On the Detectability and Observability of Discrete-Time Markov Jump Linear Systems

Eduardo F. Costa and João B. R. do Val
UNICAMP - Fac. de Engenharia Elétrica e de Computação
Depto. de Telemática, C.P. 6101, CEP 13081-970, Campinas, SP, Brazil
email: jbosco@dt.fee.unicamp.br

Abstract

This paper presents a new detectability concept for discrete-time Markov jump linear systems with finite Markov state, which generalizes the MS-detectability concept found in the literature. The new sense of detectability can similarly assure that the solution of the coupled algebraic Riccati equation associated to the quadratic control problem is a stabilizing solution. In addition, the paper introduces a related observability concept which also generalizes previous concepts. Tests for detectability or observability are derived from the corresponding definitions, that can be performed in a finite number of steps. An illustrative example is included to show that a system may be detectable in the new sense but not in the MS sense.

1 Introduction

The paper considers the discrete-time Markov jump linear system (MJLS)

\[ \Phi : \begin{cases} x(k+1) = A_{\theta(k)}x(k), & k \geq 0, \\ y(k) = C_{\theta(k)}x(k), & x(0) = x_0, \quad \theta(0) \sim \mu_0 \end{cases} \]  

(1)

where \( x \) and \( y \) are the state and the output variables, respectively. The mode \( \theta \) is the state of an underlying discrete-time homogeneous Markov chain \( \Theta = \{ \theta(k); k \geq 0 \} \) having \( S = \{ 1, \ldots, S \} \) as state space and \( P = [p_{ij}], i, j = 1, \ldots, S \) as the transition probability matrix. The initial distribution of \( \Theta \) is determined by \( \mu_{0,j} = P(\theta(0) = j), i, j = 1, \ldots, S \). Matrices \( A_i \) and \( C_i \), \( i = 1, \ldots, S \) belong to the collections of \( S \) real matrices: \( A = \{ A_1, \ldots, A_S \} \), \( \text{dim}(A_i) = s \times s \), and \( C = \{ C_1, \ldots, C_S \} \), \( \text{dim}(C_i) = q \times s \).

The theory of Markov jump linear systems is fairly complete nowadays, e.g. see \([2], [4], [5], [7], [9] \), and \([11] \), and it parallels the theory of deterministic linear systems in many aspects. One important similarity is the role of the stochastic concepts of stabilizability and detectability to guarantee existence, uniqueness and stability of closed-loop solutions of quadratic infinite horizon control problems, see \([4] \) or \([11] \).

It is widely known for deterministic linear systems that the following assertions are equivalent: (i) there exists a matrix \( G \) for the system \( (A, C) \) such that the spectral radius of \( (A + GC) \) is less than one; (ii) if the output process has “very low” energy, then the state trajectory is “fast decaying”, i.e., every non-observed state corresponds to stable modes of the system. Assertions (i) and (ii) give rise to the same detectability concept for deterministic systems. On the other hand, the single detectability concept available in the MJLS theory parallels the concept in assertion (i). The extension for the second concept has not been investigated yet and it is not known up to this date if a similar equivalence holds.

In this paper we fill this gap: we derive a detectability concept which relates the energy carried out by the output and the trajectory decay. We call this new concept W-detectability, and as a matter of fact, we show that it generalizes MS-detectability, the existing concept in the literature, see \([4] \) and \([10] \). At one stage, a simple example shows that W-detectability does not imply MS-detectability, and the paper also demonstrates that the W-detectability plays the same role as MS-detectability in the theory of coupled algebraic Riccati equations (CARE): it guarantees that the solution to the CARE is stabilizing.

The related concept of W-observability is derived from the W-detectability. It is also shown that the requirements for observability which appear in the MJLS literature, such as in \([10] \) and \([12] \), are more demanding than those for W-observability. Tests that allow computation of the structural properties of W-detectability and W-observability are presented in the paper.

