On minimal LPV state-space representations in innovation form: an algebraic characterization

Elie Rouphael, Mihaly Petreczky, Lotfi Belkoura

Abstract—In this paper a definition of the concept of minimal state-space representations in innovation form for LPV is proposed. We also present algebraic conditions for a stochastic LPV state-space representation to be minimal in forward innovation form and discuss an algorithm for transforming any stochastic LPV state-space representation to a minimal one in innovation form.

I. INTRODUCTION

Identification of Linear Parameter-Varying (LPV) models has gained significant attention, see [2], [12], [15], [22], [9], [24], [26], [28], [25] and the references therein. In particular, there is a rich literature on subspace identification of LPV state-space representations, see for instance [26], [9], [27], [28], [6], [7] and the references therein.

Despite these advances, the theoretical analysis of system identification algorithms, especially subspace methods, for stochastic LPV state-space representations remains challenging. As the history of LTI system identification indicates [13], [10], [4] that such a theoretical analysis requires a good understanding of the notion of minimal state-space representations in innovation form [13], [10]. The latter notion is not yet fully understood for LPV state-space representations.

Contribution: In this paper LPV state-space representations with affine dependence on parameters (abbreviated by LPV-SSA) are considered. We restrict attention to autonomous (without control input) stationary stochastic LPV-SSAs (asLPV-SSA for short). The main technical contributions are new algebraic conditions for an asLPV-SSA to be minimal in innovation form. These conditions depend only on the matrices of the system representation. In order to present these results, the paper also provides a systematic overview of the results on existence and minimality of asLPV-SSA in innovation form. This overview can be derived from realization theory of stochastic bilinear systems [20], but they have not been stated explicitly for LPV-SSAs. In particular, we state that any asLPV-SSA can be converted to a minimal asLPV-SSA in innovation form while preserving the output. Moreover, any two minimal asLPV-SSA in innovation form are isomorphic, if they have the same output.

Motivation for studying asLPV-SSA: Under suitable technical assumptions, any stochastic LPV-SSA can be decomposed into a noiseless deterministic LPV-SSA which is driven only by the control input, and an LPV-SSA driven by the noise, [14]. Moreover, the identification of these two subsystems can be carried out separately [14]. Note that minimality and uniqueness of deterministic LPV state-space representations is well understood [19]. Hence, in order to understand the notion of minimality and innovation representation for stochastic LPV-SSA with control input, the first step is to understand these notions for asLPV-SSA.

Motivation for minimal LPV-SSAs in innovation form: In the formulation of the identification problem for LPV-SSA [26], [9], [27], [28], [6] the stated goal is usually to find an LPV-SSA which is isomorphic or at least which is input-output equivalent to the data generating system. However, in general there may exist LPV-SSAs which generate the same output for some scheduling signal, but which are not isomorphic (see Example 3 in Section V or [11]), or which are not input-output equivalent. In the latter case the systems generate different outputs when the scheduling signal is changed. See Example 3 and Example 4 of Section V of the present paper.

This implies that in general the system identification problem is ill-posed. However, it becomes well-posed, if we add the assumption that the underlying data generating system is a minimal asLPV-SSA in innovation form.

Indeed, if the observed output has an asLPV-SSA representation, it has a minimal one in innovation form, and all such representations are related by a constant isomorphism. The necessity of this assumption is illustrated by the examples in Section V. From [16] it follows that this assumption is sufficient, as the identification algorithm in [16] returns a minimal asLPV-SSA in innovation form. We conjecture that the same will be true for most of the existing subspace identification algorithms [26], [9], [27], [28], [6], [7].

To deal only with minimal asLPV-SSA representations in innovation form, simple conditions to check minimality and being in innovation form are needed. The latter is necessary in order to check if the elements of a parametrization of asLPV-SSAs are minimal and in innovation form, or to construct such parametrizations.

Related work: As it was mentioned above, there is a rich literature on subspace identification methods for stochastic LPV-SSA representations [7], [8], [6], [26]. However, the cited papers do not deal with the problem of characterizing minimal stochastic LPV state-space representations in innovation form. In [5], [6] the existence of an LPV state-space representation in innovation form was studied, but due to the specific assumptions (deterministic scheduling) and the definition of the innovation process, the resulting LPV state-space representation in innovation form had dynamic dependence on the scheduling parameters. Moreover, [5], [6] do not address the issue of minimality of the stochastic part.

1Often, it is required that the isomorphism does not depend on the scheduling.
of LPV state-space representations.

This paper uses realization theory of stochastic generalized bilinear systems (GBS for short) of [20]. In particular, asLPV-SSAs correspond to GBSs. The existence and uniqueness of minimal asLPV-SSAs in innovation form follows from the results of [20]. The main novelty of the present paper with respect to [20] is the new algebraic characterization of minimal asLPV-SSAs in innovation form, and that the results on existence and uniqueness of minimal GBSs are spelled out explicitly for LPV-SSAs.

The paper [16] used the correspondence between GBSs and asLPV-SSAs to state existence and uniqueness of minimal asLPV-SSAs in innovation form. However, [16] did not provide an algebraic characterization of minimality or innovation form. Moreover, it considered only scheduling signals which were zero mean white noises. In contrast, in this paper more general scheduling signals are considered. The present paper is complementary to [16]. This paper explains when the assumption that the data generating system is minimal asLPV-SSA in innovation form could be true, while [16] presents an identification algorithm which is statistically consistent under the latter assumption.

Outline of the paper: In Section II we introduce the notations used and we recall [20], some technical assumptions which are necessary to define of the stationary LPV-SSA representation. In Section III some principal results on minimal asLPV-SSAs in innovation form are reviewed. In Section IV, we present the main results of the paper, namely, algebraic conditions for an asLPV-SSA to be minimal in innovation form. Finally, in Section V numerical examples are developed to illustrate the contributions.

II. Preliminaries

In the sequel, we will use the standard terminology of probability theory [3]. In particular, all the random variables and stochastic processes are understood w.r.t. to a fixed probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \(\mathcal{F}\) is a \(\sigma\)-algebra over the sample space \(\Omega\). The expected value of a random variable \(r\) is denoted by \(E[r]\) and conditional expectation w.r.t. \(\sigma\)-algebra \(\mathcal{F}\) is denoted by \(E[r | \mathcal{F}]\). All the stochastic processes in this paper are discrete-time ones defined over the time-axis \(\mathbb{Z}\). Random variables \(r\) taking values in some set \(X\).

Next, we define the class of systems studied in this paper. An autonomous stochastic linear parameter-varying state-space representation with affine dependence on scheduling parameter (aLPV-SSA) is system described by

\[
\begin{aligned}
    x(t+1) &= \sum_{i=1}^{n_r} (A_{\sigma} x(t) + K_{\sigma} v(t)) \mu_i(t) \\
y(t) &= C x(t) + F v(t)
\end{aligned}
\]

where \(A_{\sigma} \in \mathbb{R}^{n \times n}, K_{\sigma} \in \mathbb{R}^{n \times m}, C \in \mathbb{R}^{p \times n}, F \in \mathbb{R}^{n_y \times m}\), and \(x\) is the state process, \(\mu = [\mu_1, \ldots, \mu_{n_r}]^T\) is the scheduling process, \(v\) is the noise process and \(y\) is the output process.

Note that all the involved processes are defined for both negative and positive time. This may create technical problems for the existence of a solution and the role of initial state. In this paper we will circumvent this problem by considering LPV-SSA which are mean-square stable in a suitable sense and state process of which is stationary. In order to define this class of LPV-SSAs we will have to recall from [20] some notation and terminology.

A. Admissible scheduling and wide-sense stationarity (ZWSII) w.r.t. scheduling

Below we define the concept of admissible input processes and ZWSII processes w.r.t. scheduling. These concepts will be used to define the class of aLPV-SSAs we will work with. To this end, we will need the following notation from automata theory which will be used for other purposes too.

**Notation I (Sequences over \(\Sigma\))**: Consider the finite set

\[\Sigma = \{1, \ldots, n_\mu\} \]

A non empty word over \(\Sigma\) is a finite sequence of letters, i.e., \(w = \sigma_1 \sigma_2 \cdots \sigma_k\), where \(0 < k \in \mathbb{Z}\), \(\sigma_1, \sigma_2, \ldots, \sigma_k \in \Sigma\). The set of all nonempty words is denoted by \(\Sigma^+\). We denote an empty word by \(\epsilon\). Let \(\Sigma^* = \epsilon \cup \Sigma^+\). The concatenation of two nonempty words \(v = a_1 a_2 \cdots a_m\) and \(w = b_1 b_2 \cdots b_n\) is defined as \(vw = a_1 \cdots a_m b_1 \cdots b_n\) for some \(m, n > 0\). Note that if \(w = \epsilon\) or \(v = \epsilon\), then \(vw = v\) and \(cw = w\), moreover, \(\epsilon \epsilon = \epsilon\). The length of the word \(w \in \Sigma^*\) is denoted by \(|w|\), and \(|\epsilon| = 0\).

With the notation above, we can identify the process \(\mu\) with the collection \(\{\mu_\sigma\}_{\sigma \in \Sigma}\) of its components. We will say that \(\mu\) is an admissible scheduling process, if the collection \(\{\mu_\sigma\}_{\sigma \in \Sigma}\) is an admissible collection of input processes in the sense of [20, Definition 1] for \(S = \Sigma \times \mathbb{K}\). For the convenience of the reader, a version of [20, Definition 1] is presented in Appendix A. Before proceeding further we will present examples of admissible scheduling sequences.

