Memorias del XVI Congreso Latinoamericano de Control Automático, CLCA 2014
Octubre 14-17, 2014. Cancún, Quintana Roo, México

A comparative case study between the Kalman-Bucy filter and a linear Markovian filter for continuous-time Markov jump linear systems ⋆

Alfredo R. R. Narváez *, **Eduardo F. Costa** *

* *Depto. de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computacão, Universidade de São Paulo, C.P. 668, 13560-970, São Carlos, SP, Brazil (e-mail: aroanar@icmc.usp.br, efcosta@icmc.usp.br).*

Abstract: In this paper we consider two state estimators available in the literature of continuous-time Markov Jump Linear Systems with observation of the jump variable, which provide optimal estimates in different senses. We present a comparative case study of the performance of these filters both in terms of the mean square estimation error and in terms of the cost incurred by combining the estimates with the standard jump linear quadratic control. The study clarifies that the Kalman-Bucy filter attains the best performance among the considered estimators, however it is quite difficult to implement for the considered class of systems. The examples and discussions around the topic provide some insights and guidelines for selecting the most convenient filter.

Keywords: Markov Jump Linear Systems, jump linear quadratic control, linear filtering.

1. INTRODUCTION

The Kalman-Bucy filter (KBF) is a well known tool for estimation from noisy measurements, which is used in a wide range of applications due to some interesting features like optimality and stability (both under some conditions), and also to the fact that it is frequently reported as having good performance in real applications that do not correspond to the ideal setup of the KBF (e.g. in situations violating the optimality conditions, as in presence of model uncertainties), which helped to build a "reputation" of robust estimator. There is an extensive literature dealing with Kalman Filter of which include Jazwinski (1970); Anderson and Moore (1979); Kumar and Varaiya (1986); Costa and Astolfi (2009a) and in the MJLS scenario can be found the papers Miller and Runggaldier (1997); Costa and Astolfi (2009b); Gomes and Costa (2010) among others.

In spite of its simplicity, the KBF may be difficult to implement in some situations. For example, in aplications comprising continuous-time Markov jump linear systems with observation of the jump variable θ, implementing the KBF involves solving a matrix differential equation (see (9) in the variable $P(t)$ that represents the error covariance) that is parameterized by θ, making much more complex to pre-compute its solution and the Kalman gain. In fact, in such applications, at time intant t and given the current $\theta(t)$ and the last jump time instant t_{inf}, a common procedure is to pre-compute $P(t)$ and the associated KBF gain in the time interval $[t_{inf}, t^- + T]$ where T is fixed, assuming $\theta(t)$ remains fixed; however, when one observes a jump in this time interval, $\theta(t^-) \neq \theta(t)$, the procedure has to be reinitialized with a new interval $[t, t + T]$, often incurring in some time delay. This type of procedure may corrupt the performance of the filter, particularly in application with fast dynamics.

Motivated by the above limitation of the KBF, there have been recently developed some alternative filters whose gains and error covariance matrices can be pre-computed offline, making the estimate much easier to calculate in real time applications. For the class of systems considered in this paper, it has been devised in Fragoso and Costa (2005) a linear Markovian filter with some interesting features: it is optimal in a certain sense and it attains a separation principle when used in optimal control problems. We shall refer to this filter as Fragoso-Costa filter (FCF). It is worth to mention that continuous-time Markov jump systems are relatively simple (linear, Markovian, feature similarities with deterministic linear systems - Riccati and Lyapunov-like equations, etc) and yet provide good models for many applications (Athans et al. (1977); Sworder and Rogers (1983); Boukas et al. (1999); do Val and Basar (1999)), in such a manner that it is important to study and understand the main characteristics of control and filtering for this class of systems.

In this paper, the main objective is to put the KBF and FCF in perspective, making clear the differences in their conceptions and performances, and also explaining the separation principle for both filters. After a concise formulation of the control problem we present the available optimal control solution and the filters. Then, in Section 4 we develop some key case studies illustrating that the error covariances and costs incurred by the KBF are

⋆ This work was supported in part by Fapesp, by CNPq, and by CAPES

smaller than the ones by the FCF. We also clarify that the difference in performance is less marked when the distribution of the Markov chain $P(\theta(t) = i)$ remains close to one for long time intervals for some Markov state i, see Example 2 for a more concrete explanation. Indeed, the filters' performances are similar when the distribution is "concentrated" in some state.

