realization of [**14, 15**]. The construction presented here is based on an extension of the idea of Markovian splitting subspace which will be called *oblique Markovian splitting subspace*, a concept introduced in [**24, 10**]. For this reason we shall start the paper by studying oblique projections in a Hilbert space context.

2. Oblique Projections

Let $\mathcal{H}$ be a Hilbert space of real zero-mean random variables with inner product

$$\langle \boldsymbol{x}, \boldsymbol{z} \rangle := \mathrm{E}\,\{\boldsymbol{x}\boldsymbol{z}\} \tag{2.1}$$

the operator E denoting mathematical expectation. All through this paper we shall denote direct sum of subspaces by the symbol $+$. The symbol $\oplus$ will be reserved for *orthogonal* direct sum. Consider a pair of closed subspaces $\mathcal{A}, \mathcal{B}$ of $\mathcal{H}$ which are in direct sum, i.e. $\mathcal{A} \cap \mathcal{B} = \{0\}$ so that every element $\boldsymbol{v} \in \mathcal{A} + \mathcal{B}$ can be uniquely decomposed in the sum

$$\boldsymbol{v} = \boldsymbol{v}_{\mathcal{A}} + \boldsymbol{v}_{\mathcal{B}}, \quad \boldsymbol{v}_{\mathcal{A}} \in \mathcal{A} \;\; \boldsymbol{v}_{\mathcal{B}} \in \mathcal{B}$$

It follows that the orthogonal projection of a random variable $\boldsymbol{z} \in \mathcal{H}$, on $\mathcal{A}+\mathcal{B}$ admits the unique decomposition

$$\mathrm{E}\,[\,\boldsymbol{z} \mid \mathcal{A} + \mathcal{B}\,] = \boldsymbol{z}_{\mathcal{A}} + \boldsymbol{z}_{\mathcal{B}}$$

the two components $\boldsymbol{z}_{\mathcal{A}}$ and $\boldsymbol{z}_{\mathcal{B}}$ being, by definition, the oblique projection of $\boldsymbol{z}$ onto $\mathcal{A}$ along $\mathcal{B}$ and the oblique projection of $\boldsymbol{z}$ onto $\mathcal{B}$ along $\mathcal{A}$, denoted by the symbols

$$\boldsymbol{z}_{\mathcal{A}} = \mathrm{E}_{\|\mathcal{B}}\,[\,\boldsymbol{z} \mid \mathcal{A}\,] \quad , \quad \boldsymbol{z}_{\mathcal{B}} = \mathrm{E}_{\|\mathcal{A}}\,[\,\boldsymbol{z} \mid \mathcal{B}\,]$$

If $\mathcal{A}$ and $\mathcal{B}$ are orthogonal, then the oblique projection becomes orthogonal, i.e.

$$\boldsymbol{z}_{\mathcal{A}} = \mathrm{E}_{\|\mathcal{B}}\,[\,\boldsymbol{z} \mid \mathcal{A}\,] = \mathrm{E}\,[\boldsymbol{z} \mid \mathcal{A}]$$

which, trivially, does not depend on $\mathcal{B}$.

Projections of one subspace onto another subspace will be encountered frequently. We shall denote these objects by

$$\mathrm{E}\,[\,\mathcal{B} \mid \mathcal{A}\,] := \overline{\mathrm{span}}\{\,\mathrm{E}\,[\boldsymbol{z} \mid \mathcal{A}] \mid \boldsymbol{z} \in \mathcal{B}\,\}$$

and

$$\mathrm{E}_{\|\mathcal{B}}\,[\,\mathcal{C} \mid \mathcal{A}\,] := \overline{\mathrm{span}}\{\,\mathrm{E}_{\|\mathcal{B}}\,[\,\boldsymbol{z} \mid \mathcal{A}\,] \mid \boldsymbol{z} \in \mathcal{C}\,\}$$

The following lemma will be extensively used in the following.

Lemma 2.1. *Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ be closed subspaces of $\mathcal{H}$, where $\mathcal{B} \subset \mathcal{C}$. Assume that*

$$\mathcal{D} \cap \mathcal{C} = \{0\} \tag{2.2}$$

and

$$\mathrm{E}\,[\mathcal{A} \mid \mathcal{D} + \mathcal{C}] = \mathrm{E}\,[\mathcal{A} \mid \mathcal{D} + \mathcal{B}] \tag{2.3}$$

then

$$\mathrm{E}_{\|\mathcal{C}}\,[\mathcal{A} \mid \mathcal{D}] = \mathrm{E}_{\|\mathcal{B}}\,[\mathcal{A} \mid \mathcal{D}] \tag{2.4}$$

Proof. ¿From (2.2) every $\boldsymbol{a} \in \mathcal{A}$ can be uniquely decomposed as $\boldsymbol{a} = (\boldsymbol{a}_{\mathcal{D}} + \boldsymbol{a}_{\mathcal{C}}) \oplus \tilde{\boldsymbol{a}}$ where $\boldsymbol{a}_{\mathcal{D}} \in \mathcal{D}$, $\boldsymbol{a}_{\mathcal{C}} \in \mathcal{C}$, and $\tilde{\boldsymbol{a}} \perp (\mathcal{C} + \mathcal{D})$. It follows from (2.3) that $\boldsymbol{a}_{\mathcal{C}} \in \mathcal{B}$ and therefore $\boldsymbol{a}_{\mathcal{B}} = \boldsymbol{a}_{\mathcal{C}}$, or, more precisely,

$$\mathrm{E}_{\,\|\mathcal{D}}\,[\,\boldsymbol{a} \mid \mathcal{C}\,] \;=\; \mathrm{E}_{\,\|\mathcal{D}}\,[\,\boldsymbol{a} \mid \mathcal{B}\,] \qquad \boldsymbol{a} \in \mathcal{A}$$

By uniqueness of the orthogonal projection, $\mathrm{E}\,[\boldsymbol{a} \mid \mathcal{D} + \mathcal{C}] = \boldsymbol{a}_{\mathcal{D}} + \boldsymbol{a}_{\mathcal{B}} = \mathrm{E}\,[\boldsymbol{a} \mid \mathcal{D} + \mathcal{B}]$ which implies $\mathrm{E}_{\,\|\mathcal{C}}\,[\boldsymbol{a} \mid \mathcal{D}] = \mathrm{E}_{\,\|\mathcal{B}}\,[\boldsymbol{a} \mid \mathcal{D}]$. This equality obviously holds for any finite linear combinations of elements of $\mathcal{A}$. To complete the proof just take closure with respect to the inner product (2.1). $\square$

Note that in general the converse implication (2.4) ⇒ (2.3) is not true since $\mathrm{E}_{\,\|\mathcal{B}}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{D}] = 0$, does not imply that $\boldsymbol{a}_{\mathcal{C}} \in \mathcal{B}$ but only that $\mathrm{E}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{D} + \mathcal{B}] = \mathrm{E}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{B}] = \boldsymbol{a}_{\mathcal{B}}$, i.e. $\boldsymbol{a}_{\mathcal{C}} = \mathrm{E}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{B}] \oplus \tilde{\boldsymbol{a}}_{\mathcal{B}}$ where $\tilde{\boldsymbol{a}}_{\mathcal{B}} \perp \mathcal{D} + \mathcal{B}$ which is for instance always the case if $\mathcal{C} \ominus \mathcal{B} \perp \mathcal{D}$.

Also the following lemma will be of primary importance.

Lemma 2.2. *Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ be closed subspaces of $\mathcal{H}$ where*

$$\mathcal{C} \cap \mathcal{D} = \{0\} \tag{2.5}$$

If $\mathcal{B} \subset \mathcal{C}$ then the following conditions are equivalent:

(1) $\mathrm{E}_{\,\|\mathcal{D}}\,[\mathcal{A} \mid \mathcal{C}] = \mathrm{E}_{\,\|\mathcal{D}}\,[\mathcal{A} \mid \mathcal{B}]$;
(2) $\mathrm{E}\,[\mathcal{A} \mid \mathcal{C} + \mathcal{D}] = \mathrm{E}\,[\mathcal{A} \mid \mathcal{B} + \mathcal{D}]$

Proof. (1 ⇒ 2) By assumption (2.5) every $\boldsymbol{a} \in \mathcal{A}$ can be uniquely decomposed as $\boldsymbol{a} = (\boldsymbol{a}_{\mathcal{C}} + \boldsymbol{a}_{\mathcal{D}}) \oplus \tilde{\boldsymbol{a}}$ where $\boldsymbol{a}_{\mathcal{C}} \in \mathcal{C}$, $\boldsymbol{a}_{\mathcal{D}} \in \mathcal{D}$, and $\tilde{\boldsymbol{a}} \perp (\mathcal{C} + \mathcal{D})$. It follows from (1.) that $\boldsymbol{a}_{\mathcal{C}} \in \mathcal{B}$; in fact (1.) implies that $\boldsymbol{a}_{\mathcal{C}} = \mathrm{E}_{\,\|\mathcal{D}}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{C}] = \mathrm{E}_{\,\|\mathcal{D}}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{B}]$. Therefore $\boldsymbol{a}_{\mathcal{B}} = \boldsymbol{a}_{\mathcal{C}}$. This condition insures that $\mathrm{E}\,[\boldsymbol{a} \mid \mathcal{C} + \mathcal{D}] = \boldsymbol{a}_{\mathcal{B}} + \boldsymbol{a}_{\mathcal{D}} = \mathrm{E}\,[\boldsymbol{a} \mid \mathcal{B} + \mathcal{D}]$ by uniqueness of the orthogonal projection, which, taking closure, implies (2.).
(2 ⇒ 1) Making use of the same decomposition of $\boldsymbol{a}$, (2.) implies that $\mathrm{E}\,[\boldsymbol{a} \mid \mathcal{C} + \mathcal{D}] = \boldsymbol{a}_{\mathcal{C}} + \boldsymbol{a}_{\mathcal{D}} = \mathrm{E}\,[\boldsymbol{a} \mid \mathcal{B} + \mathcal{D}]$ and hence by uniqueness $\boldsymbol{a}_{\mathcal{C}} \in \mathcal{B}$. Therefore $\boldsymbol{a}_{\mathcal{B}} = \boldsymbol{a}_{\mathcal{C}}$. The above condition implies that $\mathrm{E}_{\,\|\mathcal{D}}\,[\boldsymbol{a} \mid \mathcal{C}] = \boldsymbol{a}_{\mathcal{B}} = \mathrm{E}_{\,\|\mathcal{D}}\,[\boldsymbol{a} \mid \mathcal{B}]$. To complete the proof just take closure with respect to the inner product, and everything goes trough since all subspaces are closed. $\square$

Remark 2.1. While Lemma 2.1 gives conditions for reducing the subspace *along which* we project, Lemma 2.2 gives conditions for reducing the subspace *onto which* we project. In Lemma 2.2 the conditions are equivalent, while in Lemma 2.1 only one implication holds. The reason for this is explained in the proof of lemma 2.1 and, roughly speaking, it amounts to the fact that condition (2.4) only guarantees that the component $\boldsymbol{a}_{\mathcal{C}}$ lying on $\mathcal{C}$ of any element in $\mathcal{A}$ (which is uniquely defined), splits uniquely in the orthogonal sum $\boldsymbol{a}_{\mathcal{C}} = \mathrm{E}\,[\boldsymbol{a}_{\mathcal{C}} \mid \mathcal{B}] \oplus \tilde{\boldsymbol{a}}_{\mathcal{B}}$ where $\tilde{\boldsymbol{a}}_{\mathcal{B}} \perp \mathcal{D}$ and not $\tilde{\boldsymbol{a}}_{\mathcal{B}} = 0$, which would be necessary to prove the opposite implication. In lemma 2.2 instead the condition on the oblique projection actually guarantees that the component $\boldsymbol{a}_{\mathcal{C}}$ lying on $\mathcal{C}$ of any element in $\mathcal{A}$ is indeed in $\mathcal{B}$.

How oblique projections can be computed in a finite dimensional setting is addressed in Lemma 1 of [**10**].

3. Notations and Basic Assumptions

The Hilbert space setting for the study of second-order stationary processes is standard. Here we shall work in discrete time $t = \dots, -1, 0, 1, \dots,$ and make the assumption that all processes involved are jointly (second-order) stationary and with zero mean. The $m+p$-dimensional joint process $[\boldsymbol{y}\,\boldsymbol{u}]'$ will be assumed purely non deterministic and of full rank [**26**]. Sometimes we shall make the assumption of rational spectral densities in order to work with finite-dimensional realizations, however the geometric theory described in this paper is completely general and works also in the infinite dimensional case.

For $-\infty < t < +\infty$ introduce the linear subspaces of second order random variables

$$\begin{aligned} \mathcal{U}_t^- &:= \overline{\text{span}}\,\{\boldsymbol{u}_k(s);\ k=1,\dots,p,\ s<t\,\} \\ \mathcal{Y}_t^- &:= \overline{\text{span}}\,\{\boldsymbol{y}_k(s);\ k=1,\dots,m,\ s<t\,\} \end{aligned}$$

where the bar denotes closure with respect to the metric induced by the inner product (2.1). These are the Hilbert spaces of random variables spanned by the infinite past of $\boldsymbol{u}$ and $\boldsymbol{y}$ up to time t. By convention the past spaces do not include the present. We shall call

$$\mathcal{P}_t := \mathcal{U}_t^- \vee \mathcal{Y}_t^-$$

(the $\vee$ denotes closed vector sum) the *joint past space* of the input and output processes at time t.

Subspaces spanned by random variables at just one time instant are simply denoted $\mathcal{U}_t$, $\mathcal{Y}_t$, etc. while the spaces generated by the whole time history of $\boldsymbol{u}$ and $\boldsymbol{y}$ we shall use the symbols $\mathcal{U}$, $\mathcal{Y}$, respectively.

The *shift operator* σ is a unitary map defined on a dense subset of $\mathcal{U} \vee \mathcal{Y}$ by the assignment

$$\sigma(\textstyle\sum_k a_k'\boldsymbol{y}(t_k) + \sum b_j'\boldsymbol{u}(t_j)) := (\sum_k a_k'\boldsymbol{y}(t_k+1) + \sum b_j'\boldsymbol{u}(t_j+1))$$

$$a_k \in \mathbb{R}^m,\ b_j \in \mathbb{R}^p,\ t_k,\ t_j \in \mathbb{Z}$$

Because of stationarity σ can be extended to the whole space as a unitary operator, see e.g. [**26**].

The processes $\boldsymbol{y}$ and $\boldsymbol{u}$ propagate in time by the shift operator (e.g. $\boldsymbol{y}(t) = \sigma^\top\boldsymbol{y}(0)$); this in particular implies that all relations involving random variables of $\mathcal{U} \vee \mathcal{Y}$ which are valid at a certain instant of time t, by applying the shift operator on both sides of the relation, are seen to be also automatically valid at any other time. For this reason all definitions and statements this paper involving subspaces or random variables defined at a certain time instant t are to be understood as holding also for arbitrary $t \in \mathbb{Z}$.

3.1. Conditional Orthogonality. We say that two subspaces $\mathcal{A}$ and $\mathcal{B}$ of a Hilbert space $\mathcal{H}$ are *conditionally orthogonal* given a third subspace $\mathcal{X}$ if

$$\langle \alpha - \mathrm{E}^{\mathcal{X}}\alpha, \beta - \mathrm{E}^{\mathcal{X}}\beta \rangle = 0 \quad \text{for} \quad \alpha \in \mathcal{A},\ \beta \in \mathcal{B} \tag{3.1}$$

and we shall denote this $\mathcal{A} \perp \mathcal{B} \mid \mathcal{X}$. When $\mathcal{X} = 0$, this reduces to the usual orthogonality $\mathcal{A} \perp \mathcal{B}$. Conditional orthogonality is orthogonality after subtracting the projections on $\mathcal{X}$. Using the definition of the projection operator $\mathrm{E}^{\mathcal{X}}$, it is straightforward to see that (3.1) may also be written

$$\langle \mathrm{E}^{\mathcal{X}}\alpha, \mathrm{E}^{\mathcal{X}}\beta \rangle = \langle \alpha, \beta \rangle \quad \text{for} \quad \alpha \in \mathcal{A},\ \beta \in \mathcal{B}. \tag{3.2}$$

The following lemma is a trivial consequence of the definition.

Lemma 3.1. *If* $\mathcal{A} \perp \mathcal{B} \mid \mathcal{X}$, *then* $\mathcal{A}_0 \perp \mathcal{B}_0 \mid \mathcal{X}$ *for all* $\mathcal{A}_0 \subset \mathcal{A}$ *and* $\mathcal{B}_0 \subset \mathcal{B}$.

Let $\mathcal{A} \oplus \mathcal{B}$ denote the *orthogonal* direct sum of $\mathcal{A}$ and $\mathcal{B}$. If $\mathcal{C} = \mathcal{A} \oplus \mathcal{B}$, then $\mathcal{B} = \mathcal{C} \ominus \mathcal{A}$ is the orthogonal complement of $\mathcal{A}$ in $\mathcal{C}$. The following Proposition from [**14, 15**] describes some useful alternative characterizations of conditional orthogonality.

Lemma 3.2. *The following statements are equivalent.*

(i) $\mathcal{A} \perp \mathcal{B} \mid \mathcal{X}$
(ii) $\mathcal{B} \perp \mathcal{A} \mid \mathcal{X}$
(iii) $(\mathcal{A} \vee \mathcal{X}) \perp \mathcal{B} \mid \mathcal{X}$
(iv) $\mathrm{E}^{\mathcal{A}\vee\mathcal{X}}\beta = \mathrm{E}^{\mathcal{X}}\beta$ *for* $\beta \in \mathcal{B}$
(v) $(\mathcal{A} \vee \mathcal{X}) \ominus \mathcal{X} \perp \mathcal{B}$
(vi) $\mathrm{E}^{\mathcal{A}}\beta = \mathrm{E}^{\mathcal{A}}\mathrm{E}^{\mathcal{X}}\beta$ *for* $\beta \in \mathcal{B}$

3.2. Feedback. Let $\boldsymbol{y}$ and $\boldsymbol{u}$ be two jointly stationary vector stochastic processes. In general one may express both $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$ as a sum of the best linear estimate based on the past of the other variable, plus "noise"

$$\boldsymbol{y}(t) = \mathrm{E}\,[\boldsymbol{y}(t) \mid \mathcal{U}_{t+1}^{-}] + \boldsymbol{d}(t) \tag{3.3a}$$

$$\boldsymbol{u}(t) = \mathrm{E}\,[\boldsymbol{u}(t) \mid \mathcal{Y}_{t+1}^{-}] + \boldsymbol{r}(t) \tag{3.3b}$$

so that each variable $\boldsymbol{y}(t)$ and $\boldsymbol{u}(t)$ can be expressed as a sum of a causal linear transformation of the past of the other, plus "noise". Here the noise terms are uncorrelated with the past of $\boldsymbol{u}$ and $\boldsymbol{y}$ respectively, but may in general be mutually correlated.

