Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 12, No. 1-2/2020

STATIONARY LINEAR MEAN SQUARE FILTER FOR THE OPERATION MODE OF CONTINUOUS-TIME MARKOVIAN JUMP LINEAR SYSTEMS*

Fortià Vila Vergés† Marcelo Dutra Fragoso ‡

DOI https://doi.org/10.56082/annalsarscimath.2020.1-2.501

Dedicated to Dr. Vasile Drăgan on the occasion of his 70$^{\text{th}}$ anniversary

Abstract

This paper makes a further foray on the study of the filtering problem for the class of Markov jump linear systems (MJLSs) with partial observations of the Markov parameter (the operation mode). We derive a *stationary filter* for the best linear mean square filter (BLMSF) devised in a recent paper by the authors. It amounts here to obtain the convergence of the error covariance matrix of the best linear mean square filter to a stationary value under some suitable assumptions which includes ergodicity of the Markov chain. The advantage of this scheme is that it is easier to implement since the filter gain computation can be performed offline, leading to a linear time-invariant filter.
MSC: 93E11, 93C05, 93C30, 60J74, 60J27

keywords: Stationary Linear filter, Markov Jump Linear Systems

*Accepted for publication in revised form on July 28, 2020

†`fortiav@lncc.br` National Laboratory for Scientific Computing – LNCC/MCTIC, CEP 25651-075, Av. Getúlio Vargas 333, Petrópolis, Rio de Janeiro, Brazil; Paper written with financial support of Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ, under the Grant 2016004686 and of the Brazilian National Research Council – CNPq, under the Grants 304801/2015-1 and 421486/2016-3.

‡frag@lncc.br National Laboratory for Scientific Computing – LNCC/MCTIC, CEP 25651-075, Av. Getúlio Vargas 333, Petrópolis, Rio de Janeiro, Brazil.

1 Introduction

Efficient control systems rely heavily on the premise that the model considered was devised in such a way that relevant *uncertainties* were adequately taken into account in the modelling process of the system to be controlled. For instance, accounting for *abrupt change* , such as failure, in the modelling may be crucial in the design of a reliable control system which may guarantee an acceptable behavior and meeting some performance requirements even in the presence of these abrupt events in the system dynamics. This may be crucial, for instance, either due to security reasons or efficiency necessity. With the advent of new sophisticated technologies used in the manufacture of complex systems such as airplane, nuclear power station, mobile network, robotics, among others technological symbols of modern society, this concern is still more important. Therefore, it has been widely recognized that the requirements of specific behaviors and stringent performances in complex dynamics systems call for the inclusion of possible *failure* prevention in a modern control design. In view of this, dynamical systems which are *subject to abrupt changes* have been a theme of increasing investigation in recent years and a variety of different approaches to analyze this class of systems has emerged over the last decades. In this regard, a particularly interesting class of models within this framework is the so-called *Markov jump linear systems* (MJLS), which is the subject matter of this paper. In this case, the *abrupt changes* are modeled by a Markov chain (see, e.g., [2], [3], [6], [18] and references therein).

A cursory examination of the literature reveals that there has been a steadily rising level of activity with systems which are vulnerable to abrupt changes in their structures and ramping up yet more in recent years in the case of MJLS. An initial trickle of papers using MJLS models [20], [21], soon grew to a considerable amount of papers with a sober eye towards applications, as befits a maturing field (see, e.g., [2], [3] and references therein). The development of a solid body of theoretical results on MJLS engendered, in recent years, a startling growth of application of this theory in a great variety of area, since abrupt changes can be due, for instance, to abrupt environmental disturbances, component failures or repairs, changes in subsystems interconnections, abrupt changes in the operation point for a non-linear plant, volatility, etc. This can be found, for instance, in robotic (see, e.g., [19]), communication networks (packet loss, fading channels, [14] and [26]), electromagnetic disturbances in flight systems (see, e.g., [12]), lossy sensor data (see, e.g., [8]), image-enhancement (e.g., [7]), wireless communication (see, e.g., [15]), among others. (see also, [2], [3], and references therein)

To some extent, Markov jump linear systems is by now a fully fledged theory which provides systematic tools to deal with linear systems which are subject to abrupt changes. This is certainly true for the case of *complete observations* of both, the state signal and the mode operation. This can be confirmed by the coherent body of theoretical results available in the specialized literature for this scenario (see, e.g., [2], [3], [6], [18] and references therein). However, when it comes to the scenario of MJLS with *partial observations*, the theory in this setting needs to be perfected. A salient feature here, vis-a-vis the linear case, is that the partial observations scenario in the context of MJLS has three possible settings: (i) the Markov chain is observed, but not the state signal; (ii) the state is observed, but not the chain; and (iii) none of them is observed.

One of the main hindrance in the study of MJLS with *partial observations* comes up in the context of the *stochastic optimal control problem.* This is not a surprise, since it has been really hard to make great strides in the tailoring of a general theory of optimal control with partial observations which guarantee explicit analytical solution. The rare exception in this scenario is the so-called LQG problem (see, for instance, [5]). Roughly, the difficulty lies in the fact that, within the optimal control theory framework, the approach hinges on transforming the partially observable problem in one with complete information (e.g., via filtering theory) and solve the problem via the associated Hamilton-Jacobi-Bellman equation. The main difficulty here lies on the fact that a great deal of non-linearity is introduced in the Bellman equation by the optimal filter which makes the problem intractable. Notice that, in the scenario (ii), the optimal filter for the Markov chain (also known as Wonham's filter) is nonlinear (see, [24]).

In view of the difficulty in solving the problem in the context of the optimal control theory, there has been an alternative attempt , which has come to the fore recently, in the context of the control problem for MJLS with partial observations of the Markov chain and has been dubbed in the specialized literature as *the detector based approach* (see, e.g., [10], [4]). However, in this approach the detector mechanism, $\hat{\theta}(t)$, do not make use of the information coming from the output $y(t)$. The detector do not come up as a result of an optimization problem.

Motivated by the discussions above, and in an attempt to circumvent the problems mentioned above in the context of the control problem for MJLS with partial observations of the Markov parameter, an *optimal linear filter* for the the Markov parameter was recently derived in [22]. The motivation there was at least twofold: (i) by using this linear filter, to revisit in the future the optimal control problem for the scenario of partial observations, in the context of Bellman's equation. The favorable aspect in this case is the fact that the filter is linear; (ii) to make a direct use of the available output signal.

In this paper, as a natural step forward, we derive a *stationary linear filter* associated to the best linear linear filter derived in [22] for the Markov parameter in the scenario (ii) of partial observations described above. In the spirit of what has been done in [11], the approach here amounts in obtaining the convergence of the error covariance matrix of the best linear mean square filter (BLMSF) to a stationary value under the assumption of ergodicity of the associated Markov chain. Besides the interest in its own right, in conjunction with the fact that stationary filter alleviate the computation burden of having to calculate the solution of the error covariance error matrix at each time, it may bear an important role on the analyse of the infinite horizon setting, in the spirit of the circle of ideas discussed above.

A brief outline of the content of this paper is as follows. In Section 2 we fix the notations and recall a few notions. The relevant setting and problem statement are described in section 3. The main result, the stationary linear filter, is given in section 4. Finally, a numerical simulation for the stationary linear filter deduced in this article is presented in Section 5.

