SIAM J. CONTROL OPTIM.
Vol. 54, No. 4, pp. 2063–2089




# AVERAGE REACHABILITY OF CONTINUOUS-TIME MARKOV JUMP LINEAR SYSTEMS AND THE LINEAR MINIMUM MEAN SQUARE ESTIMATOR*

ALFREDO R. R. NARVÁEZ† AND EDUARDO F. COSTA‡

**Abstract.** In this paper we study the average reachability gramian for continuous-time linear systems with additive noise and jump parameters driven by a general Markov chain. We define a rather natural reachability concept by requiring that the average reachability gramian be positive definite. Aiming at a testable condition, we introduce a set of reachability matrices for this class of systems and employ invariance properties of the null space of the noise coefficient matrices to show that the system is reachable if and only if these matrices are of full rank. We also show for reachable systems that the state second moment is positive definite. One consequence of this result in the context of linear minimum mean square state estimation for reachable systems is that the expectation of the error covariance matrix is positive definite. Moreover, the average boundedness of the error covariance matrix is invariant to a type of perturbation in the noise model, meaning that the estimates are not overly sensitive, which consists in a property that is desirable in applications and sometimes referred to as stability of the estimator.

**Key words.** Markov jump linear systems, stochastic systems, continuous-time systems, reachability of stochastic systems, filtering

**AMS subject classifications.** 93E03, 93B05, 60J27, 93E11

**DOI.** 10.1137/140972974

## 1. Introduction.

Practical systems are frequently vulnerable to abrupt changes in their behavior due to, for example, component failures or sudden environmental disturbances. In this paper, we address Markov jump linear systems (MJLSs) whose parameters are governed by the state of a Markov chain $\{\theta(t)\}_{t\geq 0}$. Applications of MJLSs include systems with random failures and plants composed of subsystems whose interaction changes abruptly, as frequently found in the context of networked control systems, which have been attracting increasing attention in recent years. Examples of applications are given in [19, 20, 30, 32].

Some structural properties of MJLSs, such as stochastic controllability and stochastic stabilizability, have been studied in great detail in the seminal paper [22] as well as in [6, 7, 9, 13, 14, 15, 23, 25]. There is a series of notions under the denomination of *weak concepts*, like weak observability and weak detectability [3, 4, 5, 12], which involve cost functionals that can be expressed in terms of the average observability gramian. The role of these concepts in control is now well known. Recently we have introduced the idea of weak controllability [26, 27] relying on the average controllability gramian, obtaining a few preliminary results for that notion. A similar notion of weak reachability is largely unexplored, both in terms of definition/characterization

*Received by the editors June 23, 2014; accepted for publication (in revised form) May 16, 2016; published electronically August 18, 2016.

http://www.siam.org/journals/sicon/54-4/97297.html

**Funding:** This work was supported by CNPq grants 306466/2010-4 and 311290/2013-2, by FAPESP grants 2010/12360-3 and 2013/19380-8, and by CAPES.

†Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, C.P. 668, 13560-970, São Carlos, SP, Brazil, and Facultad de Ingeniera, Universidad de la Costa, Barranquilla, Colombia (aroa8@cuc.edu.co).

‡Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, C.P. 668, 13560-970, São Carlos, SP, Brazil (efcosta@icmc.usp.br).

and of its role in control and filtering.

In this paper, we introduce some notions of average reachability (AR).[1] AR is defined in a control-free setup, with a clear, marked dissimilarity in its conception with available notions of stochastic reachability (SR) which play an important role in control problems aiming to bring the initial state to a prespecified set; see, e.g., [1]. The fact that AR and SR are in general not comparable is illustrated in Examples 2 and 3. Regarding the relation between AR and the aforementioned weak notions, they are not comparable since the average positiveness of the observability, controllability, reachability, and reconstructibility gramians is not comparable in the sense that the positiveness of one gramian does not imply positiveness of another gramian. It is worth mentioning that another situation may be found when comparing gramians for different systems; e.g., it is widely known in the case of linear deterministic systems that the reachability gramian of $(A, B)$ equals the observability gramian of $(A', B')$, and we employ this type of construction in section 2.3 in order to explore results for weak observability derived in [26].

This work adds to the theory of MJLSs by developing computational simple tests for the AR notions and studying their role in the filtering of MJLSs. We study the relation between AR and the positiveness of the conditional second moment matrices, with the interpretation that there is a "minimal level of noise" in the state, as detailed in section 4.1. In the context of linear minimum mean square estimation (LMMSE) of the continuous-valued component of the state $x(t)$ with observation of the jump variable, as proposed for MJLSs in [8, 17], we show that AR ensures two properties: (i) the error covariance matrix $Q(t)$ is a positive definite matrix (essentially meaning that the estimate has a minimal level of noise), and (ii) the property that $Q(t)$ is average bounded from above is invariant to certain perturbations on the noise model, which is sometimes referred to as *stability* of the filter.

In contrast to the fact that AR is defined in a simple manner in terms of the reachability gramian, the tasks of obtaining an efficient test and studying the characterization of noise diffusion are quite complex. The main issues and ideas are briefly discussed below. Consider the system defined on the fundamental space of probability $(\Omega, \mathfrak{F}, \mathcal{P})$ by

$$dx(t) = A_{\theta(t)}x(t)dt + B_{\theta(t)}dw(t), \qquad x(0) = x_0, \quad \theta(0) \sim \pi, \tag{1.1}$$

where $x(t) \in \mathbb{R}^n$, $x_0$ is a random variable, and $\{w(t)\}_{t\geq 0}$ is a Wiener process. Here, $\{\theta(t)\}_{t\geq 0}$ is a general Markov chain, and we only require finite dimension for its state space $\mathcal{S}$. At each time instant $t$, we have that $A_{\theta(t)} = A_i$ whenever $\theta(t) = i$, and similarly for the other parameter matrix in (1.1). The AR notions for the system (1.1) require that the expected value of the reachability gramian be positive, $\mathcal{E}\{\Upsilon_{\text{rch}}(0, t)\} > 0$ for some $t > 0$, $i \in \mathcal{S}$, where the reachability gramian is defined according to [10],

$$\Upsilon_{\text{rch}}(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_1, \tau)B_{\theta(\tau)}B'_{\theta(\tau)}\Phi(t_1, \tau)'d\tau. \tag{1.2}$$

We shall consider two notions of AR requiring positiveness of the gramian in different senses: AR requires positiveness for any initial distribution $\pi$ of the Markov chain; $\pi$-AR considers a particular given initial distribution $\pi$. We also devote a small portion

[1]We use the terminology *average* instead of *weak* as it is more accurate in regard to the studied notion. The word *weak* has different usages in applied mathematics, having little relation with this work; moreover, AR does not necessarily weaken previous reachability notions; see Example 2.

of the paper to $\Pi$-AR, which assumes $\pi \in \Pi$ in such a manner that transient states are never visited; see Corollary 4.7.

Checking AR directly by definition is obviously inadvisable. A first attempt to obtain a testable condition is via the second moment matrix $X_i = \mathcal{E}\{x(t)x(t)1_{\{\theta=i\}}\}$, $i \in \mathcal{S}$; however, the equation describing its evolution involves the distribution of the Markov chain at time instant $t$, $p_i(t)$, and the conditions that can be directly obtained in this fashion turn out to be dependent on $p_i(t)$. To overcome this, we study an equation in certain variables $S_i$ that dismisses terms related to $p_i(t)$; see Remark 3 for an interpretation of $S_i$. The next step is to show that the null space of $S_i$ is identical to the null space of $\mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,t)|\theta(0)=i\}$. This involves the time derivatives of $S_i$, evaluated by recasting invariance properties in [27], that in turn require using an auxiliary system involving matrices $A', B'$ and a "time-reverse" Markov chain to explore the dual relation with weak observability, as mentioned before. The time derivatives of $S_i$ allow for the construction of a collection of matrices whose rank is central to the test of AR as presented in Theorem 3.6.

The positiveness of $X$ is then tackled using a different approach. We show that $S_i$ is positive definite whenever $i$ is a recurrent Markov state. A positive lower bound for $S_i(t)$ in a finite time interval is then obtained by using the Weierstrass extreme value theorem together with a procedure of "restarting" $S$ (as in (4.10)), yielding $S(t) \geq \gamma I$. Unfortunately, $X$ is not an upper bound for $S$, obliging us to introduce a scaling factor $\delta$ to get that $\delta^{-1}X > S(t)$; see Lemma 4.4 for details. This allows us to demonstrate that $\mathcal{E}\{x(t)x'(t)\}$ and $X_i$, $i$ recurrent, are bounded from below by positive definite matrices; see Theorem 4.5.

Positiveness of $X$ has some relevant implications in filtering for $x(t)$, and in particular for the LMMSE proposed for MJLSs in [17, 8]. This estimator is attracting considerable attention in applications as it is much easier to implement than the Kalman filter due to the fact that the Kalman filter gains require solving a differential equation online, while the gains of the LMMSE can be precomputed. Apart from that, the LMMSE is optimal in the class of Markovian linear estimators. We show average positiveness for the error covariance of the LMMSE with the interpretation that its estimates contain a "minimum level of randomness" for AR systems, which can be extended to any linear Markovian filter. This could seem as a drawback for the LMMSE; however, it is precisely this property which provides the arguments to study the so-called *stability* of the estimator, a relevant feature for filters employed in practical systems [2, 28, 29]. In essence, stability means that an estimator whose error covariance is average bounded from above keeps the boundedness even under perturbations on the initial covariance matrix; otherwise the error covariance could tend to infinity, leading the estimate quality to quickly deteriorate. We show that the LMMSE for AR systems is stable.

The paper is organized as follows. Section 2 presents notation and our first results on the reachability gramian for MJLSs. The $\pi$-AR, a computational test for its characterization, and examples compared with other reachability notions are given in section 3. The AR and the positiveness of the state second moment is studied in section 4, and its application in filtering for MJLSs is presented in section 5. Section 6 gives concluding remarks.

## 2. Preliminaries.

Basic assumptions and notation used throughout the paper are presented next. Preliminary results on the reachability gramian and an invariance property of the null space of matrices $B_j$ are also presented.

**2.1. Basic assumptions.** The following hypotheses are considered in this paper.

ASSUMPTION 1. *$\{\theta(t)\}_{t\geq 0}$ is a Markov chain with finite state space denoted by $\mathcal{S} = \{1, \ldots, N\}$. The transition rate matrix of the chain is denoted by $\Lambda = [\lambda_{ij}]$, with $\lambda_{ij} = \lim_{\Delta t \to 0} \left[\mathcal{P}(\theta(t+\Delta t) = j|\theta(t) = i) - 1_{\{i=j\}}\right]/\Delta t$.*

ASSUMPTION 2. *$\{w(t)\}_{t\geq 0}$ is a Wiener process of dimension $n_w$ with incremental covariance operator $I_w dt$, where $I_w$ is the identity matrix of dimension $n_w \times n_w$.*

ASSUMPTION 3. *$x_0$ is a random variable satisfying $\mathcal{E}\{x_0\} = \bar{x}_0$ and $\mathcal{E}\{x_0 x_0'\} = \Psi$. $x_0$, $\theta(0)$, and the process $\{w(t)\}_{t\geq 0}$ are mutually independent.*

ASSUMPTION 4. *$A = (A_1, \ldots, A_N)$ is a set of matrices of dimensions $n \times n$, and $B = (B_1, \ldots, B_N)$ is made of matrices of dimensions $n \times n_w$. These sets are given and fixed, and the same is valid for $\Lambda$.*

*Remark* 1. Assumptions 1–3 represent minimal properties of the stochastic processes in the plant. Assumption 1 indicates that the variable $\theta(t)$ is the state of a standard homogeneous Markov chain with finite dimension, and it allows for rather general setups containing transient states, cemetery states, and noncommunicating groups of states, and of course *the chain is not required to be ergodic*, which significantly enlarges the set of applications for our results when compared with works that require ergodic chains. Assumptions 2 and 3 impose basic regularity hypotheses frequently found in the literature; there is no loss of generality in assuming unitary covariance for $w$ as other cases are covered by adjusting $B$. Assumption 4 is standard in the context of MJLSs with known parameters.

**2.2. Notation.** Let $\mathbb{Z}^0$ be the nonnegative integers. Let $\mathfrak{R}^{n,q}$ (respectively, $\mathfrak{R}^n$) be the linear space formed by all matrices of size $n \times q$ (respectively, $n \times n$) and $\mathfrak{R}^{r0}$ the closed convex cone of symmetric positive semidefinite matrices $\{U \in \mathfrak{R}^r : U = U' \geq 0\}$ ($\mathfrak{R}^{r+}$ when all matrices are positive definite). For $U \in \mathfrak{R}^{n,q}$, $\mathcal{N}(U)$ and $\mathcal{I}m(U)$ represent the null and the image space of $U$, respectively. $U'$ denotes the transpose of $U$; $U \geq V$ means that $U - V \in \mathfrak{R}^{r0}$, and $U > V$ means that $U - V \in \mathfrak{R}^{r+}$. The operator $1_{\{.\}}$ is the indicator function, $\mathcal{E}\{\cdot\}$ is the expectation operator, and $\mathrm{tr}\{.\}$ denotes the trace. Let $\mathcal{M}^{r,n}$ be the linear space formed by a number $N$ of matrices such that $\mathcal{M}^{r,n} = \{U = (U_1, \ldots, U_N) : U_j \in \mathfrak{R}^{r,n}, j = 1, \ldots, N\}$; also, $\mathcal{M}^r \equiv \mathcal{M}^{r,r}$. We denote by $\mathcal{M}^{r0}$ ($\mathcal{M}^{r+}$) the set $\mathcal{M}^r$ when it is formed by $U_j \in \mathfrak{R}^{r0}$ ($U_j \in \mathfrak{R}^{r+}$) for all $j = 1, \ldots, N$. $\mathcal{M}^{r,n}$ with the inner product given by

$$\langle U, V \rangle = \sum_{j=1}^{N} \mathrm{tr}\{U_j' V_j\}$$

is a Hilbert space. We use the norm $||U|| = \langle U, I \rangle$ in $\mathcal{M}^{n0}$. Following [7], we employ the linear and invertible operators vec and $\varphi$, where vec is the operator that stacks the columns of a matrix in a single column vector, and $\varphi$ is defined in such a manner that, for $U \in \mathcal{M}^n$, $\varphi(U) = [\mathrm{vec}(U_1)' \ \mathrm{vec}(U_2)' \ \cdots \ \mathrm{vec}(U_N)']'$.

Regarding the Markov chain, we consider $\mathbb{P}(t) = [p_{ij}(t)]$, $t \geq 0$, the transition semigroup on $\mathcal{S}$, where $p_{ij}(t) = \mathcal{P}(\theta(t+s) = j|\theta(s) = i)$, $i, j \in \mathcal{S}$. We set $p(t) = [p_1(t), \ldots, p_N(t)]$ with $p_i(t) = \mathcal{P}(\theta(t) = i)$ and $p(0) = \pi$ the initial distribution probability of the Markov chain. For a given $\pi$ we denote

$$\mathcal{S}_\pi = \{j \in \mathcal{S} : \exists s \geq 0 \text{ such that } \mathcal{P}(\theta(s) = j) > 0\},$$

which comprises Markov states with positive initial distribution along with the ones that can be reached from them. $\mathcal{S}_{\pi,\mathrm{rec}}$ stands for the subset of $\mathcal{S}_\pi$ formed by its recurrent Markov states.