2 Notation, Concepts and Basic Results

Let \( \mathbb{R}^s \) represent the linear space formed by all \( s \)-vectors. Let \( \mathbb{R}_{r_s} \) (respectively, \( \mathbb{R}^{r_s} \)) represent the normed linear space formed by all \( r \times s \) real matrices (respectively, \( r \times r \)) and \( \mathbb{R}^{r_s}_+ \) the closed convex cone \( \{ U \in \mathbb{R}^{r_s} : U = U' \geq 0 \} \) (the open cone \( \{ U \in \mathbb{R}^{r_s} : U = U' > 0 \} \) where
denotes the transpose of U, \( U \geq V \) (i.e., \( U - V \) is positive semidefinite) signifies that \( U - V \in \mathbb{R}^{n \times n}_{\geq} \) (\( U - V \in \mathbb{R}^{n \times n}_{+} \)). Let \( M^{0, i} \) denote the linear space formed by a number \( S \) of matrices such that \( M^{0, i} = \{ U = (U_1, \ldots, U_S) : U_i \in \mathbb{R}^{r_{xi} \times r_{xi}}, i = 1, \ldots, S \} \); also, \( M^{i} = \bigoplus M^{0, i} \). We denote by \( M^{0} (M^{i}) \) the set \( M^{i} \) when it is made up by \( U_i \in \mathbb{R}^{r_{xi} \times r_{xi}} \) for all \( i = 1, \ldots, S \). It is known that \( M^{0, i} \) equipped with the inner product

\[
< U, V > = \sum_{i=1}^{S} \text{tr} \{ U_i^t V_i \}
\]

forms a Hilbert space. We define the norm \( \| U \|^2 = < U, U > \).

Consider system \( \Phi \); for \( i = 1, \ldots, S \) we define

\[
X_i(t) = E_{x_0, \mu_0} \{ x(t) x(t)^t 1_{\theta(t) = i} \}
\]

and, for instance, with this notation we can write

\[
E_{x_0, \mu_0} \{ |x(t)|^2 \} = < X, I > = \| X(t)^{1/2} \|^2 .
\]

We also define the operators \( L : M^{0} \to M^{0} \), and the dual \( T : M^{0} \to M^{0} \) as

\[
L(U) = \sum_{j=1}^{S} p_j A_j U A_j^t
\]

and

\[
T(U) = A_i \{ \sum_{j=1}^{S} p_j U_j \} A_i, \quad i = 1, \ldots, S
\]

We denote \( T^0(U) = U \), and for \( t \geq 1 \), we can define \( T^t(U) \) recursively by \( T^t(U) = T(U^t(U)) \). Notice that \( T \) and \( L \) are linear and continuous on their arguments and the set

\[
C = \{ U \in M^{0} : \| U \| = 1, \| T^t(U) \| \geq \delta \} \quad \text{for some} \quad \delta \geq 0
\]

is a compact set.

For \( N \geq 0 \), let us introduce the functional

\[
W^N(X) = \sum_{n=0}^{N-1} X(t) C C^t
\]

whenever \( X(0) = X \). We consider the sequence \( O(n) \), \( n \geq 0 \) on \( M^{0} \), defined recursively as

\[
O_i(n) := C_i C_i + T_i(O(n - 1)), \quad n = 1, 2, \ldots
\]

for each \( i = 1, \ldots, S \), with \( O(0) = 0 \). Notice that \( O(n + 1) \geq O_n \). For each \( i = 1, \ldots, S \), with \( O(0) = 0 \). Notice that \( O(n + 1) \geq O_n \). For each \( i = 1, \ldots, S \).

The following results are adapted from [3] and [6]; the proof is omitted.

**Lemma 1.** Consider system \( \Phi \). Then,

\[
\begin{align*}
X_i(t + 1) &= L_i(X_i(t)), \quad t \geq 0 \\
X_i(0) &= x_0 + \mu_{0, i}, \quad i = 1, \ldots, S
\end{align*}
\]

and \( W^N(X) = < X, O(N) > \).