The paper is organized as follows. In Section 2 we find some notations, definitions and a description of the problem. Section 3 presents the optimal gain used to design the control law that solves the finite horizon quadratic control problem with incomplete observation based on the two filters. Numerical examples are provided in Section 4 illustrating a comparison of performance results in a table of values. Finally, Section 5 gives some conclusions.

2. PRELIMINARIES

2.1 Notations and Definitions

Let $\mathbb{R}^n$ the linear space of all n-dimensional vectors on the real field. Consider $\mathcal{R}^{n,q}$ (respectively $\mathcal{R}^n$) as the linear space formed by all matrices of size $n \times q$ (respectively $n \times n$) and $\mathcal{R}^{r0}$ ($\mathcal{R}^{r+}$) the closed convex cone of symmetric semi-definite positive matrices $\{U \in \mathcal{R}^r : U = U' \geq 0\}$, (the open cone of symmetric definite positive matrices $\{U \in \mathcal{R}^r : U = U' > 0\}$), where U' denotes the transpose of U; $U \geq V (U > V)$ means that $U - V \in \mathcal{R}^{r0}$ ($U - V \in \mathcal{R}^{r+}$). The operator $1_{\{.\}}$ is the indicator function (or characteristic function) and $\mathrm{tr}\{.\}$ denotes the trace. Let $\mathcal{M}^{r,n}$ be the linear space formed by a number N of matrices such that $\mathcal{M}^{r,n} = \{U = (U_1, \ldots, U_N) : U_i \in \mathcal{R}^{r,n}, i = 1, \ldots, N\}$; 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 $U_i \in \mathcal{R}^{r0}$ ($U_i \in \mathcal{R}^{r+}$) for all $i = 1, \ldots, N$. On a probabilistic space $(\Omega, \mathcal{F}, \mathcal{P})$, let $\{\theta(t)\}$, with $t \in \mathbb{R}$ non-negative ($t \geq 0$), be a Markov chain taking values in the finite set $\mathfrak{S} = \{1, \ldots, N\}$. The Markov chain has a transition rate matrix $\Lambda \in \mathcal{R}^N$ with elements λ_{ij}, where $\lambda_{ij} \geq 0$, $i \neq j$, and $0 \leq -\lambda_{ii} = \sum_{j \neq i} \lambda_{ij} < \infty$, $i, j \in \mathfrak{S}$, and, the probability distribution of θ at the time instant $t \geq 0$ is denoted by $\pi(t) = [\pi_1(t), \ldots, \pi_N(t)]$, with $\pi_i(t) = \mathcal{P}(\theta(t) = i)$, for each $i \in \mathfrak{S}$.

2.2 Problem Statement

Let us consider the following MJLS defined in the fundamental probabilistic space $(\Omega, \mathcal{F}, \mathcal{P})$

$$\begin{aligned} dx(t) &= (A(t,\theta(t))x(t) + B(t,\theta(t))u(t))dt + E(t,\theta(t))dw(t) \\ dy(t) &= C(t,\theta(t))x(t)dt + D(t,\theta(t))dv(t), \end{aligned} \tag{1}$$