Since both linear estimators above can be expressed as the output of linear filters, represented by causal transfer functions $F(z)$ and $H(z)$, the joint model (3.3) corresponds to a *feedback interconnection* of the type

where (in symbolic "Z-transform" notation)

$$\boldsymbol{d}(t) = G(z)\boldsymbol{w}(t) \qquad \boldsymbol{r}(t) = K(z)\boldsymbol{e}_r(t)$$

in which $G(z)$, $K(z)$ may be assumed, without loss of generality, minimum phase spectral factors of the spectra $\Phi_d(z)$ and $\Phi_r(z)$ of the two stationary "error" (or "disturbance") signals d and r.

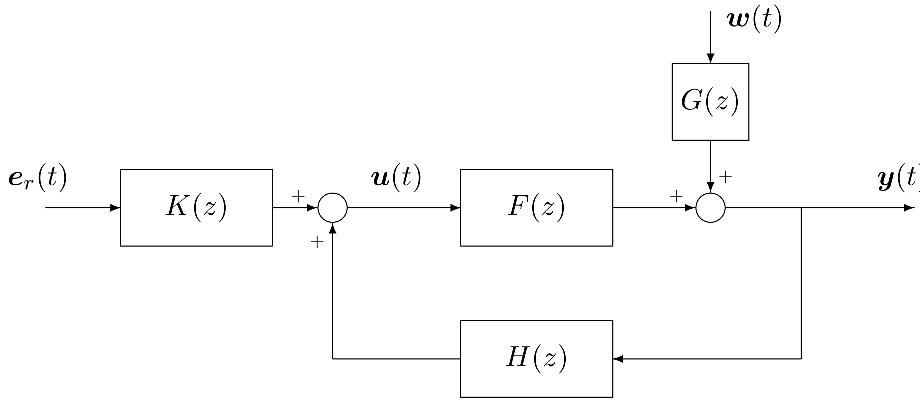

FIGURE 1. Feedback model of the processes $\boldsymbol{y}$ and $\boldsymbol{u}$.

In this scheme the "errors" $\boldsymbol{r}$ and $\boldsymbol{d}$ are in general correlated. More useful are feedback schemes which involve uncorrelated error processes, since in physical models of feedback systems this will more usually be the situation. It can be shown that any pair of jointly stationary processes can also be represented by schemes of this last type. In this case however, although the overall interconnection must be internally stable, the individual transfer functions $F(z)$ and $H(z)$ may well be unstable; see [**6**] for a detailed discussion.

Following Granger [**8**], and subsequent work by Caines, Chan, Anderson, Gevers etc. [**3, 7, 6**] we say that *there is no feedback from $\boldsymbol{y}$ to $\boldsymbol{u}$* if the future of $\boldsymbol{u}$ is conditionally uncorrelated with the past of $\boldsymbol{y}$, given the past of $\boldsymbol{u}$ itself. In our Hilbert space framework this is written as

$$\mathcal{U}_t^+ \perp \mathcal{Y}_t^- \mid \mathcal{U}_t^- \tag{3.4}$$

This condition expresses the fact that the future time evolution of the process $\boldsymbol{u}$ is not "influenced" by the past of $\boldsymbol{y}$ once the past of $\boldsymbol{u}$ is known. This captures in a coordinate free way the absence of feedback (from $\boldsymbol{y}$ to $\boldsymbol{u}$). Taking $\mathcal{A} = \mathcal{U}_t^+$ in condition (iii) of Lemma 3.2 above, the feedback-free condition is seen to be equivalent to $\mathcal{Y}_t^- \perp \mathcal{U} \mid \mathcal{U}_t^-$ and hence, from (iv), to $\mathrm{E}\,\{\mathcal{Y}_t^- \mid \mathcal{U}\} = \mathrm{E}\{\mathcal{Y}_t^- \mid \mathcal{U}_t^-\}$, so that

$$\mathrm{E}\{\boldsymbol{y}(t) \mid \mathcal{U}\} = \mathrm{E}\{\boldsymbol{y}(t) \mid \mathcal{U}_{t+1}^-\}, \quad \text{for all } t \in \mathbb{Z}, \tag{3.5}$$

meaning that $\mathrm{E}\{\boldsymbol{y}(t) \mid \mathcal{U}\}$ depends only on the past and present values of the process $\boldsymbol{u}$ but not on its future history. We take this as the definition of *causality*. In this case (and only in this case), it is appropriate to call $\boldsymbol{u}$ an *input* variable. One says that there is *causality from $\boldsymbol{u}$ to $\boldsymbol{y}$* (or $\boldsymbol{u}$ "causes" $\boldsymbol{y}$).

Let us consider the feedback interconnection of fig. 1 and let $F = N_F D_F^{-1}$, $H = D_H^{-1} N_H$ be coprime matrix fraction descriptions of the transfer functions of the direct and feedback channel. An important technical condition often used in this this paper is that the input process is "sufficiently rich", in the sense that $\mathcal{U}$ admits the direct sum decomposition $\mathcal{U} = \mathcal{U}_t^- + \mathcal{U}_t^+$ for all t. Obviously, sufficient richness is equivalent to

$$\mathcal{U}_t^- \cap \mathcal{U}_t^+ = \{0\} \tag{3.6}$$

Various conditions ensuring sufficient richness are known. For example, it is well-known that for a full-rank p.n.d. process $\boldsymbol{u}$ to be sufficiently rich it is necessary and sufficient that the determinant of the spectral density matrix Φ_u should have no zeros on the unit circle [**9**]. Using this criterion the following result follows readily.

LEMMA 3.3. *The input process is sufficiently rich, i.e. Φ_u does not have unitary zeros, if and only if both of the following conditions are satisfied:*

(1) *D_F has no zeros on the unit circle*
(2) *$N_H G$ and $D_H K$ do not have common zeros on the unit circle with coincident corresponding left zero directions (this we simply say: do not vanish simultaneously on the unit circle).*

4. Oblique Markovian Splitting Subspaces

The idea of (stochastic) state space is the fundamental concept of stochastic realization theory. In the classical setting, i.e. stochastic time series modelling, a state space is characterized by the property of being a *Markovian splitting subspace* [**14, 15**]. This idea captures in a coordinate-free way the structure of stochastic state-space models and lies at the grounds of their many useful properties. Realization of stochastic processes (without inputs) has been investigated by a number of authors, including [**21, 2, 1, 13, 17, 14, 15**].

The intuitive idea of a stochastic state space model with inputs is different from that of a state-space model of a single process since in the former type of models one wants to describe the effect of the exogenous input $\boldsymbol{u}$ on the output process *without modeling the dynamics of* $\boldsymbol{u}$. This is also in agreement with the aim of most identification experiments, where one is interested in describing the dynamics of the "open loop" system only and does not want to worry about finding a dynamic description of $\boldsymbol{u}$ at all. Hence the concept of state-space has to be generalized to the new setting. This will lead to the introduction of the concepts of oblique (conditional) Markov and oblique (conditional) splitting. The idea behind these definitions is to factor out the dynamics of the input process, which should not be modeled explicitly. When applying classical realization theory to the joint input-output process $(\boldsymbol{y}\ \boldsymbol{u})$ the dynamics of $\boldsymbol{u}$ will also be modeled.

At the end of the paper we shall see some connections between classical stochastic realizations of the joint input-output process and realization of $\boldsymbol{y}$ in terms of $\boldsymbol{u}$ as we are studying in this chapter.

Note that since the input is an observed variable, one is generally interested in realization which are causal with respect to $\boldsymbol{u}$. For this reason in this more general setting one should not expect the same mathematical symmetry between past and future as in the time-series case.

A source of difficulty in dealing with realization with inputs is the possible presence of feedback from $\boldsymbol{y}$ to $\boldsymbol{u}$. In this paper we shall strive to keep a certain level of generality without making too restrictive assumptions regarding the presence of feedback between $\boldsymbol{y}$ and $\boldsymbol{u}$. Dealing with feedback is a necessity in the design of identification algorithms and the complications we incur are not searched for just for academic sake of generality.

Later on we shall specialize to the case when there is no feedback from $\boldsymbol{y}$ to $\boldsymbol{u}$ and a much simpler and elegant theory will emerge.

All Hilbert spaces that we consider will be subspaces of an ambient Hilbert space $\mathcal{H}$ containing $\mathcal{U}$ and $\mathcal{Y}$ and equipped with a shift operator under which $\boldsymbol{y}$ and $\boldsymbol{u}$ are jointly stationary. We shall also assume that $\mathcal{H}$ has a finite number of generators i.e. there are $N < \infty$ random variables $\{h_1, \dots, h_N\}$ such that

$$\overline{\text{span}}\{\sigma^\top h_k \mid k = 1, \dots, N,\ t \in \mathbb{Z}\} = \mathcal{H}$$

This is sometimes referred to by saying that $\mathcal{H}$ (or the shift σ acting on $\mathcal{H}$) has *finite multiplicity*. The multiplicity is certainly finite in the interesting special case where

$$\mathcal{H} = \mathcal{Y} \vee \mathcal{U} \tag{4.1}$$

Let $\mathcal{X}$ be a subspace of $\mathcal{H}$ and define the *stationary family of translates*, $\{\mathcal{X}_t\}$, by: $\mathcal{X}_t := \sigma^\top \mathcal{X}, \quad t \in \mathbb{Z}$. The *past* and *future* of $\{\mathcal{X}_t\}$ at time t are

$$\mathcal{X}_t^- := \vee_{s<t} \mathcal{X}_s, \qquad \mathcal{X}_t^+ := \vee_{s \geq t} \mathcal{X}_s.$$

Generalizing a construction of stochastic realization theory (see e.g. [**14**]), we define a pair of subspaces $\left(\mathcal{S}, \bar{\mathcal{S}}\right)$ attached to a given $\mathcal{X}$, as follows:

$$\mathcal{S} = \mathcal{P}^- \vee \mathcal{X}^- \tag{4.2}$$

where $\mathcal{P}^-$ is a shorthand for the *joint past* space $\mathcal{P}_t^- := \mathcal{U}_t^- \vee \mathcal{Y}_t^-$ at time $t = 0$; $\mathcal{S}$ will be called the *incoming* subspace associated to $\mathcal{X}$, while

$$\bar{\mathcal{S}} = \mathcal{Y}^+ \vee \mathcal{X}^+ \tag{4.3}$$

will be called the corresponding *outgoing* subspace.

Recall that in classical stochastic realization the subspaces $\mathcal{S}$ and $\bar{\mathcal{S}}$ are incoming and outgoing subspaces for the shift σ, in the sense of Lax-Phillips Scattering Theory [**20**]. In the present setting $\bar{\mathcal{S}}$ is not outgoing as it does not satisfy $\cup_t \bar{\mathcal{S}}_t = \mathcal{H}$. However, it will turn out to be convenient to keep the same terminology to make connections with classical stochastic realization theory, especially when studying minimality. Later we shall also introduce "extended" versions of $\mathcal{S}$ and $\bar{\mathcal{S}}$ which will in fact be incoming and outgoing subspaces in the sense of Lax-Phillips.

Definition 4.1. The family $\{\mathcal{S}_t\}$ is (forward) *purely-non-deterministic* (p.n.d.) if

$$\bigcap_{t<0} \mathcal{S}_t = \{0\} \tag{4.4}$$

The subspace $\mathcal{X}$ is then called (forward) *purely-non-deterministic* (p.n.d.) whenever the associated incoming subspace has the p.n.d. property.

Let us define the sequence of *wandering* subspaces $\mathcal{W}_t = \sigma^\top \mathcal{W}$ associated to $\{\mathcal{S}_t\}$ as:

$$\mathcal{W}_t := \mathcal{S}_{t+1} \ominus (\mathcal{S}_t + \mathcal{U}_t)\,. \tag{4.5}$$

LEMMA 4.1. *The wandering subspaces are pairwise orthogonal, i.e.* $\mathcal{W}_t \perp \mathcal{W}_s$, $\forall\, t \neq s$.

PROOF. Let us assume $t > s$; by construction we have $\mathcal{W}_s \subseteq \mathcal{S}_{s+1}$ while clearly $\mathcal{W}_t \perp \mathcal{S}_t$. Since $\mathcal{S}_t$ is non-decreasing (backward-shift invariant), i.e. $\mathcal{S}_s \subseteq \mathcal{S}_t$, it follows that $\mathcal{W}_t \perp \mathcal{S}_s$ and hence $\mathcal{W}_t \perp \mathcal{W}_s$. □

It follows from (4.5) that the incoming subspace admits the decomposition

$$\mathcal{S}_{t+1} = (\mathcal{S}_t + \mathcal{U}_t) \oplus \mathcal{W}_t \tag{4.6}$$

For future reference we note the following fact:

FACT 4.1. *The future* $\mathcal{W}_t^+$ *is orthogonal to* $\mathcal{S}_t + \mathcal{U}_t$.

Since $\mathcal{H}$ has finite multiplicity, the wandering subspace $\mathcal{W}$ is finite-dimensional and admits an orthonormal basis $\boldsymbol{w}(0)$. It follows that $\mathcal{W}_t^- = \mathbf{H}_t^-(\boldsymbol{w})$ where $\boldsymbol{w}(t) = \sigma^\top \boldsymbol{w}(0)$ is a normalized white noise process which is called the (forward) *generating process* of $\mathcal{X}$.

The following are basic definitions which capture the notion of state space in presence of exogenous inputs.

DEFINITION 4.2. The subspace $\mathcal{X}$ is (forward) *oblique Markovian*, if $\mathcal{U}_t \cap \left(\mathcal{X}_{t+1}^- \vee \mathcal{U}_t^-\right) = \{0\}$ and the following equality holds:

$$E_{\|\mathcal{U}_t}[\,\mathcal{X}_{t+1} \mid \mathcal{X}_t^- \vee \mathcal{U}_t^-] = E_{\|\mathcal{U}_t}[\,\mathcal{X}_{t+1} \,|\, \mathcal{X}_t]. \tag{4.7}$$

We shall say that $\mathcal{X}$ is *causal* oblique Markovian if $\mathcal{X}_t \subseteq \mathcal{P}_t^-$.

Note that the condition $\mathcal{U}_t \cap \mathcal{U}_t^- = \{0\}$, necessary for the oblique projection to be well defined, is implied by the richness Assumption 3.6.

The oblique Markovian condition can be written in terms of conditional orthogonality as follows:

PROPOSITION 4.1. *The oblique Markovian property (4.7) is equivalent to*

$$E\left[\mathcal{X}_{t+1} \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{U}_{t+1}^-\right] = E[\mathcal{X}_{t+1} \,|\, \mathcal{X}_t + \mathcal{U}_t] \tag{4.8}$$

and hence to the following conditional orthogonality property

$$\mathcal{X}_{t+1} \perp \left(\mathcal{X}_t^- \vee \mathcal{U}_t^-\right) \mid (\mathcal{X}_t + \mathcal{U}_t) \tag{4.9}$$

which can be interpreted by saying that $\mathcal{X}_t$ *is* conditionally Markov *given* $\mathcal{U}_t$.

PROOF. Setting $\mathcal{A} = \mathcal{X}_{t+1}$, $\mathcal{B} = \mathcal{X}_{t+1}$, $\mathcal{C} = \mathcal{X}_t^- \vee \mathcal{U}_t^-$ and $\mathcal{D} = \mathcal{U}_t$ in Lemma 2.2 the oblique Markovian property (4.7) is seen to be equivalent to (4.8). The conditional orthogonality follows from Lemma 3.2. □

To be honest, the oblique Markovian property of the definition should be named "one-step-ahead" oblique Markovian property. For it is in general not guaranteed that the sufficient statistic property of $\mathcal{X}_t$ holds also when one wants to predict random variables in the distant future $\mathcal{X}_{t+k}$, $k > 1$. However we shall see later that the extension of the "one-step-ahead" oblique Markovian property to an arbitrary number of steps ahead, holds when there is no feedback from $\boldsymbol{x}$ to $\boldsymbol{u}$, in which case condition (4.7) is equivalent to

$$E_{\|\mathcal{U}_t^+}\left[\mathcal{X}_t^+ \mid \mathcal{X}_{t+1}^- \vee \mathcal{U}_t^-\right] = E_{\|\mathcal{U}_t^+}\left[\mathcal{X}_t^+ \mid \mathcal{X}_t\right]$$

Unfortunately in general (4.7) is not equivalent to the above. In fact, we shall see that the property of sufficient statistics for the whole future will hold only "conditionally", given the subspace generated by *all* future inputs to the realization with state space $\mathcal{X}_t$, where "inputs" now means input signal which may be observable and not. To make this precise, we shall have to introduce the "extended" or *joint future input space*

$$\mathcal{F}_t^+ := \left(\mathcal{U}_t^+ \vee \mathcal{W}_t^+\right)$$

of the p.n.d. subspace $\mathcal{X}$.

The joint future plays a role in generalizing the fundamental concept of splitting to the *oblique splitting* property defined below.

DEFINITION 4.3. A subspace $\mathcal{X}$ is (forward) *oblique splitting* for $(\mathcal{Y}, \mathcal{U})$, if

$$\mathcal{Y}_t^+ \perp \mathcal{P}_t^- \mid \left[\mathcal{X}_t \vee \mathcal{F}_t^+\right] \tag{4.10}$$

We will say that $\mathcal{X}$ is a *causal* oblique splitting subspace if $\mathcal{X}_t \subseteq \mathcal{P}_t^-$.