## 2.1 MS-Stability, MS-Detectability and W-Detectability

**Definition 1.** The system \( \Phi \) is mean square stable (MS-stable) if for each \( x_0 \) and \( \mu_0 \),

\[
\lim_{k \to \infty} E_{x_0, \mu_0} \{ |x(k)|^2 \} = 0
\]

**Remark 1.** Ji et al. in [11] have shown that MS-stability concept is equivalent to others second moment stability concepts, such as exponential stability. It is also known that: (i) if the system is not MS-stable, then there exists \( X(0) \in M^0 \) such that \( \| X(t) \| \geq p \| X(0) \| \) for some \( 0 < p \leq 1 \) and \( \zeta \geq 1 \); (ii) if \( \lim_{k \to \infty} \| X(t) \| = 0 \), then \( \| X(t) \| \leq \beta \| X(0) \| \) for some \( \beta \geq 1 \) and \( 0 < \zeta < 1 \).

**Definition 2.** Consider system \( \Phi \). We say that \( (A, C) \) is MS-detectable when there exists \( L \in M^{r_{xi}} \) for which the system \( z(t + 1) = A(t) z(t), z(0) = 0, k \geq 0 \), is MS-stable.

Notice that the functional in (4) has the physical interpretation of the accumulated energy of the output process \( y \) on the interval \( 0 \leq t \leq N - 1 \). Indeed,

\[
W^N(X) = E_{x_0, \mu_0} \{ \sum_{n=0}^{N-1} x(t)^t C_0 C_0 x(t) \}
\]

The W-detectability concept relates the energy of the output of the trajectory decay, as follows.

**Definition 3 (W-detectability).** Consider system \( \Phi \). We say that \( (A, C) \) is W-detectable when there exist integers \( N_d \) and \( \gamma > 0 \) such that \( W^N(X) \geq \gamma \| X \| \) whenever \( \| X(t_d) \| \geq \delta \| X \| \).

**Remark 2.** The above concept of detectability is based in standard concepts of detectability for linear time-varying systems, see e.g. [1] or [8]. As we shall see in Lemma 3, the concept retrieves the idea that every non-observed state corresponds to stable modes of the system. Notice that every MS-stable MJLS is W-detectable with \( N_d \) and \( \gamma > 0 \) such that \( W^N(X) \geq \gamma \| X \| \) whenever \( \| X(t_d) \| \geq \delta \| X \| \).

## 3 W-Detectability Properties

We start this section presenting some properties of the functional \( W^N(X) \) and the associated sequence \( O(n) \) in (4) and (5) respectively.

A deterministic linear representation for MJLS with finite Markov state space was proposed in [3]; see Appendix A for details. The associated dimension is

\[
\ell = s^2 S
\]
We apply the aforementioned representation and the controllability theorem for linear time-invariant systems, to obtain the results in the lemma.

Lemma 2. (i) If for some $X$, $W^f(X) = 0$, then $W^n(X) = 0$ for all $n \geq 0$;
(ii) If there exists $n$, $0 < n \leq \ell$, such that $W^f(X(n)) = 0$, then $W^f(X(t)) = 0$ for every $t \geq n$.

See Appendix A for the proof. Now we are ready to present the main result of the section.

Lemma 3. Consider system $\Phi$. $(A, C)$ is W-detectable if and only if whenever $W^f(X) = 0$ one has that $\|X(t)\| \to 0$ as $t \to \infty$.

Proof: Necessity. We shall show that under W-detectability of $(A, C)$, $\|X(t)\| \to 0$ as $t \to \infty$.

Sufficiency. Let us consider the set
$$Z = \{ Z : \|Z\| = 1, X, O(\ell) \geq 0 \}$$
and let $Z(t)$ denote as $Z(t)$ the trajectory corresponding to an initial condition $Z \in Z$. By hypothesis, $\|Z(t)\| \to 0$ as $t \to \infty$ and we can write in a similar fashion to Remark 1 that there exist $0 < \zeta < 1$ and $\beta \geq 1$ such that $\|Z(t)\| \leq \beta Z^\zeta$, $\forall Z \in Z$.