**Example 1 (White noise scheduling)**: The scheduling process \(\mu = [\mu_1, \mu_2, \ldots, \mu_{n_r}]^T\) is independent identically distributed (i.i.d.) such that for all \(i,j = 2, \ldots, n_\mu, t \in \mathbb{Z}\), \(\mu_i(t), \mu_j(t)\) are independent and \(\mu_i(t)\) is zero mean, then \(\mu\) is admissible.

**Example 2 (Discrete valued i.i.d process)**: Assume there exists an i.i.d. process \(\theta\) which takes its values from a finite set \(\Sigma = \{1, \ldots, n_\mu\}\). Let \(\mu_\sigma(t) = \chi(\theta(t) = \sigma)\) for all \(\sigma \in \Sigma, t \in \mathbb{Z}\). Then \(\mu(t) = [\mu_1(t) \ldots \mu_{n_r}]^T\) is an admissible scheduling processes.

For further examples, see [20].

We assume that the scheduling process \(\mu\) is admissible throughout the paper.

Furthermore, we will use the definition of a Zero Mean Wide Sense Stationary (abbreviated by ZMWSII) process with respect to \(\mu\) from [20, Definition 2]. For the convenience of the reader this definition is reformulated in Appendix A as Definition 5.

In order to explain the intuition behind these definitions, and because we will use them later on, we define the following products of scheduling variables along a sequence.
from $\Sigma^+$. For every word $w \in \Sigma^+$ where $w = \sigma_1\sigma_2\cdots\sigma_k$, $k \geq 1$, $\sigma_1, \ldots, \sigma_k \in \Sigma$, we define the process $\mu_w$ as follows:

$$
\mu_w(t) = \mu_{\sigma_1}(t-k+1)\mu_{\sigma_2}(t-k+2)\cdots\mu_{\sigma_k}(t) \quad (2)
$$

For an empty word $w = \epsilon$, we set $\mu_\epsilon(t) = 1$.

If $\mu$ is admissible, then $\mu_w, \mu_v$ are jointly wide-sense stationary. Moreover, $\mu_w(t)$ and $\mu_v(t)$ are uncorrelated, if the last letters of $w$ and $v$ are different, and

$$
E[\mu_w^2(t)] = p_w E[\mu_v^2(t)]
$$

for any $w \in \Sigma^+, \sigma \in \Sigma$. That is, $\mu$ uniquely determines a collection of numbers $\{p_\sigma\}_{\sigma \in \Sigma}$. This latter collection will play an important role in the sequel. In particular, for $\mu$ from Example 1, $p_i$ is the variance of $\mu_i$, and for $\mu$ from Example 2, $p_i$ is the probability $P(\theta(t) = i)$, for all $i \in \Sigma$.

In order to explain the significance of these assumptions, let $r$ be a stochastic ZMWSII process w.r.t. $\mu$, and because we will use them latter on, we define the following products. Let $\{p_\sigma\}_{\sigma \in \Sigma}$ be the constants determined by $\mu$ as explained above, and define the products

$$
p_w = p_\sigma p_\sigma_2 \cdots p_\sigma_k. \quad (3)
$$

For an empty word $w = \epsilon$, we set $p_\epsilon = 1$. For a stochastic process $r \in \mathbb{R}^{n_u}$ and for each $w \in \Sigma^*$ we define the stochastic process $z_w^r$ as

$$
z_w^r(t) = r(t - |w|)\mu_w(t-1)\frac{1}{\sqrt{p_w}}, \quad (4)
$$

where $\mu_w$ and $p_w$ are as in (2) and (3). For $w = \epsilon$, $z_w^r(\epsilon)(t) = r(t)$. The process $z_w^r$ in (4) is interpreted as the product of the past of $r$ and $\mu$. The process $z_w^r$ will be used as predictors for future values of $r$ for various choices of $r$.

Now we will explain the motivation for the concept of ZMWSII. Assume that $r$ is ZMWSII w.r.t. $\mu$. Then $z_w^r(t)$ is wide-stationary and square-integrable for all $w \in \Sigma^+$. Moreover, the covariances $E[z_w^r(t)z_v^r(t)]$ do not depend on $t$. Furthermore, $z_w^r(t)$, $z_v^r(t)$ are orthogonal, if $w$ is not a suffix of $v$ or vice versa. Recall that we say that $w$ is a suffix of $v$, if $v = sw$ for some $s \in \Sigma^*$. Moreover, when $w$ is suffix of $v$, then

$$
E[z_w^r(t)z_v^r(t)] = \begin{cases} 
E[z_w^r(t)z_v^r(t)] & \text{if } v = sw \\
E[z_w^r(t)z_v^r(t)] & \text{if } v = w = \sigma s
\end{cases}
$$

That is, $E[z_w^r(t)z_v^r(t)]$ depends only on the difference of $w$ and $v$, i.e., on the prefix of $v$. This can be viewed as a generalization of wide-sense stationarity, if the index $w$ and $v$ are viewed as additional multidimensional time instances, and $\Sigma^*$ is viewed as an additional time-axis.

### B. Stationary LPV-SSA representation

After identifying the necessary process properties and notations, we are now ready to present the definition of the stationary autonomous stochastic LPV-SSA.

**Definition 1:** A stationary autonomous stochastic LPV-SSA, abbreviated as asLPV-SSA, is a system of the form (1), such that:

1. $[x^T \ v^T]^T$ is a ZMWSII process, and for all $\sigma \in \Sigma$, $w \in \Sigma^+$, $E[z_w^\sigma(t)(z_w^\sigma(t))^T] = 0$, $E[v(t)(z_w^\sigma(t))^T] = 0$.
2. $v$ is ZMWSII and $E[v(t)(z^\sigma(t))^T] = 0$ for all $w \in \Sigma^+$.
3. The eigenvalues of the matrix $\sum_{\sigma \in \Sigma} p_\sigma A_\sigma \otimes A_\sigma$ are outside the open unit circle.

Note that, condition (2) implies that $v$ is a white noise.

In the terminology of [20], an asLPV-SSA corresponds to a stationary GBS w.r.t. inputs $\{p_\sigma\}_{\sigma \in \Sigma}$. Note that the processes $x$ and $y$ are ZMWSII, in particular, they are wide-sense stationary, and that $x$ is orthogonal to the future values of the noise process $v$. We should mention that we concentrate on wide-sense stationary processes, because it is difficult to estimate the distribution of non-stationary processes. Also, wide-sense stationary processes solve the problem of the initial state conditions.

The state of an asLPV-SSA is uniquely determined by its matrices and noise process. In order to present this relationship, we need the following notation.

**Notation 2 (Matrix Product):** Consider $n \times n$ square matrices $\{A_\sigma\}_{\sigma \in \Sigma}$. For any word $w \in \Sigma^+$ of the form $w = \sigma_1\sigma_2\cdots\sigma_k$, $k > 0$ and $\sigma_1, \ldots, \sigma_k \in \Sigma$, we define

$$
A_w = A_{\sigma_1}A_{\sigma_2-1}\cdots A_{\sigma_k}.
$$

For an empty word $\epsilon$, let $A_\epsilon = I_n$.

From [20, Lemma 2] it follows that

$$
x(t) = \sum_{w \in \Sigma^+, \sigma \in \Sigma} \sqrt{p_\sigma} A_\sigma z_w^\sigma(t)
$$

where the infinite sum on the right-hand side is absolutely convergent in the mean square sense. This prompts us to use the following notation.

**Notation 3:** We identify the asLPV-SSA $S$ of the form (1) with the tuple $S = (\{A_\sigma, K_\sigma\}_{\sigma \in \Sigma}, C, F, v)$.

Finally, we need to define what we mean by an asLPV-SSA realization of a process. An asLPV-SSA $S$ of the form (1) is a realization of a pair $(\hat{y}, \hat{\mu})$, if $\hat{y} = y$, $\hat{\mu} = \mu$. If $S$ is of the form (1), then we call the state-space dimension $n_x$ the dimension of $S$ and we denote it by $\dim S$. We say that the asLPV-SSA $S$ is a minimal realization of $(\hat{y}, \hat{\mu})$, if for any asLPV-SSA realization $S'$ of $(\hat{y}, \hat{\mu})$, the dimension of $S'$ is not smaller than $\dim S$.

### III. Existence and Minimality of asLPV-SSAs in Innovation Form

In this section we review the principal results on existence and minimality of asLPV-SSA in innovation form. To this end, in Subsection III-A we recall from [19] some results on realization theory of deterministic LPV-SSA. In Subsection III-B we present the definition of asLPV-SSAs in innovation form, and in Subsection III-C we present results on existence and uniqueness of minimal asLPV-SSAs. In Subsection III-D we present rank conditions for minimality of asLPV-SSAs and an algorithm for converting any asLPV-SSA to a minimal one in innovation form. The results presented in this section follow from [20], but they have never been formulated explicitly for LPV-SSAs.
A. Deterministic LPV-SSA representation

Recall from [19, 5] that a deterministic LPV state-space representation with affine dependence (abbreviated as dLPV-SSA) is a system of the form:

\[ x(t + 1) = \sum_{i=1}^{n_p} (A_i x(t) + B_i u(t)) \mu_i(t), \]
\[ y(t) = C x(t) + D u(t), \]

where \( A_i, B_i, C, D \) are matrices of suitable dimensions, \( x : \mathbb{Z} \rightarrow \mathbb{R}^{n_x} \) is the state trajectory \( u : \mathbb{Z} \rightarrow \mathbb{R}^{n_u} \) is the input trajectory \( y : \mathbb{Z} \rightarrow \mathbb{R}^{n_Y} \) is the output trajectory with finite support. We identify a dLPV-SSA of the form (5) with the minimal of \( S \) of parameters of \( \Sigma \) function, using Kalman decomposition [19, Corollary 1] be transformed to a minimal one with the same sub-Markov form. To this end, we need to introduce the following notation for orthogonal projection.