Now, let the operators $\mathcal{L}_A : \mathcal{M}^n \longrightarrow \mathcal{M}^n$ and $\mathcal{T}_A : \mathcal{M}^n \longrightarrow \mathcal{M}^n$ be defined as

$$\mathcal{L}_{A,j}(U) = A_j'U_j + U_jA_j + \sum_{i=1}^{N} \lambda_{ji}U_i, \qquad \mathcal{T}_{A,j}(U) = A_jU_j + U_jA_j' + \sum_{i=1}^{N} \lambda_{ij}U_i \tag{2.1}$$

for $j = 1, \ldots, N$. We denote $\mathcal{L}_A^0(U) = U$ and, for $k \geq 1$, $\mathcal{L}_A^k(U) = \mathcal{L}_A(\mathcal{L}_A^{k-1}(U))$, and we use similar notation for $\mathcal{T}_A$. $\mathcal{T}_A$ and $\mathcal{L}_A$ are self-adjoint operators. For $U, V \in \mathcal{M}^n$ it is simple to check that $\langle U, V\rangle = \varphi(U)'\varphi(V)$. As in [3], using this notation, we can write

$$\varphi(\mathcal{L}_A(U)) = \mathcal{A}\varphi(U) \quad \text{and} \quad \varphi(\mathcal{T}_A(U)) = \mathcal{A}'\varphi(U), \tag{2.2}$$

where

$$\mathcal{A} = \begin{bmatrix} \hat{A}_1 + \lambda_{11}\hat{I} & \ldots & \lambda_{1N}\hat{I} \\ \lambda_{21}\hat{I} & \ddots & \lambda_{2N}\hat{I} \\ \vdots & \vdots & \vdots \\ \lambda_{N1}\hat{I} & \ldots & \hat{A}_N + \lambda_{NN}\hat{I} \end{bmatrix}, \tag{2.3}$$

with $\hat{I} = I \in \Re^{n^2}$, $\hat{A}_j = V \otimes A_j' + A_j' \otimes V$, where $V = I \in \Re^n$ and $\otimes$ stands for the Kronecker product.

**2.3. An invariance result.** In this section we obtain an invariance property for the null space of matrix $B_j$. We proceed by exploring a connection between the reachability and observability gramians via the auxiliary system $(\tilde{A}, \tilde{B} = 0, \tilde{C}, \tilde{D} = 0, \tilde{\Lambda})$,

$$\begin{aligned} \dot{x}(t) &= \tilde{A}_{\tilde{\theta}(t)}x(t), \quad x(0) = x_0, \quad \tilde{\theta}(0) \sim \tilde{\pi}, \\ y(t) &= \tilde{C}_{\tilde{\theta}(t)}x(t), \end{aligned} \tag{2.4}$$

and the associated observability gramian

$$\Upsilon_{\mathrm{obs}}(0,t) = \int_0^t \tilde{\Phi}(\tau,0)'\tilde{C}_{\tilde{\theta}(\tau)}'\tilde{C}_{\tilde{\theta}(\tau)}\tilde{\Phi}(\tau,0)d\tau, \tag{2.5}$$

where $\tilde{\Phi}(\tau,0)$ is the state transition matrix linked with the solution of (2.4). Consider

$$W^t(x,j) = x'\mathcal{E}\left\{\Upsilon_{\mathrm{obs}}(0,t)\,|\,\tilde{\theta}(0) = j\right\}x \tag{2.6}$$

for all $x \in \mathbb{R}^n$ and $j \in \mathcal{S}$, which is usually interpreted as a cost functional. Consider also the sequence of sets of matrices $O(k) \in \mathcal{M}^n$, defined recursively by

$$\begin{aligned} O_j(k+1) &= \mathcal{L}_{\tilde{A},j}(O(k)), \qquad k \in \mathbb{Z}^0,\ j \in \mathcal{S}, \\ O_j(0) &= \tilde{C}_j'\tilde{C}_j. \end{aligned} \tag{2.7}$$

Matrices $O$ are associated to the observability of $(\tilde{A}, 0, \tilde{C}, 0, \tilde{\Lambda})$, as explained in [3]. The following is a direct adaptation of results in [26], linking $O$ with $W$ and giving an invariance property for the null space of $\tilde{C}$ that resembles a result for systems with general switching rules [31, Theorem 4.26].

PROPOSITION 2.1. *Consider the system* $(\tilde{A}, \tilde{B} = 0, \tilde{C}, \tilde{D} = 0, \tilde{\Lambda})$. *For each* $x \in \mathbb{R}^n$, *the following statements are equivalent:*

(i) $W^t(x, j) = 0$, $t \geq 0$, $j \in \mathcal{S}_\pi$.

(ii) $x' O_j(k) x = 0$, $k \in \mathbb{Z}^0$, $j \in \mathcal{S}_\pi$.

(iii) $\tilde{C}_i \tilde{A}_{i_m}^{p_0} \tilde{A}_{i_{m-1}}^{p_1} \cdots \tilde{A}_j^{p_m} x = 0$ *for any sequence of Markov states* $j, i_1, \ldots, i_m, i$ *contained in* $\mathcal{S}_\pi$ *such that* $\tilde{\lambda}_{j, i_1} \tilde{\lambda}_{i_1, i_2} \cdots \tilde{\lambda}_{i_m, i} \neq 0$ *with* $p_\ell \geq 0$, $\ell = 0, \ldots, m$.

*Proof.* (i) $\Leftrightarrow$ (ii). [3, Proposition 5(ii)] and [3, Lemma 8(i),(ii)] lead to the result. (ii) $\Rightarrow$ (iii). See [26, Corollary 3]. (iii) $\Rightarrow$ (ii). We have

$$x' \tilde{A}_j'^{q_k} \cdots \tilde{A}_{j_{k-1}}'^{q_1} \tilde{A}_{j_k}'^{q_0} \tilde{C}_i' \tilde{C}_i \tilde{A}_{i_m}^{p_0} \tilde{A}_{i_{m-1}}^{p_1} \cdots \tilde{A}_j^{p_m} x = 0 \tag{2.8}$$

for any sequences of Markov states $j, i_1, \ldots, i_m, i$ and $j, j_1, \ldots, j_k, i$ such that $\tilde{\lambda}_{j, i_1} \tilde{\lambda}_{i_1, i_2} \cdots \tilde{\lambda}_{i_m, i} \neq 0$ and/or $\tilde{\lambda}_{j, j_1} \tilde{\lambda}_{j_1, j_2} \cdots \tilde{\lambda}_{j_m, i} \neq 0$, with $p_\ell, q_\nu \geq 0$, $\ell = 0, \ldots, m$, $\nu = 0, \ldots, k$. By expanding $x' O_j(k) x$ using (2.7) one gets sums of expressions like the left-hand side of (2.8) (each one multiplied by a product of probability rates), so that employing (2.8) yields $x' O_j(k) x = 0$. ∎

In the particular case of "no jumps" where $N = 1$, $\tilde{\Lambda} = \Lambda = 0$ and $\tilde{A}_1 = A_1'$, $\tilde{C}_1 = B_1'$, it is well known that $\Upsilon_{\text{obs}}(0, t) = \Upsilon_{\text{rch}}(0, t)$ (perhaps it is more disseminated that reachability of $(A_1, B_1)$ is equivalent to observability of $(A_1', B_1')$ in the standard sense for deterministic systems). Next we extend this property to MJLSs. We will consider the system in (2.4) with the setting

$$\begin{cases} \tilde{A} = A', \ \tilde{B} = 0, \ \tilde{C} = B', \ \tilde{D} = 0, \\ \tilde{\lambda}_{\kappa\ell} = \lambda_{\ell\kappa} \text{ for } \kappa \neq \ell, \text{ and } \tilde{\lambda}_{\ell\ell} = -\sum_{\kappa \neq \ell} \tilde{\lambda}_{\ell\kappa}. \end{cases} \tag{2.9}$$

LEMMA 2.2. *Consider the system* (1.1) *and the system* (2.4) *with the parameters in* (2.9). *Then, for each realization of the Markov chain starting at* $i$ *and visiting* $j$ *at time* $t \geq 0$, *with* $i, j \in \mathcal{S}_\pi$ *communicating Markov states, we have*

$$\Upsilon_{\text{rch}}(0, t) = \Upsilon_{\text{obs}}(0, t).$$

*Proof.* Let us consider the sequence $j =: i_0, i_1, \ldots, i_{m-1}, i_m := i$ such that $\tilde{\lambda}_{j, i_1} \tilde{\lambda}_{i_1, i_2} \cdots \tilde{\lambda}_{i_{m-1}, i} \neq 0$ with its respective sequence of jump times $0 \leq s_1 \leq \ldots \leq s_m \leq t$. By denoting $s_0 := 0$ and $s_{m+1} := t$, we have

$$\begin{aligned} \Upsilon_{\text{obs}}(0, t) = &\sum_{k=0}^{q} \int_{s_k}^{s_{k+1}} \tilde{\Phi}(s, 0)' \tilde{C}_{i_k}' \tilde{C}_{i_k} \tilde{\Phi}(s, 0) ds = \int_0^{s_1} e^{\tilde{A}_j' s} \tilde{C}_j' \tilde{C}_j e^{\tilde{A}_j s} ds \\ &+ \cdots + e^{\tilde{A}_j' s_1} \cdots e^{\tilde{A}_{i_{m-1}}' (s_m - s_{m-1})} \left[ \int_{s_m}^{t} e^{\tilde{A}_i' (s - s_m)} \tilde{C}_i' \tilde{C}_i e^{\tilde{A}_i (s - s_m)} ds \right] \\ &\times e^{\tilde{A}_{i_{m-1}} (s_m - s_{m-1})} \cdots e^{\tilde{A}_j s_1}. \end{aligned} \tag{2.10}$$

Define $t_0, \ldots, t_{m+1}$ satisfying $0 = t_0 \leq (t - s_m = t_1) \leq \cdots \leq (t - s_1 = t_m) \leq t = t_{m+1}$. Since $\tilde{A} = A'$, $\tilde{C} = B'$, $\tilde{B} = 0$, $\tilde{D} = 0$, and the transition matrix $\tilde{\Lambda}$ is such that

$\tilde{\lambda}_{\kappa\ell} = \lambda_{\ell\kappa}$ for $\kappa \neq \ell$, we obtain from (2.10) with $\tau = t - s$

(2.11)
$$\begin{aligned}\Upsilon_{\text{obs}}(0,t) =& e^{A_{i_j}(t-t_m)} \cdots e^{A_{i_{m-1}}(t_2-t_1)} \left[\int_0^{t_1} e^{A_i(t_1-\tau)} B_i B_i' e^{A_i'(t_1-\tau)} d\tau\right] \\ & \times e^{A'_{i_{m-1}}(t_2-t_1)} \cdots e^{A_j'(t-t_m)} + \cdots + \int_{t_m}^{t} e^{A_j(t-\tau)} B_j B_j' e^{A_j'(t-\tau)} d\tau \\ =& \sum_{k=0}^{m} \int_{t_k}^{t_{k+1}} \Phi(t,\tau) B_{i_{m-k}} B'_{i_{m-k}} \Phi(t,\tau)' d\tau = \Upsilon_{\text{rch}}(0,t),\end{aligned}$$

where $\lambda_{i,\,i_{m-1}} \cdots \lambda_{i_1,\,j} \neq 0$. ∎

We need to compare the set $\{x : W^t(x,j) = 0\}$ with the null space of $\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\}$. One difficulty is that the initial conditions of the Markov chain may be different, $i \neq j$, so that Lemma 2.2 alone does not answer this question.

PROPOSITION 2.3. *Consider the hypothesis of Lemma* 2.2. *Then, for a given* $t > 0$ *and each* $x \in \mathbb{R}^n$ *and* $j \in \mathcal{S}_\pi$, *we have that* $W^t(x,j) = 0$ *and* $x'\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\}x = 0$ *are equivalent whenever* $i$ *and* $j$ *are communicating states.*

*Proof.* We present the proof of necessity only, as the sufficiency is quite similar. Assume $x'\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\}x = 0$, where $j$ is reached from $i$. One can check for any $t_f \geq 0$ that $x'\mathcal{E}\{\Upsilon_{\text{rch}}(0,t_f)|\theta(0) = i\}x = 0$. This leads to $x'\mathcal{E}\{\Upsilon_{\text{rch}}(0,t_f)|\theta(0) = i, \theta(t_f) = j\}x = 0$, which from Lemma 2.2 is equivalent to $x'\mathcal{E}\{\Upsilon_{\text{obs}}(0,t_f)|\tilde{\theta}(0) = j, \tilde{\theta}(t_f) = i\}x = 0$. This and the inequality $\Upsilon_{\text{obs}}(0,t_f) \geq \Upsilon_{\text{obs}}(0,t)$ (when $t_f \geq t$) yield $x'\mathcal{E}\{\Upsilon_{\text{obs}}(0,t)|\tilde{\theta}(0) = j, \tilde{\theta}(t_f) = i\}x = 0$. Since $i, j$ are communicating, any $\theta(t)$ that can be reached from $\theta(0) = j$ can in turn reach $\theta(t_f) = i$, which allows us to conclude that the information $\tilde{\theta}(t_f) = i$ is irrelevant when computing the last expected value, leading to $x'\mathcal{E}\{\Upsilon_{\text{obs}}(0,t)|\tilde{\theta}(0) = j\}x = 0$, and from (2.6), we have that $W(x,j) = 0$. ∎

Consider the sequence of sets of matrices $R(k) \in \mathcal{M}^n, k \in \mathbb{Z}^0$, satisfying

$$\begin{aligned} R_j(k+1) &= \mathcal{T}_{A,j}(R(k)), \qquad k \in \mathbb{Z}^0,\ j \in \mathcal{S}, \\ R_j(0) &= 1_{\{j \in \mathcal{S}_\pi\}} B_j B_j'. \end{aligned} \tag{2.12}$$

$R(k)$ can be interpreted as a dual of $O(k)$ defined in (2.7) and will play an important role in the computational test for AR. Proposition 2.3 allows for extending the results of Proposition 2.1 to the original system, leading to an invariance property for matrices $B_i$.

LEMMA 2.4. *Consider the system* $(A, B, C = 0, D = 0, \Lambda)$. *For each* $x \in \mathbb{R}^n$ *and* $i, j \in \mathcal{S}_\pi$ *communicating Markov states, the following statements are equivalent:*

(i) $x'\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\}x = 0$, $t \geq 0$.

(ii) $x'R_j(k)x = 0$, $k \in \mathbb{Z}^0$.