Condition (4.10) says that once the (real plus unobservable white) future inputs are given, the information in the present state space $\mathcal{X}_t$, is equivalent to the knowledge of all the (joint) past history of state input and output, for the purpose of predicting the future of $\boldsymbol{y}$. Indeed, provided that

$$\mathcal{X}_t \cap \mathcal{F}_t^+ = \{0\}, \tag{4.11}$$

by using Lemma 2.2, it follows that (4.10) can be expressed using oblique projections

$$E_{\|\mathcal{F}_t^+}\left[\mathcal{Y}_t^+ \mid \mathcal{X}_t \vee \mathcal{P}_t^-\right] = E_{\|\mathcal{F}_t^+}\left[\mathcal{Y}_t^+ \mid \mathcal{X}_t\right]$$

Unfortunately, there may be situations in which condition (4.11) does not hold, for virtually all oblique Markovian splitting subspaces $\mathcal{X}_t$. Intuitively, this will be the case when the transfer function $F(z)$ in the forward loop in Fig. 1 is not stable (which can happen, even if the feedback interconnection is internally stable).

Another difficulty related to the presence of feedback is that it may happen that $(\mathcal{U}^+ \vee \mathcal{W}^+) \cap (\mathcal{U}^- \vee \mathcal{W}^-) \neq \{0\}$. This fact would make some oblique projection formulas meaningless.

A sufficient condition for zero intersection is given in following Proposition, whose proof we shall leave to the reader.

PROPOSITION 4.2. *The joint spectral density matrix* $\Phi_{\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{w}\end{bmatrix}}$, *has no zeros on the unit circle, or, equivalently,* $(\mathcal{U}^+ \vee \mathcal{W}^+) \cap (\mathcal{U}^- \vee \mathcal{W}^-) = \{0\}$, *if and only if, with the same notation of remark 3.3,* $D_F D_H K$ *does not vanish on the unit circle.*

Note that if there are no inputs, $\mathcal{X}_t \vee \mathcal{F}_t^+ = \mathcal{X}_t \oplus \mathcal{W}_t^+$ and condition (4.10) reduces to

$$\mathcal{Y}_t^+ \perp \mathcal{Y}_t^- \mid (\mathcal{X}_t \oplus \mathcal{W}_t^+)$$

and since $\mathcal{W}_t^+ \perp \mathcal{Y}_t^-$ the latter is in turn equivalent to

$$\mathcal{Y}_t^+ \perp \mathcal{Y}_t^- \mid \mathcal{X}_t$$

which is the usual splitting property.

The oblique Markov and the splitting conditions separately are not enough, in general, to fully characterize the state space in the presence of inputs. A condition which implies both (4.10) and (4.7) is the *oblique Markovian splitting* property, defined below.

DEFINITION 4.4. A subspace $\mathcal{X}$ is an *oblique Markovian splitting subspace* for the pair $(\mathcal{Y}, \mathcal{U})$ if

$$\mathcal{U}_t \cap \left(\mathcal{X}_{t+1}^- \vee \mathcal{P}_t^-\right) = \{0\} \tag{4.12}$$

and

$$E_{\|\mathcal{U}_t}[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^-] = E_{\|\mathcal{U}_t}[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{X}_t]. \tag{4.13}$$

We shall say that $\mathcal{X}$ is a *causal* oblique Markovian splitting subspace if $\mathcal{X} \subseteq \mathcal{P}^-$

This condition is precisely what is needed for the space $\mathcal{X}$ to qualify as a state space for a stochastic model described by equations of the form (1.1)

Note that the "extended richness condition"

$$\mathcal{U}_t \cap \mathcal{P}_t^- = \{0\} \tag{4.14}$$

is a necessary condition for the oblique projection to be well-defined. This property will be necessary in order to be able to derive unique state-space equations.

PROPOSITION 4.3. *If the joint spectrum of* $\boldsymbol{y}$ *and* $\boldsymbol{u}$ *is coercive then the zero intersection property (4.14) holds.*

PROOF. If the joint spectrum is coercive, then one has that

$$(\mathcal{U}_t^+ \vee \mathcal{Y}_t^+) \cap (\mathcal{U}_t^- \vee \mathcal{Y}_t^-) = \{0\}$$

and then in particular (4.14). □

REMARK 4.5. Again, the property that we would like to hold is

$$E_{\|\mathcal{U}_t^+}[\mathcal{Y}_t^+ \vee \mathcal{X}_{t+1}^+ \mid \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^-] = E_{\|\mathcal{U}_t^+}[\mathcal{Y}_t^+ \vee \mathcal{X}_{t+1}^+ \mid \mathcal{X}_t] \tag{4.15}$$

however this condition is in general not equivalent to (4.13). In fact, when feedback is present (the past output space $\mathcal{Y}_t^-$ is not conditionally uncorrelated with the future inputs $\mathcal{U}_t^+$ given the past past inputs $\mathcal{U}_t^-$), one has

$$E[\mathcal{Y}_t \,|\, \mathcal{X}_t^- \vee \mathcal{P}_t^- \vee \mathcal{U}_t^+] \neq E[\mathcal{Y}_t \,|\, , \mathcal{X}_t^- \vee \mathcal{P}_t^- \vee \mathcal{U}_t]$$

so that requiring the stronger condition (4.15) would make the state space "unnecessarily large". We shall see later (see Lemma 7.2) that when there is no feedback from $\left[\boldsymbol{x}^\top \ \boldsymbol{y}^\top\right]$ to $\boldsymbol{u}$, condition (4.13) is equivalent to (4.15).

The result which follows is in the same spirit of Proposition 4.1.

PROPOSITION 4.4. *The oblique Markovian splitting property (4.13) is equivalent to*

$$E\left[\mathcal{X}_{t+1} \vee \mathcal{Y}_t \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{U}_t \vee \mathcal{P}_t^-\right] = E[\mathcal{X}_{t+1} \vee \mathcal{Y}_t \,|\, \mathcal{X}_t + \mathcal{U}_t]. \tag{4.16}$$

and hence to the conditional orthogonality property:

$$(\mathcal{X}_{t+1} \vee \mathcal{Y}_t) \perp \left(\mathcal{X}_t^- \vee \mathcal{P}_t^-\right) \mid (\mathcal{X}_t + \mathcal{U}_t)$$

PROOF. Letting $\mathcal{A} = (\mathcal{X}_{t+1} \vee \mathcal{Y}_t)$, $\mathcal{B} = \mathcal{X}_t$, $\mathcal{C} = \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^-$ and $\mathcal{D} = \mathcal{U}_t$ in Lemma 2.2, the oblique Markovian splitting property (4.13) is seen to be equivalent to (4.16). □

PROPOSITION 4.5. *Under condition (4.11) the oblique Markovian splitting property (4.13) is equivalent to*

$$E\left[\mathcal{X}_{t+1}^+ \vee \mathcal{Y}_t^+ \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{F}_t^+ \vee \mathcal{P}_t^-\right] = E[\mathcal{X}_{t+1}^+ \vee \mathcal{Y}_t^+ \,|\, \mathcal{X}_t + \mathcal{F}_t^+]. \tag{4.17}$$

which can also be written in form of conditional orthogonality as

$$\left(\mathcal{X}_{t+1}^+ \vee \mathcal{Y}_t^+\right) \perp \left(\mathcal{X}_t^- \vee \mathcal{P}_t^-\right) \mid \left(\mathcal{X}_t + \mathcal{F}_t^+\right)$$

PROOF. Assume (4.17) holds. It follows that

$$E[\mathcal{X}_{t+1} \vee \mathcal{Y}_t \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^- \vee \mathcal{F}_t^+] \subseteq \mathcal{X}_t + \mathcal{F}_t^+$$

However, from $\mathcal{X}_{t+1} \vee \mathcal{Y}_t \subseteq \mathcal{S}_{t+1}$ we obtain

$$E[\mathcal{X}_{t+1} \vee \mathcal{Y}_t \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^- \vee \mathcal{F}_t^+] \subseteq \mathcal{S}_t + \mathcal{U}_t \oplus \mathcal{W}_t.$$

Using the fact that (4.11) implies $\mathcal{S}_t \cap \mathcal{F}_t^+ = \{0\}$ it must be that

$$E[\mathcal{X}_{t+1} \vee \mathcal{Y}_t \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^- \vee \mathcal{U}_t \vee \mathcal{W}_t] = E[\mathcal{X}_{t+1} \vee \mathcal{Y}_t \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^- \vee \mathcal{F}_t^+] \subseteq \mathcal{X}_t + \mathcal{U}_t \oplus \mathcal{W}_t$$

which clearly implies (4.13). The proof of the converse will be given after Theorem 4.6. □

COROLLARY 4.1. *Under condition (4.11) the oblique Markovian splitting property (4.13) is equivalent to*

$$E_{\|\mathcal{F}_t^+}\left[\mathcal{X}_{t+1}^+ \vee \mathcal{Y}_t^+ \,|\, \mathcal{X}_{t+1}^- \vee \mathcal{P}_t^-\right] = E_{\|\mathcal{F}_t^+}\left[\mathcal{X}_{t+1}^+ \vee \mathcal{Y}_t^+ \,|\, \mathcal{X}_t\right]. \tag{4.18}$$

PROOF. Setting $\mathcal{A} = \left(\mathcal{X}^+_{t+1} \vee \mathcal{Y}^+_t\right)$, $\mathcal{B} = \mathcal{X}_t$, $\mathcal{C} = \mathcal{X}^-_{t+1} \vee \mathcal{P}^-_t$ and $\mathcal{D} = \mathcal{F}^+_t$ in Lemma 2.2, we have that (4.18) is equivalent to (4.17) and therefore, from Proposition 4.5 to (4.13). □

The following result is a coordinate-free version of the equivalence between oblique Markovian splitting property and representability by a state space model of the form (1.1). The Theorem holds without finite dimensionality assumptions.

THEOREM 4.6. *Let $\mathcal{X}_t$ be a p.n.d. oblique Markovian splitting subspace for $(\mathcal{Y}, \mathcal{U})$; then the following inclusions hold*

$$\mathcal{X}_{t+1} \subseteq (\mathcal{X}_t + \mathcal{U}_t) \oplus \mathcal{W}_t \tag{4.19}$$

$$\mathcal{Y}_t \subseteq (\mathcal{X}_t + \mathcal{U}_t) \oplus \mathcal{W}_t \tag{4.20}$$

PROOF. Since $\mathcal{X}_{t+1} \subseteq \mathcal{S}_{t+1}$ using the decomposition $\mathcal{S}_{t+1} = (\mathcal{S}_t + \mathcal{U}_t) \oplus \mathcal{W}_t$ we obtain :

$$\begin{array}{rcl} \mathcal{X}_{t+1} & = & E\left[\mathcal{X}_{t+1} \,|\, \mathcal{S}_{t+1}\right] \\ & = & E\left[\mathcal{X}_{t+1} \,|\, (\mathcal{S}_t + \mathcal{U}_t) \oplus \mathcal{W}_t\right] \\ & \subseteq & E\left[\mathcal{X}_{t+1} \,|\, (\mathcal{S}_t + \mathcal{U}_t)\right] \oplus \mathcal{W}_t \\ & \subseteq & \left(E_{\|\mathcal{U}_t}\left[\mathcal{X}_{t+1} \,|\, \mathcal{S}_t\right] + \mathcal{U}_t\right) \oplus \mathcal{W}_t \\ & \subseteq & (\mathcal{X}_t + \mathcal{U}_t) \oplus \mathcal{W}_t \end{array}$$

where the last equality follows from (4.13). A completely analogous derivation holds for the second inclusion. □

We can now complete the proof of Proposition 4.5. To this purpose it will be handy to introduce a notation for vector sum of subspaces of the type

$$\mathcal{U}_{[\,t,\,t+k\,)} := \mathcal{U}_t + \mathcal{U}_{t+1} + \ldots + \mathcal{U}_{t+k-1}$$

Similar notations will be used without further comments in the following.

Proof of Proposition 4.5 Conversely, assume (4.13) holds. It follows from Theorem 4.6 that, for every $k \geq 0$

$$\mathcal{X}_{t+k+1} \vee \mathcal{Y}_{t+k} \subseteq \mathcal{X}_t + \mathcal{U}_{[\,t,\,t+k\,)} + \mathcal{W}_{[\,t,\,t+k\,)}$$

which implies that

$$\mathcal{X}^+_{t+1} \vee \mathcal{Y}^+_t \subseteq \mathcal{X}_t + \mathcal{U}^+_t + \mathcal{W}^+_t$$

and therefore (4.17). □

When $\mathcal{X}$ is finite-dimensional, we can obtain state-space representations of the form (1.1) just by choosing a basis in the subspaces $\mathcal{X}$ and $\mathcal{W}$. Conversely, given a finite-dimensional state-space model of the form (1.1), it is easy to check that the subspace generated by the components of the state vector

$$\mathcal{X}_t := \text{span}\,\{\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)\,\}$$

is an oblique Markovian splitting subspace. We shall leave this check to the reader.

By using the representation Theorem 4.6 we can show that the oblique Markovian splitting property implies both the oblique Markovian and the oblique splitting property.

PROPOSITION 4.6. *The oblique Markovian splitting property implies oblique Markovian and oblique splitting, i.e. (4.13) implies both (4.7) and (4.10).*

PROOF. Projecting both members of (4.19) along $\mathcal{U}_t$ we obtain

$$E_{\|\mathcal{U}_t}\left[\mathcal{X}_{t+1} \mid \mathcal{X}_t \vee \mathcal{U}_t^-\right] \subset \mathcal{X}_t$$

which is the oblique Markovian property (4.7). Combining (4.20) and (4.19) we obtain

$$\mathcal{Y}_t^+ \subseteq \mathcal{X}_t \vee \mathcal{U}_t^+ \vee \mathcal{W}_t^+ = \mathcal{X}_t \vee \mathcal{F}_t^+$$

which implies (4.10). If (4.11) holds, the projection of any element $y^+ \in \mathcal{Y}_t^+$ onto $\mathcal{X}_t \vee \mathcal{P}_t^-$ along $\mathcal{F}_t^+$ is equal to the projection of its (unique in this case) component in $\mathcal{X}_t$. In other words

$$E_{\|\mathcal{F}_t^+}\left[y^+ \,|\, \mathcal{X}_t \vee \mathcal{P}_t^-\right] = E_{\|\mathcal{F}_t^+}\left[y^+ \,|\, \mathcal{X}_t\right], \qquad y^+ \in \mathcal{Y}_t^+.$$

□

5. Acausality of Realizations with Feedback

Stationary models for the pair $(\boldsymbol{y}, \boldsymbol{u})$ of the form (1.1), or in symbolic (z-transform) notation,

$$\boldsymbol{y}(t) = F(z)\boldsymbol{u}(t) + G(z)\boldsymbol{w}(t)$$

tend to give for granted that $\boldsymbol{y}$ depends *causally* on the input signals $\boldsymbol{u}$, $\boldsymbol{w}$. This is in general false if we are in the presence of feedback.

Certainly causality holds as long as $F(z)$ and $G(z)$ are stable, or, equivalently, $|\lambda(A)| < 1$ in the model (1.1). However, in the presence of feedback the pair $(\boldsymbol{y}, \boldsymbol{u})$ may well be stationary even if $F(z)$ is not stable. In this situation, the eigenvalues of A may lie anywhere, in particular some may be (strictly) outside of the unit circle. Then, the customary interpretation of the state space model (1.1) as a *forward difference equation*, does not make sense. Unstable modes must be integrated *backwards* see [**25**]. In general when there is feedback, past outputs may be influenced by future inputs, which, according to what has been seen above, means that the model is *not causal.* We shall briefly analyze how this acausality shows up and what are the consequences. For simplicity of exposition we analyze only the finite dimensional case.

Consider a basis for (1.1) so that the matrix A is of the block-diagonal form

$$A = \begin{pmatrix} A_- & 0 & 0 \\ 0 & A_0 & 0 \\ 0 & 0 & A_+ \end{pmatrix} \tag{5.1}$$

where $|\lambda(A_-)| < 1$, $|\lambda(A_0)| = 1$, $|\lambda(A_+)| > 1$. Correspondingly we shall denote with $\mathcal{X}_-$ the "stable manifold", with $\mathcal{X}_+$ the "unstable manifold" and with $\mathcal{X}_0$ the "central manifold" of the state space. The symbols $\boldsymbol{x}_-$, $\boldsymbol{x}_+$ and

$\boldsymbol{x}_0$ will denote the bases in the corresponding spaces. A similar meaning will be attributed to the symbols B_-, B_+, B_0, G_-, G_+ and G_0. Hence the state space equation can be rewritten in decoupled from as follows:

$$\begin{cases} \boldsymbol{x}_-(t+1) &= A_-\boldsymbol{x}_-(t) &+ B_-\boldsymbol{u}(t) &+ G_-\boldsymbol{w}(t) \\ \boldsymbol{x}_0(t+1) &= A_0\boldsymbol{x}_0(t) &+ B_0\boldsymbol{u}(t) &+ G_0\boldsymbol{w}(t) \\ \boldsymbol{x}_-(t+1) &= A_+\boldsymbol{x}_+(t) &+ B_+\boldsymbol{u}(t) &+ G_+\boldsymbol{w}(t) \end{cases} \tag{5.2}$$

The three difference equation can be thought of as running forward, forward or backward, and backward respectively. The interpretation of the equation on the "central manifold" is somewhat delicate and we shall not insist on this point here, however see [**5**, p. 105]. May it suffice to say that this component belongs to both past $\mathcal{U}_t^- \vee \mathcal{W}_t^-$ and future $\mathcal{U}_t^+ \vee \mathcal{W}_t^+$. This is not a contradiction as pointed out in remark 4.2. Concerning the first and the third block, it is trivial to recognize that these may be thought as a stable difference equation running forward and an unstable difference equation running backward in time. This implies that $\boldsymbol{x}_-(t) \in \mathcal{U}_t^- \vee \mathcal{W}_t^-$ and $\boldsymbol{x}_+(t) \in \mathcal{U}_t^+ \vee \mathcal{W}_t^+ = \mathcal{F}_t^+$. From this we see that in general one cannot assume that $\mathcal{X}_t \cap \mathcal{F}_t^+ = \{0\}$. This fact is the source of a number of complications.

To avoid these complications we shall henceforth restrict to the case $|\lambda(A)| < 1$ (i.e. we have a feedback interconnection with a stable forward loop transfer function $F(z)$) and postpone the discussion of the general case to future publications.