B. Definition of asLPV-SSA in innovation form

Let \( S \) be an asLPV-SSA of the form (1) in innovation form with \( F = I_{n_y} \). Let \( \tilde{S} = (\{A_i, K_i, Y_\sigma = 0, C, I_{n_y}, e\} \) be another asLPV-SSA of \( (y, \mu) \) in innovation form. We say that \( S \) and \( \tilde{S} \) are isomorphic, if there exists a nonsingular matrix \( T \) such that

\[ \tilde{A}_i = T A_i T^{-1}, \tilde{K}_i = T K_i, \tilde{C} = C T^{-1} \]

We will say that the process \( (y, \mu) \) is full rank, if for all \( i = 1, \ldots, n_p, Q_i = E[e(t) e(t)^T \mu_i^2(t)] \) is invertable. This is a direct extension of the classical notion of a full rank process.

Furthermore, we will say that \( y \) is Square Integrable process w.r.t. \( \mu \), abbreviated by SII, if it satisfies [20, Definition 5]. For the convenience of the reader, the definition of an SII process is presented in Appendix A. Furthermore, any dLPV-SSA can be transformed to a minimal one with the same sub-Markov function, using Kalman decomposition [19, Corollary 1]. For a more detailed discussion see [19].

B. Definition of asLPV-SSA in innovation form

Next, we define what we mean by asLPV-SSA in innovation form. To this end, we need to introduce the following notation for orthogonal projection.

Notation 4 (Orthogonal projection \( E_i \)): Recall that the set of square integrable random variables taking values in \( \mathbb{R} \) forms a Hilbert-space with the scalar product defined as \( < z_1, z_2 > = E[z_1 z_2^*] \). We denote this Hilbert-space by \( \mathcal{H}_1 \).

Let \( z \) be a square integrable vector-valued random variable taking its values in \( \mathbb{R}^k \). Let \( M \) be a closed subspace of \( \mathcal{H}_1 \).

By the orthogonal projection of \( z \) onto the subspace \( M \), denoted by \( E_i [z \mid M] \), we mean the vector-valued square-integrable random variable \( z^* = [z_1^*, \ldots, z_k^*]^T \) such that \( z_i^* \in M \) is the orthogonal projection of the \( i \)th coordinate of \( z \) onto \( M \), as it is usually defined for Hilbert spaces. Let \( \mathcal{S} \) be a subset of square integrable random variables in \( \mathbb{R}^p \) for some integer \( p \), and suppose that \( M \) is generated by the coordinates of the elements of \( \mathcal{S} \), i.e. \( M \) is the smallest (with respect to set inclusion) closed subspace of \( \mathcal{H}_1 \) which contains the set \( \{\alpha^T s : s \in \mathcal{S}, \alpha \in \mathbb{R}^p\} \). Then instead of \( E_i [z \mid M] \) we use \( E_i [z \mid \mathcal{S}] \).

This said, we define the innovation process \( e \) of \( y \) with respect to \( \mu \) as follows:

\[ e(t) = y(t) - E_i [y(t) \mid \{\mathcal{S}_i(t)\}_{w \in \Sigma^+}] \]

In other words, \( e(t) \) is the difference between the output and its projection on its past values w.r.t. the scheduling process \( \mu \), i.e., \( e \) is the best predictor of \( y \) using the product of the output and scheduling past values from (4).

Definition 2 (asLPV-SSA in innovation form): An asLPV-SSA of the form (1) is said to be in innovation form, if it is a realization of \( (y, \mu) \), \( F = I_{n_y} \), and \( v \) is the innovation process of \( y \), i.e., \( v = e \).

C. Existence and uniqueness of minimal asLPV-SSA in innovation form

Let \( S \) be an asLPV-SSA of the form (1) in innovation form with \( F = I_{n_y} \). Let \( \tilde{S} = (\{A_i, K_i, Y_\sigma = 0, C, I_{n_y}, e\} \) be another asLPV-SSA of \( (y, \mu) \) in innovation form. We say that \( S \) and \( \tilde{S} \) are isomorphic, if there exists a nonsingular matrix \( T \) such that

\[ \tilde{A}_i = T A_i T^{-1}, \tilde{K}_i = T K_i, \tilde{C} = C T^{-1} \]

We will say that the process \( (y, \mu) \) is full rank, if for all \( i = 1, \ldots, n_p, Q_i = E[e(t) e(t)^T \mu_i^2(t)] \) is invertable. This is a direct extension of the classical notion of a full rank process.

Furthermore, we will say that \( y \) is Square Integrable process w.r.t. \( \mu \), abbreviated by SII, if it satisfies [20, Definition 5]. For the convenience of the reader, the definition of an SII process is presented in Appendix A. Furthermore, any dLPV-SSA can be transformed to a minimal one with the same sub-Markov function, using Kalman decomposition [19, Corollary 1]. For a more detailed discussion see [19].

\[ 2\{19, Corollary 1\} should be applied with zero initial state \]

\[ 3\{20, Remark 2\} it follows that if \( SII \) process is presented in Appendix A, Definition 6. From [20, Remark 2] \] it follows that if \( y, \mu \) has a realization by an asLPV-SSA and \( \mu \) is bounded, then \( y \) is SII.
finding conditions for an asLPV-SSA to be minimal in innovation form and formulating algorithms for transforming an asLPV-SSA to a minimal one in innovation form. To this end, in Subsection [III-D] we present a rank condition for minimality and a minimization algorithm based on the results of [20]. However, the results of Subsection [III-D] do not allow checking that an asLPV-SSA is in innovation form. Moreover, the algebraic conditions for minimality are difficult to apply. Motivated by this, in Section [IV] we present more user-friendly characterizations of minimality and being in innovation form.

\section{D. Rank conditions and minimization algorithm}

In order to present the rank conditions for minimality of asLPV-SSA and the minimization algorithm, we need to define the \textit{dLPV-SSA associated with asLPV-SSA} as

\[ \mathcal{S}_S = \left\{ \{ A_i, G_i \}_{i=1}^{n_g}, C, I_{n_y} \right\}, \]

where

\[ G_i = \frac{1}{\sqrt{P_i}} \left( A_i P_i C^T + K_i Q_i F^T \right) \]

and \( Q_i = E[\nu(t)\nu^T(t)\mu^T_i(t)] \) for \( i = 1, \ldots, n_y \), and \( \{ P_i \}_{i=1}^{n_y} \) are the computed as follows: \( P_i = \lim_{\tau \to \infty} P_i^{\tau} \), and \( \{ P_i^{\tau} \}_{i=1}^{n_y} \) satisfy the following recursions:

\[ P_i^{\tau+1} = p_i \sum_{j=0}^{n_y} (A_j P_j T_j + K_j Q_j K_j^T) \]  \hspace{1cm} (10)

with \( P_i^0 = O_{n_x \times n_x} \) where \( O_{n_x \times n_x} \) denotes the \( n_x \times n_x \) matrix with all zero entries. Note that the existence of the limit of \( P_i^{\tau} \) when \( N \) goes to infinity follows from [20, Lemma 5]. From [20] it follows that the dLPV-SSA associated with an asLPV-SSA \( \mathcal{S} \) represents a realization of Markov-function \( \Psi_y : \Sigma^* \to \mathbb{R}^{n_y \times n_x} \)

\[ \Psi_y(w) = \begin{cases} E[y(t)(\varphi_{\theta}(t))^T] & w \in \Sigma^+ \\ I_{n_y} & w = \epsilon \end{cases} \]

computed from covariances of \( y \).

Then from [20] we can derive the following.

\textbf{Theorem 2 (Rank conditions):} An asLPV-SSA is a minimal realization of \( (y, \mu) \), if and only if the associated dLPV-SSA is minimal.

From [19] it follows that minimality of the associated dLPV-SSA can be checked using rank conditions for the corresponding extended reachability and observability matrices. Note, however, that minimal asLPV-SSAs, realizing the same output, may not be isomorphic. In fact, in Section [V] Example 3 presents a counter-example.