In this proof we shall demonstrate that there exists $\gamma > 0$ such that whenever $\|X(t)\| \geq \delta$ then $W^f(X) = X, O(\ell) \geq \gamma \|X\|$ and consequently $(A, C)$ is W-detectable. Let us deny the assertion and suppose that for every $\gamma > 0$ there exists $\|X\| = 1$ such that $X, O(\ell) > \gamma$ and $\|X(t)\| \geq \delta$, i.e., $X \in C$ where $C$ is the compact set
$$C = \{ X : \|X\| = 1, \|X(t)\| \geq \delta \}$$
which is the complement of set $\bar{C}$. Let us take a sequence $X_n \in C$ such that $X_n, O(\ell) > \gamma_n$ and $\gamma_n \to 0$ as $n \to \infty$. Hence, from the compacity of $C$ there exists a subsequence $X_m$ which converges to some $\bar{X} \in \bar{C}$ and from the continuity of $W^f(X) = X, O(\ell)$,
$$\lim_{m \to \infty}(W^f(X_m)) = W^f(\bar{X}) = 0$$
In view of (8), $\bar{X} \in Z \subset \bar{C}$, which completes the proof by contradiction. \[\square\]

3.1 W-detectability is Weaker than MS-detectability

In this section we show that MS-detectability implies W-detectability, but the reverse implication fails.

Lemma 4. Suppose system $\Phi$ is MS-stable. Then $(A - GD, D)$ is W-detectable for every $G \in M^{\ell \times \ell}$ and $D \in M^{\ell \times \ell}$.

Proof: In this proof $X(\cdot)$ and $T(\cdot)$ refers to the system $\Phi$ and $\tilde{X}(\cdot)$ and $\tilde{T}(\cdot)$ refers to the system
$$\Phi: \begin{cases} x(k + 1) = (A_0(k) - G_0(k)D_0(k))x(k), & k \geq 0 \\ y(k) = D_0(k)x(k), & x(0) = x_0, & \theta(0) \sim \mu_0 \end{cases}$$
We shall show that $\|\tilde{X}(t)\| \to 0$ as $t \to \infty$, provided that $W^f(\tilde{X}) = 0$. From Lemma 2 (i) we can write
$$0 = W^n(\tilde{X}) = \sum_{t=0}^{n-1} G_0(i) X_i D_0(i)$$
for every $n \geq 0$. Then, we have that $<\tilde{X}(t), T/D> = <DX(t)^{1/2}, \tilde{X}(t)^{1/2}> = 0$ which implies that $D_0(X_i)(t) = 0$ for all $i \geq 0$ and $i$. Now, we set $\tilde{A}_i = (A_i - G_0D_0)$ for every $i = 1, \ldots, S$ and we obtain
$$\tilde{X}_i(t) = \sum_{i=1}^{S} p_{ij} \tilde{A}_i \tilde{X}_i(t) \tilde{A}_i$$
Thus, provided that $\tilde{X}(0) = X(0)$, one has that $\tilde{X}(t) = X(t)$ for all $i \geq 1$, and since $\Phi$ is MS-stable, we conclude that $\|\tilde{X}(t)\| = \|X(t)\| \to 0$ as $t \to \infty$. Finally, from Lemma 3 we have that $(A - GD, D)$ is W-detectable. \[\square\]

Theorem 1. If $(A, C)$ is MS-detectable then $(A, C)$ is W-detectable.
Proof. Since $(A, C)$ is MS-detectable, from definition there exists $L$ such that the system

$$\Phi : \begin{cases} x(k+1) = (A_{θ(k)} - L_{θ(k)} C_{θ(k)}) x(k), & k \geq 0 \\ y(k) = D_{θ(k)} x(k), & x(0) = x_0, \quad θ(0) \sim µ_0 \end{cases}$$

is MS-stable. Let $A_i = A_i - L_i C_i$ for all $i$ and, from Lemma 4, $(A - GD, D)$ is W-detectable for every $G \in M^{r \times l}$ and $D \in M^{l \times s}$. The proof is completed by retrieving the original system $Φ$ with the choice $D = C$ and $G = -L$. □

We present here a simple example showing that the converse of Theorem 1 is not necessarily true, and hence that $W$-detectability plays the same role as the MS-detectability concept in optimal linear control problems.