We formalize this assumption below.

Assumption 5.1. The joint spectrum of $\boldsymbol{u}$ and $\boldsymbol{w}$ is coercive (see remark 4.2) and the poles of $F(z)$ lie strictly inside the unit circle.

5.1. Observability, Constructibility and Minimality. Minimality is a fundamental property of state space models. The concept can be described purely in geometrical terms as in the following definition

Definition 5.2. An oblique Markovian splitting subspace $\mathcal{X}$ is minimal if it does not contain properly other oblique Markovian splitting subspaces.

Structural properties which are instrumental in the study of minimality are *observability and constructibility* of an oblique Markovian splitting subspace $\mathcal{X}_t$. One measures the "observability" of $\mathcal{X}_t$ on the basis of its "ability" of predicting future outputs "given" (in an appropriate sense), the future inputs.

Let $\mathcal{X}_t$ be an oblique Markovian splitting subspace, and introduce the (adjoint) *observability operator*

$$\mathbb{O}^* : \mathcal{Y}^+ \to \mathcal{X} \quad , \quad \mathbb{O}^*\lambda := E_{\|\mathcal{F}^+}\left[\lambda \mid \mathcal{X}\right], \quad \lambda \in \mathcal{Y}^+ \tag{5.3}$$

The subspace

$$\mathcal{X}_t^o := \overline{\text{Range } \mathbb{O}^*} = E_{\|\mathcal{F}_t^+}\left[\mathcal{Y}_t^+ \mid \mathcal{X}_t\right] \tag{5.4}$$

will be called the *observable subspace of* $\mathcal{X}_t$ *given* $\mathcal{F}_t^+$.

DEFINITION 5.3. We shall say that $\mathcal{X}_t$ is *observable* given $\mathcal{F}_t^+$ if $\mathcal{X}_t^o = \mathcal{X}_t$ or, equivalently, if the operator $\mathbb{O}^*$ has dense range.

Similarly we may consider the "constructibility" property of $\mathcal{X}_t$. Let

$$\mathbb{K} : \mathcal{X} \to \mathcal{P}^- \quad , \quad \mathbb{K}\xi := E_{\|\mathcal{F}^+}\left[\xi \mid \mathcal{P}^-\right], \quad \xi \in \mathcal{X}. \tag{5.5}$$

be the "constructibility" operator. It measures the degree of predictability of an element of $\mathcal{X}$ based on the joint past $\mathcal{P}^-$ given the future inputs $\mathcal{F}^+$.

DEFINITION 5.4. Let $\mathcal{X}_t$ be an oblique Markovian splitting subspace. The closure of the range of the adjoint constructibility operator is called the "constructible part" of $\mathcal{X}_t$ and denoted as $\mathcal{X}_t^c$.
We shall say that $\mathcal{X}_t$ is *constructible* if $\mathcal{X}_t^c = \mathcal{X}_t$.

PROPOSITION 5.1. *The constructible part $\mathcal{X}_t^c$ of $\mathcal{X}_t$ is given by:*

$$\mathcal{X}_t^c = \mathcal{X}_t \ominus \text{Ker } \mathbb{K}$$

PROOF. This is immediate from the fact that

$$\mathcal{H} = \overline{\text{Range } \mathfrak{C}^*} \oplus (\text{Ker } \mathbb{K})^\perp$$

for any bounded linear operator. □

A central goal of this paper will be to prove the following criterion for minimality. The proof will be given in the following.

THEOREM 5.5. *An oblique Markovian splitting subspace $\mathcal{X}$ is minimal if and only if it is both observable and constructible.*

5.2. Causal Oblique Markovian Splitting Subspaces. In this section we will restrict our attention to *causal* oblique Markovian splitting subspaces, namely we will require that

$$\mathcal{X} \subseteq \mathcal{P}^-. \tag{5.6}$$

In this case clearly

$$\bigvee_{t=-\infty}^{\infty} \mathcal{X}_t \subset \mathcal{Y} \vee \mathcal{U}$$

and hence the ambient space can be taken to be

$$\mathcal{H} = \mathcal{Y} \vee \mathcal{U}. \tag{5.7}$$

The corresponding realizations are called "internal". The motivation for this restriction is that in system identification we want to construct the state space from the available data, which (ideally) generate the subspace $\mathcal{Y} \vee \mathcal{U}$.

Once we restrict to the causal situation, since (5.6) implies that $\mathcal{X}^- \subseteq \mathcal{P}^-$, the incoming subspace is given by

$$\mathcal{S} = \mathcal{X}^- \vee \mathcal{P}^- = \mathcal{P}^-.$$

Therefore the incoming subspace coincides with $\mathcal{P}^-$ in the causal situation. We shall denote by $\mathcal{E}_t$ the wandering subspace which generates $\mathcal{P}^-$

$$\mathcal{P}_{t+1}^- = \left(\mathcal{P}_t^- + \mathcal{U}_t\right) \oplus \mathcal{E}_t \tag{5.8}$$

Note that $\mathcal{P}^-$ is p.n.d. by assumption.

Assumption 5.1 allows to "iterate" (5.8) so that

$$\mathcal{P}_{t+1}^- = \left(\mathcal{E}_t^- + \mathcal{U}_{t+1}^-\right) \oplus \mathcal{E}_t$$

It is common use to take as a basis for $\mathcal{E}_t$ the *innovation* $\boldsymbol{e}(t)$ defined by

$$\boldsymbol{e}(t) := \boldsymbol{y}(t) - E\left[\boldsymbol{y}(t) \mid \mathcal{P}_t^- \vee \mathcal{U}_t\right] \tag{5.9}$$

which is a (non normalized) white noise process whose variance is positive definite by the full-rank assumption.

We are eventually in a position to give a procedure to *construct* an oblique Markovian splitting subspace. The construction is motivated by the definition of observability (5.4) [1]

Define the *oblique predictor space* $\mathcal{X}_t^{+/-}$ at time t as follows; let $\mathcal{G}_t := \mathcal{U}_t \oplus \mathcal{E}_t$ and let

$$\mathcal{X}_t^{+/-} := E_{\|\mathcal{G}_t^+}\left[\mathcal{Y}_t^+ \mid \mathcal{P}_t^-\right] \tag{5.10}$$

Obviously $\mathcal{X}_t^{+/-}$ is contained in $\mathcal{P}_t^-$ and is oblique splitting. Let us prove that it is oblique Markovian splitting.

PROPOSITION 5.2. *The predictor space $\mathcal{X}_t^{+/-}$ is a causal oblique Markovian splitting subspace*

PROOF. It suffices to prove that

$$E_{\|\mathcal{G}_t^+}\left[\mathcal{X}_{t+1}^{+/-} \mid \mathcal{P}_t^-\right] \subseteq \mathcal{X}_t^{+/-}$$

but this is trivial since

$$\begin{aligned} E_{\|\mathcal{G}_t^+}\left[\mathcal{X}_{t+1}^{+/-} \mid \mathcal{P}_t^-\right] &= E_{\|\mathcal{G}_t^+}\left[E_{\|\mathcal{G}_{t+1}^+}\left[\mathcal{Y}_{t+1}^+ \mid \mathcal{P}_{t+1}^-\right] \mid \mathcal{P}_t^-\right] \\ &= E_{\|\mathcal{G}_t^+}\left[E\left[\mathcal{Y}_{t+1}^+ \mid \mathcal{P}_{t+1}^- + \mathcal{G}_{t+1}^+\right] \mid \mathcal{P}_t^-\right] \\ &= E_{\|\mathcal{G}_t^+}\left[\mathcal{Y}_{t+1}^+ \mid \mathcal{P}_t^-\right] \\ &\subseteq \mathcal{X}_t^{+/-} \end{aligned}$$

where the last equality follows from the fact that $\mathcal{H} = \mathcal{P}_t^- + \mathcal{G}_t^+$. □

As we have anticipated in Theorem 5.5 minimality is equivalent to both constructibility and observability. Clearly in the casual case constructibility is granted for free and therefore one just need to check observability. However, for the oblique predictor space a proof of minimality can be given directly.

PROPOSITION 5.3. *The oblique predictor space is the* minimal *causal oblique Markovian splitting subspace, in the sense that*

$$\mathcal{X}_t^{+/-} \subseteq \mathcal{X}_t$$

for every causal oblique Markovian splitting subspace $\mathcal{X}$.

[1]Note that any causal oblique Markovian splitting subspace will be constructible by construction.

PROOF. From the fact that $\mathcal{S}_t = \mathcal{X}_t^- \vee \mathcal{P}_t^- = \mathcal{P}_t^-$ we obtain

$$\mathcal{X}_t^{+/-} = E_{\|\mathcal{G}_t^+}\left[\mathcal{Y}_t^+ \mid \mathcal{P}_t^-\right] \subseteq \mathcal{X}_t$$

from which the statement follows. □

Note that in order to construct the oblique predictor space we have used the "innovation" space $\mathcal{E}_t$. Theoretically one could construct the innovation space $\mathcal{E}$ starting form the "data" $\mathcal{Y}$ and $\mathcal{U}$, using (5.8), and after that construct $\mathcal{X}_t^{+/-}$.

There is, however, a direct construction which, although somewhat complicated, permits to skip the first step of this procedure.

PROPOSITION 5.4. *Define the k step ahead (oblique) predictor space*

$$\begin{aligned}\mathcal{X}_t^k &:= E_{\|\mathcal{G}_t^+}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}_t^-\right] \\ &= E_{\|\mathcal{U}_t}\left[E_{\|\mathcal{U}_{t+1}}\left[\cdots E_{\|\mathcal{U}_{t+k}}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}_{t+k}^-\right]\cdots \mid \mathcal{P}_{t+1}^-\right] \mid \mathcal{P}_t^-\right]\end{aligned} \tag{5.11}$$

Then the oblique predictor space can be computed as the (closed) infinite vector sum

$$\mathcal{X}_t^{+/-} = \bigvee_{k\geq 0} \mathcal{X}_t^k \tag{5.12}$$

PROOF. We just need to show that (5.11) holds true. From Theorem 4.6 we have

$$\mathcal{Y}_{t+k} \subseteq \left(\mathcal{X}_{t+k}^{+/-} + \mathcal{U}_{t+k}\right) \oplus \mathcal{E}_{t+k}$$

and

$$\mathcal{X}_{t+h}^{+/-} \subseteq \left(\mathcal{X}_{t+h-1}^{+/-} + \mathcal{U}_{t+h-1}\right) \oplus \mathcal{E}_{t+h-1}.$$

These two conditions imply that

$$E_{\|\mathcal{U}_{t+k}}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}_{t+k}^-\right] \subseteq \mathcal{X}_{t+k}^{+/-}$$

and

$$E_{\|\mathcal{U}_{t+h}}\left[\mathcal{X}_{t+h+1} \mid \mathcal{P}_{t+h}^-\right] \subseteq \mathcal{X}_{t+h}^{+/-}$$

which, together with

$$\mathcal{Y}_{t+k} \subseteq \mathcal{X}_t^{+/-} + \mathcal{U}_{[t,t+k)} + \mathcal{E}_{[t,t+k)}$$

imply (5.11). □

REMARK 5.6. Note that in the finite dimensional case the sum (5.12) can be limited to n terms, where n is the dimension of a minimal causal realization.

6. Scattering Representations of Oblique Markovian Splitting Subspaces

In this section we shall establish some general properties of oblique Markovian splitting subspaces in order to facilitate the study of minimality.

Let $\mathcal{X}$ be an oblique Markovian splitting subspace and let $\mathcal{S}$ defined by (4.2) and $\bar{\mathcal{S}}$, defined by (4.3) be the associated incoming-outgoing pair. The oblique Markovian splitting property (4.13) can be written as

$$E_{\|\mathcal{F}^+}[\bar{\mathcal{S}} \,|\, \mathcal{S}] = E_{\|\mathcal{F}^+}[\bar{\mathcal{S}} \,|\, \mathcal{X}]. \tag{6.1}$$

The following Lemma gives a formal characterization of any oblique Markovian splitting subspace as the oblique predictor space of the outgoing subspace.

Lemma 6.1. *Let* $(\mathcal{S}, \bar{\mathcal{S}})$ *and* $\mathcal{F}^+$ *be as defined above. Then*

$$\mathcal{X} = E_{\|\mathcal{F}^+}[\bar{\mathcal{S}} \,|\, \mathcal{S}]$$

hence every $\mathcal{X}$ *is the oblique predictor space of* $\bar{\mathcal{S}}$*, given* $\mathcal{S}$*, along* $\mathcal{F}^+$*.*

Proof. Every element $\bar{\boldsymbol{s}}$ of $\bar{\mathcal{S}}$ has the form $\bar{\boldsymbol{s}} = \boldsymbol{y} + \boldsymbol{x}$, $\boldsymbol{y} \in \mathcal{Y}^+$, $\boldsymbol{x} \in \mathcal{X}^+$ so that $E_{\|\mathcal{U}^+}[\boldsymbol{y} \,|\, \mathcal{S}] = E_{\|\mathcal{F}^+}[\boldsymbol{y} \,|\, \mathcal{X}] \in \mathcal{X} \subseteq \mathcal{S}$. On the other hand, by definition of oblique splitting we have

$$E_{\|\mathcal{F}^+}[\boldsymbol{x} \,|\, \mathcal{S}] = E_{\|\mathcal{F}^+}[\boldsymbol{x} \,|\, \mathcal{X}] \qquad \boldsymbol{x} \in \mathcal{X}^+,$$

therefore

$$\overline{\text{span}}\{E_{\|\mathcal{F}^+}[\bar{\boldsymbol{s}} \,|\, \mathcal{S}] \,|\, \bar{\boldsymbol{s}} \in \bar{\mathcal{S}}\} = \mathcal{X}.$$

This implies that $\mathcal{X}$ is the oblique predictor space of $\bar{\mathcal{S}}$ given $\mathcal{S}$ along $\mathcal{F}^+$. □

Lemma 6.2. *Let* $(\mathcal{S}, \bar{\mathcal{S}})$ *be as above. Then*

$$\mathcal{X} = \bar{\mathcal{S}} \cap \mathcal{S} \tag{6.2}$$

Proof. The fact that $\mathcal{X} \subseteq \bar{\mathcal{S}} \cap \mathcal{S}$ is trivial. Let us show the other inclusion. Let $\bar{\boldsymbol{s}} \in \bar{\mathcal{S}} \cap \mathcal{S}$; then $\bar{\boldsymbol{s}} \in \bar{\mathcal{S}}$ and $\bar{\boldsymbol{s}} \in \mathcal{S}$. Therefore

$$\bar{\boldsymbol{s}} = E_{\|\mathcal{F}^+}[\bar{\boldsymbol{s}} \,|\, \mathcal{S}] = E_{\|\mathcal{F}^+}[\bar{\boldsymbol{s}} \,|\, \mathcal{X}] \in \mathcal{X}.$$

□

The following proposition is a generalization of the perpendicular intersection property known for ”orthogonal” splitting subspaces.

Proposition 6.1. *Let* $\mathcal{X}$ *be an oblique Markovian splitting subspace and let* $\mathcal{S}$*,* $\bar{\mathcal{S}}$ *be the relative incoming-outgoing pair of subspaces. Then the following* oblique intersection property *holds:*

$$\bar{\mathcal{S}} \perp \mathcal{S} \mid \left(\mathcal{X} + \mathcal{F}^+\right) \tag{6.3}$$

Proof. Condition (7.3), is equivalent to (see Lemma 2.2)

$$E[\bar{\mathcal{S}} \,|\, \mathcal{S} + \mathcal{F}^+] = E[\bar{\mathcal{S}} \,|\, \mathcal{X} + \mathcal{F}^+]$$

which by (6.2) is precisely the oblique intersection property (6.3). □

The following theorem gives an "almost" one-to-one correspondence between oblique Markovian splitting subspaces and "scattering pairs".

THEOREM 6.1. *Let $\mathcal{H}$ be a Hilbert space of random variables with shift operator σ and let $\mathcal{X}$ be a subspace of $\mathcal{H}$ such that*

$$\mathcal{H} = \mathcal{Y} \vee \mathcal{U} \vee \left(\bigvee_t \mathcal{X}_t \right).$$

Then $\mathcal{X}$ is an oblique Markovian splitting subspace, if and only if

$$\mathcal{X} = \bar{\mathcal{S}} \cap \mathcal{S}$$

for some pair of subspaces $\mathcal{S}$, $\bar{\mathcal{S}}$ such that the following properties hold

(1) *Extended past and future property*

$$\begin{cases} \mathcal{Y}^+ \subseteq \bar{\mathcal{S}} \\ \mathcal{P}^- \subseteq \mathcal{S} \end{cases} , \quad \mathcal{S} \cap \mathcal{F}^+ = \{0\}$$

(2) *Shift-invariance*

$$\begin{cases} \sigma\bar{\mathcal{S}} \subseteq \bar{\mathcal{S}} \\ \sigma^*\mathcal{S} \subseteq \mathcal{S} \end{cases}$$

(3) *Oblique intersection at $\mathcal{X}$*

$$\bar{\mathcal{S}} \perp \mathcal{S} \mid \left((\bar{\mathcal{S}} \cap \mathcal{S}) + \mathcal{F}^+ \right)$$

Conversely, given an oblique Markovian splitting subspace $\mathcal{X}$, a pair of subspaces satisfying conditions 1), 2), 3), *can be constructed as follows*

$$\mathcal{S} = \mathcal{P}^- \vee \mathcal{X}^- \quad , \quad \mathcal{Y}^+ \vee \mathcal{X}^+ \subseteq \bar{\mathcal{S}} \subseteq \mathcal{Y}^+ \vee \mathcal{X}^+ \vee \mathcal{U}^+. \tag{6.4}$$

The minimal subspace $\bar{\mathcal{S}}$ satisfying 1), 2), *and* 3) *(i.e. contained in any other $\bar{\mathcal{S}}$ satisfying* 1), 2), *and* 3)*), is given by*

$$\bar{\mathcal{S}} = \mathcal{Y}^+ \vee \mathcal{X}^+$$

PROOF. Let $\mathcal{X}$ be an oblique Markovian splitting subspace, then $\mathcal{S} = \mathcal{P}^- \vee \mathcal{X}^-$ and $\bar{\mathcal{S}} = \mathcal{Y}^+ \vee \mathcal{X}^+$ satisfy the assumptions above and $\mathcal{X} = \bar{\mathcal{S}} \cap \mathcal{S}$. Conversely, let $\mathcal{S}$, $\bar{\mathcal{S}}$ be subspaces of $\mathcal{H}$ which satisfies the conditions above. Define the subspace $\mathcal{X} = \bar{\mathcal{S}} \cap \mathcal{S}$; then by assumptions 1) and 2) we have that $\mathcal{P}^- \vee \mathcal{X}^- \subseteq \mathcal{S}$ and $\mathcal{Y}^+ \vee \mathcal{X}^+ \subseteq \bar{\mathcal{S}}$, which by the oblique intersection property 3) implies that

$$\left(\mathcal{Y}^+ \vee \mathcal{X}^+\right) \perp \left(\mathcal{P}^- \vee \mathcal{X}^-\right) \mid \left(\mathcal{X} + \mathcal{F}^+\right). \tag{6.5}$$

By lemma 2.2 condition (6.5) is equivalent to the oblique Markovian splitting property (4.13).