Vice versa, with any dLPV-SSA realization of \( \Psi_y \) we can associate an asLPV-SSA in innovation form. This latter relationship is useful for formulating a minimization algorithm. More precisely, consider a dLPV-SSA

\[ \mathcal{S} = \left\{ \{ A_i, K_i \}_{i=1}^{n_y}, C, I_{n_y} \right\} \]

which is a minimal realization of \( \Psi_y \). Define the \textit{asLPV-SSA} \( S_{\mathcal{S}} \) associated with \( \mathcal{S} \) as

\[ S_{\mathcal{S}} = \left\{ \{ \frac{1}{\sqrt{P_i}} A_i, K_i \}_{i=1}^{n_y}, C, I_{n_y}, e \right\}, \]

where \( \hat{K}_i = \lim_{\tau \to \infty} \hat{K}_i^{\tau} \), and \( (\hat{K}_i^{\tau})_{\sigma \in \Sigma, \tau \in \mathbb{N}} \) satisfies the following recursion

\[ \hat{P}_\sigma^{\tau+1} = \sum_{\sigma_i \in \Sigma} p_{\sigma_i} \left( \frac{1}{\sqrt{P_{\sigma_i}}} \hat{A}_{\sigma_i} \hat{P}_{\sigma_i} (\hat{A}_{\sigma_i})^T + \hat{K}_{\sigma_i} \hat{Q}_{\sigma_i} (\hat{K}_{\sigma_i})^T \right) \]

\[ \hat{Q}_{\sigma_i} = p_{\sigma_i} T_{\sigma_i, \sigma} y - \hat{C} \hat{P}_{\sigma_i} (\hat{C})^T \]

\[ \hat{K}_{\sigma_i} = \left( \hat{G}_{\sigma_i} \sqrt{p_{\sigma_i}} - \frac{1}{\sqrt{p_{\sigma_i}}} \hat{A}_{\sigma_i} \hat{P}_{\sigma_i} (\hat{C})^T \right) \left( \hat{Q}_{\sigma_i} \right)^{-1} \]  \hspace{1cm} (11)

where \( \hat{P}_{\sigma_i}^{0} \) is a \( n_x \times n_x \) zero matrix and \( T_{\sigma_i, \sigma} y = E[y(t)(\varphi_{\theta}(t))^T] \). From [20] it follows that if \( \mathcal{S} \) is a minimal dLPV-SSA realization of \( \Psi_y \), then the associated asLPV-SSA \( S_{\mathcal{S}} \) is an asLPV-SSA of \( (y, \mu) \).

The discussion above suggests the following realization algorithm for transforming an asLPV-SSA \( \mathcal{S} \) to a minimal one, which can be deduced from [20].

\textbf{Algorithm 1 Minimization algorithm}

\textbf{Input :} The asLPV-SSA \( \mathcal{S} = \left\{ \{ A_{\sigma}, K_{\sigma} \}_{\sigma=1}^{n_y}, C, F, v \right\} \) representation matrices.

1. Compute the dLPV-SSA \( \mathcal{S}_S \) associated with \( \mathcal{S} \) and compute \( T_{\sigma, \sigma} y = \frac{1}{\sqrt{p_{\sigma}}} (CP_{\sigma} C^T + F Q_{\sigma} F^T) \).

2. Transform the dLPV-SSA \( \mathcal{S}_S \) to a minimal dLPV-SSA \( \mathcal{S}_m \) using [19, Corollary 1].

3. Construct the asLPV-SSA \( S_{\mathcal{S}_m} \) associated with \( \mathcal{S}_m \).

\textbf{Output :} The asLPV-SSA \( S_{\mathcal{S}_m} \).

From [20, Theorem 3], it follows that Algorithm 1 returns a minimal realization of \( (y, \mu) \), if \( S \) is an asLPV-SSA realization of \( (y, \mu) \) in innovation form. Note that Algorithm 1 requires only the knowledge of the matrices of \( \mathcal{S} \) and the noise covariance matrix \( Q_{\sigma} = E[\nu(t)\nu^T(t)\mu^T_i(t)] \). Also note that all the steps above are computationally efficient, however, they require finding the limits of \( P_i^{\tau} \) and \( K_i^{\tau} \) for \( \tau \rightarrow \infty \) respectively. Also note that, in [16], it exists another minimization algorithm which uses covariances matrices.

\textbf{Remark 1 (Challenges):} The main disadvantage of verifying the rank condition of Theorem 2 or applying Algorithm 1 is the necessity of constructing a dLPV-SSA and the necessity to find the limit of the matrices in (11). The latter represents an extension of algebraic Riccati equations [20, Remark 7] and even for the linear case requires attention. Moreover, the rank conditions of Theorem 2 are not easy to apply to parametrizations: even if the dependence of the matrices \( \{ A_i, K_i \}_{i=1}^{n_y} \) and \( C \) on a parameter \( \theta \) are linear or polynomial, the dependence of the matrices of the associated dLPV-SSA need not remain linear or polynomial, due to the definition of \( G_i \) in (9). For the same reason, it is difficult to analyze the result of applying Algorithm 1 to elements of a parametrization. Moreover, the conditions of Theorem 2 do not allow us to check if the elements of a parametrizations are in innovation form. These shortcomings motivate the contribution of Section [V].
IV. Main results: algebraic conditions for an asLPV-SSA to be minimal in innovation form

Motivated by the challenges explained in Remark 1 in this section we present sufficient conditions for an asLPV-SSA to be minimal and in innovation form. These conditions depend only on the matrices of the asLPV-SSA in question and do not require any information on the noise processes.

The first result concerns an algebraic characterization of asLPV-SSA in innovation form. This characterization does not require any knowledge of the noise process, only the knowledge of system matrices. In order to streamline the discussion, we introduce the following definition.

Definition 3 (Stably invertable w.r.t. $\mu$): Assume that $S$ is an asLPV-SSA of the form (1) and $F = I_{n_y}$. We will call $S$ stably invertable with respect to $\mu$, or stably invertable if $\mu$ is clear from the context, if the matrix

$$
\sum_{i=1}^{n_\mu} p_i (A_i - K_i C) \otimes (A_i - K_i C) \tag{12}
$$

is stable (all its eigenvalues are inside the complex unit disk).

Note that a system can be stably invertable w.r.t. one scheduling process, and not to be stably invertable w.r.t. another one. We can now state the result relating stable invertability to asLPV-SSAs in innovation forms.

Theorem 3 (Innovation form condition): Assume that $y$ is SII and $(y,\mu)$ is full rank. If an asLPV-SSA realization of $(y,\mu)$ is stably invertable, then it is in innovation form.

The proof of Theorem 3 can be found in Appendix B.

Stably invertable asLPV-SSAs can be viewed as optimal predictors. Indeed, let $S$ be the asLPV-SSA of the form (1) which is in innovation form, and let $x$ be the state process of $S$. It then follows

$$
x(t+1) = \sum_{i=1}^{n_\mu} (A_i - K_i C)x(t) + K_i y(t)\mu_i(t), \tag{13}
$$

$$
y(t) = Cx(t)
$$

where $\hat{y}(t) = E_t[y(t) | \{z^w_w(t)\}_{w \in \Sigma^+}]$, i.e., $\hat{y}$ is the best linear prediction of $y(t)$ based on the predictors $\{z^w_w(t)\}_{w \in \Sigma^+}$. Intuitively, (13) could be viewed as a filter, i.e., a dynamical system driven by past values of $y$ and generating the best possible linear prediction $\hat{y}(t)$ of $y(t)$ based on $\{z^w_w(t)\}_{w \in \Sigma^+}$.

However, the solution of (13) is defined on the whole time axis $\mathbb{Z}$ and hence cannot be computed exactly. For stably invertable asLPV-SSA we can approximate $\hat{y}(t)$ as follows.

Lemma 1: With the assumptions of Theorem 3 if $S$ of the form (1) is a stably invertable realization of $(y,\mu)$, and we consider the following dynamical system:

$$
x(t+1) = \sum_{i=1}^{n_\mu} (A_i - K_i C)x(t) + K_i y(t)\mu_i(t), \tag{14}
$$

$$
y(t) = Cx(t), \quad \bar{x}(0) = 0
$$

then $\lim_{t \to \infty} (\bar{x}(t) - x(t)) = 0$, and $\lim_{t \to \infty} (\hat{y}(t) - y(t)) = 0$, where the limits are understood in the mean square sense.

The proof of Lemma 1 is found in Appendix B. That is, the output $\hat{y}(t)$ of the recursive filter (14) is an approximation of the optimal prediction $\bar{y}(t)$ of $y(t)$ for large enough $t$. That is, stably invertable asLPV-SSA not only result in asLPV-SSAs in innovation form, but they represent a class of asLPV-SSAs for which recursive filters of the form (14) exist.

Next, we present algebraic conditions for minimality of an asLPV-SSA in innovation form.

Theorem 4 (Minimality condition in innovation form): Assume that $S$ is an asLPV-SSA of the form (1) and $S$ is a realization of $(y,\mu)$ in innovation form. Assume that $(y,\mu)$ is full rank and $y$ is SII. Then $S$ is a minimal realization of $(y,\mu)$, if and only if the dLPV-SSA $\mathcal{D}_S = \{(A_i, K_i)\}_{i=0}^{n_\mu}, C, I_{n_y}\}$ is minimal.

The proof of Theorem 4 can be found in Appendix B.

Theorem 4 in combination with Theorem 3 leads to the following corollary.

Corollary 1 (Minimality and innovation form): With the assumptions of Theorem 4 if $\mathcal{D}_S$ is minimal and $S$ if stably invertable, then $S$ is a minimal asLPV-SSA realization of $(y,\mu)$ in innovation form.

Remark 2 (Checking minimality and innovation form): We recall that $\mathcal{D}_S$ is minimal, if and only if it satisfies the rank conditions for the extended $n$-step reachability and observability matrices [19, Theorem 2], which can easily be computed from the matrices of $S$. Checking that $S$ is stably invertable boils down to checking the eigenvalues of the matrix (12). That is, Corollary 1 provides effective procedure for verifying that an asLPV-SSA is minimal and in innovation form. Note that in contrast to the rank condition of Theorem 2 which required computing the limit of (10), the procedure above uses only the matrices of the system.