where $t \geq 0$, $(x(t), \theta(t))$ is the state process with $x(t) \in \mathbb{R}^n$; $u(t) \in \mathbb{R}^m$ is the input process, $y(t) \in \mathbb{R}^q$ the measurement process and, $\{w(t)\}$ and $\{v(t)\}$ are independent, standard Wiener processes with incremental covariance matrix Idt, where I is the identity matrix of appropriated dimension (more generally we could assume here that $\{w(t)\}$ and $\{v(t)\}$ are zero mean processes with uncorrelated increments). Let us assume $A(t, \theta(t)) := A_{\theta(t)}(t) = A_i(t)$, whenever $\theta(t) = i$, where $A_i(t)$ at each time instant t is taken from a collection $A(t) = (A_1(t), \ldots, A_N(t)) \in \mathcal{M}^n$ and similarly for the remaining parameters of the system $B(t)$, $E(t)$, and so. $\theta(0)$ has initial distribution $\pi(0)$ and, the initial condition $(x(0), \theta(0))$ and the processes $\{w(t)\}$ and $\{v(t)\}$ are mutually independent; $x(0)$ is such that $\mathcal{E}(x(0)) = \bar{x}_0$ and $\mathcal{E}(x(0)x(0)') = \Sigma$, where $\mathcal{E}(\cdot)$ denotes the mathematical expectation operator. Assume also that $D_{\theta(t)}(t)D_{\theta(t)}(t)' > 0$ (nonsingular measurement noise) for each $t \geq 0$. Let $\mathcal{F}_t = \sigma\{y(s), 0 \leq s \leq t\}$ and $\mathcal{G}_t = \sigma\{\theta(s), 0 \leq s \leq t\}$ the sigma-algebras generated by the values of the output process $\{y(s), 0 \leq s \leq t\}$ and the values of the jump process $\{\theta(s), 0 \leq s \leq t\}$ respectively. The variable x is not observed (not available for control). Both x and the measurement variable y are driven by Wiener processes, which allow for modeling noise and innacuracies in the devices and sensors. Consider the quadratic cost,of (1),

$$\begin{aligned} J(x(0), \theta(0), u) = \mathcal{E}\Big\{ \int_0^T &\big[x(t)'Q_{\theta(t)}(t)x(t) \\ &+ u(t)'R_{\theta(t)}(t)u(t)\big]dt + x(T)'Q_{\theta(T)}(T)x(T)\Big\}, \end{aligned} \tag{2}$$

where T is a finite positive real number ($T > 0$) and $Q_{\theta(t)}(t) \geq 0$, $R_{\theta(t)}(t) > 0$, for each $t \geq 0$. The finite horizon quadratic optimal control problem for continuous-time MJLS with imperfect or partial state vector information roughly consists of finding a control law belonging to the family of admissible functions $u(t) = \mu(t, y(s), \theta(s), 0 \leq s \leq t)$, $0 \leq t \leq T$, measurable on $(\mathcal{F}_t, \mathcal{G}_t)$, that minimizes (2) subject to (1); we will denote the minimal value of (2) by J^*.

In order to implement control strategies that depend on the state, a natural approach to solve this problem can be taken by constructing a linear estimator (or observer, see Kwakernaak and Sivan (1972)) given as,

$$\begin{aligned} d\hat{x}(t) &= (A_{\theta(t)}(t)\hat{x}(t) + B_{\theta(t)}(t)u(t))dt + K(t) \\ &\quad \times (dy(t) - C_{\theta(t)}(t)\hat{x}(t)dt), \\ \hat{x}(0) &= \hat{x}_0, \end{aligned} \tag{3}$$

which would become the best one among all those linear estimators by finding the appropriated $\hat{x}_0$ and $K(t)$ minimizing the associated quadratic estimation error for each $t \geq 0$. Since this is a problem involving a linear system, we would like to interconnect the optimal linear estimator in (3) with the optimal control law in a linear fashion such as,

$$u(t) = L(t)\hat{x}(t), \tag{4}$$

and, then, by substituting (4) in (3), to obtain the closed-loop controller,

$$\begin{aligned} d\hat{x}(t) &= (A_{\theta(t)}(t) + B_{\theta(t)}(t)L(t) - K(t)C_{\theta(t)}(t))\hat{x}(t)dt \\ &\quad + K(t)dy(t), \\ u(t) &= L(t)\hat{x}(t). \end{aligned} \tag{5}$$

A natural approach to the presented control problem is to construct, separately, a filter to estimate the variable x and a control law to minimize the cost as if x were observed, and to combine them by replacing, in the control law, x with the estimate. This procedure leads to the optimal solution in situations where the Separation Principle (SP) holds, however it is commonly used (irrespectively of checking the SP) and good results are frequently reported. Note that this control law is generated by the output of a closed-loop controller as in (5) based on the past and present values of the observed variables, taking advantage of the information about the plant (the next sections are dedicated to further describe the performance of two controllers and we will discuss a bit about the fact of use all the available information coming from (1) and its direct relation with the optimal solution of the control problem presented here). In fact, the best linear estimate used to design the optimal controller is the KBF appearing e.g. in Ji and Chizeck (1992).