2 Notation and Preliminaries

We will denote by $\mathbb{R}^n$ the n-dimensional Euclidean space and by $\mathbb{B}(\mathbb{R}^n, \mathbb{R}^m)$ the normed bounded linear space of all $n \times m$ matrices with $\mathbb{B}(\mathbb{R}^n) : = \mathbb{B}(\mathbb{R}^n, \mathbb{R}^n)$. For $L \in \mathbb{B}(\mathbb{R}^n)$, $L^{'}$ will indicate the transpose of L. As usual, $L \leq 0 \quad (L > 0)$ will mean that the symmetric matrix $L \in \mathbb{B}(\mathbb{R}^n)$ is positive semi-definite (positive definite), respectively. In addition, we set $\mathbb{B}(\mathbb{R}^n)^+ : = \{L \in \mathbb{B}(\mathbb{R}^n); L = L^{'} \geq 0\}$. We use $\mathbb{R}^+$ to denote the interval $[0, \infty)$ and define by $L \otimes K \in \mathbb{B}(\mathbb{R}^{sn}, \mathbb{R}^{rm})$, the Kronecker product for any $L \in \mathbb{B}(\mathbb{R}^s, \mathbb{R}^r)$ and $K \in \mathbb{B}(\mathbb{R}^n, \mathbb{R}^m)$. In that sense $L \otimes z \in \mathbb{B}(\mathbb{R}^s, \mathbb{R}^{rn})$, for any $L \in \mathbb{B}(\mathbb{R}^s, \mathbb{R}^r)$ and $z \in \mathbb{R}^n$. For $D_i \in \mathbb{B}(\mathbb{C}^n), \quad i = 1, \dots, N$, $diag(D_i)$ stands for an $Nn \times Nn$ matrix where the matrices D_i are put together corner-to-corner diagonally, with all other entries being zero. For $V \in \mathbb{C}^n$, $diag(V)$ stands for an $n \times n$ matrix where the components v_i of the vector V are put together corner-to-corner diagonally, with all other entries being zero. Also, V^m stands for an $nm \times n$ matrix where V is concatenated m times, that is $V^m = [V_1 \dots V_m]$.

As usual, we define $\mathcal{H} := \mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$, the Hilbert space of all square integrable random variables (r.v.) in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with the inner product $\langle x, y\rangle = \mathrm{E}\left(x^{'}y\right)$. Convergence here will be in the quadratic mean (q.m.) sense, i.e., a sequence $\{x(n)\}$ converges to x if $\|x(n)-x\|_2 \longrightarrow 0$. We define also $\mathcal{H}_0 = \{x \in \mathcal{H} | \mathrm{E}x = 0\}$, the closed subspace of all centred r.v.'s of $\mathcal{H}$ are said to be orthogonal (from now on $x \perp y$) if $\langle x, y\rangle = 0$. Furthermore, for $y \in \mathcal{H}_0$ we consider the subspace $\mathcal{H}_t^y \subset \mathcal{H}_0$ defined by $\mathcal{H}_t^y = \mathcal{L}\{y(s), 0 \le s \le t\}$ which consists of all linear combinations $\sum_i \alpha_i^{'} y(t_i)$, where $t_i < t$, and q.m. limits of these combinations (a closed subspace) such that $\mathcal{H}_t^y \subset \mathcal{H}_{t'}^y \subset \mathcal{H}^y := \mathcal{H}_\infty^y$ for $t < t^{'}$. We recall that if $\{y(t)\}$ is q.m. continuous then $\mathcal{H}^y$ is a separable Hilbert space and, as a fundamental property of a Hilbert space, any $z \in \mathcal{H}_0$ has a unique decomposition $z = \hat{z} + \tilde{z}$ where $\hat{z} = \mathcal{P}_t^y z \in \mathcal{H}_t^y$ and $\tilde{z} \perp \mathcal{H}_t^y$. Here $\mathcal{P}_t^y$ denotes the projection operator which projects each element of $\mathcal{H}_0$ onto $\mathcal{H}_t^y$. Moreover, we can have the following properties:

- $\|z - \mathcal{P}_t^y z\| = min_{\upsilon \in \mathcal{H}_t^y} \|z - \upsilon\|$ and, therefore, $\hat{z} = \mathcal{P}_t^y z$ is the linear least-square estimator of z given $\mathcal{H}_t^y$, i.e., the best linear estimator is the projection of z onto $\mathcal{H}_t^y$;
- $\tilde{z} = z - \hat{z} \perp \mathcal{P}_t^y$.

A vector process $\left\{x(t) = \left(x^1(t), \ldots, x^n(t)\right); t \in \mathbb{R}^+\right\} \in \mathbb{R}^n$ has orthogonal increments (o.i.) if for each i, j and any non-overlapping intervals (u, r) and (s, t), $\left(x^i(t) - x^i(s)\right) \perp \left(x^i(r) - x^i(u)\right)$, that is $\left(x^i(t) - x^i(s)\right) \perp \mathcal{H}_s^x$. For a second order vector process $\{x(t)\}$, $\mathrm{Cov}(x(t))$ will refer to its covariance function, and $\delta_{\{.\}}$ stands for the Dirac measure.

3 Problem Description and Auxiliary Results

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with its natural filtration $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$, as usual augmented by all null sets in the $\mathbb{P}$-completion of $\mathcal{F}$, carrying the following mutually independent objects:

O.1) An standard Wiener process $W = \{(\omega(t), \mathcal{F}_t), t \in \mathbb{R}^+\}$ in $\mathbb{R}^p$.

O.2) An homogeneous Markov process with right continuous trajectories $\theta = \{(\theta_t, \mathcal{F}_t), t \in \mathbb{R}^+\}$ taking values on the finite set $\mathrm{S} := \{1, 2, \ldots, N\}$. In addition:

$$\mathbb{P}(\theta_t = j | \theta_0 = i) = \begin{cases} \lambda_{ij} t + o(t), i \neq j \\ 1 + \lambda_{ii} t + o(t), i = j \end{cases} \tag{1}$$

where $[(\lambda_{ij})]$ is the stationary $N \times N$ transition rate matrix of θ with $\lambda_{ij} \geq 0, i \neq j$ and $\lambda_i = -\lambda_{ii} = \sum_{j:j\neq i} \lambda_{ij} < \infty$. We define $p_{ij}(t) := \mathbb{P}\left(\theta_{t+s} = j \mid \theta_s = i\right), i, j = 1, \ldots, N$, and denote $p_i(t) := \mathbb{P}(\theta_t = i) > 0$, for any $i \in \mathrm{S}$. Notice that, in this setting, $P(t) := [p_1(t), \ldots, p_N(t)]^{'}$, satisfies the Kolmogorov forward differential equation $dP(t)/dt = \Lambda P(t); P(0) = P_0, t \in \mathbb{R}^+$, where $\Lambda := \left[(\lambda_{ij})^{'}\right]$.

O.3) A random variable y, with $E\left[y^2\right] < \infty$.