Remark 3 (Parametrizations of asLPV-SSAs): Below we will sketch some ideas for applying the above results to parametrizations of asLPV-SSAs. A detailed study of these issues remains a topic for future research.

For all the elements of a parametrization of asLPV-SSAs to be minimal and in innovation form, by Corollary 1 it is necessary that (A) all elements of the parametrization, when viewed as dLPV-SSA, are minimal, and that (B) they are stably invertable and satisfy condition (3) of Definition 1. In order to deal with (A), the techniques used in [1] or [18] could be used. The condition (B) is equivalent to stability of a suitable parametrization of LTI systems, as we could view the matrices (12) and $\sum_{i=1}^{n_\mu} p_i A_i$ as matrices of an LTI state-space representation. For the latter we can use standard techniques, see [23] and the references therein.

Finally, note that extension of argument of [1, Theorem 2] would also lead to identifiability conditions for parametrizations which satisfy the conditions (A) and (B) described above.

Corollary 1 suggests the following minimization algorithm.

\footnote{In order to use [18] the relationship between minimality of dLPV-SSA and that of switched systems [19] should be exploited}
Algorithm 2 Minimization algorithm

Input: A stably invertible asLPV-SSA \( S \)

1. Transform the dLPV-SSA \( \mathcal{D}_S \) from Theorem 2 to a minimal dLPV-SSA \( \mathcal{D}_m = (\{A_i^m, K_i^m\}_{i=1}^{\mu}, C^m, I_{n_y}) \) using [19, Corollary 1].

Output: asLPV-SSA \( S_m = (\{A_i^m, K_i^m\}_{i=1}^{\mu}, C^m, I_{n_y}, e) \).

**Lemma 2 (Correctness of Algorithm 2):** The asLPV-SSA \( S_m \) is stably invertible and it is a minimal asLPV-SSA realization of \( (y, \mu) \) in innovation form.

That is, Algorithm 2 is a simpler counterpart of Algorithm 1 which does not require computing the limits (10) and (11), which can create difficulties (see Remark 1). Lemma 2 implies that stable invertability is preserved by minimization.

V. NUMERICAL EXAMPLES

In this section, we present numerical examples in order to illustrate the main results.

**Example 3:** Consider an asLPV-SSA of the form (1), where \( n_\mu = 2 \) and

\[
A_1 = \begin{bmatrix} 0.4 & 0.4 \\ 0.2 & 0.1 \end{bmatrix}, \quad K_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [10 \ 0 \ 0]
\]

\[
A_2 = \begin{bmatrix} 0.1 & 0.3 \\ 0.2 & 0.3 \end{bmatrix}, \quad K_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \text{and } F = 1
\]

Note that this representation is not in innovation form. The scheduling signal process is defined as \( \mu = [\mu_1, \mu_2] \) such that \( \mu_1(t) = 1 \) and \( \mu_2(t) = 1 \) is a white-noise process with uniform distribution \( \mathcal{U}(-1.5, 1.5) \). This corresponds to the parameters values \( \{p_\sigma\}_{\sigma \in [1, 2]} \) to be \( p_1 = E[\mu_2^2(t)] = 1 \) and \( p_2 = E[\mu_2^2(t)] = 0.75 \). The noise process is a white Gaussian noise with a variance equal to 1, i.e., \( v \sim \mathcal{N}(0, 1) \).

Using Algorithm 1 we can find a minimal representation in innovation form with the following matrices:

\[
A_1^m = \begin{bmatrix} 0.4007 & -0.3997 \\ -0.1997 & 0.0993 \end{bmatrix}, \quad K_1^m = \begin{bmatrix} -0.046 \\ 0.0541 \end{bmatrix}
\]

\[
A_2^m = \begin{bmatrix} 0.1003 & -0.1002 \\ -0.2002 & 0.2997 \end{bmatrix}, \quad K_2^m = \begin{bmatrix} -0.0116 \\ 0.0578 \end{bmatrix}
\]

\[
C^m = \begin{bmatrix} -10 \\ -0.0116 \end{bmatrix}, \quad \text{and } F^m = 1
\]

The system above is not in innovation form, in fact, it is not stably invertable. As before, we use Algorithm 1 to obtain a minimal asLPV-SSA innovation form with the following matrices:

\[
A_1^m = \begin{bmatrix} 0.4007 & -0.3997 \\ -0.1997 & 0.0993 \end{bmatrix}, \quad K_1^m = \begin{bmatrix} -0.046 \\ 0.0541 \end{bmatrix}
\]

\[
A_2^m = \begin{bmatrix} 0.1003 & -0.1002 \\ -0.2002 & 0.2997 \end{bmatrix}, \quad K_2^m = \begin{bmatrix} -0.0116 \\ 0.0578 \end{bmatrix}
\]

\[
C^m = \begin{bmatrix} -10 \\ -0.0116 \end{bmatrix}, \quad \text{and } F^m = 1
\]

The two asLPV-SSA systems \( S = (\{A_\sigma, K_\sigma\}_{\sigma = 1}^{\mu}, C, F, v) \) and \( S^m = (\{A_\sigma^m, K_\sigma^m\}_{\sigma = 1}^{\mu}, C^m, F^m, e) \) are not isomorphic.

In fact, they have different output trajectories, when using the scheduling process \( \mu^* \) from Example 3. This is due to the fact that the system \( S \) is not in innovation form.

From Examples 3 and 4 we can conclude that minimality alone does not provide uniqueness.

**Example 5:** We present a minimal asLPV-SSA in innovation form, where the state dimension \( n_x = 2 \) and its matrices are as follows:

\[
A_1 = \begin{bmatrix} 0.4 & 0.4 \\ 0.2 & 0.1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.1 & 0.1 \\ 0.2 & 0.3 \end{bmatrix}
\]

\[
K_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad K_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1 \ 0], \quad \text{and } F = 1
\]

The asLPV-SSA above is stably invertable, and hence in innovation form. We use the same scheduling process \( \mu \) and noise process \( v \) as Example 3. If we apply Algorithm 1 to the system above, we get another asLPV-SSA with the matrices

\[
A_1^m = \begin{bmatrix} 0.4642 & -0.3581 \\ -0.1581 & 0.0358 \end{bmatrix}, \quad K_1^m = \begin{bmatrix} -0.1143 \\ 0.9934 \end{bmatrix}
\]

\[
A_2^m = \begin{bmatrix} 0.1367 & -0.1188 \\ -0.2188 & 0.2633 \end{bmatrix}, \quad K_2^m = \begin{bmatrix} -0.1143 \\ 0.9934 \end{bmatrix}
\]

\[
C^m = \begin{bmatrix} -0.9934 \\ -0.1143 \end{bmatrix}, \quad \text{and } F^m = 1
\]

As expected, the two systems are isomorphic, the corresponding matrix is

\[
T = \begin{bmatrix} -0.9934 & -0.1143 \\ -0.1143 & 0.9934 \end{bmatrix}
\]

Finally, we realize that the output trajectories of both systems are indeed the same not only for the chosen scheduling sequence, but also for any other scheduling process.

VI. CONCLUSION

This paper formulates conditions for a LPV state-space representation to be minimal and in innovation form. These conditions depend only the matrices of the LPV representation. A minimization algorithm for transforming any LPV representation to a minimal one in innovation form is formulated too. These results are expected to be useful for system identification, in particular, for making the identification problem mathematically well-posed. In the future, we will explore the application of the proposed results to concrete system identification algorithms.
REFERENCES

[1] Z. Alkhoury, M. Petreczky, and G. Mercère. Structural properties of affine LPV to LFR transformation: minimality, input-output behavior and identifiability. In Proceedings of the IEEE Conference on Decision and Control, Las Vegas, USA, 2016.

[2] B. A. Bamieh and L. Giarré. Identification of linear parameter-varying systems. International Journal of Robust Nonlinear Control, 12(9):841–853, 2002.

[3] P. Bilingsley. Probability and measure. Wiley, 1986.

[4] Alessandro Chiuso and Giorgio Picci. Consistency analysis of some closed-loop subspace identification methods. Automatica, 41(3):377–391, 2005. Data-Based Modelling and System Identification.

[5] P. Cox, M. Petreczky, and R. Tóth. Towards efficient maximum likelihood estimation of LPV-SS models. Automatica, 97(9):392–403, 2018.

[6] Pepijn Bastiaan Cox and Roland Tóth. Linear parameter-varying subspace identification: A unified framework. Automatica, 123:109296, 2021.

[7] P.L. dos Santos, J.A. Ramos, and J.L.M. de Carvalho. Identification of bilinear systems with white noise inputs: An iterative deterministic-stochastic subspace approach. IEEE Transactions on Control Systems Technology, 17(5):1145–1153, Sept 2009.

[8] W. Favorov, B. De Moor, and P. Van Overschee. Subspace identification of bilinear systems subject to white inputs. IEEE Transactions on Automatic Control, 44(6):1157–1165, Jun 1999.

[9] F. Felici, J. W. van Wingerden, and M. Verhaegen. Subspace identification of MIMO LPV systems using a periodic scheduling sequence. Automatica, 43:1684–1697, 2007.

[10] T. Katayama. Subspace Methods for System Identification. Springer-Verlag, 2005.

[11] N. Kulcsár and Tóth R. On the similarity state transformation for linear Subspace Methods for System Identification.