Let us prove that $\mathcal{S}$ and $\bar{\mathcal{S}}$ are given by (6.4). We have already pointed out that $\mathcal{S}^m := \mathcal{P}^- \vee \mathcal{X}^- \subseteq \mathcal{S}$. Assume the inclusion is strict; then since $\mathcal{S} \subseteq \mathcal{H} = \mathcal{S}^m \vee \bar{\mathcal{S}} \vee \mathcal{U}^+$, $\boldsymbol{s} \in \mathcal{S}$ can be written as: $\boldsymbol{s} = \boldsymbol{s}^m + \bar{\boldsymbol{s}} + \boldsymbol{u}$ where $\boldsymbol{s}^m \in \mathcal{S}^m$, $\bar{\boldsymbol{s}} \in \bar{\mathcal{S}}$, $\boldsymbol{u} \in \mathcal{U}^+$. Therefore we have:

$$\boldsymbol{s} = E_{\|\mathcal{F}^+}[\boldsymbol{s}|\mathcal{S}] = \boldsymbol{s}^m + E_{\|\mathcal{F}^+}[\bar{\boldsymbol{s}}|\mathcal{S}] = \boldsymbol{s}^m + \boldsymbol{x}$$

where $\boldsymbol{x} \in \mathcal{X} \subseteq \mathcal{S}^m$ which contradict the hypothesis that $\boldsymbol{s} \notin \mathcal{S}^m$.
Similarly, we have seen that $\bar{\mathcal{S}}^m := \mathcal{Y}^+ \vee \mathcal{X}^+ \subseteq \bar{\mathcal{S}}$. Assume $(\mathcal{Y}^+ \vee \mathcal{X}^+ \vee \mathcal{U}^+) \subset \bar{\mathcal{S}}$ strictly. Then, since $\mathcal{Y}^+ \vee \mathcal{X}^+ \vee \mathcal{U}^+ \vee \mathcal{S}^m = \mathcal{H}$, there exists $\bar{\boldsymbol{s}} \in \bar{\mathcal{S}}$, $\boldsymbol{s} \notin (\mathcal{Y}^+ \vee \mathcal{X}^+ \vee \mathcal{U}^+)$ which lies in $\mathcal{S}^m$. Therefore $\bar{\boldsymbol{s}} \in \mathcal{X}$, and hence $\bar{\boldsymbol{s}} \in \bar{\mathcal{S}}^m$ which contradicts the hypothesis. Requiring $\bar{\mathcal{S}}$ to be minimal implies that $\bar{\mathcal{S}}^m = \bar{\mathcal{S}}$ since $\bar{\mathcal{S}}^m \subseteq \bar{\mathcal{S}}$. □

The question of minimality of oblique Markovian splitting subspaces can be rephrased as a question of minimality for the subspaces $\mathcal{S}$ and $\bar{\mathcal{S}}$. In fact, given a Markovian splitting subspace $\mathcal{X}$ and the corresponding pair $(\mathcal{S}, \bar{\mathcal{S}})$, reducing $(\mathcal{S}, \bar{\mathcal{S}})$ without violating the properties 1), 2) and 3) of theorem 6.1 amounts to constructing an oblique Markovian splitting subspace which is contained in $\mathcal{X}$.

6.1. Scattering Pairs and Minimality. Our aim in this section is to adapt to oblique splitting subspaces a construction inspired by a similar procedure in stochastic realization theory, [**14**], which allows to construct a minimal Markovian splitting subspace starting from an arbitrarily "large" scattering pair $\left(\mathcal{S}, \bar{\mathcal{S}}\right)$ of perpendicularly intersecting subspaces. As it will be clear in a little while, we will only be able to draw a completely parallel construction in case of absence of feedback.

The construction of a minimal Markovian splitting subspace can be done in principle by reducing (in the sense of subspace inclusion) the subspaces $\left(\mathcal{S}, \bar{\mathcal{S}}\right)$ without violating properties 1), 2) and 3) of Theorem 6.1.

Before doing so we shall clarify the geometric meaning of constructibility and observability.

Proposition 6.2. *Let $\mathcal{X}$ be an oblique Markovian splitting subspace and let $\left(\mathcal{S}, \bar{\mathcal{S}}\right)$ be the scattering pair associated to it. Let us introduce the* extended scattering pair, $\mathcal{S}_e := \mathcal{S} + \mathcal{F}^+$ *and* $\bar{\mathcal{S}}_e := \bar{\mathcal{S}} \vee \mathcal{F}^+$. *Then $\mathcal{X}$ is observable if and only if*

$$\bar{\mathcal{S}}_e = \mathcal{S}_e^\perp \vee \mathcal{Y}^+ \vee \mathcal{F}^+ \tag{6.6}$$

and constructible if and only if

$$\mathcal{S}_e = \bar{\mathcal{S}}_e^\perp \vee \mathcal{P}^- \vee \mathcal{F}^+. \tag{6.7}$$

Proof. Assume (6.6) holds. Since by definition $\mathcal{S}_e^\perp \perp (\mathcal{S} + \mathcal{F}^+)$ then

$$\begin{aligned} \mathcal{X} &= E_{\|\mathcal{F}^+}\left[\bar{\mathcal{S}} \mid \mathcal{S}\right] \\ &= E_{\|\mathcal{F}^+}\left[\bar{\mathcal{S}}_e \mid \mathcal{X}\right] \\ &= E_{\|\mathcal{F}^+}\left[\mathcal{Y}^+ \mid \mathcal{X}\right] \end{aligned}$$

which is observability. Conversely, if $\mathcal{X}$ is observable

$$\mathcal{X} + \mathcal{F}^+ = E\left[\mathcal{Y}^+ \vee \mathcal{F}^+ \mid \mathcal{X} + \mathcal{F}^+\right]$$

which is in turn equivalent to

$$\left(\mathcal{Y}^+ \vee \mathcal{F}^+\right)^\perp \cap \left(\mathcal{X} + \mathcal{F}^+\right) = \{0\}$$

Taking orthogonal complements we can rewrite

$$\mathcal{Y}^+ \vee \mathcal{F}^+ \vee \left(\mathcal{X} + \mathcal{F}^+\right)^\perp = \mathcal{H};$$

since

$$\left(\mathcal{X} + \mathcal{F}^+\right) = \mathcal{S}_e \cap \bar{\mathcal{S}}_e$$

and $\mathcal{S}_e^\perp \subseteq \bar{\mathcal{S}}_e$ the following orthogonal decomposition holds

$$\mathcal{H} = \bar{\mathcal{S}}_e^\perp \oplus \left(\mathcal{S}_e^\perp \vee \mathcal{Y}^+ \vee \mathcal{F}^+\right)$$

from which the conclusion follows.
As far as constructibility is concerned, assume $\mathcal{X}$ is not constructible. There exist $\boldsymbol{x} \in \mathcal{X}$ such that $E_{\|\mathcal{F}^+}[x \mid \mathcal{P}^-] = 0$, i.e. $\boldsymbol{x} \in \mathcal{F}^+ \oplus (\mathcal{P}^- + \mathcal{F}^+)^\perp$ and therefore can be uniquely decomposed as $\boldsymbol{x} = \boldsymbol{x}_f \oplus \tilde{\boldsymbol{x}}_f$ where $\boldsymbol{x}_f \in \mathcal{F}^+$ and $\tilde{\boldsymbol{x}}_f \in (\mathcal{P}^- + \mathcal{F}^+)^\perp$. Note that $\tilde{\boldsymbol{x}}_f \neq 0$ since $\mathcal{X} \cap \mathcal{F}^+ = \{0\}$. It follows that $\tilde{\boldsymbol{x}}_f \in \mathcal{S}_t^e \cap \bar{\mathcal{S}}_t^e$. This condition insures that $\tilde{\boldsymbol{x}}_f \perp \left(\bar{\mathcal{S}}^e\right)^\perp \vee \mathcal{P}^- \vee \mathcal{F}^+$ and therefore $\tilde{\boldsymbol{x}}_f \notin \left(\bar{\mathcal{S}}^e\right)^\perp \vee \mathcal{P}^- \vee \mathcal{F}^+$ which implies $\mathcal{S}^e \subset \left(\bar{\mathcal{S}}^e\right)^\perp \vee \mathcal{P}^- \vee \mathcal{F}^+$ strictly.
Conversely, assume $\mathcal{S}^e \subset \left(\bar{\mathcal{S}}^e\right)^\perp \vee \mathcal{P}^- \vee \mathcal{F}^+$ strictly. Then there exists $\boldsymbol{s} \in \mathcal{S}^e$ and $\boldsymbol{s} \in \left[\left(\bar{\mathcal{S}}^e\right)^\perp \vee \mathcal{P}^- \vee \mathcal{F}^+\right]^\perp$ or, alternatively, $\boldsymbol{s} \in \bar{\mathcal{S}}^e \cap (\mathcal{P}^- + \mathcal{F}^+)^\perp$. Therefore $\boldsymbol{s} \in \mathcal{S}_t^e \cap \mathcal{S}^e \cap (\mathcal{P}^- + \mathcal{F}^+)^\perp = (\mathcal{X} + \mathcal{F}^+) \cap (\mathcal{P}^- + \mathcal{F}^+)^\perp$. The last condition insures that $\boldsymbol{s} = \boldsymbol{s}_x + \boldsymbol{s}_f$, $\boldsymbol{s}_x \in \mathcal{X}$, $\boldsymbol{s}_f \in \mathcal{F}^+$, and, for obvious reasons $\boldsymbol{s}_x \neq 0$. Writing $\boldsymbol{s}_x = \boldsymbol{s} - \boldsymbol{s}_f$ we have that $\boldsymbol{s}_x \in \mathcal{X} \cap \mathcal{F}^+ \oplus (\mathcal{P}^- + \mathcal{F}^+)^\perp$ contradicting constructibility, which concludes the proof. □

We shall now introduce an orthogonal intersection property which is implied by the oblique intersection.

LEMMA 6.3. *Let* $\left(\mathcal{S}, \bar{\mathcal{S}}\right)$ *satisfy the oblique intersection property*

$$\mathcal{S} \perp \bar{\mathcal{S}} \mid \left(\mathcal{X} + \mathcal{F}^+\right).$$

Then the extended subspaces $\mathcal{S}_e = \mathcal{S} \vee \mathcal{F}^+$ *and* $\bar{\mathcal{S}}_e = \bar{\mathcal{S}} \vee \mathcal{F}^+$ *intersect perpendicularly, i.e.*

$$\mathcal{S}_e \perp \bar{\mathcal{S}}_e \mid \left(\mathcal{S}_e \cap \bar{\mathcal{S}}_e\right).$$

PROOF. The proof follows readily from Theorem 2.1 in [**14**] □

Making use of this lemma we obtain immediately the following *orthogonal decomposition of the ambient space* $\mathcal{H}$ which is analogous to the one valid in stochastic realization for time series, and plays an important role in many structural questions in stochastic systems theory.

THEOREM 6.2. *Let* $\mathcal{X}$ *be an oblique Markovian splitting subspace and* $(\mathcal{S}_e, \bar{\mathcal{S}}_e)$ *the associated extended scattering pair. Then the following orthogonal decomposition holds*

$$\mathcal{H} = \mathcal{S}_e^\perp \oplus \left(\mathcal{X} + \mathcal{F}^+\right) \oplus \bar{\mathcal{S}}_e^\perp \tag{6.8}$$

The characterization of oblique Markovian splitting subspaces in terms of their scattering pair is a fundamental tool to study minimality. As we have seen an oblique Markovian splitting subspace can always be represented as the intersection of $\mathcal{S}$ and $\bar{\mathcal{S}}$. These subspaces, or more precisely their extended versions are related to observability and constructibility. Apparently, failing either of them, these subspace are not "minimal", in the sense of proposition 6.2. At this point, following classical stochastic realization theory, we would need a procedure to reduce, if possible, the subspaces $\mathcal{S}$ and $\bar{\mathcal{S}}$. Unfortunately such a procedure is not yet available and at this point the analogy with the classical theory seems to halt.

Nevertheless a proof of theorem 5.5 can still be given.

Proof of Theorem 5.5 According to theorem 6.1, if (6.6) and (6.7) hold, it is not possible to reduce these subspaces and therefore $\mathcal{X}$ must be minimal. Conversely, if $\mathcal{X}$ is minimal, it cannot be possible to reduce $\mathcal{S}$ and $\bar{\mathcal{S}}$ any further. This implies that (6.6) and (6.7) hold and therefore $\mathcal{X}$ is both observable and constructible. □

7. Stochastic Realization in the Absence of Feedback

When there is no feedback from from $\boldsymbol{y}$ to $\boldsymbol{u}$, some of the results presented above simplify considerably. For instance the construction of the oblique predictor space, somewhat complicated in the general setting, can be simplified when there is no feedback. Moreover, specializing some definitions, we shall also be able in this case to give a procedure to reduce the incoming and outgoing subspaces in order to achieve minimality. The following lemma will be useful in this respect.

LEMMA 7.1. *Assume there is no feedback from $\boldsymbol{y}$ to $\boldsymbol{u}$. Let $\mathcal{E}$ be the innovation space of $\boldsymbol{y}$ defined by (5.9), then:*

$$E_{\|\mathcal{U}_t^+}\left[\mathcal{E}_t^+ \,|\, \mathcal{P}_t^-\right] = \{0\}$$

PROOF. Decomposition (4.5) can be rewritten in this causal case as:

$$\mathcal{P}_{t+1}^- = \left(\mathcal{P}_t^- + \mathcal{U}_t\right) \oplus \mathcal{E}_t$$

Since $\mathcal{E}_t = \text{span}\,\{\boldsymbol{e}(t)\}$ it suffices to show that

$$E_{\|\mathcal{U}_t^+}\left[\boldsymbol{e}(t+k) \,|\, \mathcal{P}_t^-\right] = 0$$

for all $k \geq 0$. As we have already pointed out $\boldsymbol{e}(t) \perp \mathcal{U}_{t+1}^+$ and obviously $\boldsymbol{e}(t) \perp \mathcal{P}_t^- \vee \mathcal{U}_t$, therefore the oblique projection is zero.

This fact can be verified directly since by absence of feedback

$$\begin{aligned} E\left[\boldsymbol{y}(t) \mid \mathcal{P}_t^- \vee \mathcal{U}_t^+\right] &= E\left[\boldsymbol{y}(t) \mid (\mathcal{Y}_s)_t^- \oplus \left(\mathcal{U}_t^- \vee \mathcal{U}_t^+\right)\right] \\ &= E\left[\boldsymbol{y}(t) \mid (\mathcal{Y}_s)_t^-\right] \oplus E\left[\boldsymbol{y}(t) \mid \mathcal{U}\right] \\ &= \hat{\boldsymbol{y}}_s(t) \oplus E\left[\boldsymbol{y}(t) \mid \mathcal{U}_{t+1}^-\right] \\ &= E\left[\boldsymbol{y}(t) \mid \mathcal{P}_t^- \vee \mathcal{U}_t\right] \end{aligned}$$

which proves that $\boldsymbol{e}(t)$ is orthogonal to $\mathcal{U}_{t+1}^+$ and hence $\boldsymbol{e}(t) \perp \mathcal{U}$. □

We have seen that the wandering subspace generated by the innovation is orthogonal to the whole input history in the absence of feedback. Recall that the feedback-free property was defined from an input-output point of view, apparently putting no restrictions on the state $\boldsymbol{x}$. However it is straightforward to see that:

PROPOSITION 7.1. *In the causal case, absence of feedback from* $\boldsymbol{y}$ *to* $\boldsymbol{u}$ *implies absence of feedback from* $\boldsymbol{x}$ *to* $\boldsymbol{u}$.

PROOF. Since $\mathcal{X}_t^- \subseteq \mathcal{P}_t^-$ and $\mathcal{P}_t^- \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^-$, it is also true that

$$\mathcal{X}_t^- \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^- \tag{7.1}$$

holds. □

Even in a non-causal situation, it is useful to restrict our attention to realizations whose state space satisfies the condition (7.1). Let us consider the subspace $\mathcal{Z}_t^- = \mathcal{Y}_t^- \vee \mathcal{X}_t^-$. Henceforth we shall only consider state spaces such that

$$\mathcal{Z}_t^- \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^- \tag{7.2}$$

and we say that the corresponding realizations are *feedback free*. This extended notion of absence of feedback guarantees not only that the innovation $\boldsymbol{e}$ is orthogonal to future inputs, but that so will be any wandering subspace $\mathcal{W}_t$. The following proposition states this formally.

PROPOSITION 7.2. *Let* $\mathcal{X}$ *be an oblique Markovian splitting subspace. The wandering subspace* $\mathcal{W}$ *which generates* $\mathcal{X}$ *is orthogonal to the whole input history if and only if the feedback-free condition (7.2) is satisfied.*

PROOF. (if) Recall that

$$\mathcal{S}_t = \mathcal{X}_t^- \vee \mathcal{Y}_t^- \vee \mathcal{U}_t^-,$$

and, by (7.2),

$$\mathcal{S}_t \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^-.$$

From

$$\mathcal{S}_{t+1} = (\mathcal{S}_t + \mathcal{U}_t) \oplus \mathcal{W}_t$$

we have that $\mathcal{W}_t$ is contained in $\mathcal{S}_{t+1}$ and therefore $\mathcal{W}_t \perp \mathcal{U}_{t+1}^+ \mid \mathcal{U}_{t+1}^-$; since $\mathcal{W}_t \perp \mathcal{U}_{t+1}^-$ by construction, we obtain $\mathcal{W}_t \perp \mathcal{U}_{t+1}^+$ and hence

$$\mathcal{W}_t \perp \mathcal{U}$$

which is the thesis.