4 W-Detectability and the Linear Quadratic Control

Consider the closed loop version of system $Φ$:

$$\Phi_c : x(k+1) = (A_{θ(k)} - B_{θ(k)} G_{θ(k)}) x(k)$$

where $B \in M^{r \times r}$ and $G \in M^{r \times s}$ can be regarded as a linear state feedback control; consider also the associated cost functional

$$J^N(X) = E_{x_0, P_0} \left\{ \sum_{t=0}^{N-1} x(t)'(Q_{θ(t)} + C_{θ(t)} R_{θ(t)} C_{θ(t)}') x(t) \right\}$$

where $Q \in M^{0 \times 0}$ and $R \in M^{+}$ are weighting matrices. Notice that the functionals $J^N$ and $W^N$ are closed related. In fact, it is easy to check that when $C = (Q + G'RG)^{1/2}$, the expressions for $W^N(X)$ and $J^N(X)$ are identical.

Consider the coupled equation in the unknown $P \in M^{0 \times 0}$

$$P_i = \hat{J}_i(P) + G_i' R_i G_i + Q_i, \quad i = 1, \ldots, S$$

where $\hat{J}_i(U) = A_i' E_i(U) A_i$ and $A_i = A_i - B_i G_i$ for each $i$. It can be shown from [4] or [11] that when $G_i, i = 1, \ldots, S$ satisfies (13), then

$$J^N(X) = \lim_{N \to \infty} J^N(X) \leq \langle X, P \rangle$$

In this section we shall show that system $Φ_c$ is MS-stable provided that $G$ satisfies (13) and $(A, Q^{1/2})$ is W-detectable. In particular, we conclude that the solution to the CARE arising in optimal linear quadratic control problems is a stabilizing solution under the W-detectability assumption, in such a manner that W-detectability plays the same role as the MS-detectability concept in optimal linear quadratic problems, cf. [4] and [7].

Lemma 5. Assume $(A, Q^{1/2})$ is W-detectable. Then, $(A - BG, (Q + G'RG)^{1/2})$ is W-detectable for any $G \in M^{r \times s}$.

Proof. In this proof, $L$ and $W$ refer to the system $Φ_c$ and $L \hat{=} \hat{Φ}$ refers to the trajectory of system $Φ_c$. We show that $\|X(t)\| \to 0$ as $t \to \infty$ whenever $\hat{L} \hat{=} \hat{Φ}(X(\cdot))$.

Consider the system $Φ_c$ and assume that $(A, Q^{1/2})$ is W-detectable and $G$ satisfies (13) in the unknown $P \in M^{0 \times 0}$. Then, the system is MS-stable.

Proof: In this proof $X(\cdot)$ refers to the trajectory of system $Φ_c$. Notice that, from Lemma 5 and the assumption in the theorem, we have that $(\hat{A}, \hat{C})$ is W-detectable, where
**5 W-Observability Concept**

A particularization of the W-detectability concept in Definition 3 is obtained by setting $\delta = 0$. This leads to the following observability concept.

**Lemma 6.** Consider system $\Phi$. $\Phi$ is W-observable if and only if $O_i(\ell) > 0$ for each $i = 1, \ldots, S$.

The observability condition in Definition 4 or Lemma 6 is more general than similar conditions for MJLS in the literature. We can mention the concept found in [12], namely, observability in the deterministic sense of each pair $(p_i, A_i, C_i)$, $i = 1, \ldots, S$. This condition is sufficient for the W-observability to hold, since from (5) we can write

$$O_i(s) \geq C_i' C_i + p_i A_i' C_i A_i + \cdots + p_i^{s-1} A_i' C_i A_i^{-1} > 0$$

In another comparison, Ji and Chizeck introduce in [10] the idea of sample path controllability, for which, there should exist a number $N$ such that for each sample path $\theta(0), \ldots, \theta(N)$, the matrix $[C_{\theta(0)} A_{\theta(1)}' C_{\theta(1)} \cdots A_{\theta(N)}' C_{\theta(N)}]$ has a.s. full rank. It is simple to check that this condition assures observability in Definition 4.