(iii) $B_i' A_{\iota_1}'^{p_0} A_{\iota_2}'^{p_1} \cdots A_j'^{p_m} x = 0$ *for any sequence of Markov states* $i, \iota_1, \ldots, \iota_m, j$ *contained in* $\mathcal{S}_\pi$ *such that* $\lambda_{i,\,\iota_1}\lambda_{\iota_1,\,\iota_2} \cdots \lambda_{\iota_m,\,j} \neq 0$ *with* $p_\ell \geq 0$, $\ell = 0, \ldots, m$.

*Proof.* (i) ⇔ (iii). Considering the system (2.4) with the parameters in (2.9), we conclude that statement (i) of Proposition 2.1 holds true. Consequently, we have that (iii) of Proposition 2.1 holds true, and, in turn, it can be rewritten as (iii) above.

(iii) ⇒ (ii). This is similar to the proof of implication (iii) ⇒ (ii) of Proposition 2.1. (ii) ⇒ (i). Considering the system (2.4) with the parameters in (2.9), from (2.12) we can write

$$
\begin{aligned}
R_j(k+1) &= \tilde{A}_j' R_j(k) + R_j(k)\tilde{A}_j + \sum_{i\in\mathcal{S}} \tilde{\lambda}_{ji} R_i(k) + \lambda_{jj} R_j(k) - \tilde{\lambda}_{jj} R_j(k) \\
&= \bar{A}_j' R_j(k) + R_j(k)\bar{A}_j + \sum_{i\in\mathcal{S}} \tilde{\lambda}_{ji} R_i(k) \\
&= \mathcal{L}_{\bar{A},j}(R(k)),
\end{aligned}
\tag{2.13}
$$

where $\bar{A}_j = \tilde{A}_j + \kappa_j I$, with $\kappa_j = \frac{1}{2}(\lambda_{jj} - \tilde{\lambda}_{jj}) = \frac{1}{2}\sum_{i\in\mathcal{S}} \lambda_{ij}$. This means that $R(k)$ satisfies (2.7) (with matrix $\tilde{A}$ replaced with $\bar{A}$), and we can use Proposition 2.1 to conclude that

$$
\tilde{C}_i \bar{A}_{i_m}^{p_0} \bar{A}_{i_{m-1}}^{p_1} \cdots \bar{A}_j^{p_m} x = 0 \tag{2.14}
$$

for any sequence of Markov states $j, i_1, \ldots, i_m, i$ such that $\tilde{\lambda}_{j,i_1}\tilde{\lambda}_{i_1,i_2}\cdots\tilde{\lambda}_{i_m,i} \neq 0$ with $p_\ell \geq 0$, $\ell = 0, \ldots, m$. Expanding the expression $\tilde{C}_i \tilde{A}_{i_m}^{q_0} \tilde{A}_{i_{m-1}}^{q_1} \cdots \tilde{A}_j^{q_m} x$ with $q_\ell \geq 0$, $\ell = 0, \ldots, m$, by substituting $\tilde{A}_j = \bar{A}_j - \kappa_j I$ for each $j \in \mathcal{S}_\pi$, we obtain a sum of expressions exactly as in the left-hand of (2.14), and each one of them equals to zero. This yields $\tilde{C}_i \tilde{A}_{i_m}^{q_0} \tilde{A}_{i_{m-1}}^{q_1} \cdots \tilde{A}_j^{q_m} x = 0$, and from Proposition 2.1 we have $W^t(x,j) = 0$. The result follows directly by Proposition 2.3. ∎

COROLLARY 2.5. *Consider $j \in \mathcal{S}_\pi$ and $v \in \mathbb{R}^n$. Then, for all $k \in \mathbb{Z}^0$,*

$$
v' R_j(k) v = 0 \quad \textit{is equivalent to} \quad R_j(k) v = 0.
$$

*Proof.* The implication ($\Leftarrow$) is trivial, and the reverse one can be obtained using the equivalence in (ii) and (iii) of Lemma 2.4. ∎

*Remark* 2. Reasoning similar to that in the proof of (ii) $\Rightarrow$ (i) of Lemma 2.4 allows us to conclude that AR is invariant to perturbations of the form $\kappa_j I$ in matrix $A_j$, with $\kappa_j \in \mathbb{R}$. In particular, $(A, B, \Lambda)$ is AR if and only if $(A + \kappa I, B, \Lambda)$ is AR.

**3. $\pi$-AR and a computational test.** When seeking conditions for positiveness of the average second moment of $x$, it is important to note that this quantity depends not only on the structure of $A$ and $\Lambda$ but also on the initial probability distribution $\pi$, because the visited Markov states are "selected" essentially by $\pi$ and $\Lambda$, as illustrated in the next example.

*Example* 1. Consider the system $(A, B, \Lambda)$ with $A_1 = A_2 = B_2 = 0$, $B_1 = 1$, $\Lambda = 0$, and $x_0 = 0$. For any $0 \leq \alpha \leq 1$ we have $E\{x(t)x'(t)|\pi = [\alpha \ 1-\alpha]\} = \alpha E\{w(t)w'(t)\} = \alpha$. Then $E\{x(t)x(t)'|\pi\} > 0$ if and only if $\alpha \neq 0$.

DEFINITION 3.1 ($\pi$-AR). *The triple $(A, B, \Lambda)$ is $\pi$-average reachable ($\pi$-AR) if there exists $t > 0$ such that $\mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,t)|\theta(0) = i\} > 0$ for all $i \in \mathcal{S}_\pi$.*

Next we introduce a quite standard ordinary differential equation (ODE) whose solution $S_j(t)$, $j \in \mathcal{S}$, $t \geq 0$, is linked with reachability gramian via $R$ defined in (2.12) and Lemma 3.2. $S(t)$ satisfies

$$
\dot{S}_j(t) = \mathcal{T}_{A,j}(S(t)) + 1_{\{j\in\mathcal{S}_\pi\}} B_j B_j', \quad S_j(0) = 0, \ j \in \mathcal{S}. \tag{3.1}
$$

The easy-to-handle exponential solution of (3.1) is employed as a tool in what follows. The interpretation for $S$ is postponed to section 4.1, when we study its relation to the conditional second moment $X_i = \mathcal{E}\{x(t)x(t)1_{\{\theta=i\}}\}$. Using the operators vec and

$\varphi$, we can vectorize (3.1) to obtain a standard linear time invariant system of the form $\dot{\sigma}(t) = \mathcal{A}'\sigma(t) + b$, where $\mathcal{A}$ is as in (2.3) and $b = \varphi(B_\pi B_\pi')$, where $B_\pi B_\pi'$ is the set formed by the matrices $1_{\{j\in\mathcal{S}_\pi\}}B_jB_j'$. By noting that there are $n(n-1)N/2$ repeated entries in both $\sigma$ and $b$ (arising from the symmetry of $S_j(t)$, $\mathcal{T}_A(\cdot)$, and $B_jB_j'$ in (3.1)), we set $\bar{N} := n^2N - \frac{n(n-1)N}{2} = \frac{n(n+1)N}{2}$ and consider the projection matrix $\mathfrak{P} \in \mathfrak{R}^{n^2N,\bar{N}}$ that eliminates these entries. We define $\hat{\sigma} \in \mathbb{R}^{\bar{N}}$, $\hat{b} \in \mathbb{R}^{\bar{N}}$, and $\hat{\mathcal{A}} \in \mathfrak{R}^{\bar{N},\bar{N}}$ given as $\hat{\sigma} = \mathfrak{P}\sigma$, $\hat{b} = \mathfrak{P}b$, and $\hat{\mathcal{A}} = \mathfrak{P}'\mathcal{A}\mathfrak{P}$, respectively, leading to the reduced order system

$$\dot{\hat{\sigma}}(t) = \hat{\mathcal{A}}'\hat{\sigma}(t) + \hat{b}, \quad \hat{\sigma}(0) = 0. \tag{3.2}$$

Note that the $k$th order derivatives, $k \geq 1$, of $\hat{\sigma}(t)$ calculated at $t = 0$ are given by

$$\frac{d^k\hat{\sigma}}{dt^k}(0) = \hat{\mathcal{A}}'^{k-1}\hat{b}. \tag{3.3}$$

The next result gives the relation between the function $S$ and its derivatives.

LEMMA 3.2. *Let $j \in \mathcal{S}_{\pi,\text{rec}}$ and $v \in \mathbb{R}^n$. Define $V = (V_1, \ldots, V_N)$ such that $V_j = vv'$ and, if $i \neq j$, $V_i = 0$ and $\hat{w} = \mathfrak{P}\varphi(V)$. The following assertions are equivalent:*

(i) *$v'S_j(s)v = 0$ or, equivalently, $\hat{w}'\hat{\sigma}(s) = 0$ for some $s > 0$.*

(ii) *$\hat{w}'\frac{d^{k+1}\hat{\sigma}}{dt^{k+1}}(0) = 0$ for $k \in \mathbb{Z}^0$.*

(iii) *$v'S_j(t)v = 0$ or, equivalently, $\hat{w}'\hat{\sigma}(t) = 0$ for all $t \geq 0$.*

(iv) *$v'\mathcal{E}\{\Upsilon_{rch}(0,t)|\theta(0) = i\}v = 0$ for all $i \in \mathcal{S}_\pi$ communicating with $j$ and all $t \geq 0$.*

*Proof.* (i) $\Rightarrow$ (ii). Suppose that there exists $s > 0$ such that $v'S_j(s)v = 0$. Then,

$$\begin{aligned} \hat{w}'\hat{\sigma}(s) = 0 \Rightarrow \hat{w}'\int_0^s e^{\hat{\mathcal{A}}'(s-\tau)}\hat{b}\, d\tau = 0 &\Rightarrow \hat{w}'e^{\hat{\mathcal{A}}'(s-\tau)}\hat{b} = 0, \quad 0 \leq \tau \leq s, \\ &\Rightarrow \hat{w}'\hat{\mathcal{A}}'^k e^{\hat{\mathcal{A}}'(s-\tau)}\hat{b} = 0, \quad 0 \leq \tau \leq s. \end{aligned} \tag{3.4}$$

In particular, for $\tau = s$ we have $\hat{w}'\hat{\mathcal{A}}'^k\hat{b} = 0$, or equivalently $\hat{w}'\frac{d^{k+1}\hat{\sigma}}{dt^{k+1}}(0) = 0$, for all $k = 0, 1, \ldots$. Note that $e^{\hat{\mathcal{A}}'(s-\tau)}\hat{b} = \mathfrak{P}\varphi(S_{B_\pi}(t))$, where $S_{B_\pi}$ is a set of matrix functions defined in a manner that satisfies $\varphi(\dot{S}_{B_\pi}(t)) = \mathcal{A}'\varphi(S_{B_\pi}(t))$, with $S_{B_\pi}(0) = B_\pi B_\pi'$. Thus, $S_{B_\pi} \in \mathcal{M}^{n0}$, and hence $v'S_{B_\pi,j}(s-\tau)v = \langle V, S_{B_\pi}(s-\tau)\rangle \geq 0$, leading to $\hat{w}'e^{\hat{\mathcal{A}}'(s-\tau)}\hat{b} \geq 0$, which brings us to the second implication in (3.4).

(ii) $\Rightarrow$ (iii). Using the power series expansion, we obtain $e^{\hat{\mathcal{A}}'(t-\tau)} = \sum_{k=0}^\infty \hat{\alpha}_k(t-\tau)\hat{\mathcal{A}}'^k$, where $\hat{\alpha}_k(t-\tau) = \frac{(t-\tau)^k}{k!}$ for each $k \in \mathbb{Z}^0$. Then,

$$\hat{w}'\hat{\sigma}(t) = \int_0^t \hat{w}'\sum_{k=0}^\infty \hat{\alpha}_k(t-\tau)\hat{\mathcal{A}}'^k\hat{b}\, d\tau = \sum_{k=0}^\infty \left(\int_0^t \hat{\alpha}_k(t-\tau)\, d\tau\right)\hat{w}'\hat{\mathcal{A}}'^k\hat{b}, \tag{3.5}$$

and, since (3.3) and (ii) together yield $\hat{w}'\hat{\mathcal{A}}'^k\hat{b} = 0$, we obtain $\langle S(t), V\rangle = \hat{w}'\hat{\sigma}(t) = 0$. Equivalently, $v'S_j(t)v = 0$ for all $t \geq 0$.

(iii) $\Rightarrow$ (i). This part of the proof is trivial.

(iii) $\Leftrightarrow$ (iv). This part of the proof follows from Lemma 2.4 and the equivalence between (ii) and (iii) above. $\square$

Now we can employ Lemma 3.2 in the study of the null space of $R$ and its role in testing $\pi$-AR. The link between the function $S$ and $R$ follows quite directly from (2.12) and (3.1):

$$R_j(k) = \frac{d^{k+1}S_j}{dt^{k+1}}(0), \quad k \in \mathbb{Z}^0,\ j \in \mathcal{S}. \tag{3.6}$$

In order to ease notation, let us define the *reachability matrices* $\mathbf{R}_j \in \Re^{n,n\bar{N}}$ by

$$\mathbf{R}_j = \begin{bmatrix} R_j(0) & R_j(1) & \cdots & R_j\left(\bar{N}-1\right) \end{bmatrix} \tag{3.7}$$

for all $j \in \mathcal{S}_{\pi,\text{rec}}$. We write $\mathbf{R}_j(A, B, \Lambda, \pi)$ in some passages to emphasize the parameters required to calculate $\mathbf{R}_j$.