(only if) Assume $\mathcal{W}_t \perp \mathcal{U}$, since $\mathcal{S}_t = \mathcal{U}_t^- + \mathcal{W}_t^-$ it follows that

$$\mathcal{S}_t \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^-.$$

Therefore, since $\mathcal{Z}_t^- = \left(\mathcal{X}_t^- \vee \mathcal{Y}_t^-\right) \subseteq \mathcal{S}_t$, the thesis follows

$$\mathcal{Z}_t^- \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^-.$$

□

REMARK 7.1. Note that, under hypothesis (7.2), also the incoming subspace $\mathcal{S}_t$ satisfies $\mathcal{S}_t \perp \mathcal{U}_t^+ \mid \mathcal{U}_t^-$. It follows that the richness condition

$$\mathcal{U}_t^+ \cap \mathcal{S}_t = \{0\}$$

is automatically satisfied as long as the input is coercive (Assumption 3.6).

The following result gives a somehow simpler geometric characterization of the oblique Markovian splitting property in the absence of feedback.

THEOREM 7.2. *Let the symbols have the same meaning as above. Assume there is no feedback from* $\boldsymbol{y}$ *to* $\boldsymbol{u}$ *and that (7.2) holds. The subspace* $\mathcal{X}$ *is oblique Markovian splitting if and only if*

$$E_{\|\mathcal{U}^+}[\bar{\mathcal{S}} \mid \mathcal{S}] = E_{\|\mathcal{U}^+}[\bar{\mathcal{S}} \mid \mathcal{X}]. \tag{7.3}$$

PROOF. The condition is obviously sufficient since $\mathcal{Y}_t \vee \mathcal{X}_{t+1} \subseteq \bar{\mathcal{S}}_t$ and, by absence of feedback (7.2):

$$E[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{S}_t + \mathcal{U}_t^+] = E[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{S}_t + \mathcal{U}_t]$$

which by lemma (2.1) implies that

$$E_{\|\mathcal{U}_t^+}[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{S}_t] = E_{\|\mathcal{U}_t}[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{S}_t]$$

and therefore

$$E_{\|\mathcal{U}_t}[\mathcal{Y}_t \vee \mathcal{X}_{t+1} \mid \mathcal{S}_t] \subseteq \mathcal{X}_t.$$

To prove the other implication just note that by Theorem 4.6

$$\mathcal{Y}_{t+k} \subseteq \big(\mathcal{X}_t + \mathcal{U}_{[t,\,t+k)}\big) \oplus \mathcal{W}_{[t,\,t+k)}$$

and

$$\mathcal{X}_{t+k+1} \subseteq \big(\mathcal{X}_t + \mathcal{U}_{[t,\,t+k)}\big) \oplus \mathcal{W}_{[t,\,t+k)}$$

where the last sum is orthogonal from Proposition 7.2. It follows that for every $k \geq 0$

$$E_{\|\mathcal{U}_t^+}[\mathcal{Y}_{t+k} \vee \mathcal{X}_{t+k+1} \mid \mathcal{S}_t] \subseteq \mathcal{X}_t$$

which is equivalent to (7.3). □

The following lemma is the equivalent of Lemma 6.1

LEMMA 7.2. *Let* $(\mathcal{S}, \bar{\mathcal{S}})$ *and* $\mathcal{U}^+$ *be as defined above. Then*

$$\mathcal{X} = E_{\|\mathcal{U}^+}[\bar{\mathcal{S}} \mid \mathcal{S}]$$

Hence every oblique MArkovian splitting subspace $\mathcal{X}$ *is the oblique predictor space for* $\bar{\mathcal{S}}$*, given* $\mathcal{S}$*, along* $\mathcal{U}^+$*.*

PROOF. In the feedback free case, from Proposition 7.2we get $\mathcal{W}^+ \perp (\mathcal{S} + \mathcal{U}^+)$. Therefore, since $\mathcal{F}^+ = \mathcal{U}^+ \vee \mathcal{W}^+$,

$$E_{\|\mathcal{U}^+}[\mathcal{X} \mid \mathcal{S}] = E_{\|\mathcal{F}^+}[\mathcal{X} \mid \mathcal{S}] = \mathcal{X}$$

where the last equality follows from Lemma 6.1. □

The following proposition specializes the concept of oblique intersection to the case when there is no feedback.

PROPOSITION 7.3. *Assume there is no feedback from $\boldsymbol{y}$ to $\boldsymbol{u}$. Let $\mathcal{X}$ be an oblique Markovian splitting subspace and let $\mathcal{S}$, $\bar{\mathcal{S}}$ be the incoming and outgoing subspaces attached to it. Then the following holds:*

$$\bar{\mathcal{S}} \perp \mathcal{S} \mid \left((\bar{\mathcal{S}} \cap \mathcal{S}) + \mathcal{U}^+\right) \tag{7.4}$$

PROOF. Condition (7.3), is equivalent to (see Lemma 2.2)

$$E[\bar{\mathcal{S}} \mid \mathcal{S} + \mathcal{U}^+] = E[\bar{\mathcal{S}} \mid \mathcal{X} + \mathcal{U}^+]$$

which by (6.2) is precisely the oblique intersection property (7.4). □

REMARK 7.3. It is worth to stress, at this point, that condition (3) (oblique intersection) of Theorem 6.1 can be replaced by condition (7.4).

The following theorem gives a characterization of the oblique predictor space in absence of feedback.

THEOREM 7.4. *In absence of feedback the oblique predictor space can be computed by the formula*

$$\mathcal{X}^{+/-} := E_{\|\mathcal{U}^+}[\mathcal{Y}^+ \mid \mathcal{P}^-]. \tag{7.5}$$

PROOF. To show this let us just note that by (7.2) and by lemma 2.1

$$E_{\|\mathcal{U}_{t+h}}\left[\mathcal{Y}_{t+h} \mid \mathcal{P}^-_{t+h}\right] = E_{\|\mathcal{U}^+_{t+h}}\left[\mathcal{Y}_{t+h} \mid \mathcal{P}^-_{t+h}\right] \subseteq \mathcal{X}_{t+h}$$

for any causal oblique markovian splitting subspace $\mathcal{X}_{t+h}$ and

$$E_{\|\mathcal{U}_{t+h}}\left[\mathcal{X}_{t+h+1} \mid \mathcal{P}^-_{t+h}\right] = E_{\|\mathcal{U}^+_{t+h}}\left[\mathcal{X}_{t+h+1} \mid \mathcal{P}^-_{t+h}\right]$$

in the absence of feedback; this implies the following:

$$\begin{aligned}\left(\mathcal{X}^{+/-}_t\right)^k &= E_{\|\mathcal{U}_t}\left[E_{\|\mathcal{U}_{t+1}}\left[\cdots E_{\|\mathcal{U}_{t+k}}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}^-_{t+k}\right]\cdots \mid \mathcal{P}^-_{t+1}\right] \mid \mathcal{P}^-_t\right]\\ &= E_{\|\mathcal{U}^+_t}\left[E_{\|\mathcal{U}^+_{t+1}}\left[\cdots E_{\|\mathcal{U}^+_{t+k}}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}^-_{t+k}\right]\cdots \mid \mathcal{P}^-_{t+1}\right] \mid \mathcal{P}^-_t\right]\\ &= E_{\|\mathcal{U}^+_t}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}^-_t\right]\end{aligned}$$

and hence

$$\begin{aligned}\mathcal{X}^{+/-}_t &= \bigvee_{k\geq 0} E_{\|\mathcal{U}^+_t}\left[\mathcal{Y}_{t+k} \mid \mathcal{P}^-_t\right]\\ &= E_{\|\mathcal{U}^+_t}\left[\mathcal{Y}^+_t \mid \mathcal{P}^-_t\right]\end{aligned} \tag{7.6}$$

□

7.1. Scattering Pairs and Minimality (without Feedback). In this section we shall give a procedure to reduce the state space when it is not minimal, by reducing the incoming and outgoing subspaces using a two-steps procedure similar to that described in [**14**].

The construction of a minimal Markovian splitting subspace can be done by reducing (in the sense of subspace inclusion) the subspaces $(\mathcal{S}, \bar{\mathcal{S}})$ without violating properties (1), (2) and (3) (which in this case is equivalent to (7.4)) of Theorem 6.1. We shall first state some technical results which will be needed throughout the section.

First of all, we shall introduce an orthogonal intersection property which is implied by the oblique intersection.

Lemma 7.3. *Let* $(\mathcal{S}, \bar{\mathcal{S}})$ *satisfy the oblique intersection property*

$$\mathcal{S} \perp \bar{\mathcal{S}} \mid \left((\mathcal{S} \cap \bar{\mathcal{S}}) + \mathcal{U}^+ \right).$$

Then the extended subspaces $\mathcal{S}_{eu} := \mathcal{S} \vee \mathcal{U}^+$ *and* $\bar{\mathcal{S}}_{eu} := \bar{\mathcal{S}} \vee \mathcal{U}^+$ *intersect perpendicularly, i.e.*

$$\mathcal{S}_{eu} \perp \bar{\mathcal{S}}_{eu} \mid \left(\mathcal{S}_{eu} \cap \bar{\mathcal{S}}_{eu} \right).$$

Proof. The proof follows readily from Theorem 2.1 in [**14**] □

Using this lemma we obtain the orthogonal decomposition of the ambient space $\mathcal{H}$ valid in the feedback-free case

$$\mathcal{H} = \mathcal{S}_{eu}^{\perp} \oplus \left(\mathcal{X} + \mathcal{U}^+ \right) \oplus \bar{\mathcal{S}}_{eu}^{\perp} \tag{7.7}$$

Note that $\mathcal{Y}^+ \vee \mathcal{U}^+ \subset \bar{\mathcal{S}}_{eu}$ and $\mathcal{P}^- \vee \mathcal{U}^+ \subset \mathcal{S}_{eu}$ must hold for every pair of subspaces $(\mathcal{S}, \bar{\mathcal{S}})$ attached to an oblique Markovian splitting subspace.

We shall construct a pair of subspaces $(\mathcal{S}^1, \bar{\mathcal{S}}^1)$, which are contained in $(\mathcal{S}, \bar{\mathcal{S}})$ and which satisfies all the conditions of Theorem 6.1 by defining their "extended" version and then by properly reducing them.

Define

$$\bar{\mathcal{S}}^1_{eu} := \mathcal{S}_{eu}^{\perp} \vee \mathcal{Y}^+ \vee \mathcal{U}^+ \tag{7.8a}$$

$$\mathcal{S}^1_{eu} := \left(\bar{\mathcal{S}}^1_{eu}\right)^{\perp} \vee \mathcal{P}^- \vee \mathcal{U}^+ \tag{7.8b}$$

and also the related state space

$$\mathcal{X}^1_{eu} := \bar{\mathcal{S}}^1_{eu} \cap \mathcal{S}^1_{eu} \tag{7.9}$$

Lemma 7.4. *The pair of subspaces defined in (7.8) intersect perpendicularly, i.e.*

$$\bar{\mathcal{S}}^1_{eu} \perp \mathcal{S}^1_{eu} \mid \bar{\mathcal{S}}^1_{eu} \cap \mathcal{S}^1_{eu}. \tag{7.10}$$

Proof. Clearly $\mathcal{H} = \bar{\mathcal{S}}^1_{eu} \vee \mathcal{S}^1_{eu}$ and $\left(\bar{\mathcal{S}}^1_{eu}\right)^{\perp} \subset \mathcal{S}^1_{eu}$, which together implies that $\bar{\mathcal{S}}^1_{eu}$ and $\mathcal{S}^1_{eu}$ intersect perpendicularly. □

Of course from (7.8) we can see that $\bar{\mathcal{S}}^1_{eu} \subseteq \bar{\mathcal{S}}_{eu}$ and from the fact that $\left(\bar{\mathcal{S}}^1_{eu}\right)^{\perp} = \mathcal{S}_{eu} \cap (\mathcal{Y}^+)^{\perp} \cap (\mathcal{U}^+)^{\perp}$ also $\mathcal{S}^1_{eu} \subseteq \mathcal{S}_{eu}$, which implies that

$$\mathcal{U}^+ \subseteq \bar{\mathcal{S}}^1_{eu} \cap \mathcal{S}^1_{eu} \subseteq \mathcal{X} + \mathcal{U}^+.$$

Let us define the subspace $\mathcal{X}^1$ as follows:

$$\mathcal{X}^1 := \mathcal{X} \cap \left(\bar{\mathcal{S}}^1_{eu} \cap \mathcal{S}^1_{eu}\right) = \mathcal{X} \cap \mathcal{X}^1_{eu}. \tag{7.11}$$

We shall show that $\mathcal{X}^1$ is a minimal Markovian splitting subspace (clearly contained in $\mathcal{X}$). Before doing so we state the following technical lemma which will be used in the following:

LEMMA 7.5. *There holds*

$$\mathcal{X}^1_{eu} = \mathcal{X}^1 + \mathcal{U}^+ \tag{7.12}$$

PROOF. Since $\mathcal{X}^1 \subset \mathcal{X}^1_{eu}$ and $\mathcal{U}^+ \subset \mathcal{X}^1_{eu}$, clearly $\mathcal{X}^1 + \mathcal{U}^+ \subset \mathcal{X}^1_{eu}$. Conversely, since $\mathcal{X}^1_{eu} \subset \mathcal{X} + \mathcal{U}^+$, any $\xi \in \mathcal{X}^1_{eu}$ can be written as $\xi = \boldsymbol{x} + \boldsymbol{u}^+$ with $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{u}^+ \in \mathcal{U}^+$. However since $\mathcal{X}^1_{eu} \supset \mathcal{U}^+$, then $\boldsymbol{u}^+ \in \mathcal{X}^1_{eu}$ as well, and hence $\boldsymbol{x}$ belongs to both $\mathcal{X}^1_{eu}$ and $\mathcal{X}$, so that $\boldsymbol{x} \in \mathcal{X}^1$. □

THEOREM 7.5. *The pair of subspaces* $(\mathcal{S}^1, \bar{\mathcal{S}}^1)$ *defined as*

$$\begin{aligned} \mathcal{S}_1 &:= \mathcal{S} \cap \mathcal{S}^1_{eu} \\ \bar{\mathcal{S}}_1 &:= \bar{\mathcal{S}} \cap \bar{\mathcal{S}}^1_{eu} \end{aligned} \tag{7.13}$$

satisfy the conditions of theorem 6.1 and $\mathcal{X}^1 = \mathcal{S}^1 \cap \bar{\mathcal{S}}^1$*. Therefore* $\mathcal{X}^1$ *is oblique Markovian splitting subspace and is contained in* $\mathcal{X}$*.*

PROOF. First of all let us just note that since $\mathcal{X}^1 \subseteq \mathcal{X}$ then $\mathcal{X}^1 \subseteq \mathcal{S}$ and $\mathcal{X}^1 \subseteq \bar{\mathcal{S}}$, which implies that $\mathcal{X}^1 \subseteq \mathcal{S}^1$ and $\mathcal{X}^1 \subseteq \bar{\mathcal{S}}^1$. Moreover $\mathcal{P}^- \subseteq \mathcal{S}^1$, $\mathcal{Y}^+ \subseteq \bar{\mathcal{S}}^1$ and $\mathcal{U}^+ \cap \mathcal{S}^1 = \{0\}$. Clearly $\mathcal{X}^1 \subseteq \mathcal{S}^1 \cap \bar{\mathcal{S}}^1$, let us show the converse. Since, from Lemma 7.5,

$$\mathcal{S}^1_{eu} \cap \bar{\mathcal{S}}^1_{eu} = \mathcal{X}^1 + \mathcal{U}^+ \supseteq \left(\mathcal{S}^1 \cap \bar{\mathcal{S}}^1\right) + \mathcal{U}^+,$$

by the direct sum property we have that

$$\mathcal{S}^1 \cap \bar{\mathcal{S}}^1 \subseteq \mathcal{X}^1$$

and therefore

$$\mathcal{X}^1 = \mathcal{S}^1 \cap \bar{\mathcal{S}}^1 \tag{7.14}$$

Proving the shift invariance properties

$$\begin{cases} \sigma\bar{\mathcal{S}} \subseteq \bar{\mathcal{S}} \\ \sigma^*\mathcal{S} \subseteq \mathcal{S} \end{cases}$$

is just an easy check. The oblique intersection property is immediate from (7.10) and (7.14). □

We have seen a construction which allows us to reduce an oblique Markovian splitting subspace. We shall now prove that indeed this procedure yields a minimal one.

THEOREM 7.6. *The subspace* $\mathcal{X}^1$ *defined above is a minimal oblique Markovian splitting subspace.*

PROOF. The proof follows the same lines as in the stochastic case. Namely we assume that there exists an oblique Markovian splitting subspace, say $\mathcal{X}^o$, properly contained in $\mathcal{X}^1$. If such a subspace exists, then we could attach to it a pair of subspaces $(\mathcal{S}^o, \bar{\mathcal{S}}^o)$, which obviously satisfy $\mathcal{S}^o \subseteq \mathcal{S}^1$ and $\bar{\mathcal{S}}^o \subseteq \bar{\mathcal{S}}^1$. Then by the same argument we have already used we get that $\left(\bar{\mathcal{S}}^o_{eu}\right)^\perp \subseteq \mathcal{S}^o_{eu}$

and therefore $\mathcal{S}^o_{eu} \supseteq \left(\bar{\mathcal{S}}^o_{eu}\right)^\perp \vee \bar{\mathcal{P}}^- \vee \mathcal{U}^+$. But from the very definition of $\mathcal{S}^1_{eu}$ we have

$$\mathcal{S}^1_{eu} = \left(\bar{\mathcal{S}}^1_{eu}\right)^\perp \vee \bar{\mathcal{P}}^- \vee \mathcal{U}^+ \subseteq \mathcal{S}^o_{eu}$$

which implies that $\mathcal{S}^1_{eu} = \mathcal{S}^o_{eu}$. Moreover since $\mathcal{S}^1_{eu} \subseteq \mathcal{S}_{eu}$, $\bar{\mathcal{S}}^1_{eu} = \mathcal{S}^\perp_{eu} \vee \mathcal{Y}^+ \vee \mathcal{U}^+ \subseteq \bar{\mathcal{S}}^o_{eu}$ and therefore $\bar{\mathcal{S}}^1_{eu} = \bar{\mathcal{S}}^o_{eu}$ which guarantees that $\mathcal{S}^1 = \mathcal{S}^o$ and $\bar{\mathcal{S}}^1 = \bar{\mathcal{S}}^o$ and therefore $\mathcal{X}^o = \mathcal{X}^1$. □

7.2. Splitting property and Hankel Operators in the Absence of Feedback. In this section we shall study observability and constructibility of an oblique Markovian splitting subspace in the absence of feedback. The state space property of $\mathcal{X}$ will be interpreted in terms of factorization of a certain Hankel operator defined on the data space.