Consider now the class of hybrid dynamical systems modeled by the following *Markov jump linear systems*:

$$dy(t) = H_{\theta(t)} y(t)dt + rd\omega(t), \quad y(0) = y. \tag{2}$$

where $y(t) \in \mathbb{R}^m$ denotes the observation vector of the Markov chain $\theta(t)$, $H_{\theta(t)}$ is a random matrix such that, for $\theta(t) = i$ assumes the value $H_i \in, \mathbb{R}^m$ and $r \in \mathbb{B}(\mathbb{R}^p, \mathbb{R}^m)$ is the diffusion coefficient.

A central piece in the development of our approach is the following representation result for $\theta(t)$ (see, e.g. [16]):

$$d\delta_i(t) = \sum_{j=1}^{N} \lambda_{ji}\delta_i(t)dt + dm_i(t), \tag{3}$$

where $\delta_i(t) = \delta_{\{\theta_t = i\}}$ and $m_i(t)$ is a square integrable $\mathcal{F}_t^\theta$-martingale with right-continuous trajectories and bounded variation. Define now $\rho(t) = [\delta_0(t), \ldots, \delta_N(t)]^{'} \in \mathbb{R}^N$ and $\bar{m}(t) = [m_i(t), \ldots, m_N(t)]^{'} \in \mathbb{R}^N$. Then from (3) we have

$$d\rho(t) = \Lambda\rho(t)dt + d\bar{m}(t). \tag{4}$$

Setting now $\bar{H} = [H_1 \ldots H_N] \in \mathbb{B}(\mathbb{R}^N, \mathbb{R}^m)$ and $Y(t) = I_N \otimes y(t) \in \mathbb{B}(\mathbb{R}^{Nm}, \mathbb{R}^m)$, we can rewrite (2) as

$$dy(t) = \bar{H}Y(t)\rho(t)dt + rd\omega(t). \tag{5}$$

Define also $\bar{y}(t) = [\bar{y}_1(t)^{'}, \ldots, \bar{y}_N(t)^{'}]^{'}$ where $\bar{y}_i(t) = \delta_i(t)y(t)$. Notice that $\bar{y}(t) = Y(t)\rho(t)$, and from Proposition 5.3 in [9]

$$d\mathrm{E}[\bar{y}(t)] = F\mathrm{E}[\bar{y}(t)]dt, \qquad \mathrm{E}[\bar{y}(0)] = \mu \tag{6}$$

for $F = diag(H_i) + \Lambda \otimes I_m$. From now on assume that:

A.1) y and $\theta(t)$ (and therefore $\delta_i(t) \quad \forall i \in \mathrm{S}$ and $\rho(t)$) are independent from $\omega(t)$. Consequently we have $\langle \bar{m}, W \rangle \equiv 0$.

A.2) $rr^{'} > 0$.

A.3) $\theta(t)$ is ergodic. Therefore, $P(t)$ converges to a stationary probability vector $\Pi = [\pi_1, \dots, \pi_N]^{'}$, where $\pi_i = \lim_{t\to\infty} p_i(t)$ (see [1] Section 6.2).

A.4) $\mathbb{R}_e\{\lambda(F)\} \leq 0$ with a zero eigenvalue (no imaginary part).

Remark. $\mathbb{R}_e\{\lambda(F)\} \leq 0$ prevents instability for $y(t)$ and the zero eigenvalue avoids $[\bar{y}(t)] \to 0$ as $t \to \infty$. Notice that if $[\bar{y}(t)] \to 0$ the signal to noise ratio (SNR) of $y(t)$ will be too low to "detect" $\theta(t)$ after the early stage.

Before going deeper into the problem formulation we need to introduce some definitions and auxiliary results.

Name now $\hat{\rho}(t)$ as the best linear mean square estimator of $\rho(t)$ given $\mathcal{H}_t^y$, that is $\hat{\rho}(t) = \mathcal{P}_t^y \rho(t)$ which minimizes the square error $\|\tilde{\rho}(t)\|_2^2$, where $\tilde{\rho}(t) = \rho(t) - \hat{\rho}(t)$. Bearing in mind that $\theta(t) = \sum_{i=1}^N i\delta(t)$ and $\rho(t) = [\delta_1(t), \dots, \delta_N(t)]^{'}$, the best linear mean square estimation for $\theta(t)$ is given by $\hat{\theta}(t) = \sum_{i=1}^N i\hat{\delta}(t)$. Define also

$$\begin{cases} \bar{y} = \lim_{t\to\infty} \mathrm{E}[y(t)] \\ D = (rr^{'})^{-1/2} \\ C = -(I_N \otimes \bar{y})^{'} \bar{H}^{'} D^{'} D \bar{H} (I_N \otimes \bar{y}) \end{cases} \tag{7}$$

and

$$\begin{cases} \bar{P}(t) = \mathrm{E}[\rho(t)\rho(t)^{'}] \\ \hat{P}(t) = \mathrm{E}[\hat{\rho}(t)\hat{\rho}(t)^{'}] \\ \tilde{P}(t) = \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}]. \end{cases} \tag{8}$$

Notice that $\tilde{P}(t)$ is the error covariance matrix of the estimator $\hat{\rho}(t)$. Finally define the innovation process $\upsilon(t)$ as

$$d\upsilon(t) = dy(t) - \bar{H}Y(t)\hat{\rho}(t)dt = \bar{H}Y(t)\tilde{\rho}(t)dt + rd\omega(t). \tag{9}$$

Theorem 1. *Consider system* (5) *with assumptions A1) and A2). Then the best linear mean square estimator* $\hat{\rho}(t)$ *is given by the following filter,*

$$d\hat{\rho}(t) = \Lambda\hat{\rho}(t)dt + \mathrm{E}[\rho(t)\tilde{\rho}(t)^{'}Y(t)^{'}]\bar{H}^{'}D^{'}Dd\upsilon(t). \tag{10}$$

Proof. See Theorem 1 in [22]. □

The Problem: The subject matter of this paper is to derive a stationary linear filter for the *best linear mean square filter* (BLMSF) derived in Theorem 1, above. It amounts here to obtain the convergence of the error covariance matrix of the BLMSF to a stationary value under the assumption of ergodicity of the associated Markov chain $\theta(t)$, which is tantamount to prove convergence of $\mathrm{E}[\rho(t)\tilde{\rho}(t)^{'}Y(t)^{'}]$ to a stationary value.

4 Main Result

In this section, we obtain our main result concerning the stationary filter for $\rho(t)$, which consists in a convergence analysis of the behavior of $\mathrm{E}[\rho(t)\tilde{\rho}(t)^{'}Y(t)^{'}]$, as mentioned above. The main result reads as follows.

Theorem 2. *For* (5) *with assumptions A1)-A4) and assuming that* $(\Lambda^{'}, C)$ *is controllable, the stationary linear filter associated to* (10) *is given by:*

$$d\hat{\rho}(t) = \Lambda\hat{\rho}(t)dt + \tilde{P}(I_N \otimes \bar{y})^{'}\bar{H}^{'}D^{'}Dd\upsilon(t) \tag{11}$$

where

$$\tilde{P} = \lim_{t\to\infty} \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}], \tag{12}$$

satisfies the Riccati equation

$$\Lambda\tilde{P} + \tilde{P}\Lambda^{'} - \tilde{P}(I_N \otimes \bar{y})^{'}\bar{H}^{'}D^{'}D\bar{H}(I_N \otimes \bar{y})\tilde{P} + \Psi(\Pi) = 0, \tag{13}$$

with $\Psi(\Pi) = diag(\Lambda\Pi) - \Lambda diag(\Pi) - diag(\Pi)\Lambda^{'}$.