PROPOSITION 3.3. *Let $j \in \mathcal{S}_{\pi,\text{rec}}$. We have that $v \in \mathcal{N}(R_j(k))$, $k = 0, \ldots, \bar{N}-1$, is equivalent to $v \in \mathcal{N}(R_j(k))$, $k = 0, \ldots, m$, for $m \geq \bar{N}$.*

*Proof.* We shall only show the direct equivalence implication, as the converse one is trivial. From the hypothesis and (3.6) it follows that $v'\frac{d^{k+1}S_j}{dt^{k+1}}(0)v = 0$, $k = 0, \ldots, \bar{N}-1$, which is equivalent to $w'\hat{\mathcal{A}}'^k\hat{b} = 0$, $k = 0, \ldots, \bar{N}-1$, where $\hat{w}$ is defined as in Lemma 3.2. Using the Cayley–Hamilton theorem, we obtain that $w'\hat{\mathcal{A}}'^m\hat{b} = 0$ for $m \geq \bar{N}$. That is, $v'R_j(m)v = 0$ for $m \geq \bar{N}$. Corollary 2.5 completes the proof. ∎

PROPOSITION 3.4. *$\mathcal{I}m\,(\mathbf{R}_j) = \mathcal{I}m\,([R_j(0)\,R_j(1) \cdots R_j(m)])$ for all $m \geq \bar{N}$.*

*Proof.* For brevity denote $\mathbf{Q}_m = [R_j(0)\,R_j(1) \cdots R_j(m)]$, $m \geq \bar{N}$. Assume $v \in \mathcal{I}\text{m}(\mathbf{Q}_m)$, so that $v = \mathbf{Q}_m\xi$ for some $\xi \in \mathbb{R}^{n(m+1)}$. Now, we take an arbitrary $\eta \in \mathcal{N}(\mathbf{R}_j')$, and, by definition of $\mathbf{R}_j$, we have $R_j(k)\eta = 0$, $k = 0, \ldots, \bar{N}-1$. Then, Proposition 3.3 produces $R_j(k)\eta = 0$, $k = 0, \ldots, m$, where $m \geq \bar{N}$, yielding $\eta'\mathbf{Q}_m = 0$ and $\eta'v = 0$, in such a manner that $v \in (\mathcal{N}(\mathbf{R}_j'))^\perp$. From the fundamental theorem of linear algebra we have that $v \in \mathcal{I}\text{m}(\mathbf{R}_j)$. The converse implication can be shown in a similar manner. ∎

PROPOSITION 3.5. *For any $i$ and $j$ communicating Markov states in $\mathcal{S}_\pi$ we have*

$$\mathcal{I}m\,(\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\}) \;=\; \mathcal{I}m(\mathbf{R}_j).$$

*Proof.* Let us assume $v \in \mathcal{I}m(\mathbf{R}_j)$. Then there exists $\xi \in \mathbb{R}^{n\bar{N}}$ such that $v = \mathbf{R}_j\xi$. Moreover, for any $\eta \in \mathcal{N}(\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\})$, Lemma 3.2 implies $\eta'S_j(t)\eta = 0$. From Lemma 3.2 and Corollary 2.5 we obtain $\eta'R_j(k) = 0$, $k = 0, \ldots, \bar{N}-1$. Consequently, we have that $\eta'\mathbf{R}_j = 0$ and hence $\eta'\mathbf{R}_j\xi = 0$, leading in turn to $\eta'v = 0$. Thus, by the fact that $\eta$ is arbitrary in $\mathcal{N}(\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\})$ we conclude that $v \in (\mathcal{N}(\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(t) = i\}))^\perp$. Therefore, the fundamental theorem of linear algebra yields $v \in \mathcal{I}\text{m}\,(\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\})$. The converse implication can be shown in a similar manner, replacing $\mathbf{R}_j$ by $\mathcal{E}\{\Upsilon_{\text{rch}}(0,t)|\theta(0) = i\}$ (and vice-versa) in the above. ∎

THEOREM 3.6. *$(A, B, \Lambda)$ is $\pi$-AR if and only if rank$(\mathbf{R}_j) = n$ for all $j \in \mathcal{S}_{\pi,\text{rec}}$.*

*Proof. Necessity.* It is straightforward from the definition of $\pi$-AR and Proposition 3.5 that $\mathcal{I}m(\mathbf{R}_j) = \mathbb{R}^n$ for any $j$ such that $i, j$ are communicating and $i \in \mathcal{S}_\pi$. This yields rank$(\mathbf{R}_j) = n$, $j \in \mathcal{S}_{\pi,\text{rec}}$.

*Sufficiency.* Since rank$(\mathbf{R}_j) = n$, we have from Proposition 3.5 that

$$\mathcal{E}\{\Upsilon_{\text{rch}}(0,\bar{t})|\theta(0) = i\} > 0 \tag{3.8}$$

for any $\bar{t} > 0$ and $i \in \mathcal{S}_\pi$ communicating with $j$. It remains to show positiveness of the gramian for initial Markov states that do not communicate with $j$, that is, for $\theta(0) = \ell$ being a transient state reaching $j$. Let $T^*$ be the first visit time to the recurrent class in $\mathcal{S}_{\pi,\mathrm{rec}}$ containing $j$, and let $s \geq 0$. Then,

$$\begin{aligned}\mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,\bar{t}+s)|\theta(0)=\ell\} &= \mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,\bar{t}+s)|\theta(0)=\ell, T^*\leq s\}\mathcal{P}(T^*\leq s|\theta(0)=\ell)\\ &\quad + \mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,\bar{t}+s)|\theta(0)=\ell, T^*> s\}\mathcal{P}(T^*> s|\theta(0)=\ell).\end{aligned} \tag{3.9}$$

It is evident that $\mathcal{P}(T^* \leq s|\theta(0) = \ell) > 0$; moreover, $\mathcal{E}\{\Upsilon_{\mathrm{rch}}(0, \bar{t} + s)|\theta(0) = \ell, T^* \leq s\} \geq \mathcal{E}\{\Upsilon_{\mathrm{rch}}(T^*, \bar{t} + s)|\theta(0) = \ell, T^* \leq s\} = \mathcal{E}\{\Upsilon_{\mathrm{rch}}(T^*, \bar{t} + s)|\theta(T^*) = i \in \mathcal{S}_{\pi,\mathrm{rec}}\} \geq \mathcal{E}\{\Upsilon_{\mathrm{rch}}(T^*, T^* + \bar{t})|\theta(T^*) = i \in \mathcal{S}_{\pi,\mathrm{rec}}\} > 0$, where the last inequality comes from the use of time translation and homogeneity properties in (3.8). Substituting these inequalities into (3.9) yields $\mathcal{E}\{\Upsilon_{\mathrm{rch}}(0, \bar{t} + s)|\theta(0) = \ell\} > 0$ for arbitrary $s, \bar{t} > 0$, concluding the proof. $\square$

*Example* 2. Consider the system $(A, B, \Lambda)$ with

$$\begin{gathered} A_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \Lambda = \begin{pmatrix} 0 & 0 \\ 0.1 & -0.1 \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \pi = \begin{pmatrix} 0 & 1 \end{pmatrix}. \end{gathered} \tag{3.10}$$

Note that $\mathcal{S}_{\pi,\mathrm{rec}} = \{1\}$. We compute the reachability matrices (3.7) to check that $\mathrm{rank}(\mathbf{R}_1) = 2$ and conclude from Theorem 3.6 that the system is $\pi$-AR. We have simulated the variable $x(t)$ for 1.000 Markov state realizations, and the result is depicted in Figure 1, which illustrates that $\pi$-AR does not mean that $x(t)$ eventually reaches the entire state space (dissimilarly to linear systems with no jumps; see [11]). We also note from Figure 1 that if we consider as a "target state" $x_{tgt} = (10 \;\; 0)'$, then the ball $\{x : \|x - x_{tgt}\|_\infty < 1/2\}$ is visited with *zero* probability by $x(t)$, $t \geq 0$, meaning that the system is not SR in the usual sense (e.g., as defined in [1, 22, 24]).

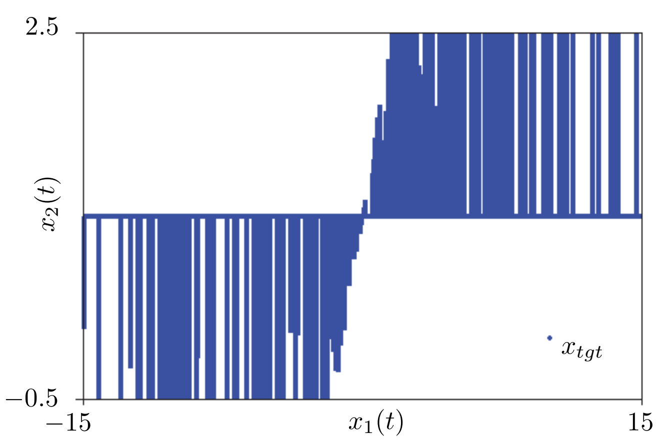


FIG. 1. *Portion of phase plots of* 1000 *realizations for the process* $\{x(t)\}$, $t \geq 0$, *of the system in Example* 2.

*Example* 3. Consider the system $(A, B, \Lambda)$ with $A_1 = A_2 = B_1 = 0$, $B_2 = 1$, and $\Lambda$ and $\pi$ as in Example 2. By implementing the rank test of Theorem 3.6 we conclude that the system is not $\pi$-AR, as $\mathbf{R}_1 = 0$. Now, considering the SR problem of driving $x(t)$ to an $\epsilon$-neighborhood of a "target state" at a fixed time instant, $\|x(t) - x_{tgt}\| < \epsilon$, in finite time with positive probability, as studied in [1], we can show that the system is SR. In fact, since the subsystem $(A_2, B_2)$ is controllable in the standard sense for deterministic linear systems, it is simple to design a deterministic input function $w(t)$, $0 \le t \le t_f$, resulting in $x(t_f) = x_{tgt}$ provided there is no jump in this time interval, thus yielding $\mathcal{P}(\|x(t) - x_{tgt}\| < \epsilon) \ge \mathcal{P}(\theta(t) = 2, 0 \le t \le t_f) = e^{-t_f}$.

**4. AR and positiveness of the state second moment.** At times, the initial probability distribution $\pi$ can be unknown, forcing us to consider a general $\pi$.

DEFINITION 4.1 (AR). *The triple $(A, B, \Lambda)$ is average reachable (AR) if there exists $t > 0$ such that $\mathcal{E}\{\Upsilon_{\mathrm{rch}}(0, t)|\theta(0) \sim \pi\} > 0$ for every initial distribution $\pi$.*

A rank test for AR can be obtained in a rather direct manner based on the test for $\pi$-AR given in Theorem 3.6. The idea is to check if the system is $\pi^{(\ell)}$-AR for certain initial distributions $\pi^{(\ell)}$, each of them "exciting" one set of recurrent Markov states. Formally, let the irreducible sets of recurrent Markov states be denoted by

$$\mathcal{S}_{\mathrm{rec},1}, \ldots, \mathcal{S}_{\mathrm{rec},\bar{m}} \tag{4.1}$$

(so that the Markov chain is composed by a total of $\bar{m}$ blocks of recurrent states plus the transient states) and take associated initial distributions

$$\pi^{(1)}, \ldots, \pi^{(\bar{m})}, \tag{4.2}$$

where each $\pi^{(\ell)}$ is concentrated on an arbitrary state belonging to $\mathcal{S}_{\mathrm{rec},\ell}$, i.e., $\pi^{(\ell)}_{i_\ell} = P(\theta(0) = i_\ell) = 1$ for some $i_\ell \in \mathcal{S}_{\mathrm{rec},\ell}$.

THEOREM 4.2. *Consider the initial distributions $\pi^{(1)}, \ldots, \pi^{(\bar{m})}$ as defined above. The following statements are equivalent:*

(i) $(A, B, \Lambda)$ *is AR.*

(ii) $(A, B, \Lambda)$ *is $\pi^{(\ell)}$-AR for all $\ell = 1, \ldots, \bar{m}$.*

(iii) $(A, B, \Lambda)$ *is $\pi$-AR for all probability distribution $\pi \in \mathbb{R}^N$.*

*Proof.* The facts that (iii) and (i) are equivalent and that (iii) implies (ii) are trivial. We show (ii) $\Rightarrow$ (iii) by contradiction. First, we deny (iii) by assuming that the system is not $\pi$-AR for some $\pi$, and from Theorem 3.6 we have

$$\mathrm{rank}(\mathbf{R}_{j^*}(A, B, \Lambda, \pi)) < n \quad \text{for some } j^* \in \mathcal{S}_{\pi,\mathrm{rec}}. \tag{4.3}$$

Now, assume that (ii) is valid, and consider $\ell \in \{1, \ldots, \bar{m}\}$ for which $j^* \in \mathcal{S}_{\mathrm{rec},\ell}$. Theorem 3.6 yields $\mathrm{rank}(\mathbf{R}_j(A, B, \Lambda, \pi^{(\ell)})) = n$ for all $j \in \mathcal{S}_{\pi^{(\ell)},\mathrm{rec}}$, and, in particular for $j = j^*$,

$$\mathrm{rank}(\mathbf{R}_{j^*}(A, B, \Lambda, \pi^{(\ell)})) = n. \tag{4.4}$$

Since $j^* \in \mathcal{S}_{\mathrm{rec},\ell}$, we have $\mathcal{S}_{\pi^{(\ell)},\mathrm{rec}} \subset \mathcal{S}_{\pi,\mathrm{rec}}$, leading to

$$\mathrm{rank}(\mathbf{R}_{j^*}(A, B, \Lambda, \pi^{(\ell)})) \le \mathrm{rank}(\mathbf{R}_{j^*}(A, B, \Lambda, \pi)). \tag{4.5}$$

Combining (4.4) and (4.5) produces $\mathrm{rank}(\mathbf{R}_{j^*}(A, B, \Lambda, \pi)) = n$, which is absurd in view of (4.3). ∎

### 4.1. The function $S(t)$ and the positiveness of the second moment.

LEMMA 4.3. *The following assertions hold:*

(i) *For a given initial probability distribution $\pi$ we have that $(A, B, \Lambda)$ is $\pi$-AR if and only if $S_j(t) > 0$ for all $t > 0$ and $j \in \mathcal{S}_{\pi,\mathrm{rec}}$.*

(ii) *$(A, B, \Lambda)$ is AR if and only if for each initial probability distribution $\pi$ we have that $S_j(t) > 0$ for all $t > 0$ and $j \in \mathcal{S}_{\pi,\mathrm{rec}}$.*

*Proof.* (i) We prove necessity by contradiction. By supposing that there exist $j \in \mathcal{S}_{\pi,\mathrm{rec}}$ and $\tilde{t} > 0$ such that $S_j(\tilde{t})$ is not positive definite, we have $v' S_j(\tilde{t}) v = 0$ for a certain $v \in \mathbb{R}^n$ different from zero. Then, the equivalence between (i) and (iii) of Lemma 3.2 yields $v' S_j(t) v = 0$ for all $t \geq 0$, which in turn, by the equivalence between (iii) and (iv) of Lemma 3.2, leads to $v' \mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,t) | \theta(0) = i\} v = 0$ whenever $i \in \mathcal{S}_\pi$ and $j$ are communicating Markov states. Thus, we have that $(A, B, \Lambda)$ is not $\pi$-AR. The proof of sufficiency is similar.