We first state, without proof, a lemma which provides a representation of the observability and constructibility operators in the absence of feedback.

LEMMA 7.6. *Assume there is non feedback from* $\boldsymbol{y}$ *to* $\boldsymbol{u}$. *Then*

$$\mathbb{O}^* \lambda = E_{\|\mathcal{U}^+}\left[\lambda \mid \mathcal{X}\right], \quad \lambda \in \mathcal{Y}^+ \tag{7.15}$$

and

$$\mathbb{K}\xi = E\left[\xi \mid \mathcal{P}^-\right], \quad \xi \in \mathcal{X}. \tag{7.16}$$

Let us consider also the *Hankel operator* $\mathbb{H} : \mathcal{Y}^+ \to \mathcal{P}^-$ defined as

$$\mathbb{H}\lambda := E_{\|\mathcal{U}^+}\left[\lambda \mid \mathcal{P}^-\right] \quad , \quad \lambda \in \mathcal{Y}^+.$$

PROPOSITION 7.4. *The splitting property of* $\mathcal{X}$ *is equivalent to the factorization* $\mathbb{H} = \mathbb{K}\mathbb{O}^*$.

PROOF. For every $\lambda \in \mathcal{Y}^+$,

$$E\left[\lambda \mid \mathcal{P}^- + \mathcal{U}^+\right] = E\left[E\left[\lambda \mid \left(\mathcal{P}^- + \mathcal{X}\right) + \mathcal{U}^+\right] \mid \mathcal{P}^- + \mathcal{U}^+\right]$$

and by the splitting property,

$$E\left[\lambda \mid \mathcal{P}^- + \mathcal{U}^+\right] = E\left[E\left[\lambda \mid \mathcal{X} + \mathcal{U}^+\right] \mid \mathcal{P}^- + \mathcal{U}^+\right]$$

which implies that

$$E_{\|\mathcal{U}^+}\left[\lambda \mid \mathcal{P}^-\right] = E_{\|\mathcal{U}^+}\left[E\left[\lambda \mid \mathcal{X} + \mathcal{U}^+\right] \mid \mathcal{P}^-\right]$$

and since $\mathcal{X} \cap \mathcal{U}^+ = 0$ in the feedback-free situation, we have

$$E\left[\lambda \mid \mathcal{X} + \mathcal{U}^+\right] = E_{\|\mathcal{U}^+}\left[\lambda \mid \mathcal{X}\right] + E_{\|\mathcal{X}}\left[\lambda \mid \mathcal{U}^+\right]$$

which implies that

$$E_{\|\mathcal{U}^+}\left[\lambda \mid \mathcal{P}^-\right] = E\left[E_{\|\mathcal{U}^+}\left[\lambda \mid \mathcal{X}\right] \mid \mathcal{P}^-\right]$$

This is the factorization $\mathbb{H} = \mathbb{K}\mathbb{O}^*$. □

As usual we say that this factorization is *canonical* if $\mathbb{O}^*$ has dense range and if $\mathbb{K}$ is injective. Using the well known relations, valid for every bounded linear operator [2]

$$\mathcal{X} = \overline{\text{Range } \ \mathbb{O}^*} \oplus \text{Ker } \ \mathbb{O} \qquad \mathcal{X} = \overline{\text{Range } \ \mathbb{K}^*} \oplus \text{Ker } \ \mathbb{K}$$

we see that

$$\overline{\text{Range } \ \mathbb{O}^*} = E_{\|\mathcal{U}^+}\left[\mathcal{Y}^+ \mid \mathcal{X}\right] \quad , \quad \text{Ker } \ \mathbb{O} = \mathcal{X} \cap \left(E_{\|\mathcal{U}^+}\left[\mathcal{Y}^+ \mid \mathcal{X}\right]\right)^{\perp}$$

and

$$\overline{\text{Range } \mathbb{K}^*} = E\left[\mathcal{P}^- \mid \mathcal{X}\right] \quad , \quad \text{Ker } \ \mathbb{K} = \mathcal{X} \cap \left(\mathcal{P}^-\right)^{\perp}$$

Therefore we shall call $\overline{\text{Range } \ \mathbb{O}^*}$ the *observable component of* $\mathcal{X}$ and its orthogonal complement Ker $\mathbb{O}$ *the unobservable subspace*. Similarly, $\overline{\text{Range } \ \mathbb{K}^*}$ is the *constructible component* of the state and Ker $\mathbb{K} = \mathcal{X} \cap (\mathcal{P}^-)^{\perp}$ is the *unconstructible* subspace.

As we can see, in stochastic realization with inputs one is led to consider a "mixture" of the concepts of constructibility and reachability[3]. Let us look at the expression for the unconstructible component. Since by the feedback free property $\mathcal{P}^- = \mathcal{Y}_s^- \oplus \mathcal{U}^-$ we obtain:

$$\text{Ker } \ \mathbb{K} = \left[\mathcal{X} \cap (\mathcal{Y}_s^-)^{\perp}\right] \cap \left[\mathcal{X} \cap (\mathcal{U}^-)^{\perp}\right]. \tag{7.17}$$

We anticipate here that this concept of constructibility, which, as we shall see later, is the one which is linked to minimality (in the sense of subspace inclusion) does not in general imply constructibility of the stochastic component and hence the condition is not strong enough to characterize stochastic minimality.

In order to get a deeper understanding of the situation, let us define the restricted operators

$$\mathbb{K}_r^* := \mathbb{K}^*_{|\mathcal{Y}_s^-} \tag{7.18}$$

and

$$\mathbb{R}^* := \mathbb{K}^*_{|\mathcal{U}^-} \tag{7.19}$$

the former being related to the "stochastic" component and the latter to the "deterministic" component.

Let us define

$$\mathcal{X}_d := E\left[\mathcal{X} \mid \mathcal{U}^-\right] = \mathbb{R}\,\mathcal{X}, \quad , \quad \mathcal{X}_s := E\left[\mathcal{X} \mid \mathcal{W}^-\right] = \mathbb{R}_w\,\mathcal{X}, \tag{7.20}$$

[2]Note that the norm of the oblique projection satisfies $\|\mathbb{O}^*\|^2 = \left(1 - \sigma_{max}^2\right)^{-1}$ where σ_{max} is the maximum canonical correlation coefficient between $\mathcal{U}^+$ and $\mathcal{X}$, which is strictly less than 1 by the zero intersection property $\mathcal{X} \cap \mathcal{U}^+ = \{0\}$.

[3]This is the reason why in the beginning of the section we have used the word "constructibility" between quotes.

where $\mathbb{R}_w$ is the "stochastic" reachability operator

$$\begin{array}{rcccl} \mathbb{R}_w & : & \mathcal{X} & \rightarrow & \mathcal{W}^- \\ & & \boldsymbol{x} & \rightarrow & E\left[\boldsymbol{x} \mid \mathcal{W}^-\right] \end{array}$$

Recall that $\mathcal{S} = \mathcal{U}^- \oplus \mathcal{W}^-$, which implies that $\mathcal{X}_d \perp \mathcal{X}_s$. Let us also note that $\mathcal{X} \subseteq \mathcal{X}_s \oplus \mathcal{X}_d$. From

$$E\left[\boldsymbol{x} \mid \mathcal{Y}_s^-\right] = E\left[E\left[\boldsymbol{x} \mid \mathcal{W}^-\right] \mid \mathcal{Y}_s^-\right] \quad , \quad \forall \boldsymbol{x} \in \mathcal{X}$$

the restricted constructibility operator $\mathbb{K}_r$ can be factorized as follows:

$$\mathbb{K}_r = \mathbb{K}_s \mathbb{R}_w \tag{7.21}$$

where $\mathbb{K}_s$ is the usual "stochastic" constructibility operator.

In general neither $\mathbb{K}_r^*$ nor $\mathbb{R}^*$ will have dense range, or equivalently, neither $\mathbb{K}_r$ nor $\mathbb{R}$ will have a trivial kernel. The meaning of the constructibility condition for the state space, is that the intersection of the two kernels must be the zero random variable.

On the other hand, since all processes involved are assumed to be p.n.d., the joint system is reachable, (recall that $\mathcal{X} \subset \mathcal{S}$),[4] and therefore Ker $\mathbb{R}_w \cap$ Ker $\mathbb{R} = \{0\}$.

The following fact clarifies the link between constructibility of the joint model and constructibility of the "stochastic" component.

PROPOSITION 7.5. *If the "stochastic component" is constructible, i.e.* Ker $\mathbb{K}_s = \{0\}$*, then the joint model is so. Geometrically this condition reads as*

$$\mathcal{X}_s \cap \left(\mathcal{Y}_s^-\right)^\perp = \{0\}.$$

PROOF. This is immediate since Ker $\mathbb{K} =$ Ker $\mathbb{K}_r \cap$ Ker $\mathbb{R}$. Therefore, if $\boldsymbol{x} \in$ Ker $\mathbb{K}$ then, $\boldsymbol{x} \in$ Ker $\mathbb{R}$, $\boldsymbol{x} \notin$ Ker $\mathbb{R}_w$ and hence $\mathbb{R}_w \boldsymbol{x} \in$ Ker $\mathbb{K}_s$ form (7.21). □

REMARK 7.7. Note that in general Ker $\mathbb{K} = \{0\}$ does not imply Ker $\mathbb{K}_s = \{0\}$. In fact, assume $\boldsymbol{x} = \boldsymbol{x}_s + \boldsymbol{x}_d$. It might well happen that $\mathbb{K}_r \boldsymbol{x} = \mathbb{K}_s \mathbb{R}_w \boldsymbol{x} = \mathbb{K}_s \mathbb{R}_w \boldsymbol{x}_s = 0$ while $\mathbb{R}\boldsymbol{x} = \mathbb{R}\boldsymbol{x}_d \neq 0$. Geometrically

$$\mathcal{X} \cap \left(\mathcal{P}^-\right)^\perp = \{0\}$$

does not imply

$$\mathcal{X}_s \cap \left(\mathcal{Y}_s^-\right)^\perp = \{0\}.$$

In fact we might have $\boldsymbol{x}_s \in (\mathcal{Y}_s^-)^\perp$, without $\boldsymbol{x} \in (\mathcal{P}^-)^\perp$.

In some sense this shows that the definition we have given of "minimality" is not quite complete. Since however the concept of minimality is historically linked to "dimension", or more generally, to inclusion in the infinite dimensional case, we shall introduce a further definition.

[4] Here we assume that there are no p.n.d. components, see [**12**] for a discussion on this topic.

DEFINITION 7.8. An oblique Markovian splitting subspace $\mathcal{X}$ is *strongly minimal* if it is minimal and $\mathcal{X}_s$ defined in (7.20) is constructible, i.e. Ker $\mathbb{K}_s = \{0\}$, or, equivalently,

$$\mathcal{X}_s \cap \left(\mathcal{Y}_s^-\right)^\perp = \{0\}$$

It is apparent that minimality and strong minimality are equivalent in the causal case, as the following proposition states.

PROPOSITION 7.6. *Let $\mathcal{X}$ be a minimal causal oblique Markovian splitting subspace, then it is strongly minimal.*

PROOF. In the causal case $\mathcal{W}^-$ is the space spanned by the past innovation $\mathcal{E}^-$, which is nothing but $\mathcal{Y}_s^-$; therefore $\mathcal{X}_s \cap (\mathcal{Y}_s^-)^\perp = \{0\}$. □

Using decomposition (7.20) one can define the observability operators

$$\mathbb{O}_s^* := E_{\|\mathcal{U}^+}\left[\mathcal{Y}^+ \mid \mathcal{X}_s\right] = E\left[\mathcal{Y}^+ \mid \mathcal{X}_s\right] = E\left[\mathcal{Y}_s^+ \mid \mathcal{X}_s\right] \tag{7.22}$$

and

$$\mathbb{O}_d^* := E_{\|\mathcal{U}^+}\left[\mathcal{Y}^+ \mid \mathcal{X}_d\right] = E_{\|\mathcal{U}^+}\left[\mathcal{Y}_d^+ \mid \mathcal{X}_d\right]. \tag{7.23}$$

It is easy to see that the following factorizations hold

$$\mathbb{O}_s^* = \mathbb{R}_w \mathbb{O}^*$$

and

$$\mathbb{O}_d^* = \mathbb{R}\mathbb{O}^*.$$

Note that the observability conditions for the “deterministic” and “stochastic” component, i.e. Ker $\mathbb{O}_d = \{0\}$ and Ker $\mathbb{O}_s = \{0\}$ do not in general imply that Ker $\mathbb{O} = \{0\}$, while the converse is always true since $(\mathbb{R}_w)_{|\mathcal{X}_s}$ and $\mathbb{R}_{|\mathcal{X}_d}$ have trivial kernel by construction.

The geometric characterizations of minimality given above does not address the question of strong minimality. This has clearly to do only with the “stochastic” component and therefore it can be expressed geometrically in the usual way as

$$\mathcal{Y}_s^- \vee \mathcal{X}_s^- = \left(\mathcal{Y}_s^+ \vee \mathcal{X}_s^+\right)^\perp \vee \mathcal{Y}_s^-.$$

where orthogonal complement is taken in $H(\boldsymbol{w})$.

THEOREM 7.9. *Let $\mathcal{X}$ be an oblique Markovian splitting subspace and let $\mathcal{S}_{eu} = \mathcal{S} + \mathcal{U}^+$, $\bar{\mathcal{S}}_{eu} = \bar{\mathcal{S}} \vee \mathcal{U}^+$. The following conditions are equivalent:*

- i) *$\mathcal{X}$ is strongly minimal*
- ii) *$\mathcal{X}$ is minimal and $\mathbb{K}_s$ is injective*
- iii) *$\bar{\mathcal{S}}_{eu} = \mathcal{S}_{eu}^\perp \vee \mathcal{Y}^+ \vee \mathcal{U}^+$ and $\mathcal{Y}_s^- \vee \mathcal{X}_s^- = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{Y}_s^-$*
- iv) *$\bar{\mathcal{S}}_{eu} = \mathcal{S}_{eu}^\perp \vee \mathcal{Y}^+ \vee \mathcal{U}^+$ and $\mathcal{S}_{eu} = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{P}^- \vee \mathcal{U}^+$*

PROOF. *i)* and *ii)* are equivalent by definition. The fact that $\mathcal{X}$ is minimal implies that it is observable, i.e. $\bar{\mathcal{S}}_{eu} = \mathcal{S}_{eu}^\perp \vee \mathcal{Y}^+ \vee \mathcal{U}^+$, and $\mathbb{K}_s$ injective is equivalent to $\mathcal{Y}_s^- \vee \mathcal{X}_s^- = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{Y}_s^-$, from which condition *iii)*. To show that *iii)* implies *iv)* just note that since $\mathcal{Y}_s^- \vee \mathcal{X}_s^- \vee \mathcal{U} = \mathcal{S}_{eu}$,

$\mathcal{Y}_s^- \vee \mathcal{X}_s^- = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{Y}_s^-$ implies that $\mathcal{S}_{eu} = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{Y}_s^- \vee \mathcal{U} = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{P}^- \vee \mathcal{U}^+$. Conversely, if *iv*) holds, $\mathcal{S}_{eu} \ominus \mathcal{U} = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{P}^- \vee \mathcal{U}^+ \ominus \mathcal{U}$, i.e. $\mathcal{Y}_s^- \vee \mathcal{X}_s^- = [H(\boldsymbol{w}) \ominus (\mathcal{Y}_s^+ \vee \mathcal{X}_s^+)] \vee \mathcal{Y}_s^-$, which concludes the proof. $\square$

8. Reconciliation with Stochastic Realization Theory

So far we have studied state space construction in the presence of exogenous inputs, based on the concept of *Oblique Markovian splitting subspace*. In this section we shall examine the relation between oblique splitting and the classical construction of stochastic realization theory based on (orthogonal) splitting. We shall show that, in the absence of feedback, stochastic realizations with inputs can be constructed directly from stochastic realizations of the joint input-output process.

Let $(\mathcal{S}_J, \bar{\mathcal{S}}_J)$ be an orthogonal scattering pair for the joint process $\begin{bmatrix}\boldsymbol{y}^\top & \boldsymbol{u}^\top\end{bmatrix}^\top$, where the subscript J stands for joint, and let us assume that there is no feedback form $\boldsymbol{y}$ to $\boldsymbol{u}$. The following technical lemmas will be useful.