Corollary 2.1. *From* $\hat{\rho}(t)$ *we obtain* $\hat{\theta}(t) = \sum_{i=1}^{N} i\hat{\delta}(t)$ *reminding that* $\hat{\delta}_i(t)$ *are the components of* $\hat{\rho}(t)$*.*

Before going into the proof of our main result, let us tarry for a while and consider results which play a central role in the arguments of the proof. First we need to work out a differential equation for $\mathrm{E}[\rho(t)\tilde{\rho}(t)^{'}Y(t)^{'}]$ concerning (10). Observe that,

$$\begin{aligned}
\mathrm{E}[\rho(t)\tilde{\rho}(t)^{'}Y(t)^{'}] &= \mathrm{E}[\hat{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] + \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] \\
&= \mathrm{E}[\mathrm{E}[\hat{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'} \mid \mathcal{H}_t^y]] + \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] \\
&= \mathrm{E}[\mathrm{E}[\hat{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'} \mid \mathcal{H}_t^y]] + \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] \\
&= \mathrm{E}[\hat{\rho}(t)\mathrm{E}[\tilde{\rho}(t)^{'} \mid \mathcal{H}_t^y]Y(t)^{'}] + \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] \\
&= \mathrm{E}[\hat{\rho}(t)\mathrm{E}[\tilde{\rho}(t)^{'}]Y(t)^{'}] + \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] \\
&= \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}].
\end{aligned} \tag{14}$$

Also notice that,

$$\begin{aligned}
\mathrm{E}[Y(t)\tilde{\rho}(t)] &= \mathrm{E}[Y(t)\rho(t)] - \mathrm{E}[Y(t)\hat{\rho}(t)] \\
&= \mathrm{E}[\mathrm{E}[Y(t)\rho(t) \mid \mathcal{H}_t^y]] - \mathrm{E}[Y(t)\hat{\rho}(t)] \\
&= \mathrm{E}[Y(t)\mathrm{E}[\rho(t) \mid \mathcal{H}_t^y]] - \mathrm{E}[Y(t)\hat{\rho}(t)] \\
&= \mathrm{E}[Y(t)\hat{\rho}(t)] - \mathrm{E}[Y(t)\hat{\rho}(t)] = 0,
\end{aligned} \tag{15}$$

therefore $Y(t)$ and $\tilde{\rho}(t)$ are orthonormals. Furthermore, as $\mathrm{E}[\tilde{\rho}(t)] = 0$

$$\mathrm{Cov}(Y(t), \tilde{\rho}(t)) = \mathrm{E}[Y(t)\tilde{\rho}(t)] - \mathrm{E}[Y(t)]\mathrm{E}[\tilde{\rho}(t)] = 0, \tag{16}$$

and they are also uncorrelated. Hence, using L'Hopital

$$\mathrm{E}[Y(t) \mid \tilde{\rho}(t)] = \frac{\mathrm{E}[Y(t)\tilde{\rho}(t)]}{\mathrm{E}[\tilde{\rho}(t)]} = \mathrm{E}[Y(t)], \tag{17}$$

and

$$\begin{aligned}
\mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'}] &= \mathrm{E}[\mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}Y(t)^{'} \mid \tilde{\rho}(t)]] = \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}\mathrm{E}[Y(t)^{'} \mid \tilde{\rho}(t)]] \\
&= \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}\mathrm{E}[Y(t)^{'}]] = \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}]\mathrm{E}[Y(t)^{'}] \\
&= \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}]\left(I_N \otimes \mathrm{E}[y(t)]\right)^{'}.
\end{aligned} \tag{18}$$

Lemma 1. $\lim_{t\to\infty} \mathrm{E}[y(t)] = \bar{y} = I_m^N \dfrac{u^{'}\mu}{u^{'}v} v$ *where* u *and* v *are he left and right eigenvectors corresponding to the zero eigenvalue of* F*, respectively.*

Proof. First notice that $\mathrm{E}[y(t)] = \sum_{j=1}^{N} \mathrm{E}[\bar{y}_j(t)] = I_m^N \mathrm{E}[\bar{y}(t)]$. From (6), and assumption A.4), $\mathrm{E}[\bar{y}(t)]$ converges to a constant value and therefore

$\dot{\mathrm{E}}[\bar{y}(t)] = F\mathrm{E}[\bar{y}(t)]$ converges to zero. Let v be the normalized right eigenvector corresponding to the zero eigenvalue of F, then for any $\bar{v} \in \mathcal{V} = \{\beta v : \beta \in \mathbb{R}\}$, $F\bar{v} = 0$. We need now to find the specific $\bar{v} \in \mathcal{V}$ to which $\mathrm{E}[\bar{y}(t)]$ converges. For that, set now u as the normalized left eigenvector corresponding to the zero eigenvalue of F. Observe that $d(u^{'}\mathrm{E}[\bar{y}(t)])/dt = u^{'}d\mathrm{E}[\bar{y}(t)]/dt = u^{'}F\mathrm{E}[\bar{y}(t)] = 0$, consequently $u^{'}\mathrm{E}[\bar{y}(t)]$ has a constant value $\forall t \in \mathbb{R}^{+}$. Let $w = \alpha v \in \mathcal{V}$, $\alpha \in \mathbb{R}$ be the specific point to where $\mathrm{E}[\bar{y}(t)]$ converges. Then, it follows that $u^{'}\mu = u^{'}w = u^{'}\alpha v = \alpha u^{'}v$ and therefore we have that $\alpha = \frac{u^{'}\mu}{u^{'}v}$ and finally

$$w = \frac{u^{'}\mu}{u^{'}v}v. \tag{19}$$

Considering that $\mathrm{E}[\bar{y}(t)]$ converges to $\frac{u^{'}\mu}{u^{'}v}v$ and as $\mathrm{E}[y(t)] = I_m^N\mathrm{E}[\bar{y}(t)]$,

$$\lim_{t\to\infty} \mathrm{E}[y(t)] = I_m^N\frac{u^{'}\mu}{u^{'}v}v, \tag{20}$$

and this completes the proof.