[12] M. Petreczky, R. Tóth, and G. Mercère. Realization theory for LPV systems. In 3rd IFAC Workshop on Linear Parameter-Varying Systems, Eindhoven, Netherlands, 2019.

[13] M. Mejari and M. Petreczky. Realization and identification algorithm for stochastic LPV state-space models with exogenous inputs. In 3rd IFAC Workshop on Linear Parameter-Varying Systems, Eindhoven, Netherlands, 2019.

[14] A. Lindquist and G. Picci. Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification. Springer Berlin, 2015.

[15] M. Coex, M. Petreczky, and R. Tóth. Towards efficient maximum likelihood estimation of LPV-SS models. Automatica, 97(9):392–403, 2018.

A. Technical definitions

Definition 4 (Admissible scheduling sequences): A stochastic process \( \{ \mu(k) \}_{k \in \mathbb{Z}} \) is called an Admissible scheduling sequence if it satisfies the following properties:

1. Denote by \( \mathcal{F}^\mu_t \) the \( \sigma \)-algebra generated by the random variables \( \{ \mu(k) \}_{k \leq t} \). There exists positive numbers \( \{ p_\sigma \}_{\sigma \in \Sigma} \) such that for any \( w, v \in \Sigma^+ \), \( \sigma, \sigma' \in \Sigma, t \in \mathbb{Z} \):

\[
E[\mu_{w\sigma}(t)\mu_{v\sigma'}(t) | \mathcal{F}^\mu_t] = \begin{cases} p_{\sigma \sigma'} & \text{if } \sigma = \sigma' \\ 0 & \text{otherwise} \end{cases}
\]

2. There exist real numbers \( \{ \alpha_\sigma \}_{\sigma \in \Sigma} \) such that \( \sum_{\sigma \in \Sigma} \alpha_\sigma \mu_{w\sigma}(t) = 1 \) for all \( t \in \mathbb{Z} \).

3. For each \( w, v \in \Sigma^+ \), the process \( \{ \mu_{w\sigma}, \mu_{v\sigma} \} \) is wide-sense stationary.

Definition 5 (ZMWSSI, [20, Definition 2]): A stochastic process \( (r, \mu) \) is Zero Mean Wide Sense Stationary (ZMWSSI) if

1. For \( t \in \mathbb{Z} \), the \( \sigma \)-algebras generated by the variables \( \{ r(k) \}_{k \leq t} \), \( \{ \mu_{w\sigma}(k) \}_{k < t, \sigma \in \Sigma} \) and \( \{ \mu_{v\sigma'}(k) \}_{k \geq t, \sigma \in \Sigma} \), denoted by \( \mathcal{F}^r_t, \mathcal{F}^{\mu_t}_{t-}, \mathcal{F}^{\mu_t}_t \) and \( \mathcal{F}^{\mu_t}_{t+} \) respectively, are such that \( \mathcal{F}^r_t \) and \( \mathcal{F}^{\mu_t}_{t+} \) are conditionally independent w.r.t. \( \mathcal{F}^{\mu_t}_t \).

2. The processes \( \{ r, \{ z_{w\sigma}^r(t) \}_{w \in \Sigma^+} \} \) are zero mean, square integrable and are jointly wide sense stationary, i.e., \( \forall t, s, k \in \mathbb{Z}, v \in \Sigma^+ \):

\[
E[r(t)] = 0, \quad E[z_{w\sigma}^r(t)] = 0
\]

\[
E[r(t + s)](z_{w\sigma}(s + k)) = E[r(t)](z_{w\sigma}^r(s))^T,
\]

\[
E[r(t + s)](r(s + k)) = E[r(t)](r(s))^T,
\]

\[
E[z_{w\sigma}^r(t + s)](z_{w\sigma}(s + k)) = E[z_{w\sigma}^r(t)](z_{w\sigma}^r(s))^T.
\]

Definition 6 (SII process [20, Definition 5]): A process \( r \) is said to be Square Integrable w.r.t. \( \mu \) (abbreviated as SII when \( \mu \) is clear from the context), for all \( w \in \Sigma^+, t \in \mathbb{Z} \), the random variable \( z_{w\sigma}^r(t) = r(t + |w|) \mu_{w\sigma}(t + |w| - 1) \) is square integrable.
B. Proofs of Theorems 3 and 4

First we show the following technical result, which states the following.

**Lemma 3 (Mean-square stability block-diagonal matrices):** Consider matrices \( n_S \times n \) \( \{B_i\}_{i=1}^{n_S} \) and \( n \times n \) matrices \( \{F_i\}_{i=1}^{n_S} \) such that

\[
\sum_{i=1}^{n_S} p_i(F_i \otimes F_i) \text{ is stable.}
\]

Then \( \sum_{i=1}^{n_S} p_i(F_i \otimes F_i) \subseteq \mathbb{D} \) is is stable. Let \( \mathbf{v} \) be a ZMWSII such that \( E[\mathbf{v}(t)(\mathbf{z}_S^R(t))^T] = 0, \) \( w \in \Sigma^+ \). Then the infinite sum

\[
r(t) = \sum_{w \in \Sigma^+, \sigma \in \Sigma} \sqrt{\sigma \omega_w} C G_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)
\]

is absolutely convergent in the mean-square sense, i.e.,

\[
\sum_{w \in \Sigma^+, \sigma \in \Sigma} E[|C G_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)|^2]
\]

and the process \( r(t) \) is ZMWSII and it the unique state process of the asLPV-SSA \( \{\{F_i, B_i\}_{i=1}^{n_S}, C, I, v\} \). From Lemma 5 it follows that \( \mathbf{z}_w(t)^R(t), \mathbf{z}_w^R(t) \) are uncorrelated. Hence

\[
E[|| \sum_{w \in \Sigma^+, |w| \leq N, \sigma \in \Sigma} C G_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)||_2^2] =
\]

\[
\text{trace} \left( \sum_{w \in \Sigma^+, |w| \leq N, \sigma \in \Sigma} C G_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t) \right)^T C G_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)
\]

is stable, and as \( E[|| \sum_{w \in \Sigma^+, |w| \leq N, \sigma \in \Sigma} C G_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)||_2^2] \) is convergent by mean-square convergence of (17), it follows that (18) is convergent.

**Lemma 6:** Assume that \( (\mathbf{y}, \mathbf{\mu}) \) has a asLPV-SSA realization. Consider matrices \( n \times n \) \( \{B_i\}_{i=1}^{n_S} \) and \( n \times n \) matrices \( \{F_i\}_{i=1}^{n_S} \) such that \( \sum_{i=1}^{n_S} p_i(F_i \otimes F_i) \text{ is stable.} \)

Then the infinite sum

\[
\sum_{w \in \Sigma^+, \sigma \in \Sigma} F_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)
\]

converges absolutely in the mean-square sense and \( r(t) = \sum_{w \in \Sigma^+, \sigma \in \Sigma} F_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t) \) is a ZMWSII process, and \( r \) is the unique process \( \bar{r} \) which satisfies

\[
\bar{r}(t + 1) = \sum_{i=1}^{n_S} (F_i \bar{r}(t) + B_i \mathbf{y}(t)) \mu_i(t)
\]

and \( \bar{r}^T \mathbf{y}^T \) is ZMWSII and the components of \( \bar{r} \) belong to the Hilbert-space \( H^2 \) generated by \( \{\mathbf{z}_{\sigma w}(t)\}_{w \in \Sigma^+} \). From Lemma 5 it follows that \( \sum_{i=1}^{n_S} p_i(F_i \otimes F_i) \subseteq \mathbb{D} \) is is stable, and hence by Lemma 5

\[
\bar{s}(t) = \sum_{w \in \Sigma^+, \sigma \in \Sigma} \sqrt{\sigma \omega_w} F_w \mathbf{B}_\sigma \mathbf{z}_{\sigma w}(t)
\]
is absolutely convergent in the mean square sense and it is the unique state process of \( \{ (\bar{F}_i, \bar{B}_i) \}_{i=1}^{n_u}, \tilde{C}, I, \nu \).

From [20, Lemma 2] and Lemma 9 it follows that
\[
\begin{align*}
\bar{x}_{w}(t) = & \sum_{w \in \Sigma^*, \sigma \in \Sigma} CA_w K_{\sigma} z_{w,\sigma}(t) + D z_{w}(t) \\
\end{align*}
\]
and hence, by using (22) it follows that \( H_t^\nu \) is a subspace of the Hilbert-space \( H_{v}^\nu \) generated by \( \{ z_{w,\sigma}(t) \}_{w \in \Sigma^*, \sigma \in \Sigma} \). That is, \( H_{v}^\nu = H_t^\nu \).

Then \( \bar{x} = [\bar{X}^T \bar{Y}^T]^T \) is such that \( \bar{x} \) satisfies the conditions of [20, Lemma 10] for \( r = \nu \). \([\bar{X}^T \bar{Y}^T]^T \) is ZMSII. Moreover, \( E[\nu(t)|\bar{x}(t)\bar{Y}^T] = 0 \) as \( \nu(t) \) is orthogonal to \( H_t^\nu \) and the components of \( z_{w,\sigma}(t) \) belong to \( H_t^\nu \) for all \( w \in \Sigma^* \).

Finally \( \bar{x}(t+1) = \sum_{i=1}^{n_u} (\bar{F}_i \bar{x}(t) + \bar{B}_i \nu(t)) \mu_i(t) \).