(ii) Here we prove sufficiency by contradiction. Let us suppose that $(A, B, \Lambda)$ is not AR. Then, by Theorem 4.2 there exists $\pi^{(\ell)}$ as in (4.2) such that $(A, B, \Lambda)$ is not $\pi^{(\ell)}$-AR. Thus, by Definition 3.1, for all $t > 0$ there exists $v \neq 0$ in $\mathbb{R}^n$ such that $v' \mathcal{E}\{\Upsilon_{\mathrm{rch}}(0,t) | \theta(0) = i\} v = 0$ for some $i \in \mathcal{S}_{\pi^{(\ell)}} = \mathcal{S}_{\mathrm{rec},\ell}$. Consider a fixed $t > 0$ and $j \in \mathcal{S}_{\pi^{(\ell)},\mathrm{rec}} = \mathcal{S}_{\mathrm{rec},\ell}$. The equivalence (iii)–(iv) of Lemma 3.2 leads to $v' S_j(t) v = 0$, $t > 0$, which concludes the proof of the sufficiency. Necessity follows in a straightforward manner from Lemma 3.2 (iii), (iv). ∎

Next we extend the positiveness of $S$ to the function $X(t) \in \mathcal{M}^{n0}$. Recall that $X(t)$ is associated with the second moment of the process $x(t)$ by

$$X_j(t) = \mathcal{E}\{x(t)x(t)' 1_{\{\theta(t)=j\}}\}, \;\; t \geq 0, \; j \in \mathcal{S}, \tag{4.6}$$

and satisfies the linear matrix differential equations (see, e.g., [16])

$$\dot{X}_j(t) = \mathcal{T}_{A,j}(X(t)) + p_j(t) B_j B_j', \quad \text{with} \quad X_j(0) = p_j(0)\Psi. \tag{4.7}$$

*Remark* 3. By direct comparison between (3.1) and (4.7) one can conclude that, if $x(0) = 0$ and the noise $w(t)$ is multiplied by $1/\sqrt{p_{\theta(t)}(t)}$ whenever $\theta(t) \in \mathcal{S}_\pi$ (otherwise it is multiplied by zero), then $S$ coincides with $X$. This gives an interpretation for $S$ as the conditional second moment of the system when the noise amplitude is conveniently "modulated" and the initial condition is set to zero.

In view of Remark 3 , it is not a surprise that $p_j(t)$ is required in our analysis (so that $1/\sqrt{p_{\theta(t)}(t)}$ is well posed). It is a well-known property of finite-dimension Markov chains that by the continuity of its probability distribution $p(t)$ there exist $\delta > 0$ and a finite closed interval $[T_1, T_2]$, $0 < T_1 < T_2$, such that

$$p_j(t) \geq \delta, \quad j \in \mathcal{S}_\pi, \; t \in [T_1, T_2] \tag{4.8}$$

(and $p_j(t) = 0$, $j \notin \mathcal{S}_\pi$), for Markov states that can be reached from the initial state. For $t \in [T_1, T_2]$ we define $U(t) \in \mathcal{M}^{n0}$ satisfying

$$\dot{U}_j(t) = \mathcal{T}_{A,j}(U(t)) + p_j(t) B_j B_j', \quad \text{with} \quad U_j(T_1) = 0. \tag{4.9}$$

Also, for an arbitrary fixed time instant $\tilde{t} > 0$, we define an auxiliary variable $V(t) \in \mathcal{M}^{n0}$ for $t \geq \tilde{t}$ via

$$\dot{V}_j(t) = \mathcal{T}_{A,j}(V(t)) + 1_{\{j \in \mathcal{S}_\pi\}} B_j B_j', \quad \text{with} \quad V_j(\tilde{t}) = 0. \tag{4.10}$$

The relation between $S$, $V$, $U$, and $X$ is given next.

LEMMA 4.4. *Let $j \in \mathcal{S}_{\pi,\text{rec}}$. For $\delta$ and $[T_1, T_2]$ as in (4.8) the following assertions hold:*

(i) *$V_j(\tilde{t}+s) = S_j(s)$ for all $s \geq 0$.*

(ii) *If $S_j(t) > 0$ for all $t > 0$, then for any fixed $\tilde{t} > 0$ there exists $\gamma > 0$ such that $S_j(t) \geq \gamma I$ for all $t \geq \tilde{t}$.*

(iii) *$U_j(t) \geq \delta V_j(t)$ for all $t \in [T_1, T_2]$.*

(iv) *$X_j(t) \geq U_j(t)$ for all $t \in [T_1, T_2]$.*

(v) *If for any fixed $T^* \in \mathbb{R}$ such that $0 < T^* < T_2 - T_1$ there exists $\gamma > 0$ such that $S_j(s) \geq \gamma I$ for all $s \geq T^*$, then there exists $\rho > 0$ such that $X_j(t) \geq \rho I$ for all $t$ belonging to an interval $\mathcal{J} \subset [T_1, T_2]$.*

*Proof.* (i) This part of the proof is immediate from (3.1) and (4.10).

(ii) Consider the function $V_j$ in (4.10). The function defined as $\tilde{S}_j = S_j - V_j$ evolves, for $t \geq \tilde{t}$, according to

$$\dot{\tilde{S}}_j(t) = \mathcal{T}_{A,j}(\tilde{S}(t)), \text{ with } \tilde{S}_j(\tilde{t}) = S_j(\tilde{t}). \tag{4.11}$$

From (4.11) and the fact that $S(\tilde{t}) \in \mathcal{M}^{n0}$ we have that $\tilde{S}_j(t) \geq 0$. This means that $S_j(t) \geq V_j(t)$ for all $t \geq \tilde{t}$, and, consequently,

$$S_j(\tilde{t}+s) \geq S_j(s) \tag{4.12}$$

for all $s \geq 0$. Now, let us consider $j \in \mathcal{S}_{\pi,\text{rec}}$ and the compact set $D_0 := \mathbb{S}^1 \times [\tilde{t}, 2\tilde{t}]$ where $\mathbb{S}^1 = \{v \in \mathbb{R}^n : ||v|| = 1\}$ in order to define for each $j$ the function $g_j : D_0 \longrightarrow \mathbb{R}^+$ as $g_j(v,s) = v' S_j(s) v$. The function $g_j$ is continuous, and, by the Weierstrass extreme value theorem, there exists $(\bar{v}, \bar{s}) \in D_0$ such that $\gamma = \bar{v}' S_j(\bar{s}) \bar{v} = \min_{(v,s)\in D_0} g_j(v,s)$. Recall that $S_j(t) > 0$ for all $t > 0$; then we have that $\gamma > 0$. This yields

$$S_j(s) \geq \gamma I, \quad s \in [\tilde{t}, 2\tilde{t}]. \tag{4.13}$$

From (4.12) and (4.13) we have that $S_j(2\tilde{t}+\zeta) \geq S_j(\tilde{t}+\zeta) \geq \gamma I$, where $\zeta$ is any real number such that $0 \leq \zeta \leq \tilde{t}$. Consequently, we obtain $S_j(s) \geq \gamma I$ for all $s \in [2\tilde{t}, 3\tilde{t}]$, and recursively we obtain the result.

(iii) Consider $V_j$ in (4.10) with $\tilde{t} = T_1$. For $t \in [T_1, T_2]$, we define $Z(t) = \delta V(t)$, which evolves according to

$$\dot{Z}_j(t) = \mathcal{T}_{A,j}(Z(t)) + \delta 1_{\{j \in \mathcal{S}_\pi\}} B_j B_j', \quad \text{with} \quad Z_j(T_1) = 0. \tag{4.14}$$

Now, the function defined as $\tilde{Z}_j = U_j - Z_j$ satisfies

$$\dot{\tilde{Z}}_j(t) = \mathcal{T}_{A,j}(\tilde{Z}(t)) + (p_j(t) - \delta 1_{\{j \in \mathcal{S}_\pi\}}) B_j B_j', \quad \text{with} \quad \tilde{Z}_j(T_1) = 0, \tag{4.15}$$

where $p_j(t) - \delta 1_{\{j \in \mathcal{S}_\pi\}} \geq 0$. Then we have that $\tilde{Z}_j(t) \geq 0$ for all $t \in [T_1, T_2]$. Hence $U_j(t) \geq \delta V_j(t)$ for all $t \in [T_1, T_2]$.

(iv) The function defined as $\tilde{U}_j = X_j - U_j$ evolves for $t \in [T_1, T_2]$ according to

$$\dot{\tilde{U}}_j(t) = \mathcal{T}_{A,j}(\tilde{U}(t)), \text{ with } \tilde{U}_j(T_1) = X_j(T_1), \tag{4.16}$$

yielding $\tilde{U}_j(t) \geq 0$ for all $t \in [T_1, T_2]$. Thus, we obtain that $X_j(t) \geq U_j(t)$ for all $t \in [T_1, T_2]$.

(v) Given a fixed $T^* \in \mathbb{R}$ such that $0 < T^* < T_2 - T_1$, assume that there exists $\gamma > 0$ such that $S_j(s) \geq \gamma I$ for all $s \geq T^*$. Now by considering the function $V_j$ as in (4.10) with $\tilde{t} = T^*$ and using (i), we obtain $V_j(T_1 + s) = S_j(s) \geq \gamma I$ for all $s \in [T^*, T_2 - T_1]$, noticing that $t = T_1 + s \in \mathcal{J} = [T_1 + T^*, T_2] \subset [T_1, T_2]$ whenever $s \in [T^*, T_2 - T_1]$. Then by multiplying $V$ and $\gamma I$ in the last inequality by $\delta$ and using (iii), we obtain $U_j(t) \geq \delta V_j(t) \geq \rho I$ for all $t \in \mathcal{J}$ where $\rho = \delta\gamma$. Finally, using (iv), we conclude the proof. ∎

*Remark* 4. Note that, when considering an initial probability distribution $\pi^{(\ell)}$ as in (4.2) for some $\ell \in \{1, \dots, \bar{m}\}$, we have $X_j(t) = 0$ for all $t \geq 0$ and $j \notin \mathcal{S}_{\text{rec},\ell}$. Indeed, for such a $j$, $p_j(t) = 0$ for all $t \geq 0$, and from (4.6) we have $X_j(t) = \mathcal{E}\{x(t)x(t)' | \theta(t) = j\} p_j(t) = 0$.

THEOREM 4.5. *The following statements are equivalent:*

(i) *The tuple $(A, B, \Lambda)$ is AR.*

(ii) *There exists $\rho > 0$ such that for each initial condition $x_0$ and $\pi$ we have $X_j(t) \geq \rho I$ for all $t$ belonging to an interval $\mathcal{J} \subset [T_1, T_2]$ and all $j \in \mathcal{S}_{\pi,\text{rec}}$.*

(iii) *There exists $\rho > 0$ such that for each initial condition $x_0$ and $\pi$ we have $\mathcal{E}\{x(t)x(t)'\} \geq \rho I$ for all $t$ belonging to an interval $\mathcal{J} \subset [T_1, T_2]$.*

*Proof.* (i) $\Rightarrow$ (ii). Suppose that $(A, B, \Lambda)$ is AR. From Lemma 4.3(ii) we obtain $S_j(t) > 0$ for all $t > 0$ and $j \in \mathcal{S}_{\pi,\text{rec}}$, for each initial probability distribution $\pi$. Statement (ii) of Lemma 4.4 implies that for any fixed $\tilde{t} > 0$ there exists $\gamma > 0$ such that $S_j(s) \geq \gamma I$ for all $s \geq \tilde{t}$. In particular, consider $\tilde{t}$ such that $0 < \tilde{t} < T_2 - T_1$. By Lemma 4.4(iv) there exists $\rho > 0$ such that $X_j(t) \geq \rho I$ for all $t \in \mathcal{J} \subset [T_1, T_2]$.

(ii) $\Rightarrow$ (iii). The fact that $\mathcal{E}\{x(t)x(t)'\} = \sum_{i \in \mathcal{S}} X_i(t) \geq X_j(t)$ leads to the result.

(iii) $\Rightarrow$ (i). We proceed by contradiction, assuming that $(A, B, \Lambda)$ is not AR. Then, by Theorem 4.2 there exists $\pi^{(\ell)}$ as in (4.2) such that $(A, B, \Lambda)$ is not $\pi^{(\ell)}-$AR. Lemma 4.3(i) implies that $S_j(\bar{t})$ is not positive definite for a certain $\bar{t} > 0$ and some $j \in \mathcal{S}_{\pi^{(\ell)},\text{rec}} = \mathcal{S}_{\text{rec},\ell}$; that is, there exists $v \in \mathbb{R}^n$ different from zero such that $v' S_j(\bar{t}) v = 0$. This equality can be extended to a general $t \geq 0$ using the equivalence of (i) and (ii) of Lemma 3.2, and differentiation yields $v' \dot{S}_j(t) v = 0$ for all $t \geq 0$. Note that, by the extension of the Cholesky decomposition result for positive semidefinite matrices, we have that for each $t > 0$ there exists an appropriate matrix $Y_t \in \mathfrak{R}^n$ such that $S_j(t) = Y_t Y_t'$, which can be employed to show that $S_j(t) v = 0$. In this manner, the terms $v'(A_j S_j(t) + S_j(t) A_j')v$ and $\lambda_{jj} v' S_j(t) v$ become zero for all $t \geq 0$ in the derivative $v' \dot{S}_j(t) v$ (see (3.1)). Hence we have $\sum_{i \in \mathcal{S}_{\text{rec},\ell}, i \neq j} v' \lambda_{ij} S_i(t) v + v' B_j B_j' v = 0$, which is a sum involving positive semidefinite matrices. Then we obtain $v' S_i(t) v = 0$ for all $t \geq 0$ and all $i \in \mathcal{S}_{\text{rec},\ell}$. Now, assume $x_0 = 0$, yielding $X(0) = 0$. Thus $X(t) \leq S(t)$, which in turn implies $v' X_i(t) v = 0$ for all $t \geq 0$ and all $i \in \mathcal{S}_{\text{rec},\ell}$. Finally, we obtain $X_j(t) = 0$ for all $t \geq 0$ and $j \notin \mathcal{S}_{\text{rec},\ell}$ (see Remark 4), and then $v' \mathcal{E}\{x(t)x(t)'\} v = \sum_{i \in \mathcal{S}_{\text{rec},\ell}} v' X_i(t) v = 0$ for all $t \geq 0$. ∎

The single reason we cannot take $T_2$ to infinity in the previous results is that (4.8) holds for finite $T_2$ only. However, if we restrain the analysis to the situation where only recurrent Markov states are visited, we can replace $\mathcal{S}_\pi$ by $\mathcal{S}_{\pi,\text{rec}}$ in (4.8) (and in other relevant passages including (3.1)), which now holds for all $t \geq T_1$. In the particular case where we only require positiveness (not uniform) of the second moment average, we could set $T_1 = 0$. Joining these arguments, we get the following result. Let $\Pi$ be the set of distributions "exciting recurrent Markov states only";

formally, if we set $\mathcal{S}_{\text{rec}} = \bigcup_{\ell=1}^{m} \mathcal{S}_{\text{rec},\ell}$, with $\mathcal{S}_{\text{rec},\ell}$ as in (4.1), then

$$\Pi = \left\{ \pi \in \mathbb{R}^N : \sum_{i \in \mathcal{S}_{\text{rec}}} \pi_i = 1 \right\}. \tag{4.17}$$

DEFINITION 4.6. *The system* $(A, B, \Lambda)$ *is* $\Pi$*-AR if and only if it is* $\pi$*-AR for all* $\pi \in \Pi$.

COROLLARY 4.7. *The following statements are equivalent:*

(i) *The tuple* $(A, B, \Lambda)$ *is* $\Pi$*-AR.*

(ii) *There exist* $\rho > 0$ *and* $T > T_1$ *such that for each initial condition* $x_0$ *and* $\pi \in \Pi$ *we have* $X_j(t) \geq \rho I$ *for all* $t \geq T$ *and all* $j \in \mathcal{S}_{\pi,\text{rec}}$.