□

We also need to find a differential equation for $\tilde{P}(t) = \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}]$. Observe that

$$\begin{aligned}
\tilde{P}(t) &= \mathrm{E}\left[(\rho(t)-\hat{\rho}(t))\,(\rho(t)-\hat{\rho}(t))^{'}\right] = \mathrm{E}\left[\rho(t)\rho(t)^{'}\right] + \mathrm{E}\left[\rho(t)\hat{\rho}(t)^{'}\right] \\
&+ \mathrm{E}\left[\hat{\rho}(t)\rho(t)^{'}\right] + \mathrm{E}\left[\hat{\rho}(t)\hat{\rho}(t)^{'}\right] = \mathrm{E}\left[\rho(t)\rho(t)^{'}\right] - \mathrm{E}\left[\mathrm{E}\left[\rho(t)\hat{\rho}(t)^{'} \mid \mathcal{H}_t^y\right]\right] \\
&- \mathrm{E}\left[\mathrm{E}\left[\hat{\rho}(t)\rho(t)^{'} \mid \mathcal{H}_t^y\right]\right] + \mathrm{E}\left[\hat{\rho}(t)\hat{\rho}(t)^{'}\right] = \mathrm{E}\left[\rho(t)\rho(t)^{'}\right] \\
&- \mathrm{E}\left[\mathrm{E}\left[\rho(t) \mid \mathcal{H}_t^y\right]\hat{\rho}(t)^{'}\right] - \mathrm{E}\left[\hat{\rho}(t)\mathrm{E}\left[\rho(t)^{'} \mid \mathcal{H}_t^y\right]\right] \\
&+ \mathrm{E}\left[\hat{\rho}(t)\hat{\rho}(t)^{'}\right] = \mathrm{E}\left[\rho(t)\rho(t)^{'}\right] - \mathrm{E}\left[\hat{\rho}(t)\hat{\rho}(t)^{'}\right] = \bar{P}(t) - \hat{P}(t).
\end{aligned} \tag{21}$$

Now, bearing in mind that $\delta_i(t)$,

$$\delta_i(t)\delta_j(t) = \begin{cases} \delta_i(t), & i = j \\ 0, & i \neq j, \end{cases}$$

we obtain that $\bar{P}(t) = \mathrm{E}\left[\rho(t)\rho(t)^{'}\right] = \mathrm{E}\left[diag(\rho(t))\right]$. But $\mathrm{E}[\delta_i(t)] = p_i(t)$, so $\bar{P}(t) = diag(P(t))$. Consequently,

$$d\bar{P}(t) = diag(\Lambda P(t))dt. \tag{22}$$

In addition, from (18), Theorem 1 and bearing in mind that $v(t)$ is an o.i. process, we have that

$$\begin{aligned} d\hat{P}(t) &= d\mathrm{E}\left[\hat{\rho}(t)\hat{\rho}(t)^{'}\right] = \mathrm{E}\left[d\hat{\rho}(t)\hat{\rho}(t)^{'}\right] + \mathrm{E}\left[\hat{\rho}(t)d\hat{\rho}(t)^{'}\right] + \mathrm{E}\left[d\hat{\rho}(t)d\hat{\rho}(t)^{'}\right] \\ &= \Lambda\hat{P}(t)dt + \hat{P}(t)\Lambda^{'}dt + \tilde{P}(t)\mathrm{E}[Y(t)^{'}]\bar{H}^{'}D^{'}D\bar{H}\mathrm{E}[Y(t)]\tilde{P}(t)dt. \end{aligned} \tag{23}$$

So, from (21), (22) and (23), it follows that

$$\begin{aligned} d\tilde{P}(t) &= diag(\Lambda P(t))dt - \Lambda\hat{P}(t)dt - \hat{P}(t)\Lambda^{'}dt \\ &\quad - \tilde{P}(t)\mathrm{E}[Y(t)^{'}]\bar{H}^{'}D^{'}D\bar{H}\mathrm{E}[Y(t)]\tilde{P}(t)dt = \Lambda\tilde{P}(t)dt + \tilde{P}(t)\Lambda^{'}dt \\ &\quad - \tilde{P}(t)\mathrm{E}[Y(t)^{'}]\bar{H}^{'}D^{'}D\bar{H}\mathrm{E}[Y(t)]\tilde{P}(t)dt + \Psi(P(t))dt, \end{aligned} \tag{24}$$

where $\Psi(P(t)) = diag(\Lambda P(t)) - \Lambda diag(P(t)) - diag(P(t))\Lambda^{'}$.

Lemma 2. $\Psi(P(t))$ *is a positive semi-definite matrix.*

Proof. Lets develop $\Psi(P(t))$ term by term,

$$\left\{ \begin{array}{ll} diag(\Lambda P(t)) = [(\alpha_{ij})] \quad \text{for} \quad \alpha_{ij} = \left\{ \begin{array}{ll} \sum_{k=1}^{N} \lambda_{ki}p_k(t) & i = j \\ 0 \quad i \neq j \end{array} \right. \\ \\ \Lambda diag(P(t)) = [(\beta_{ij})] \quad \text{for} \quad \beta_{ij} = \lambda_{ji}p_j(t) \\ \\ diag(P(t))\Lambda^{'} = [(\gamma_{ij})] \quad \text{for} \quad \gamma_{ij} = \lambda_{ij}p_i(t). \end{array} \right. \tag{25}$$

Now,

$$\begin{aligned} \Psi(P(t)) &= diag(\Lambda P(t)) - \Lambda diag(P(t)) - diag(P(t))\Lambda^{'} = [(a_{ij})] \\ &\text{for} \quad a_{ij} = \left\{ \begin{array}{ll} \sum_{k=1}^{N} \lambda_{ki}p_k(t) - 2\lambda_{ij}p_j(t) & i = j \\ -\lambda_{ij}p_i(t) - \lambda_{ji}p_j(t) & i \neq j. \end{array} \right. \end{aligned} \tag{26}$$

Because λ_{ij} are the components of a Markov transition rate matrix and $p_i(t) > 0 \forall i \in N, \quad \forall t \in \mathbb{R}^+$, $\Psi(P(t))$ has the following properties, P1) $a_{ij} > 0$ for $i = j$ and $a_{ij} < 0$ for $i \neq j$

P2) $\sum_{j=1}^{N} a_{ij} = 0 \quad \forall i \in N$ and $\sum_{i=1}^{N} a_{ij} = 0 \quad \forall j \in N$.

Notice that $[-\Psi(P(t))]$ has exactly the same properties of a symmetric Markov transition rate matrix, and therefore its eigenvalues are non-positive. Immediately then, $\Psi(P(t))$ is symmetric and has its eigenvalues non-negative and therefore it is positive semi-definite.

$\square$

In the sequel we do need the following definition:

$$M = \begin{bmatrix} \Lambda^{'} & C \\ -\Psi(\Pi) & -\Lambda \end{bmatrix}. \tag{27}$$

Lemma 3. *Assume that* $(\Lambda^{'}, C)$ *is controllable. For* (24) *with assumptions A3)-A4) and* $\tilde{P}(0) \geq 0$ *we have that* $\tilde{P}(t) \rightarrow \tilde{P}$ *with* $\tilde{P}$ *the unique positive semidefinite solution of the algebraic Riccati equation (ARE):*

$$\Lambda\tilde{P} + \tilde{P}\Lambda^{'} - \tilde{P}(I_N \otimes \bar{y})^{'}\bar{H}^{'}D^{'}D\bar{H}(I_N \otimes \bar{y})\tilde{P} + \Psi(\Pi) = 0. \tag{28}$$

Proof. The idea of the proof runs as follows. First we have to prove the existence and uniqueness of a positive semidefinite solution, $\tilde{P}$, for the algebraic Riccati equation (28). Now, proving that $\tilde{P}(t) \rightarrow \tilde{P}$ is tantamount to proving that there exist lower and upper bound functions $\tilde{P}_l(t)$ and $\tilde{P}_u(t)$ for $\tilde{P}(t)$, i.e., $\tilde{P}_l(t) \leq \tilde{P}(t) \leq \tilde{P}_u(t)$, and these functions squeeze $\tilde{P}(t)$ to $\tilde{P}$. According to [23], to prove existence and uniqueness of (28) it suffices that the following conditions are satisfied:

R1) The matrices $\tilde{P}$, Λ, C and $\Psi(\Pi)$ must be $N \times N$ and real.