**6 Conclusions**

The paper introduces the concept of W-detectability for MJLS, which generalizes the previous concept of MS-detectability, and plays the same role in the quest for stabilizing solutions of quadratic control problems. The corresponding concept of W-observability also generalizes previous ideas on observability encountered in the literature of discrete-time MJLS. The W-detectability concept carries along a computational test for this structural property as presented in Remark 3. The test of W-observability in Lemma 6 for MJLS is similar to the test for deterministic time-invariant linear systems with dimension $\ell$.

**A Appendix: Proof of Lemma 2**

For $V \in \mathcal{R}^q$, let us identify the columns of $V = [v_1; v_2; \ldots; v_q]$. For $U = (U_1, \ldots, U_5)$ and following [3], let us introduce the following linear and invertible operators

$$\hat{\phi}(V) = \begin{bmatrix} v_1 \\ \vdots \\ v_q \end{bmatrix}, \quad \hat{\phi}(U) = \begin{bmatrix} \phi(U_1) \\ \vdots \\ \phi(U_5) \end{bmatrix}$$

Proposition 4a) of [3] yields that

$$\hat{\phi}(T(U)) = \hat{A}\hat{\phi}(U)$$ (19)
Proof of Lemma 2 (ii).

Let us consider $O(n) = (O_1(n), \ldots, O_S(n))$, $n \geq 0$ as in (5), and define $o^n \in \mathbb{R}^{k \times 1}$ and $q \in \mathbb{R}^{k \times 1}$ by

$$o^n = \hat{\phi}(O(n)), \quad q = \hat{\phi}(C'C)$$

and we write (5) equivalently as

$$o^n = \hat{\phi}(C_1'C_1 + T_1(O(n-1)), \ldots, C_S'C_S + T_S(O(n-1))) = q + \hat{\phi}(T_1(O(n-1)), \ldots, T_S(O(n-1)))$$

the last identity is provided by the linearity of $\hat{\phi}$. From (19), we have that

$$o^n = q + A\hat{\phi}(O_1(n-1), \ldots, O_S(n-1)) = q + A o^{n-1}, \quad n \geq 1$$

and (20) is equivalent to

$$o^n = q + A q + \cdots + A^{n-1} q, \quad n \geq 1. \quad (21)$$

Now let us consider the expression $<X, O(n)>$. We denote $x_v = \hat{\phi}(X) \in \mathbb{R}^{k \times 1}$ and it is straightforward to check that

$$<X, O(n)> = x_v^T o^n \quad (22)$$

and since $<X, T^n(C'C)> \geq 0$, one has that $x_v^T A^n q \geq 0$.

Proof of Lemma 2 (i).

From Lemma 1, we have that $W^X(X) = <X, O(\ell)>$. If there exists $X$ such that $W^X(X) = 0$ then, from (21) and (22), we obtain

$$0 = x_v^T o^\ell = x_v^T (q + A q + \cdots + A^{\ell-1} q) \quad (23)$$

Since $x_v^T A^n q \geq 0$ for each $0 \leq n \leq \ell - 1$, we have that

$$x_v^T A^n q = 0, \quad \text{for each} \ 0 \leq n \leq \ell - 1$$

Thus, using the Cayley-Hamilton theorem, we can conclude that $x_v^T o^k = 0$ for each $k \geq 0$, and from (22) it follows that

$$<X, O(k)> = 0, \quad k \geq 0 \quad \square$$

Proof of Lemma 2 (ii).

From Lemma 1, we have that $W^X(X(t)) = <X(t), O(\ell)>$. If $<X(n), O(\ell)> = 0$ for some $0 \leq n \leq \ell$, then from Lemma 2 (i) we have that $<X(n), O(\ell+n)> = 0$, for each $t \geq n$, and we can write

$$<X(t), O(\ell)> = <L^{\ell-n}(X(n)), O(\ell)>$$

$$= <X(n), T^{\ell-n}(O(\ell)>$$

$$\leq <X(n), O(\ell+n)> = 0 \quad \square$$