(iii) *There exist* $\rho > 0$ *and* $T > T_1$ *such that for each initial condition* $x_0$ *and* $\pi \in \Pi$ *we have* $\mathcal{E}\{x(t)x(t)'\} \geq \rho I$ *for all* $t \geq T$.

(iv) *For each initial condition* $x_0$ *and* $\pi \in \Pi$ *we have* $X_j(t) > 0$ *for all* $t > 0$ *and all* $j \in \mathcal{S}_{\pi,\text{rec}}$.

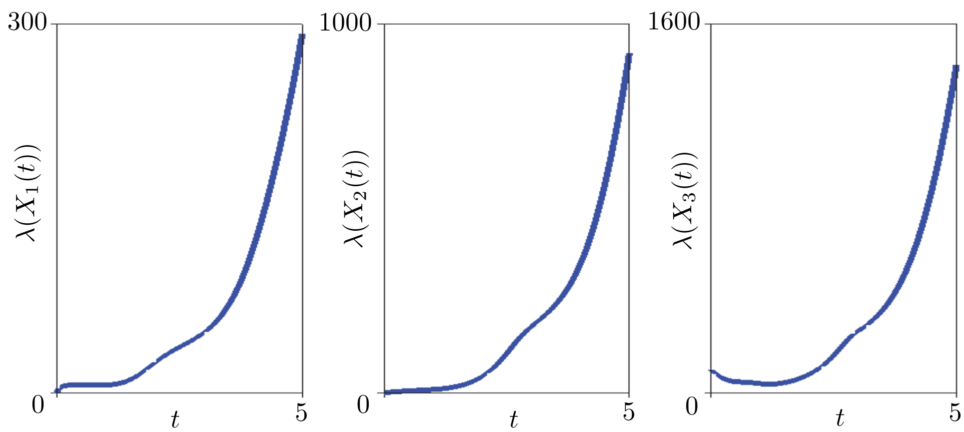


FIG. 2. *Smallest eigenvalue of* $X_j(t)$, $j \in \mathcal{S} = \{1, 2, 3\}$, *for the system in Example* 4(a).

*Example* 4 (AR and the Markov chain structure). Consider the system $(A, B, \Lambda)$ with

$$A_1 = \begin{pmatrix} -1 & 0.05 \\ 10 & -1 \end{pmatrix}, \; A_2 = \begin{pmatrix} 1 & -0.9 \\ 1.1 & 0.6 \end{pmatrix}, \; A_3 = \begin{pmatrix} 0 & -1.7 \\ 1.4 & -0.5 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 1 & 0.2 & -1.9 \\ -0.1 & 1.4 & -0.3 \end{pmatrix}, \quad B_2 = B_3 = 0,$$

$x(0)$ with Gaussian distribution with zero mean and covariance matrix $100I$. We shall consider different values for $\pi$ and $\Lambda$. Let $\pi_a = \pi_c = (0\ 0\ 1)$, $\pi_b = (0\ 1\ 0)$,

$$\Lambda_a = \begin{pmatrix} -1 & 0.5 & 0.5 \\ 0.5 & -2 & 1.5 \\ 0.9 & 0.1 & -1 \end{pmatrix}, \qquad \Lambda_b = \begin{pmatrix} -1 & 1 & 0 \\ 2 & -2 & 0 \\ 0.9 & 0.1 & -1 \end{pmatrix}, \qquad \Lambda_c = \begin{pmatrix} -1 & 1 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(a) Let $\pi = \pi_a$ and $\Lambda = \Lambda_a$. The Markov chain is ergodic so that we have only one irreducible set ($\bar{m} = 1$), given by $\mathcal{S}_{\text{rec},1} = \{1, 2, 3\}$. We take, e.g., $\pi^{(1)} = (1\ 0\ 0)$ and compute the reachability matrices by (3.7) to find that $\text{rank}(\mathbf{R}_j) = 2$ for $j \in \mathcal{S} = \{1, 2, 3\}$. From Theorem 4.2 we conclude that the system is both $\pi^{(1)}$-AR and AR. Note from Figure 2 that $X_j(t)$ is positive for $t > 0$ and $j \in \mathcal{S}$, and consequently $\mathcal{E}\{x(t)x(t)'\} > 0$, confirming Theorem 4.5. AR implies $\Pi$-AR, and Figure 2 also confirms Corollary 4.7 (we have obtained $\rho \approx 45$ and $T = 2.5$).

(b) Consider $\pi = \pi_b$ and $\Lambda = \Lambda_b$. We have $\bar{m} = 1$, $\mathcal{S}_{\text{rec},1} = \{1, 2\}$. For $\pi^{(1)} = [1\ 0\ 0]$ we compute the reachability matrices by (3.7) and conclude that $(A, B, \Lambda)$ is $\pi^{(1)}$-AR and AR. We obtain that $X_3(t) = 0$ and $X_j$ is positive definite for recurrent Markov states only (as seen in Figure 3), confirming Theorem 4.5 and Corollary 4.7.

(c) Let $\pi = \pi_c$ and $\Lambda = \Lambda_c$. In this case we have $\bar{m} = 2$, $\mathcal{S}_{\text{rec},1} = \{1, 2\}$, and $\mathcal{S}_{\text{rec},2} = \{3\}$; we can set $\pi^{(1)} = [1\ 0\ 0]$ and $\pi^{(2)} = [0\ 0\ 1]$ and use (3.7) to obtain that $\text{rank}(\mathbf{R}_j) = 2$ for $j \in \mathcal{S}_{\pi^{(1)},\text{rec}} = \mathcal{S}_{\text{rec},1}$, and $\text{rank}(\mathbf{R}_3) = 0$, thus concluding that the system is $\pi^{(1)}$-AR and not $\pi^{(2)}$-AR. Consequently, $(A, B, \Lambda)$ is not AR. Note also that $\mathcal{E}\{x(t)x(t)'\}$ tends to zero when $t$ increases, as illustrated in Figure 4. In this example, $\pi \notin \Pi$; hence Corollary 4.7 does not apply.

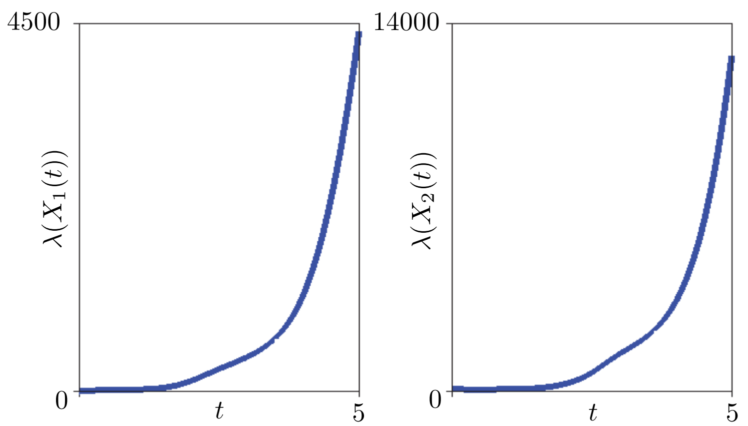


FIG. 3. *Smallest eigenvalue of $X_j(t)$, $j \in \{1, 2\}$, for the system in Example* 4(b).

**5. Error boundedness of linear state estimation.** The concept of AR can be applied in state estimation problems to characterize stability of the estimator and the positiveness of the covariance matrix of the estimation error. Therefore, in this section we assume that the variable $x$ is indirectly observed as described by the system

$$\begin{aligned} dx(t) &= A_{\theta(t)}x(t)dt + B_{\theta(t)}dw(t), \\ dy(t) &= C_{\theta(t)}x(t)dt + D_{\theta(t)}dv(t), \end{aligned} \tag{5.1}$$

where $y(t) \in \mathbb{R}^{n_y}$ is the measurement process and $D_{\theta(t)}D'_{\theta(t)} > 0$ (nonsingular measurement noise) for each $t \geq 0$. In this section we study the LMMSE proposed in [8], which is one of the most successful filters for continuous-time MJLSs. To be consistent with [8] we consider the following.

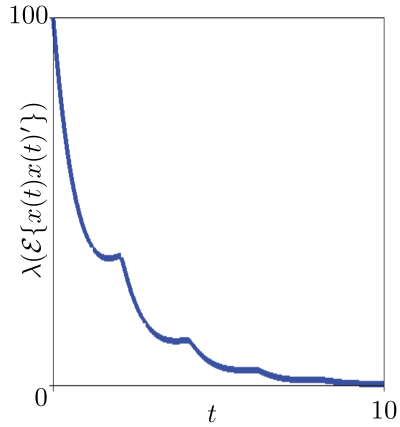


FIG. 4. *Smallest eigenvalue of* $\mathcal{E}\{x(t)x(t)'\}$, $t \in [0,5]$, *for the system in Example* 4(c).

ASSUMPTION 5. *The Markov chain is ergodic, and* $p_j(t)$ *is positive for all* $j \in \mathcal{S}$ *and* $t \geq 0$.

ASSUMPTION 6. $\{v(t)\}_{t\geq 0}$ *is a Wiener process of dimension* $n_v$ *with incremental covariance operator* $I_v dt$ *(*$I_v$ *is the identity matrix of dimension* $n_v \times n_v$*) satisfying* $\mathcal{E}\{w(t)v(t)'\} = 0$ *and* $\mathcal{E}\{x_0 v(t)'\} = 0$.

*Remark* 5. Assumptions 5 and 6 appear in the derivation of the LMMSE in [8]. Although some results and most of the elements in the proofs of this section are irrespective of these assumptions, or may be adapted to situations where they are void, we are inevitably tied to them when using the LMMSE, because Assumption 5 ensures existence of the estimator and Assumption 6 is linked with the optimality.

By defining $\hat{x}(t)$ as the state estimate at time instant $t$ and setting $\tilde{x}(t) = x(t) - \hat{x}(t)$, it is shown in [8] that

$$d\tilde{x}(t) = (A_{\theta(t)} - K_{\theta(t)}(t)C_{\theta(t)})\tilde{x}(t)dt + B_{\theta(t)}dw(t) - K_{\theta(t)}(t)D_{\theta(t)}dv(t), \tag{5.2}$$

where $K(t) = (K_1(t), \ldots, K_N(t))$ is a certain set of gains of appropriate dimensions. We define $Q(t) = \mathcal{E}\{\tilde{x}(t)\tilde{x}(t)'|\mathcal{F}_t, \mathcal{G}_t\}$, where $\mathcal{F}_t = \sigma\{y(s) : 0 \leq s \leq t\}$ and $\mathcal{G}_t = \sigma\{\theta(s) : 0 \leq s \leq t\}$ are the $\sigma$-fields generated by the measurement process and the Markov chain until the time instant $t$, respectively. Using standard machinery from filtering theory for time varying linear systems theory, we can derive the following linear differential equation related to $Q$:

$$\begin{aligned}\dot{Q}(t) = {} & (A_{\theta(t)} - K_{\theta(t)}(t)C_{\theta(t)})Q(t) + Q(t)(A_{\theta(t)} - K_{\theta(t)}(t)C_{\theta(t)})' + B_{\theta(t)}B'_{\theta(t)} \\ & + K_{\theta(t)}(t)D_{\theta(t)}D'_{\theta(t)}K'_{\theta(t)}(t), \quad t \geq 0,\end{aligned} \tag{5.3}$$

with $Q(0) = \Psi - \bar{x}_0\bar{x}_0' \in \mathfrak{R}^{n0}$. We also introduce the average matrix

$$Q_j(t) = \mathcal{E}\{Q(t)1_{\{\theta(t)=j\}}\}, \quad j \in \mathcal{S}.$$

Now, when differentiating the average matrix $Q_j(t)$ we obtain

$$\begin{aligned}\dot{Q}_j(t) = {} & (A_j - K_j(t)C_j)Q_j(t) + Q_j(t)(A_j - K_j(t)C_j)' + \sum_{i=1}^{N} \lambda_{ij} Q_i(t) \\ & + (B_j B_j' + K_j(t) D_j D_j' K_j'(t)) p_j(t).\end{aligned} \tag{5.4}$$

Proposition 5.1. *If $(A, B, \Lambda)$ is AR, then $Q_j(t) > 0$ for all $t > 0$ and $j \in \mathcal{S}$, or, equivalently, there is no nontrivial projection operator $\mathfrak{P}$ such that $\mathfrak{P}(\tilde{x}(t)) = 0$ (a.s.).*

*Proof.* In this proof, for brevity we write $E$ (a.s. $t, j$) for any event $E$ such that $\mathcal{P}(E|\theta(t) = j) = 1$. We proceed by contradiction supposing that there exist $t > 0$ and $v \neq 0$ such that $vQ_j(t)v = 0$ for some $j \in \mathcal{S}$. Recalling that we have assumed $p_j(t) > 0$, this is equivalent to $v'\mathcal{E}\{Q(t)|\theta(t) = j\}v = 0$, yielding $\mathcal{P}(v'Q(t)v = 0|\theta(t) = j) = 1$, that is, $v'Q(t)v = 0$ (a.s. $t, j$). Consider $\Gamma(\cdot, \cdot)$ the transition matrix of $\dot{z} = (A_{\theta(s)} - K_{\theta(s)}(s)C_{\theta(s)})z$. For each realization of $\{\theta(t)\}_{t \geq 0}$ the estimation error matrix $Q(t)$ can be written as

$$Q(t) = \Gamma(t, 0)Q(0)\Gamma(t, 0)' + \int_0^t \Gamma(t, \tau)\tilde{B}(\tau)\tilde{W}(\tau)\tilde{B}(\tau)'\Gamma(t, \tau)' d\tau,$$

where $\tilde{B}(t) = \begin{bmatrix} B_{\theta(t)} & -K_{\theta(\tau)}(\tau)D_{\theta(t)} \end{bmatrix}$ and

$$\tilde{W}(t) = \begin{bmatrix} I_w & 0 \\ 0 & I_v \end{bmatrix}, \tag{5.5}$$

where 0 represents a null matrix of appropriated dimensions. Thus, $v'Q(t)v = 0$ (a.s. $t, j$) yields $\tilde{W}(\tau)\tilde{B}(\tau)'\Gamma(t, \tau)'v = 0$ (a.s. $t, j$) for all $\tau \in [0, t]$, which in turn leads to $Q(0)\Gamma(t, 0)'v = 0$ and

$$B'_{\theta(\tau)}\Gamma(t, \tau)'v = 0 \quad \text{and} \quad K_{\theta(\tau)}(\tau)'\Gamma(t, \tau)'v = 0 \text{ (a.s. } t, j) \text{ for } 0 \leq \tau \leq t, \tag{5.6}$$

due to the positiveness of both $\tilde{W}(t)$ and $D_{\theta(t)}D'_{\theta(t)}$ for each $t$. Noting that

$$\frac{d\Gamma(t, \tau)}{d\tau} = -\Gamma(t, \tau)(A_{\theta(\tau)} - K_{\theta(\tau)}(\tau)C_{\theta(\tau)})$$

and using (5.6), one obtains

$$v'\frac{d\Gamma(t, \tau)}{d\tau} = -v'\Gamma(t, \tau)A_{\theta(\tau)} \text{ (a.s. } t, j) \quad \text{for } 0 \leq \tau \leq t. \tag{5.7}$$