R2) $\Psi(\Pi)$ and $-C$ must be symmetric positive semidefinite.

R3) M can not have eigenvalues on the imaginary axis.

The condition R1) is obviously satisfied. By definition of C and from Lemma 2, R2) is also satisfied. Finally, we need to proof that M has no eigenvalues on the imaginary axis. We will proof that by contradiction, as follows:

Suppose $i\alpha$, α real, is an eigenvalue of M with eigenvecor $\begin{bmatrix} v \\ u \end{bmatrix} \neq \mathbf{0}$, that is

$$M \begin{bmatrix} v \\ u \end{bmatrix} = i\alpha \begin{bmatrix} v \\ u \end{bmatrix}. \tag{29}$$

Multiplying now the left hand side of (29) by $[u^* v^*]$, we obtain

$$u^* \Lambda^{'} v - v^* \Lambda u + u^* C u - v^* \Psi(\Pi) v = i\alpha(u^* v + v^* u) \tag{30}$$

Separating now real and imaginary part we get that $u^*Cu - v^*\Psi(\Pi)v = 0$ and as C and $-\Psi(\Pi)$ are symmetric negative semidefinite matrix $Cu = 0$ and $\Psi(\Pi)v = 0$. Consequently from (29) yields $\Lambda^{'} v = i\alpha v$ and $\Lambda u = -i\alpha u$. But this implies that $i\alpha$ is an eigenvalue of $\Lambda^{'}$ which is a contradiction because $\Lambda^{'}$ is the transition rate matrix of an ergodic Markov chain. Therefore M has no eigenvalues on the imaginary axis and R3) is also satisfied. As R1)-R3) are all satisfied, (28) has a unique solution $\tilde{P}$.

Because $(\Lambda^{'}, C)$ is controllable and $\tilde{P}$ is the unique solution of (28), then by ([25] Theorem 4.1) $\Lambda + \tilde{P}C$ is stable. Finally, the proof of

$$\tilde{P}_l(t) \le \tilde{P}(t) \le \tilde{P}_u(t) \tag{31}$$

and

$$\lim_{t\to\infty} \tilde{P}_l(t) = \lim_{t\to\infty} \tilde{P}_u(t) = \tilde{P} \tag{32}$$

follows, *mutatis mutandis*, from ([11] Lemmas 6.1 and 6.4 in the appendix). And consequently $\lim_{t\to\infty} \tilde{P}(t) = \tilde{P}$ which completes the proof.

□

We are now ready to prove the main result, as a straightforward consequence of the previous auxiliary results.

Proof of Theorem 2. From (18), in conjunction with Lemma 1 and Lemma 3 , we have that $\mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'} Y(t)^{'}]$ converge to a stationary value, as follows:

$$\lim_{t\to\infty} \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'} Y(t)^{'}] = \lim_{t\to\infty} \mathrm{E}[\tilde{\rho}(t)\tilde{\rho}(t)^{'}] \left(I_N \otimes \mathrm{E}[y(t)]\right)^{'} = \tilde{P}(I_N \otimes \bar{y})^{'}, \tag{33}$$

and this completes the proof.

□

5 Simulation

To illustrate the effectiveness of the result obtained in Theorem 2 exhaustive simulations has been carried out. For this purpose, we have used the algorithm developed in [17] that generates random Markov chain path's from its discretized transition rate matrix. The numerical method chosen for the simulation of the stochastic equations is the Euler-Murayama (see, [13]), mainly because of Yuan and Mao in [17] prove its convergence for the MJLS setting. The programming language used for algorithm development is the latest version of Python 3. The simulation examples presented in this section will have a simulation time of $t = 50s$ and a time step of $\Delta t = 0.001s$ unless a different value is specified.

Consider as an example of system (5), $\theta = \{1, 2, 3, 4\}$ with Markov transition rate matrix,

$$\Lambda^{'} = \begin{bmatrix} -1.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & -1.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & -1.5 & 0.5 \\ 0.2 & 0.2 & 0.2 & -0.6 \end{bmatrix}, \tag{34}$$

and initial probabilities

$$p(0) = [0.4 \quad 0.3 \quad 0.2 \quad 0.1]^{'}. \tag{35}$$

Figure 1 shows a Markov chain example generated from (34) and (35) where the y axis represents the Markov chain states.

The $H_{\theta(t)}$ of the MJLS (2) are given by

$$H_1 = \begin{bmatrix} -4.5 & 0.5 & 0.5 & 0.5 \\ 1.5 & -1.5 & 0.5 & 0.5 \\ 1.5 & 0.5 & -1.5 & 0.5 \\ 1.5 & 0.5 & 0.5 & -1.5 \end{bmatrix} \quad H_2 = \begin{bmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$
$$H_3 = \begin{bmatrix} -3 & 1 & 5 & 1 \\ 1 & -3 & 5 & 1 \\ 1 & 1 & -15 & 1 \\ 1 & 1 & 5 & -3 \end{bmatrix} \quad H_4 = \begin{bmatrix} -3 & 1 & 2.5 & 2 \\ 1 & -3 & 2.5 & 2 \\ 1 & 1 & -7.5 & 2 \\ 1 & 1 & 2.5 & -6 \end{bmatrix}, \tag{36}$$

with initial condition

$$y(0) = [3 \quad 4 \quad 3 \quad 5]^{'}, \tag{37}$$

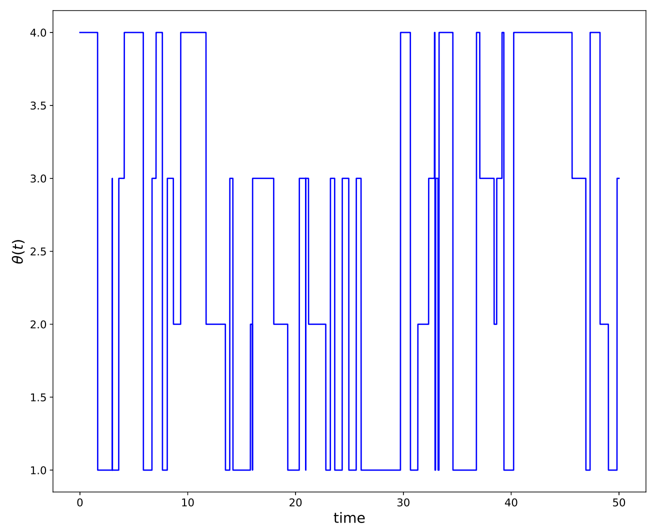

Figure 1: Markov chain from the transition rate matrix in (34).

and diffusion coefficient $r = gI_4$ for $g = 0.1$. Figure 2 shows the results corresponding to the stationary linear filter $\hat{\rho}(t)$ deduced in Theorem 2, where each subplot represents a component of $\hat{\rho}(t)$, that is $\hat{\delta}_i(t)$. Two important characteristics are also noticible in Figure 2, the curl on the stabilized states, and the state transition track. The first indicates the amplification of the noise $\omega(t)$ to the filter and the second the probability of missing some state jumps. Both depend in opposite ways on the gain of the innovation process and the diffusion coefficient, so, the noisier the signal system $y(t)$ is, the slower the filter will track the state jumps.