Lemma 8.1. *Let $\begin{bmatrix}\boldsymbol{y}^\top & \boldsymbol{u}^\top\end{bmatrix}^\top$ be a stationary process, and assume that there is no feedback form $\boldsymbol{y}$ to $\boldsymbol{u}$. Then there exist joint Markovian splitting subspaces $\mathcal{X}_J \equiv (\mathcal{S}_J, \bar{\mathcal{S}}_J)$ such that*

$$\mathcal{S}_J \perp \mathcal{U}^+ \mid \mathcal{U}^- \tag{8.1}$$

Proof. We just need to show that there exists at least one. Let us consider any causal realization, and let $\mathcal{X}_J$ be its state space. It is clear that by causality $\mathcal{X}_J \subset \mathcal{Y}^- \vee \mathcal{U}^-$, which implies that $\mathcal{S}_J = \mathcal{Y}^- \vee \mathcal{U}^-$ and therefore (8.1) follows. $\square$

Definition 8.1. In the sequel we shall say that joint Markovian splitting subspaces (realizations) which satisfy (8.1) are *feedback free.*

Lemma 8.2. *Let $\mathcal{X}_J \equiv (\mathcal{S}_J, \bar{\mathcal{S}}_J)$ be feedback free, then*

$$\mathcal{S}_J \cap \mathcal{U}^+ = \{0\}.$$

Proof. From (8.1) $\mathrm{E}\left[\mathcal{S}_J \mid (\mathcal{U}^-)^\perp\right] \perp \mathcal{U}$, so that any element $\boldsymbol{s} \in \mathcal{S}_J$ can be uniquely decomposed as $\boldsymbol{s} = \hat{\boldsymbol{s}} + \tilde{\boldsymbol{s}}$ where $\hat{\boldsymbol{s}} := E\left[\boldsymbol{s} \mid \mathcal{U}^-\right] \in \mathcal{U}^-$ and $\tilde{\boldsymbol{s}} \perp \mathcal{U}$. Let us assume that $\boldsymbol{s} \in \mathcal{U}^+$; it follows that $\hat{\boldsymbol{s}} = E\left[E\left[\hat{\boldsymbol{s}} \mid \mathcal{U}^+\right] \mid \mathcal{U}^-\right]$ and therefore $\hat{\boldsymbol{s}} \in \mathcal{U}^+ \cap \mathcal{U}^-$ which, recalling the sufficiently rich assumption (3.6), implies $\hat{\boldsymbol{s}} = 0$ and therefore $\tilde{\boldsymbol{s}} \in \mathcal{U}^+$; hence, $\tilde{\boldsymbol{s}} = 0$ and therefore $\boldsymbol{s} = 0$, which concludes the proof. $\square$

We are now ready to state the following result.

Theorem 8.2. *Let $\mathcal{X}_J := (\mathcal{S}_J, \bar{\mathcal{S}}_J)$ be a feedback free realization of the stationary process $\begin{bmatrix}\boldsymbol{y}^\top & \boldsymbol{u}^\top\end{bmatrix}^\top$. Then $\mathcal{X}_J$ is an oblique Markovian splitting subspace.*

PROOF. We just need to verify the conditions of Theorem 6.1; $\mathcal{Y}^+ \subseteq \bar{\mathcal{S}}_J$ and $\mathcal{P}^- \subseteq \mathcal{S}_J$ by construction; the fact that $\mathcal{S}_J \cap \mathcal{U}^+ = 0$ follows from Lemma 8.2; forward and backward shift invariance follow from the fact that $(\mathcal{S}_J, \bar{\mathcal{S}}_J)$ is a scattering pair and the oblique intersection property holds since, in particular, also

$$\mathcal{S}_J \perp \bar{\mathcal{S}}_J \mid \mathcal{S}_J \cap \bar{\mathcal{S}}_J$$

holds. □

However, as one may expect, $\mathcal{X}_J$ is not, in general, a minimal oblique Markovian splitting subspace.

In order to construct a minimal oblique Markovian splitting subspace we can follow the procedure described in the previous section.

Let us assume that $\mathcal{X}_J := \mathcal{S}_J \cap \bar{\mathcal{S}}_J$ is a minimal Markovian splitting subspace for the joint process. The reason for $\mathcal{X}_J$ being not minimal oblique splitting, is that it includes the dynamics of the input process $\boldsymbol{u}$. We want to "factor out" this dynamics.

The basic idea in the reduction process is to consider first an "extended" state space

$$\mathcal{X}_{Je} := \mathcal{X}_J + \mathcal{U}^+$$

together with the associated "extended" pair $(\mathcal{S}_{Je}, \bar{\mathcal{S}}_{Je})$, $\mathcal{S}_{Je} := \mathcal{S}_J + \mathcal{U}^+$, $\bar{\mathcal{S}}_{Je} := \bar{\mathcal{S}}_J \vee \mathcal{U}^+$. We proceed to reduce this subspace, subject to the constraint that it must always contain $\mathcal{U}^+$.

Denote, as in the previous section, by $(\mathcal{S}^1_{Je}, \bar{\mathcal{S}}^1_{Je})$ the reduced pair, let $\mathcal{X}^1_{Je} := \mathcal{S}^1_{Je} \cap \bar{\mathcal{S}}^1_{Je}$ and let

$$\mathcal{X}^1 := \mathcal{X}_J \cap \left(\mathcal{S}^1_{Je} \cap \bar{\mathcal{S}}^1_{Je}\right) = \mathcal{X}_J \cap \mathcal{X}^1_{Je}$$

Introduce the generating process, $\mathcal{W}_t$, for $\mathcal{S}^1_{Je}$, as

$$\left(\mathcal{S}^1_{Je}\right)_{t+1} = \left(\mathcal{S}^1_{Je}\right)_t \oplus \mathcal{W}_t.$$

Of course the subspaces $\{\mathcal{W}_t\}$ are pairwise orthogonal, i.e. $\mathcal{W}_t \perp \mathcal{W}_s \quad t \neq s$.

Since $\mathcal{X}^1_{Je}$ is (orthogonally) Markovian splitting, the usual state update equation in geometric form holds

$$\left(\mathcal{X}^1_{Je}\right)_{t+1} \subseteq \left(\mathcal{X}^1_{Je}\right)_t \oplus \mathcal{W}_t.$$

Using Lemma 7.5, the last equation can be written as

$$\mathcal{X}^1_{t+1} + \mathcal{U}^+_{t+1} \subseteq \left(\mathcal{X}^1_t + \mathcal{U}^+_t\right) \oplus \mathcal{W}_t. \tag{8.2}$$

THEOREM 8.3. *The subspace $\mathcal{X}^1$ is a minimal oblique splitting subspace. In fact we have*

$$\begin{cases} \mathcal{X}^1_{t+1} & \subseteq & \left(\mathcal{X}^1_t + \mathcal{U}_t\right) \oplus \mathcal{W}_t \\ \mathcal{Y}_t & \subseteq & \left(\mathcal{X}^1_t + \mathcal{U}_t\right) \oplus \mathcal{W}_t. \end{cases}$$

PROOF. We show that (8.2) is equivalent to

$$\begin{cases} \mathcal{X}^1_{t+1} & \subseteq & \left(\mathcal{X}^1_t + \mathcal{U}_t\right) \oplus \mathcal{W}_t \\ \mathcal{U}^+_{t+1} & \subseteq & \mathcal{U}^+_t. \end{cases}$$

In fact, if this was not the case, there would be elements $\boldsymbol{x}^1_{t+1} \in \mathcal{X}^1_{t+1} \subseteq (\mathcal{S}_J)_{t+1}$ which could be written as $\boldsymbol{x}^1_{t+1} = (\boldsymbol{x}^1_t + \boldsymbol{u}_t + \boldsymbol{w}_t) + \boldsymbol{u}^+_{t+1}$ where $(\boldsymbol{x}^1_t + \boldsymbol{u}_t + \boldsymbol{w}_t) \in (\mathcal{S}_J)_{t+1}$, $\boldsymbol{u}^+_{t+1} \in \mathcal{U}^+_{t+1}$, which would imply that $\boldsymbol{u}^+_{t+1} = \boldsymbol{x}^1_{t+1} - (\boldsymbol{x}^1_t + \boldsymbol{u}_t + \boldsymbol{w}_t) \in (\mathcal{S}_J)_{t+1}$ contradicting the fact that $(\mathcal{S}_J)_{t+1} \cap \mathcal{U}^+_{t+1} = \{0\}$. Similarly, since $\mathcal{Y}_t \subseteq (\mathcal{S}_J)_{t+1}$ it follows that

$$\begin{cases} \mathcal{Y}_t & \subseteq & \left(\mathcal{X}^1_t + \mathcal{U}_t\right) \oplus \mathcal{W}_t \\ \mathcal{U}_t & \subseteq & \mathcal{U}^+_t . \end{cases}$$

Since $\boldsymbol{u}$ is known, we can "drop" the second part of the state equations, i.e. the component lying on $\mathcal{U}^+$ and end up with the equation state din the Theorem. This equation obcviously implies the oblique splitting property of $\mathcal{X}^1$.

Note that, by construction, $\mathcal{X}^1 + \mathcal{U}^+$ is the minimal splitting subspace containing the future of the input process, which, as it turns out, corresponds to the fact that $\mathcal{X}^1$ is the minimal subspace of $\mathcal{X}_J$ which is oblique splitting. □

It is natural to ask when a minimal (feedback free) Markovian splitting subspace for the joint process is a minimal oblique Markovian splitting subspace. It turns out that this is true if and only if the input predictor space is contained in the oblique Markovian splitting state.

Theorem 8.4. *Let $\mathcal{X}_J$ be a (feedback free) minimal Markovian splitting subspace for the joint process $\begin{bmatrix}\boldsymbol{y}^\top & \boldsymbol{u}^\top\end{bmatrix}^\top$ and let $\mathcal{X}^{+/-}_u := E\left[\mathcal{U}^+ \mid \mathcal{U}^-\right]$ be the predictor space of the input process. Then $\mathcal{X}^1 = \mathcal{X}_J$ if and only if $\mathcal{X}^{+/-}_u \subseteq \mathcal{X}^1$.*

Proof. Since $\mathcal{X}_J$ is minimal it is constructible. Therefore, $\mathcal{S}_J = \bar{\mathcal{S}}^\perp_J \vee \mathcal{U}^- \vee \mathcal{Y}^-$, which implies that $\mathcal{S}_J$ is minimal. By observability we have that

$$E\left[\mathcal{U}^+ \vee \mathcal{Y}^+ \mid \mathcal{S}_J\right] = \mathcal{X}_J. \tag{8.3}$$

On the other hand

$$E\left[\mathcal{U}^+ \vee \mathcal{Y}^+ \mid \mathcal{S}_J + \mathcal{U}^+\right] = \mathcal{X}^1 + \mathcal{U}^+. \tag{8.4}$$

Equation (8.3) can be rewritten as

$$\begin{aligned} E\left[E\left[\mathcal{U}^+ \vee \mathcal{Y}^+ \mid \mathcal{S}_J + \mathcal{U}^+\right] \mid \mathcal{S}_J\right] &= E\left[\mathcal{X}^1 + \mathcal{U}^+ \mid \mathcal{S}_J\right] = \\ &= \mathcal{X}^1 \vee E\left[\mathcal{U}^+ \mid \mathcal{S}_J\right] = \\ &= \mathcal{X}^1 \vee E\left[\mathcal{U}^+ \mid \mathcal{U}^-\right] = \\ &= \mathcal{X}^1 \vee \mathcal{X}^{+/-}_u . \end{aligned} \tag{8.5}$$

Therefore $\mathcal{X}_J = \mathcal{X}^1 \vee \mathcal{X}^{+/-}_u$ which implies that $\mathcal{X}_J = \mathcal{X}^1$ if and only if $\mathcal{X}^{+/-}_u \subseteq \mathcal{X}^1$. □

It is natural to ask what kind of situations may lead to such degeneracy. In order to address this problem we shall make a finite dimensionality assumption and work with the spectral representations of these spaces. We shall have to refer the reader to [**23**] for details.

Let $d\hat{\boldsymbol{z}} := \begin{bmatrix} d\hat{\boldsymbol{u}}^\top & d\hat{\boldsymbol{w}}^\top \end{bmatrix}^\top$ be the spectral measure (i.e. the Fourier transform [**26**]) of the joint stationary process $\begin{bmatrix}\boldsymbol{u}^\top(t) \ \boldsymbol{w}^\top(t)\end{bmatrix}^\top$. Let also (A, B, C, D, K)

be a minimal (oblique) realization of $\boldsymbol{y}$ with state space $\mathcal{X}^1$ and let (A_u, K_u, C_u, I) be a minimal innovation representation (with state space $\mathcal{X}_u^{+/-}$) of $\boldsymbol{u}$. The spectral representations, $\hat{\mathcal{X}}^1$, and $\hat{\mathcal{X}}_u^{+/-}$, of $\mathcal{X}^1$ and $\mathcal{X}_u^{+/-}$ with respect to the spectral measure $d\hat{\boldsymbol{z}}$ are given by:

$$\hat{\mathcal{X}}^1 = \overline{\text{row-span}} \left\{ \begin{bmatrix} (zI - A)^{-1} B & \quad (zI - A)^{-1} K \end{bmatrix} \right\} \tag{8.6}$$

and

$$\hat{\mathcal{X}}_u^{+/-} = \overline{\text{row-span}} \left\{ \begin{bmatrix} (zI - (A_u - K_u C_u))^{-1} K_u & \quad 0 \end{bmatrix} \right\}. \tag{8.7}$$

We are now able to give precise conditions for $\mathcal{X}^1$ and $\mathcal{X}_J$ to be the same space.

PROPOSITION 8.1. *Let (A, B, C, D, K) be a minimal (oblique) realization of $\boldsymbol{y}$ with state space $\mathcal{X}^1$ and let (A_u, K_u, C_u, I) be a minimal innovation representation of $\boldsymbol{u}$ (with state space $\mathcal{X}_u^{+/-}$). Then $\mathcal{X}^1 = \mathcal{X}_J$ if and only if there exist a nonsingular change of basis T such that:*

$$TAT^{-1} = \begin{bmatrix} A_u - K_u C_u & 0 \\ * & * \end{bmatrix}$$

and

$$TB = \begin{bmatrix} K_u \\ * \end{bmatrix} \quad , \quad TK = \begin{bmatrix} 0 \\ * \end{bmatrix}$$

The proof of this Proposition is a bit lengthy and will be given in another publication.

9. Conclusions

In this paper we have presented the basic ideas for a comprehensive theory of stochastic realization in the presence of exogenous inputs. The central concept of *Oblique Markovian Splitting Subspace* leads in principle to state-space construction and to a coordinate-free analysis of stochastic models with inputs. Most of the ideas are applicable to the general case where feedback is present, however there seem to be still some gaps to be filled, in particular we need to understand better how to deal with mixed causality structures as they occur in feedback interconnections where $F(z)$ may be unstable. Some new idea and additional work are needed.

Bibliography

[1] H. Akaike. Stochastic theory of minimal realization. *IEEE Trans. Automat. Contr.*, 19(6):667–674, 1974.

[2] H. Akaike. Markovian representation of stochastic processes by canonical variables. *SIAM J. Control*, 13:162–173, 1975.

[3] P.E. Caines and C.W. Chan. Estimation, identification and feedback. In R. Mehra and D. Lainiotis, editors, *System Identification: Advances and Case Studies*, pages 349–405. Academic, 1976.

[4] A. Chiuso and G. Picci. On the ill-conditioning of subspace identification with inputs. Tech. Report TRITA/MATH-01-OS5, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 2001. submitted for pubblication.

[5] H. Dym and H.P. McKean. *Gaussian Processes, Function Theory and the Inverse Spectral Problem.* Academic Press, New York, 1976.

[6] M.R. Gevers and B.D.O. Anderson. Representation of jointly stationary feedback free processes. *Intern. Journal of Control*, 33:777–809, 1981.

[7] M.R. Gevers and B.D.O. Anderson. On jointly stationary feedback-free stochastic processes. *IEEE Trans. Automat. Contr.*, 27:431–436, 1982.

[8] C.W.J. Granger. Economic processes involving feedback. *Information and Control*, 6:28–48, 1963.

[9] E.J. Hannan and D.S. Poskitt. Unit canonical correlations between future and past. *The Annals of Statistics*, 16:784–790, 1988.

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

[11] W.E. Larimore. System identification, reduced-order filtering and modeling via canonical variate analysis. In *Proc. American Control Conference*, pages 445–451, 1983.

[12] A. Lindquist and G. Picci. *Linear Stochastic Systems.* ?, ?

[13] A. Lindquist and G. Picci. On the stochastic realization problem. *SIAM Journal on Control and Optimization*, 17(3):365–389, 1979.

[14] A. Lindquist and G. Picci. Realization theory for multivariate stationary gaussian processes. *SIAM Journal on Control and Optimization*, 23(6):809–857, 1985.

[15] A. Lindquist and G. Picci. A geometric approach to modelling and estimation of linear stochastic systems. *Journal of Mathematical Systems, Estimation and Control*, 1:241–333, 1991.

[16] A. Lindquist and G. Picci. Canonical correlation analysis approximate covariance extension and identification of stationary time series. *Automatica*, 32:709–733, 1996.

[17] A. Lindquist, G. Picci, and G. Ruckebush. On minimal splitting subspaces and markovian representation. *Mathematical Systems Theory*, 12:271–279, 1979.

[18] P. Van Overschee and B. De Moor. Subspace algorithms for the stochastic identification problem. *Automatica*, 29:649–660, 1993.

[19] P. Van Overschee and B. De Moor. N4SID: Subspace algorithms for the identification of combined deterministic– stochastic systems. *Automatica*, 30:75–93, 1994.

[20] P.D.Lax and R.S.Phillips. *Scattering Theory.* Academic Press, NewYork, 1967.

[21] G. Picci. Stochastic realization of gaussian processes. *Proc. of the IEEE*, 64:112–122, 1976.

[22] G. Picci. Geometric methods in stochastic realization and system identification. In *CWI Quarterly special Issue on System Theory*, volume 9, pages 205–240, 1996.

[23] G. Picci. Oblique splitting susbspaces and stochastic realization with inputs. In D. Prätzel-Wolters U. Helmke and E. Zerz, editors, *Operators, Systems and Linear Algebra*, pages 157–174, Stuttgart, 1997. Teubner,.

[24] G. Picci. Stochastic realization and system identification. In T. Katayama and I. Sugimoto, editors, *Statistical Methods in Control and Signal Processing*, pages 205–240, N.Y., 1997. M. Dekker.

[25] G. Picci and S. Pinzoni. Acausal models and balanced realizations of stationary processes. *Linear Algebra and its Applications*, 205-206:957–1003, 1994.

[26] Y. A. Rozanov. *Stationary Random Processes*. Holden-Day, San Francisco, 1967.

[27] M. Verhaegen. Identification of the deterministic part of mimo state space models given in innovations form from input-output data. *Automatica*, 30:61–74, 1994.