Hence, \( \bar{x}(t) \) is a state process of \( \{ (\bar{F}_i, \bar{B}_i) \}_{i=1}^{n_u}, \tilde{C}, I, \nu \) and hence it is unique and equals \( \bar{x} \) and \( \bar{r} = r \).

**Proof:** [Proof of Lemma 1] From Lemma 6 it follows that \( \bar{x}(t) = \sum_{w \in \Sigma^*, \sigma \in \Sigma} F_w K_{\sigma} z_{w,\sigma}(t) \), where \( F_i = (A_i - \tilde{K}_i C) \), \( i = 1, \ldots, n_u \) is absolutely convergent in the mean-square sense. Hence, \( \bar{x}(t) \) converges to zero in the mean square sense. It remains to show that \( \bar{x}(t) = \bar{x}(t) \).

To this end, notice that \( [\bar{X}^T \bar{Y}^T]^T \) is ZMSII by Lemma 6 the elements of \( \bar{x} \) belong to \( H_t^\nu \) and it satisfies
\[
\bar{x}(t+1) = \sum_{i=1}^{n_u} (A_i - \tilde{K}_i C) \bar{x}(t) + \tilde{K}_i \nu(t) \mu_i(t) \]

At the same time, \( [\bar{X}^T \bar{Y}^T]^T \) is ZMSII. Note that the components of \( \bar{x} \) belong to \( H_t^\nu \) as it was pointed out in the proof of Lemma 6 \( H_t^\nu \) equals the Hilbert-space generated by \( \{ z_{w,\sigma}(t) \}_{w \in \Sigma^*} \) and the components of \( \bar{x}(t) \) belong to the latter Hilbert-space. Hence, \( \bar{x}(t) \) satisfies (13), hence by Lemma 6 \( \bar{x}(t) \).

**Proof:** [Proof of Theorem 3] Note that we can write \( \nu(t) = \bar{y}(t) - C \bar{x}(t) \) and hence the first equation of (11) holds, i.e., \( \bar{x}(t+1) = \sum_{i=1}^{n_u} (A_i - \tilde{K}_i C) \bar{x}(t) + \tilde{K}_i \bar{y}(t) \mu_i(t) \). Since the matrix \( \sum_{i=1}^{n_u} \mu_i (A_i - \tilde{K}_i C) \otimes (A_i - \tilde{K}_i C) \) is stable, then by repeating the steps of the proof of [20, Lemma 1] it can be shown that \( \bar{x}(t) = \sum_{w \in \Sigma^*, \sigma \in \Sigma} \sqrt{\mu_{w,\sigma}} \bar{F}_w K_{\sigma} z_{w,\sigma}(t) \) and hence the elements of \( \bar{x}(t) \) belong to the Hilbert-space generated by \( \{ z_{w,\sigma}(t) \}_{w \in \Sigma^*} \). Note that \( E[\nu(t) | \{ z_{w,\sigma}(t) \}_{w \in \Sigma^*}] = 0 \), see the proof of [20, eq. (37), proof of Theorem 4], hence, \( E[\bar{y}(t) | \{ z_{w,\sigma}(t) \}_{w \in \Sigma^*}] = C \bar{x}(t) \) and therefore \( \bar{e}(t) = \nu(t) \).

**Proof:** [Proof of Theorem 3] Note that \( S \) is minimal if and only if the observability and reachability matrices satisfy the following rank conditions rank \( R_{n-1}(S) = n \) and rank \( R_{n-1}(S) = n \). Note that the rows of the extended observability matrix \( \bar{C}_{n-1} \) of the associated \( d \)-LPV-SSA \( D_{S} \) are either zero or they coincide with the rows of the observability matrix \( \bar{C}_{n-1} \), i.e., rank \( \bar{C}_{n-1}(S) = \) rank \( \bar{C}_{n-1} \). That is, \( S \) satisfies the observability rank condition if and only if the \( d \)-LPV-SSA \( D_{S} \) is observable. We will show that \( \text{Im} \bar{R}_{n-1}(S) = \text{Im} \bar{R}_{n-1} \). Since \( \bar{R}_{n-1} \) is the extended controllability matrix of the \( d \)-LPV-SSA \( D_{S} \). From this, it follows that \( S \) satisfies the reachability rank condition if and only if \( LS \) is span-reachable.

Now we will show that \( \text{Im} \bar{R}_{n-1}(S) = \text{Im} \bar{R}_{n-1} \). To this end, we recall that \( \bar{e}(t) \) belongs to the linear space generated by the columns of \( \bar{B}_w K_{w} \), \( w \in \Sigma^*, \sigma \in \Sigma \). Since \( \bar{B}_w = E[\bar{y}(t) | \{ z_{w,\sigma}(t) \}_{w \in \Sigma^*}] \), it then follows that the columns of \( \bar{B}_w \) also belong to the linear space generated by the columns...
of $A_vK_{v\sigma}$, $v \in \Sigma^*$, $\sigma \in \Sigma$. Therefore, the columns of $A_vB_{v\sigma}$, $v \in \Sigma^*$, $\sigma \in \Sigma$ also belong to the linear space generated by the columns of $A_vK_{v\sigma}$, $v \in \Sigma^*$, $\sigma \in \Sigma$. In turn, it is easy to see that the latter subspace equals $\text{Im } R_{n-1}$. That is, $\text{Im } A_vB_{v\sigma}$ is a subspace of $\text{Im } R_{n-1}$, and therefore $\text{Im } R_{n-1}(S) \subseteq \text{Im } R_{n-1}$. Conversely, from [20, eq. (37), proof of Theorem 4] it follows that $E[x(t)] = \sqrt{\eta}A_vB_{v\sigma}$, i.e., for every $w \in \Sigma^+$, the columns $E[x(t)]z_{w}(t)$ belong to the space generated by $A_vB_{v\sigma}$, $\sigma \in \Sigma$, $v \in \Sigma^n$. Notice that by [17, Theorem 2 and Remark 1] applied to the LSS $\Sigma_S$, the latter space equals $\text{Im } R_{n-1}(S)$. Since the elements of $z_{w}(t)$ are limits of finite linear combinations of the rows of $\{z_{w}(t)\}_{w \in \Sigma^+}$, then it follows that the columns of $E[x(t)]$ are limits of finite linear combinations of columns of $E[x(t)]$, $w \in \Sigma^+$, and hence the columns of $E[x(t)]z_{w}(t)^T$, $w \in \Sigma^+$ also belong to $\text{Im } R_{n-1}(S)$. From [20, Proof of Theorem 4] it follows that $K \text{Im } Q_v = E[x(t)]z_{w}(t)^T$, where $Q_v = E[e(t)e(t)^T]z_{w}(t)^T$, and hence the columns of $K \text{Im } Q_v$ belong to $\text{Im } R_{n-1}(S)$. Since $Q_v$ is non-singular, then it follows that the columns of $K \text{v\sigma}$ belong to $\text{Im } R_{n-1}(S)$. Since $\text{Im } R_{n-1}(S)$ is $A_v$-invariant for all $\sigma \in \Sigma$ and $A_v = 0$, it then follows that $\text{Im } A_vK_{v\sigma} \subseteq \text{Im } R_{n-1}(S)$ for all $\sigma \in \Sigma$, $v \in \Sigma^n$, $\sigma \in \Sigma^n$, and thus $\text{Im } R_{n-1}(S) \subseteq \text{Im } R_{n-1}(S)$.}

C. Proof of Lemma 2

Definition 7 (Necessary Maps): Define, for the sequel, the following maps:

$$\mathcal{Y} : (u, \mu) \mapsto y = \mathcal{Y}(u, \mu)$$

$$\mathcal{F} : D \mapsto \mathcal{F}(D)$$

$$\mathcal{D} : D \mapsto \mathcal{D}(D)$$

where $\mathcal{Y}$ is the input-output map, $D = \{A_1, K_1\}_{i=1}^{n_u}, C, I$ is a dLPV-SSA and $\mathcal{F}(D)$ and $\mathcal{D}(D)$ are two transformed dLPV-SSAs with $\mathcal{F}(D) = \{A_{1-}, K_1\}_{i=1}^{n_u}, C, I$ and $\mathcal{D}(D) = \{A_{1-}, K_1\}_{i=1}^{n_u}, -C, I$.

Note that, if $X$ is a dLPV-SSA representation, then it is clear that $\mathcal{F}(\mathcal{F}(X)) = X$.

Lemma 7: Consider the following dLPV-SSA: $D = \{A_1, K_1\}_{i=1}^{n_u}, C, I$ where it is a realization of the sub-Markov parameters $M_D = CA_w K_{w\sigma}$, where $w \in \Sigma^*$, $\sigma \in \Sigma$. Subsequently, $\mathcal{F}(D)$ is a realization of $M_{\mathcal{F}(D)} = CA_w K_{w\sigma}$ with $\bar{A}_1 = A_1K_1C$. If a dLPV-SSA $D' = \{A', K'\}_{i=1}^{n_u}, C, I$ is a realization of $M_D$, then the $\mathcal{F}(D')$ dLPV-SSA is a realization of $M_{\mathcal{F}(D)}$.