Figure 3 compares the operation mode $\theta(t)$ with its corresponding stationary linear filter $\hat{\theta}(t)$ from Corollary 2.1. The top subplot shows $\hat{\theta}(t)$ signal and the mid subplot $\theta(t)$. The bottom subplot depicts a superposition of both signals offering a clearer impression. The tracking is outstandingly good and, despite the noise curl being noticeable, it does not disturb or mistake any of the states.

As expected, the stationary linear filter looses precision in the early stage in comparison with the BLMSF. Figure 4 compares the stationary linear filter in red, to the BLMSF in doted blue, for the first two seconds. There exist some difference only in the first jump but is not significant and do not affect the tracking.

We have also carried out some numerical experiments in order to verify

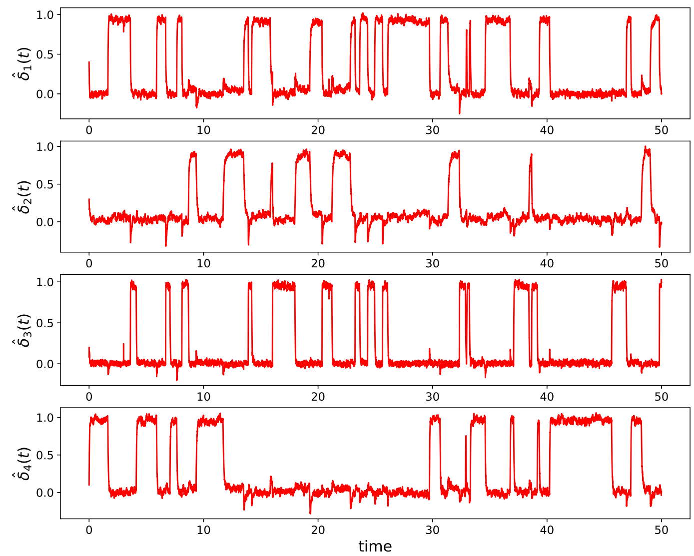

Figure 2: Stationary linear filter $\hat{\rho}(t)$.

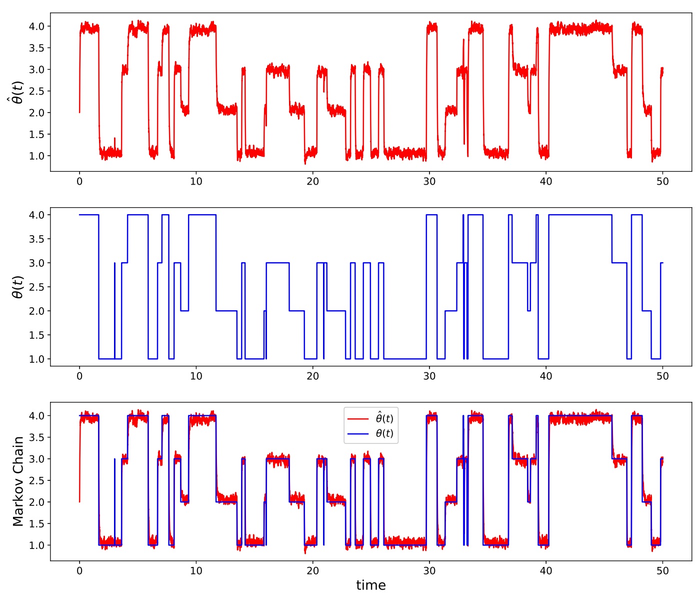

Figure 3: Operation mode $\theta(t)$ and stationary linear filter $\hat{\theta}(t)$ comparison.

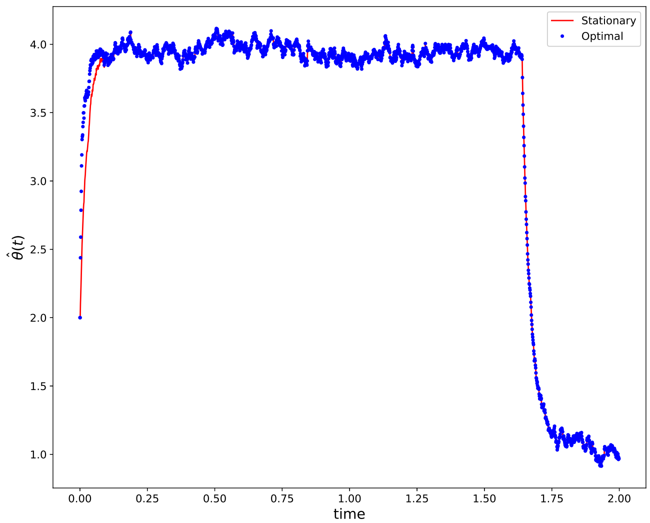

Figure 4: Stationary linear filter and BLMSF comparison on the early stage.

the stationary filter sensitivity regarding the noise intensity. The same system (36) affected by the Markov chain of Figure 1 has been simulated for a noisy scenario and a small noise scenario. The idea is to have a glimpse on how the signal to noise ratio (SNR) is important on the filter performance. Thus, it is expected that the variation of the noise factor g affects directly on the jump misses and the curl amplitude. Let $g = 0.5$, that is, an SNR five times smaller. Now the stationary linear filter for $\rho(t)$ has a bigger curl on the stabilized states and a slower tracking in the jumps. That implies that the stationary linear filter for $\theta(t)$ has more jump misses and a more disturbed stabilized states that can cause some state misinterpretation (observe Figure 5 in the time interval $30 < t < 40$).

On the other hand, for the small noise scenario, that is, for instance, $g = 0.01$, the innovation process gain can be bigger helping the jump tracks, and because the noise amplification to the filter remains low, avoid perceptible curl in the stabilized states. Figure 6 shows that the filter's signal for the small noise scenario is clean and smooth with almost zero miss tracks and an inconspicuous curl in the stabilized states.

Finally, although we have not included here, several other numerical simulations were carried out in different setting for the parameters and the results were comparable to the ones depicted in this section.

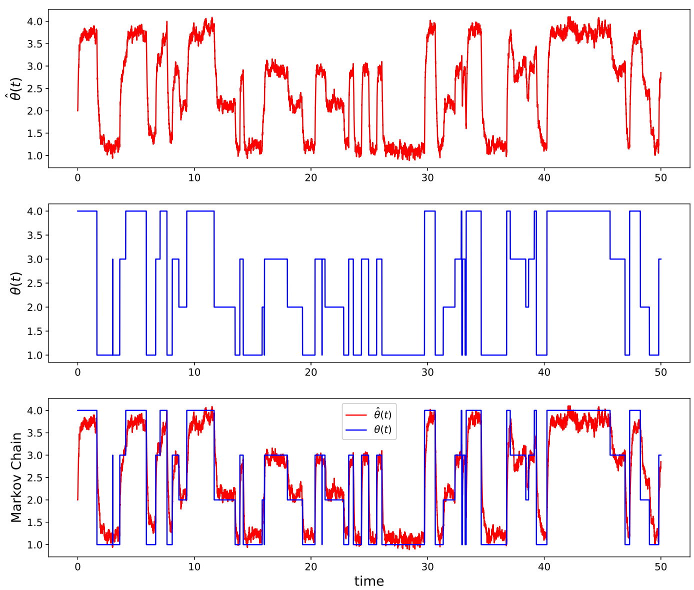

Figure 5: Stationary linear filter $\hat{\theta}(t)$ for a noisy signal system $y(t)$.