Now consider the state transition matrix $\Phi$ linked with the solution of (1.1). It is simple to check for each realization of the chain that

$$v'\frac{d\Phi(t, \tau)}{d\tau} = -v'\Phi(t, \tau)A_{\theta(\tau)} \quad \text{for } 0 \leq \tau \leq t. \tag{5.8}$$

From (5.7) and (5.8), and the fact that $\Gamma(t, t) = \Phi(t, t) = I$, we obtain

$$\Gamma(t, \tau)'v = \Phi(t, \tau)'v \text{ (a.s. } t, j) \quad \text{for } 0 \leq \tau \leq t. \tag{5.9}$$

Next we shall show that $(A, B, \Lambda)$ is not AR, considering a version of system (1.1) with different statistics for the initial condition; in order to avoid any confusion with (1.1), we will introduce the system

$$d\zeta(t) = A_{\theta(t)}\zeta(t)dt + B_{\theta(t)}dw(t), \tag{5.10}$$

with initial condition satisfying $\mathcal{E}\{\zeta(0)\} = 0$ and $\mathcal{E}\{\zeta(0)\zeta(0)'\} = \Psi - \bar{x}_0\bar{x}_0'$. Note that (5.10) is AR because the AR notion is irrespective of the statistics of the initial condition. We also introduce the process $\mathcal{X}(t) = \mathcal{E}\{\zeta(t)\zeta(t)'|\mathcal{G}_t\}$, which satisfies for each realization of the chain

$$\mathcal{X}(s) = \Phi(s,0)(\Psi - \bar{x}_0\bar{x}_0')\Phi(s,0)' + \int_0^s \Phi(s,\tau)B_{\theta(\tau)}B'_{\theta(\tau)}\Phi(s,\tau)'d\tau \quad \text{for all } s > 0.$$

Then, since $Q(0)\Gamma(t,0)'v = 0$, the first equality in (5.6) and that in (5.9) produce $v'\mathcal{X}(t)v = 0$ (a.s. $t, j$). Hence we have $0 = v'\mathcal{E}\{\mathcal{X}(t)|\theta(t) = j\}v = v'\mathcal{E}\{\zeta(t)\zeta(t)'|\theta(t) = j\}v = v'\mathcal{E}\{\zeta(t)\zeta(t)'1_{\{\theta(t)=j\}}\}v$, which is a contradiction with the AR of system (5.10) (see Corollary 4.7(iv)). Thus, we have shown that

$$Q_j(t) > 0 \quad \text{for all } t > 0,\ j \in \mathcal{S}. \tag{5.11}$$

It remains to prove that for any nontrivial projection operator $\mathfrak{P}$ acting on subspaces of $\mathbb{R}^n$ we obtain $\mathfrak{P}\tilde{x}(t) \neq 0$ for all $t > 0$. In fact, let us suppose that there exists $\mathfrak{P}$ such that $\mathfrak{P}\tilde{x}(t) = 0$ (a.s.) for some $t > 0$. That is, there exists $v \in \mathbb{R}^n$, $v \neq 0$, such that $(v'\tilde{x}(t))v = vv'\tilde{x}(t) = 0$ (a.s.) for some $t > 0$ and $v \neq 0$, which is equivalent to $v'\tilde{x}(t) = 0$ (a.s.). This is equivalent to $v'Q_j(t)v = v'\mathcal{E}\{\tilde{x}(t)\tilde{x}(t)'|\theta(t) = j\}p_j(t)v = 0$ for some $t > 0$ and $v \neq 0$, which is absurd in view of (5.11). ∎

Apart from characterizing positiveness of the average error covariance, a natural question refers to its boundedness. [17] provides an answer by showing that $Q_j(t)$ converges as time tends to infinity whenever the system is mean square detectable, irrespective of AR of $(A, B, \Lambda)$. In the remainder of this section we assume boundedness of $Q$.

Assumption 7. *There exists a matrix $Q^* \in \mathfrak{R}^{n0}$ such that $Q_j(t) \leq Q^*$ for all $t \geq 0$ and $j \in \mathcal{S}$.*

Expanding on the boundedness of $Q$, we study its invariance under perturbations on the model. This issue has been addressed in the filtering literature, with some variations, under the general denomination of *filter stability* and *error bound filter* or *error sensitivity*; see, e.g., [21, 28] for systems without Markov jumps. For discrete-time MJLSs some preliminary studies were presented in [18]. Here, we introduce perturbations on the initial condition of (5.3) and study its new solution, denoted by $P(t)$, which satisfies

$$\begin{aligned}\dot{P}(t,\Sigma) &= (A_{\theta(t)} - K_{\theta(t)}(t)C_{\theta(t)})P(t,\Sigma) + P(t,\Sigma)(A_{\theta(t)} - K_{\theta(t)}(t)C_{\theta(t)})' \\ &\qquad + B_{\theta(t)}B'_{\theta(t)} + K_{\theta(t)}(t)D_{\theta(t)}D'_{\theta(t)}K_{\theta(t)}(t)', \\ P(0,\Sigma) &= \Sigma,\end{aligned} \tag{5.12}$$

where $\Sigma \in \mathfrak{R}^{n0}$. If $P$ is average bounded for any $\Sigma \in \mathfrak{R}^{n0}$, we say that *the filter is stable*. Stability is relevant when implementing the filter because the modeled initial condition $\Psi - \bar{x}_0\bar{x}_0'$ may contain errors, and even infinitesimal errors may cause fast divergence of $P$; see Example 6. $Q(t)$ can be interpreted as the "modeled error covariance" linked with (1.1), based on which the gain function $K(t)$ is designed (in fact, $K_j(t) = Q_j(t)C_j'(D_jD_jp_j(t))^{-1}$, $j \in \mathcal{S}$, as prescribed in [8]), while $P(t)$ is interpreted as the "actual error covariance" of the estimates obtained with this gain function.

Definition 5.2. *We say that the LMMSE is stable if, for each $\Sigma \in \mathfrak{R}^{n0}$, there exists $P^* \geq 0$ such that $\mathcal{E}\{P(t)\} \leq P^*$ for all $t \geq 0$, where $P(t)$ is given by* (5.12).

THEOREM 5.3. *If Assumption 7 holds and $Q_j(t) > 0$ for all $t > 0$ and $j \in \mathcal{S}$, then the LMMSE is stable.*

*Proof.* For a fixed $t^* > 0$ we have $Q(t^*) \geq \mu I$ for some $\mu > 0$. By differentiating $\mathcal{E}\{P(t)1_{\{\theta(t)=j\}}\}$ and denoting $\dot{P}_j(t) = \frac{d}{dt}\left(\mathcal{E}\{P(t)1_{\{\theta(t)=j\}}\}\right)$, $j \in \mathcal{S}$, we obtain

$$\dot{P}_j(t,\Sigma) = (A_j - K_j(t)C_j)P_j(t,\Sigma) + P_j(t,\Sigma)(A_j - K_j(t)C_j)' + \sum_{i=1}^{N} \lambda_{ij} P_i(t,\Sigma) + (B_j B_j' + K_j(t) D_j D_j' K_j(t)')p_j(t). \tag{5.13}$$

For each $\Sigma$ there exists a sufficiently small $\epsilon > 0$ such that the following inequalities hold: $\epsilon\mu < 1$ and $\epsilon P_j(t^*,\Sigma) \leq I$; the later inequality allows us to write

$$\epsilon\mu P_j(t^*,\Sigma) \leq \mu I \leq Q_j(t^*). \tag{5.14}$$

Now, consider the functions $Z_j(t) = Q_j(t) - \epsilon\mu P_j(t,\Sigma)$ for each $j \in \mathcal{S}$ and $t \geq t^*$ satisfying

$$\dot{Z}_j(t) = \mathcal{T}_{A-K(t)C,j}(Z(t)) + (B_j B_j' + K_j(t) D_j D_j' K_j(t)')(1-\epsilon\mu)p_j(t). \tag{5.15}$$

From (5.14) and $1-\epsilon\mu > 0$, one can check that the solution of (5.15) is positive definite for $t \geq t^*$ and $j \in \mathcal{S}$. This fact and Assumption 7 lead to $P_j(t) \leq P^*$ for $t \geq t^*$ and all $j \in \mathcal{S}$, where $P^* = (\epsilon\mu)^{-1}Q^*$. Finally we have $\mathcal{E}\{P(t)\} = \sum_{i\in\mathcal{S}} P_i(t) \leq NP^*$, concluding the proof. ∎

The above result and Proposition 5.1 lead to the following.

COROLLARY 5.4. *If Assumption 7 holds and $(A, B, \Lambda)$ is AR, then the LMMSE is stable.*

*Remark* 6. The arguments in the proof of Theorem 5.3 are irrespective of Assumptions 5 and 6, so the result can be directly extended to other linear estimators.

*Remark* 7. It is shown in [8] for a mean square- (MS-) detectable and MS-stabilizable system that $Q(t)$ converges and the associated gains given by $K_i^* = \lim_{t\to\infty} K_i(t)$, $i \in \mathcal{S}$, yield an MS-stable limiting dynamics $A_i - K_i^* C_i$, which can be employed to show that the LMMSE is stable. This means that MS-detectability and MS-stabilizability work together as a sufficient condition for the stability of the filter, in this sense *competing* with AR; in fact, there are systems that are AR and not MS-stabilizable; see Example 5. Note that the proof of Theorem 5.3 is irrespective of the limiting dynamics of the system, which may allow for future extensions. As a simple example, we can replace matrices $C_i$ by periodic matrices $C_i(t+\delta) = C_i(t)$ (with period $\delta > 1$), thus avoiding the convergence of $Q(t)$, and still keep Theorem 5.3 unaltered. Another example is given in Remark 9 and Example 7.

*Remark* 8 (plant with correlated noise). The noise processes affecting variables $x$ and $y$ may be correlated in many applications. With this motivation, and as an illustration of a simple extension of our results, we address here the case of correlated noise processes (weakening Assumption 6). The same formulation for the estimator given in [8] is considered here because it is still appealing due to its relative simplicity (gains can be precomputed and stored) and the fact that the degree of suboptimality may be relatively small depending on how correlated the noises are. Let $\mathcal{E}\{w(t)v(t)'\} = W(t)$. The single change we have in Proposition 5.1 is that instead of (5.5) we have

$$\tilde{W}(t) = \begin{bmatrix} I_w & W(t) \\ W(t)' & I_v \end{bmatrix}.$$

Then, a sufficient condition for the result of Proposition 5.1 to remain valid is

$$W(t) \text{ is such that } \tilde{W}(t) > 0, \quad t > 0.$$

Note that positiveness of $\tilde{W}$ was automatically granted when $W(t) = 0$; this is also true when $\|W(t)\|$ is small enough. Theorem 5.3 is valid in this setup, as mentioned in Remark 5. Then, we conclude that the filter is stable provided the MJLS is AR and the above additional condition is valid.

*Remark* 9 (plant with time-varying noise covariance). We give here an extension of our stability result to a class of systems featuring time-varying intensity of the additive noise in the variable $x$, modeled as

$$\begin{aligned} dx(t) &= A_{\theta(t)}x(t)dt + F_{\theta(t)}(t)dw(t), \\ dy(t) &= C_{\theta(t)}x(t)dt + D_{\theta(t)}dv(t). \end{aligned}$$

Consider the same filter formulation of the LMMSE, except that $F$ is now time dependent, and consider Assumptions 1–3 and 5–7 (Assumption 4 is relaxed to cope with the above setup). It is well known that the error covariance matrix increases with the intensity of the additive noise, or, formally, if we denote by $Q(t, F)$ the error covariance to emphasize the dependence on $F$, and if we denote by $G$ a set of matrices of the same dimension as $F$, then it can be shown that $Q(t, F) \geq Q(t, G)$ whenever $FF' \geq GG'$. Then, provided we find a set of matrices $G$ such that $F_i(t)F_i(t)' \geq G_iG_i'$ for all $t \geq 0$ and check that $(A, G, \Lambda)$ is AR, from the above relation and Proposition 5.1 we obtain $Q(t, F(t)) \geq Q(t, G) > 0$. Moreover, one can easily check that Theorem 5.3 remains valid for the above system, allowing us to conclude that the filter is stable whenever $(A, G, \Lambda)$ is AR and $G$ fulfils $F_i(t)F_i(t)' \geq G_iG_i'$, $t \geq 0$. An illustration is given in Example 7.

*Example* 5. Let $A$ and $B$ be as in Example 4(a), and let

$$C_1 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 & -0.5 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad D_1 = D_2 = D_3 = 1,$$

$$\Lambda = \begin{pmatrix} -1 & 0.5 & 0.5 \\ 0.5 & -2 & 1.5 \\ 0.01 & 0.99 & -1 \end{pmatrix}, \ \pi = 1/3\,(1 \ \ 1 \ \ 1)\,, \ \bar{x}_0 = \begin{pmatrix} 1.1 \\ -1.5 \end{pmatrix}, \ \text{ and } \Psi = I \in \Re^2.$$

First, we consider $\Psi - \bar{x}_0\bar{x}_0'$ as the initial covariance matrix for the LMMSE design ("nominal setup"), and, in a second step, we employ this filter for estimating the state of a system with initial covariance $\Sigma$ ("actual setup"). We have computed $Q_i(t)$ via (5.4) and checked Assumption 7 by direct inspection of the "nominal" error $Q(t) = \sum_{i=1}^{N} Q_i(t)$. AR of $(A, B, \Lambda)$ has been checked in Example 4, so that Theorem 5.3 ensures that the filter is stable. Figure 5 illustrates the norm of the "actual" error $P(t)$ for five different initial conditions $P(0) = \Sigma$, confirming that $P(t)$ is bounded. We mention in connection with Remark 7 that, by applying the linear matrix inequalities test appearing in [8], the system is not MS-stabilizable.

*Example* 6. Consider the system $(A, B, C, D, \Lambda)$ with

$$A_1 = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 = 0, \quad A_3 = I, \quad C_1 = C_3 = 0, \quad C_2 = \begin{pmatrix} 0 & 0 \\ -10 & 0 \end{pmatrix},$$

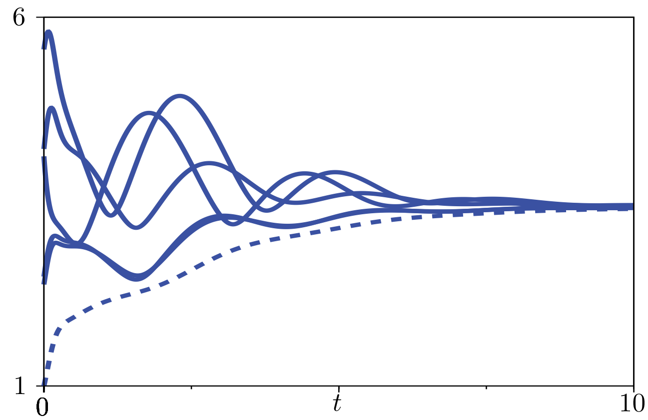


FIG. 5. *Norm of the matrices* $\mathcal{E}\{P(t)\}$ *for different values of* $\Sigma$ *(solid lines) and* $\mathcal{E}\{Q(t)\}$ *(dashed line) on the interval* $[0, 10]$ *for the LMMSE of Example* 5.