Proof: [Proof of Lemma 7] Consider that $(x, u, \mu, y)$ is a solution of the system $\mathcal{D}$ if it then follows that

$$x(t+1) = \sum_{i=1}^{n_u} (A_i x(t) + K_i y(t) + u(t)) \mu_i(t)$$

$$y(t) = C x(t)$$

It can then be modified to

$$x(t+1) = \sum_{i=1}^{n_u} (A_i x(t) + K_i (C x(t) - C x(t) + u(t))) \mu_i(t)$$

Subsequently, it is safe to say that $(x, v, \mu, y)$ is a solution of the system $\mathcal{F}(D)$ with $v = y + u$. It can also be shown, with the same demonstration, that if $(x, v, \mu, y)$ is a solution of $\mathcal{F}(D)$, then $(x, u, \mu, y)$ is a solution of $\mathcal{D}$, with $u = v - y$. That is, it can be said that $\mathcal{F}(D), 0(\mu, u) = y = \mathcal{F}(D), 0(\mu, u) = \mathcal{F}(D), 0(\mu, u)$.

Consider now that $(x, u, \mu, y)$ is a solution of $D'$, it is safe to say that $(x, v, \mu, y)$ is a solution of $\mathcal{F}(D)$, with $v = u + y$. Lemma 2 assumes that $D$ and $D'$ are both realization of the sub-Markov parameters $M_D$. This leads to the following conclusion:

$$\mathcal{F}(D), 0(v, \mu) = \mathcal{F}(D), 0(u, \mu) = \mathcal{F}(D), 0(\mu, u) = \mathcal{F}(D), 0(\mu, u)$$

In other words, $\mathcal{F}(D')$ is a realization of $M_{\mathcal{F}(D)}$. □

Corollary 2: An dLPV-SSA representation $\Sigma$ is a minimal realization of $M_{\Sigma}$ if and only if $\mathcal{F}(\Sigma)$ is a minimal realization of $M_{\mathcal{F}(\Sigma)}$.

Proof: Suppose that $\Sigma$ is minimal and $\mathcal{F}(\Sigma)$ is not. Which means that, there exists a minimal dLPV-SSA representation $\Sigma'$ such that $\dim \Sigma' < \mathcal{F}(\Sigma) = \dim \Sigma$. Define a dLPV-SSA representation $\Sigma'' := \mathcal{F}(\mathcal{F}(\Sigma'))$. It is known that $\Sigma''$ and $\mathcal{F}(\Sigma)$ are both a realization of $M_{\Sigma''} = M_{\mathcal{F}(\Sigma)}$, then, by applying $\mathcal{F}(\mathcal{F}(\cdot))$ on both sides, we will get the following:

$$M_{\mathcal{F}(\mathcal{F}(\Sigma'))} = M_{\mathcal{F}(\mathcal{F}(\Sigma)))} \Rightarrow M_{\Sigma} = M_{\Sigma''}$$

Which means that $\Sigma''$ is a realization of $M_{\Sigma}$ and $\mathcal{F}(\Sigma) = \dim \mathcal{F}(\Sigma) > \dim \Sigma'' = \dim \Sigma''$, which implies that $\Sigma$ is not minimal, and results to a contradiction. □

Proof: [Proof of Lemma 2] In order to prove Lemma 2 two steps are needed: (1) proving that $S_m$ is stably invertible, and (2) proving that $S_m$ is a realization of $(y, \mu)$ in innovation form.

1. From Lemma 7 it follows that $\mathcal{F}(D)$ and $\mathcal{D}(D)$ are both a realization of $M_{\mathcal{F}(D)}$, this is due to the fact that $D$ and $D_m$ are both a realization of $M_{\mathcal{F}(D)}$. Recall from [21, Lemma 6] that if the matrix $\sum_{i=1}^{n_u} (\sqrt{\mu_i} \sigma_i A_i) \otimes (\sqrt{\mu_i} \sigma_i B_i) = \sum_{i=1}^{n_u} \mu_i \sigma_i A_i \otimes \sigma_i A_i$ is stable then the matrix $\sum_{i=1}^{n_u} \mu_i \sigma_i A_i \otimes \sigma_i A_i$ is also stable, where $\sigma_i A_i$ are the state matrices of an dLPV-SSA $\Sigma_1$ and $\sigma_i A_i$ are the state matrices of the minimized dLPV-SSA $\Sigma_1'. This said, recall that Algorithm 2 assumes that its input $S = \{A_{1-}, K_1\}_{\sigma=1}^{n_u}, C, I_n\}$ is stably invertible. It follows that the matrix $\sum_{i=1}^{n_u} \mu_i (A_i - K_1C) \otimes (A_i - K_1C)$ is stable, where $(A_i - K_1C)$ are the state matrices of $\mathcal{F}(D)$. Note that $\mathcal{F}(D_m)$ is minimal due to the minimality of $D_m$. It follows, from [21, Lemma 6], that the matrix $\sum_{i=1}^{n_u} \mu_i (A_i - K_{m}C) \otimes (A_i - K_{m}C)$ is stable, where $(A_i - K_{m}C)$ are the state matrices of $\mathcal{F}(D_m)$. Therefore, the algorithm’s output $S_m$ is indeed stably invertible.

2. ($S_m$ is a realization of $(y, \mu)$ in innovation form, then there exists a process $x$ such that (1) holds with $v = y$. This

5 The terminology of [21] refers to the switch systems representations not the LPV-SSA representations.
said, let us apply the linear transformation $T$ as described in [19, Corollary 1] to $S$. Let $\tilde{x}(t) = T x(t)$, $\dot{A}_i = TA_i$, $\dot{K}_i = TK_i$, for $i = 1, \ldots, n_\mu$, and $\dot{C} = CT^{-1}$. Then the asLPV-SSA $\tilde{S} = (\dot{A}_1, \dot{K}_1)_{\mu=1}^n, \dot{C}, I, e$ is also a realization of $(y, \mu)$ in innovation form. Now, in order to get the Kalman decomposition, recall form [19, Corollary 1] that:

$$\dot{A}_i = \begin{bmatrix} A_{i}^m & 0 & A_i^\prime \nonumber \\ A_i^\prime & A_i'' & A_i''' \\ 0 & 0 & A_i^{uc} \nonumber \end{bmatrix}, \quad \dot{K}_i = \begin{bmatrix} \dot{K}_i^m \\ \dot{K}_i^\prime \\ 0 \nonumber \end{bmatrix}, \quad \dot{C} = [C^m \ 0 \ C']$$

for $i = 1, \ldots, n_\mu$, and for suitable block matrices. In particular, for all $w \in \Sigma^*$ we found:

$$\hat{A}_w = \begin{bmatrix} A_{w}^m & 0 & * \\ * & A_{w}'' & * \\ 0 & 0 & A_{w}^{uc} \nonumber \end{bmatrix}$$

(27)

where $*$ refers to a block matrix. Indeed, it can be shown that for $w = e$, $\hat{A}_w = I$. Now if (27) holds for $w = v$, then for $w = vd$:

$$\hat{A}_w = \begin{bmatrix} A_{vd}^m & 0 & * \\ * & A_{vd}'' & * \\ 0 & 0 & A_{vd}^{uc} \nonumber \end{bmatrix}$$

Hence, by induction of the length of $w$, (27) holds. It the follows that:

$$\hat{A}_w \hat{K}_\sigma = \begin{bmatrix} A_{w}^m \hat{K}_\sigma^m \\ * \\ 0 \nonumber \end{bmatrix}$$

Consider the following decomposition:

$$\tilde{x}(t) = T x(t) = \begin{bmatrix} x^m(t) \\ x^e(t) \\ x^{uc}(t) \nonumber \end{bmatrix}$$

From Lemma $\tilde{x}(t)$ can be expressed as:

$$\tilde{x}(t) = \sum_{w \in \Sigma^*, \sigma \in \Sigma} \sqrt{p_{\sigma w}} \hat{A}_w \hat{K}_\sigma z_{\sigma w}(t)$$

It follows that:

$$\tilde{x}(t) = \sum_{w \in \Sigma^*, \sigma \in \Sigma} \sqrt{p_{\sigma w}} \begin{bmatrix} A_{w}^m \hat{K}_\sigma^m z_{\sigma w}(t) \\ * \\ 0 \nonumber \end{bmatrix} = \begin{bmatrix} \sum_{w \in \Sigma^*, \sigma \in \Sigma} \sqrt{p_{\sigma w}} A_{w}^m \hat{K}_\sigma^m z_{\sigma w}(t) \\ * \\ 0 \nonumber \end{bmatrix}$$

(28)

From (28) it follows that $x^m(t) = \sum_{w \in \Sigma^*, \sigma \in \Sigma} \sqrt{p_{\sigma w}} A_{w}^m \hat{K}_\sigma^m z_{\sigma w}(t)$. Then from [20, Lemma 3], it follows that $x^m(t)$ is the state process of $S_m$. Finally, notice that, as $\tilde{S}$ is a realization of $(y, \mu)$, $y(t) = C x(t) + e(t) = CT \tilde{x}(t) + e(t) = \tilde{C} \tilde{x}(t) + e(t). The output can then be expressed as:

$$y(t) = \begin{bmatrix} C^m & 0 & C' \nonumber \end{bmatrix} \begin{bmatrix} x^m(t) \\ x^e(t) \\ x^{uc}(t) \nonumber \end{bmatrix} + e(t)$$

(29)

From (28) it follows that $x^{uc}(t) = 0$. Hence, from (29) it follows that $y(t) = C^m x^m(t) + e(t)$. Since $x^m$ is the state process of $S_m$, it then follows that $S_m$ is a realization of $(y, \mu)$. In addition, because $S_m$ is stably invertible, as proven in (1), then $S_m$ is a realization of $(y, \mu)$ in innovation form.