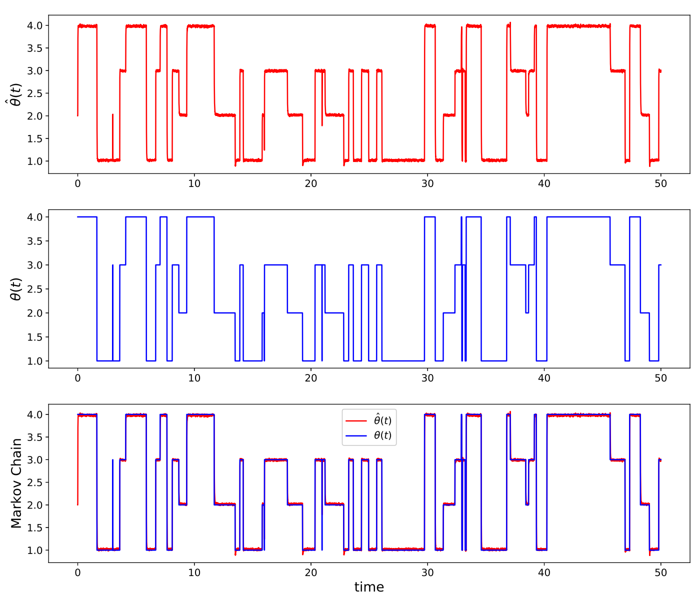

Figure 6: Stationary linear filter $\hat{\theta}(t)$ for the small noise scenario.

6 Conclusions

In this work we have dealt with the problem of derive a stationary linear filter for MJLS in the scenario of partial observations of the jump parameter $\theta(t)$. It amounted here to derive a stationary filter associated to the best linear mean square filter (BLMSF) derived in [22] via a convergence analysis of the error covariance matrix. A major hindrance in solving the problem comes up from the fact that the error covariance matrix is not standard, in the context of the optimal filtering theory. Besides the interest in its own right, the stationary filter has the advantage to alleviate the computational burden vis-a-vis the BLMSF. The simulation results have shown that the stationary linear filter performance is outstandingly good. We have also shown that the difference between the BLMSF and the stationary linear filter appears only in the early stages, but has not any significant consequences in the long run. We have also carried out a noise analysis. Despite the evident influence of the SNR, the stationary linear filter does not lose its efficiency for different diffusion coefficient values. As far as the authors are aware this is the first stationary linear filter for the operation mode $\theta(t)$ of a MJLS in the scenario of partial observations of this parameter.

Future directions of work is certainly to use the results derive here (and the one in [22]) to analysis the control problem of MJLS with partial observations of the Markov parameter.

References

[1] W. J. Anderson. *Continuous-Time Markov Chains.* Springer-Verlag, New-York, 1991.

[2] O. L. V. Costa, M. D. Fragoso, R. P. Marques. *Discrete-Time Markov Jump Linear Systems.* Springer-Verlag, London, 2004.

[3] O. L. V. Costa, M. D. Fragoso, M. G. Todorov. *Continuous-Time Markov Jump Linear Systems.* Springer-Verlag, Berlin Heidelberg, 2013.

[4] O. L. V. Costa, M. D. Fragoso, M. G. Todorov. A Detector-Based Approach for the H_2 Control of Markov Jump Linear Systems With Partial Information. *Journal IEEE Transactions on Automatic Control.* 60: 1219-1234, 2015.

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

[6] V. Dragan, T. Morozon, A. M. Stoica. *Mathematical Methods in Robust Control of Linear Stochastic Systems.* Springer-Verlag, New-York, 2010.

[7] J. S. Evans, R. J. Evans. Image-enhanced multiple model tracking. *Automatica.* 35: 1769-1786, 1999.

[8] A. K. Fletcher, S. Rangan, V. K. Goyal. Estimation from Lossy Sensor Data: Jump Linear Modeling and Kalman Filtering. *Proc. Third Int. Symp. on Information Processing in Sensor Networks (IPSN).* 251-258, 2000.

[9] M. D. Fragoso, O. L. V. Costa. A unified approach for stochastic and mean square stability of continuous-time linear systems with Markovian jump parameters and additive disturbances. *SIAM Journal of Control Optimization.* 44(4):1165–1191, 2005.

[10] M. D. Fragoso, O.L.V. Costa. Mean square stabilizability of continuous-time linear systems with partial information on the Markovian jumping parameters. *Stochastic Analysis and Applications.* 22: 99-111, 2004.

[11] M. D. Fragoso, C. Rocha. Stationary filter for continuous-time Markovian jump linear systems. *SIAM Journal of Control Optimization.* 44(3):801–815, 2005.

[12] W. S. Gray, O. R. González, M. Doğan. Stability Analysis of Digital Linear Flight Controllers Subject to Electromagnetic Disturbances. *IEEE Transactions on Aerospace and Electronic Systems.* 36: 1204-1218, 2000.

[13] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *Society for Industrial and Applied Mathematics.* 43(3):525–546, 2001.

[14] S. Hu, W.-Y. Yan. Stability robustness of networked control systems with respect to packet loss. *Automatica.* 43: 1243-1248, 2007.

[15] P. A. Kawka, A. G. Alleyne. Robust wireless servo control using a discrete-time uncertain Markovian jump linear model. *IEEE Transactions on Control Systems Technology.* 17: 733-742, 2009.

[16] R. S. Lipster, A. N. Shiryayev. *Statistics of Random Processes, volume I-II.* Springer-Verlag, Berlin Heidelberg, 1977.

[17] X. Mao, C. Yuan. Convergence of the Euler-Murayama method for stochastic differential equations with Markovian switching. *Mathematics and Computers in Simulation.* 64:223-235, 2003.

[18] M. Mariton. *Jump Linear Systems in Automatic Control.* Marcel Dekker, New York and Basel, 1990.

[19] A. A. G. Siqueira, M. H. Terra. A fault-tolerant manipulator robot based on $\mathcal{H}_2$, $\mathcal{H}_\infty$, and mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Markovian controls. *IEEE-ASME Transactions on Mechatronics.* 14: 257-263, 2009.

[20] D. D. Sworder. On the stochastic maximum principle. *Journal of Mathematics Analysis and Applications.* 24:627–640, 1968.

[21] D. D. Sworder. Feedback control of a class of linear systems with jump parameters. *IEEE Transactions on Automatic Control.* 14:9–14, 1969.

[22] F. V. Vergés, M. D. Fragoso. Optimal linear mean square filter for the operation mode of continuous-time Markovian jump linear systems. *IET Control Theory and Applications.* 13:1309–1319, 2019.

[23] H. K. Wimmer. On the algebraic Riccati equation. *Bulletin of the Australian Mathematical Society.* 14:457-461, 1976.

[24] W. M. Wonham. Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM Journal of Control Optimization.* 2(3):347–369, 1965.

[25] W. M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM Journal of Control Optimization.*6(4):681-697,1968.

[26] Q. Zhang, S. A. Kassam. Finite-state Markov model for Rayleigh fading channels. *IEEE Transactions on Communications.* 47: 1688-1692, 1999.