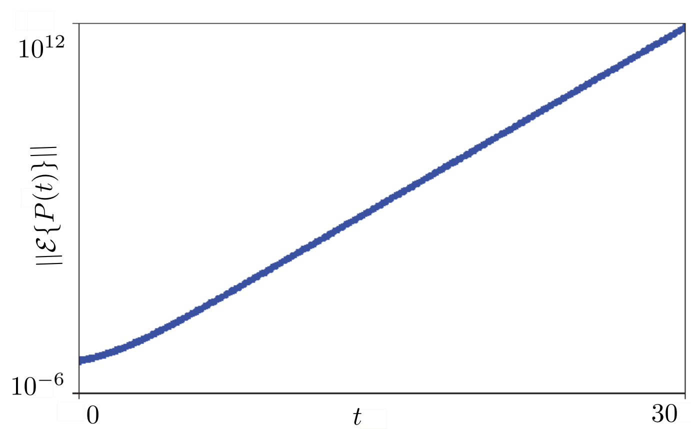


FIG. 6. *Norm of the matrix* $\mathcal{E}\{P(t)\}$ *(logarithmic scale) on the interval* $[0, 30]$ *for the LMMSE of Example* 6(a).

$B_j = 0 \in \Re^{2,3}$, $D_j = I \in \Re^2$, $j \in \mathcal{S}$, and $\Lambda$, $\pi$, and $\bar{x}_0$ are as in Example 5. By applying the rank test in Theorem 3.6, we check that the system is not AR. Now we present the following two cases.

(a) $Q(0) = 0 \in \Re^2$. The filter presents bounded nominal error $Q(t) = 0$, while the actual error $P(t)$ diverges exponentially as illustrated in Figure 6, where we have considered $\Sigma = 10^{-6}I$.

(b) $Q(0) = I$. We note from simulations with $\Sigma = 1.7I$ that both the nominal error $Q(t)$ and the actual error $P(t)$ are bounded; see Figure 7.

*Example* 7. We illustrate the stability conditions given in Remark 9 in a numerical example. Consider parameters as in Example 5, and assume that $F$ changes

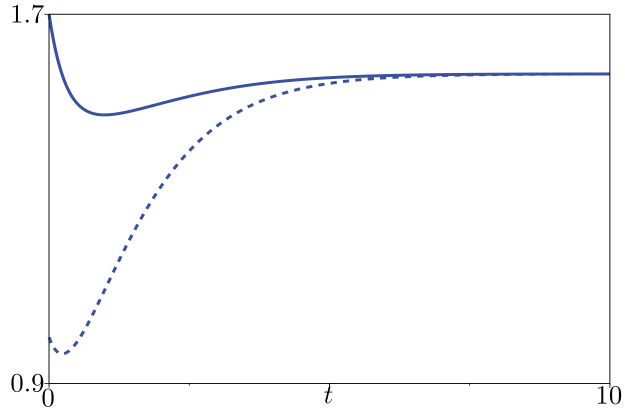


FIG. 7. *Norm of the matrix* $\mathcal{E}\{P(t)\}$ *(solid line) and* $\mathcal{E}\{Q(t)\}$ *(dashed line) on the interval* $[0, 10]$ *for the LMMSE of Example* 6(b).

periodically with period 5 according to the following rule. For each $\ell = 0, 1, \dots,$

$$F_1(t) = \begin{cases} B_1, & 10\ell \le t \le 10\ell + 5, \\ 2B_1 & \text{otherwise,} \end{cases}$$

where $B$ is as given in Example 5 (also Example 4(a)). Moreover, $F_i(t) = B_i$ for $i = 2, 3$, $t \ge 0$. Recall from Example 5 that $(A, B, \Lambda)$ is AR, and note that $F(t) \ge B > 0$, so that from the results given in Remark 9 we conclude that the filter is stable. In order to check this fact, similarly to Example 5, we show in Figure 8 the norm of $\mathcal{E}\{Q(t)\}$ and $\mathcal{E}\{P(t)\}$ with different initial conditions, confirming that $P$ is bounded, as expected for a stable filter. Note that one would not be able to check stability of the filter via its limiting behavior, as $Q(t)$ does not converge.

**6. Conclusions.** We have studied the average reachability gramian for MJLSs with emphasis on its nonsingularity. Systems featuring nonsingular average gramian are referred to as AR (average reachable), and we introduce three notions of AR depending on the initial distribution of the Markov state, ranging from the weaker $\pi$-AR for situations when $\pi$ is known, $\Pi$-AR when $\pi$ excites recurrent Markov states only, and AR when no information is available on $\pi$. A testable, necessary, and sufficient condition for $\pi$-AR has been obtained in Theorem 3.6, based on which a testable condition for AR is derived; see Theorem 4.2. The role of the AR notion in the structure of the system excited by a Wiener process is explored, seeking positiveness of the state second moment $E\{x(t)x(t)'\}$ and of the conditional second moment $X_j(t)$ for recurrent Markov states $j$. One difficulty is that $X_j(t)$ is described by the nonhomogeneous differential equation (4.7) involving the distribution $\pi_t$, and we do not assume any hypothesis on the Markov chain, which has forced us to introduce a "modulated" second moment $S_j(t)$ (see Remark 3) that plays a key role in linking the definition of AR with the rank test of Theorem 3.6 and with the positiveness of $X$, as stated in Theorem 4.5. Positiveness of the state second moment of an MJLS does not mean that the state $x(t)$ eventually reaches any neighborhood of $\mathbb{R}^n$ as it

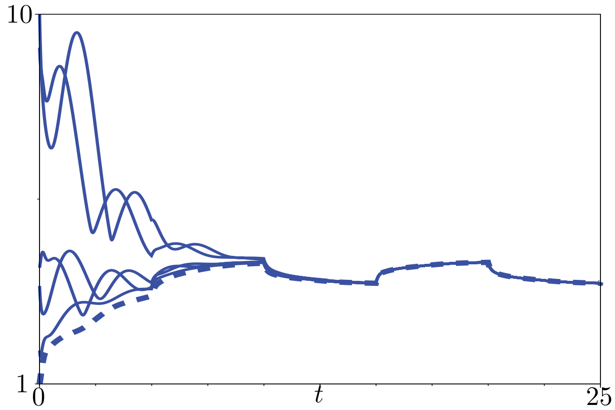


Fig. 8. *Norm of the matrices* $\mathcal{E}\{P(t)\}$ *for different values of* $\Sigma$ *(solid lines) and* $\mathcal{E}\{Q(t)\}$ *(dashed line) on the interval* $[0, 25]$ *for the LMMSE of Example* 7.

does in linear deterministic systems [11]; an illustration of a possible behavior of $x(t)$ is given in Example 2. The role of AR is studied in the context of linear filtering to show that the error covariance matrix is positive with the interpretation that the estimate is "never precise" in the sense of Proposition 5.1. We also show that AR ensures stability of the LMMSE in [8], as stated in Corollary 5.4, meaning that the error covariance matrix remains bounded under arbitrary perturbations on the initial condition $\Psi$. The key passages in our analysis of the LMMSE are (5.6), which may be interpreted as the filter gain not acting on subspaces where there is no noise in the variable $x$, and the order preserving property of the differential equation (5.15), which enabled us to employ the positiveness of $Q$ in a productive way in (5.14); note also that ergodicity of the Markov chain has never been employed in the arguments. As illustrated in Remarks 8 and 9, these properties may remain unaltered or be adapted when dealing with other problems, so that future work may look into extensions to other systems and filters.

## REFERENCES

[1] M. L. Bujorianu and J. Lygeros, *New insights on stochastic reachability*, in Proceedings of the 46th IEEE Conference on Decision and Control CDC'07, New Orleans, LA, 2007, pp. 6172–6177, doi:10.1109/CDC.2007.4434784.

[2] E. F. Costa and A. Astolfi, *Characterization of exponential divergence of the Kalman filter for time-varying systems*, SIAM J. Control Optim., 48 (2010), pp. 2917–2944, doi:10.1137/080717055.

[3] E. F. Costa and J. B. R. do Val, *On the observability and detectability of continuous-time Markov jump linear systems*, SIAM J. Control Optim., 41 (2002), pp. 1295–1314, doi:10.1137/S0363012901385460.

[4] E. F. Costa and J. B. R. do Val, *Weak detectability and the linear-quadratic control problem of discrete-time Markov jump linear systems*, Internat. J. Control, Special Issue on Switched and Polytopic Linear Systems, 75 (2002), pp. 1282–1292, doi:10.1080/002071702000023717.

[5] E. F. Costa, J. B. R. do Val, and M. D. Fragoso, *A new approach to detectability of discrete-time infinite Markov jump linear systems*, SIAM J. Control Optim., 43 (2005), pp. 2132–2156, doi:10.1137/S036301290342992X.

[6] O. L. V. Costa and M. D. Fragoso, *Discrete-time LQ-optimal control problems for infinite Markov jump parameter systems*, IEEE Trans. Automat. Control, 40 (1995), pp. 2076–2088, doi:10.1109/9.478328.

[7] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-Time Markov Jump Linear Systems*, Springer-Verlag, New York, 2005, doi:10.1007/b138575.

[8] O. L. V. Costa, M. D. Fragoso, and M. G. Todorov, *Continuous-Time Markov Jump Linear Systems*, Springer-Verlag, New York, 2013, doi:10.1007/978-3-642-34100-7.

[9] O. L. V. Costa and R. P. Marques, *Robust $H_2$-control for discrete-time Markovian jump linear systems*, Internat. J. Control, 73 (2000), pp. 11–21, doi:10.1080/002071700219894.

[10] J. H. Davis, *Foundations of Deterministic and Stochastic Control*, Birkhäuser, Basel, 2002, doi:10.1007/978-1-4612-0071-0.

[11] M. H. A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall, London, 1977.

[12] J. B. R. do Val and E. F. Costa, *Numerical solution for linear-quadratic control problems of Markov jump linear systems and weak detectability concept*, J. Optim. Theory Appl., 114 (2002), pp. 69–96, doi:10.1023/A:1015412121001.

[13] V. Dragan, T. Morozan, and A. M. Stoica, *Mathematical Methods in Robust Control of Discrete-Time Linear Stochastic Systems*, Springer, New York, 2009, doi:10.1007/978-1-4419-0630-4.

[14] V. Dragan, T. Morozan, and A. M. Stoica, *Mathematical Methods in Robust Control of Linear Stochastic Systems*, 2nd ed., Springer, New York, 2013, doi:10.1007/978-1-4614-8663-3.

[15] M. D. Fragoso and O. L. V. Costa, *Mean square stabilizability of continuous-time linear systems with partial information on the Markovian jumping parameters*, Stoch. Anal. Appl., 22 (2004), pp. 99–111, doi:10.1081/SAP-120028025.

[16] M. D. Fragoso and O. L. V. Costa, *A unified approach for stochastic and mean square stability of continuous-time linear systems with Markovian jumping parameters and additive disturbances*, SIAM J. Control Optim., 44 (2005), pp. 1165–1191, doi:10.1137/S0363012903434753.

[17] M. D. Fragoso and O. L. V. Costa, *A separation principle for the continuous-time LQ-problem with Markovian jump parameters*, IEEE Trans. Automat. Control, 55 (2010), pp. 2692–2707, doi:10.1109/TAC.2010.2048056.

[18] M. J. F. Gomes and E. F. Costa, *On the stability of the recursive Kalman filter with Markov jump parameters*, in Proceedings of the American Control Conference ACC'10, Baltimore, MD, 2010, pp. 4159–4163, doi:10.1109/ACC.2010.5530604.

[19] A. P. Gonçalves, A. R. Fioravanti, and J. C. Geromel, *$H_\infty$ robust and networked control of discrete-time MJLS through LMIs*, J. Franklin Inst., 349 (2012), pp. 2171–2181, doi:10.1016/j.jfranklin.2012.03.002.

[20] J. Huang, *Analysis and Synthesis of Semi-Markov Jump Linear Systems and Networked Dynamic Systems*, Ph.D. thesis, University of Victoria, Victoria, BC, Canada, 2013, http://hdl.handle.net/1828/4592.

[21] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[22] Y. Ji and H. J. Chizeck, *Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control*, IEEE Trans. Automat. Control, 35 (1990), pp. 777–788, doi:10.1109/9.57016.

[23] Y. Ji and H. J. Chizeck, *Jump linear quadratic Gaussian control in continuous time*, IEEE Trans. Automat. Control, 37 (1992), pp. 1884–1892, doi:10.1109/9.182475.

[24] M. Mariton, *Stochastic controllability of linear systems with Markovian jumps*, Automatica J. IFAC, 23 (1987), pp. 783–785, doi:10.1016/0005-1098(87)90039-2.

[25] T. Morozan, *Stability and control for linear systems with jump Markov perturbations*, Stoch. Anal. Appl., 13 (1995), pp. 91–110, doi:10.1080/07362999508809384.

[26] A. R. R. Narváez and E. F. Costa, *On the observability of continuous time linear systems with Markov jump parameters*, in Proceedings of the American Control Conference ACC'10, Baltimore, MD, 2010, pp. 4170–4174, doi:10.1109/ACC.2010.5530597.

[27] A. R. R. Narváez and E. F. Costa, *On the controllability of continuous-time Markov jump linear systems*, in Proceedings of the IFAC 18th World Congress of the International Federation of Automatic Control, Milan, Italy, 2011, pp. 9103–9108, doi:10.3182/20110828-6-IT-1002.03706.

[28] T. Nishimura, *Error bounds of continuous Kalman filters and the application to orbit determination problems*, IEEE Trans. Automat. Control, 12 (1967), pp. 268–275, doi:10.1109/TAC.1967.1098597.

[29] S. Sangsuk-Iam and T. E. Bullock, *Analysis of discrete-time Kalman filtering under incorrect noise covariances*, IEEE Trans. Automat. Control, 35 (1990), pp. 1304–1309, doi:10.1109/9.61006.

[30] A. A. G. Siqueira and M. H. Terra, *Nonlinear and Markovian $H_\infty$ controls of underactuated manipulators*, IEEE Trans. Control Syst. Tech., 12 (2004), pp. 811–826, doi:10.1109/TCST.2004.833626.

[31] Z. Sun and S. S. Ge, *Switched Linear Systems: Control and Design*, Springer, New York, 2005, doi:10.1007/1-84628-131-8.

[32] D. D. Sworder and R. O. Rogers, *An LQ-solution to a control problem associated with a solar thermal central receiver*, IEEE Trans. Automat. Control, 28 (1983), pp. 971–978, doi:10.1109/TAC.1983.1103151.