Fragoso and Costa (2005) proposed optimal estimates (called here FCF) which will be addressed in this paper through a comparative study based on some tests that allow us to observe their performance in terms of the control problem posed above. That is, we compare the cost incurred when using the estimates given by the KBF and FCF.

The dissimilarities among the FCF and the KBF are not much discussed and explored, for instance it is not much disseminated that the only distinction between them is that the FCF has a built in "markovianity constraint". Indeed, the matrix functions $L(t)$ and $K(t)$ given by (3) and (4) respectively, do not only depend on the time-parameter t but also depend on the values of the markovian parameter $\theta(s)$, for $0 \leq s \leq t$, which prevents the Markov property to hold true for KBF. In contrast, the optimal gain matrix of the FCF filter is defined as a deterministic function with values $K(t, \theta(t)) := K_{\theta(t)}(t)$.

3. SEPARATION PRINCIPLE AND THE FILTERS

In general, the separation principle states that, under some conditions, the optimal control law and the optimal state estimate can be calculated separately and then combined to produce an optimal solution. Following this idea and assuming that both the variables x and θ are observed, we have that the optimal law is given by $u(t) = L(t, \theta(t))x(t)$, where $L(t, \theta(t)) := L_{\theta(t)}(t)$ is the optimal gain control matrix for the problem with perfect state information. The optimal control gain can be calculated off-line backward in time from $t = T$ up to $t = 0$, for each $i \in \mathfrak{S}$. In fact, the solution of a coupled matrix Riccati differential equation allows to obtain the optimal gain matrix. By defining the terminal condition $\Gamma_{\theta(T)}(N) = Q_{\theta(T)}(T)$, the equation

$$\dot{\Gamma}_i(t) = -A_i(t)'\Gamma_i(t) - \Gamma_i(t)A_i(t) - \sum_{j=1}^{N} \lambda_{ij}\Gamma_j(t) - Q_i(t) + \Gamma_i(t)B_i(t)R_i^{-1}(t)B_i(t)'\Gamma_i(t), \tag{6}$$

is solved for each $i \in \mathfrak{S}$, and then we obtain the optimal collection $(L_1, \ldots, L_N)$ by

$$L_i(t) = -R_i^{-1}(t)B_i(t)'\Gamma_i(t). \tag{7}$$

In both Ji and Chizeck (1992) and Fragoso and Costa (2005) it is shown that (6) and (7) provide the optimal control gain.

When implementing the above control in the partial information problem considered in this paper, we use the separation principle and replace $x(t)$ in the linear state feedback control by an estimate $\hat{x}(t)$, that is, we set

$$u(t) = L(t, \theta(t))\hat{x}(t).$$

Estimates $\hat{x}$ of the vector state process can be obtained by solving optimization problems involving the mean integral of $||e(t)||^2$ along the horizon $T > 0$, where $e(t) = x(t) - \hat{x}(t)$ is the estimation error associated to the estimate $\hat{x}(t)$. The optimization problem may incorporate constraints to make it easier to find a solution, or to reduce the complexity of the filter, eventually deteriorating the quality of the estimate. As we will see in the sequel, the KBF requires linearity and the FCF requires both linearity and "markovianity".

3.1 Kalman-Bucy Filter

It is a well known fact that the linear optimal estimate for MJLS when θ is observed is given by the standard KBF Ji and Chizeck (1992). A simple interpretation is that, at time instant $t > 0$ and given $\theta(s), 0 \leq s \leq t$, equation (1) for $s \leq t$ becomes a standard differential equation of a linear time-varying system. Then, the optimal linear estimate is given by the observer,

$$d\hat{x}_{KB}(t) = (A_{\theta(t)}(t)\hat{x}_{KB}(t) + B_{\theta(t)}(t)u(t))dt + K(t) \times (dy(t) - C_{\theta(t)}(t)\hat{x}_{KB}(t)dt), \tag{8}$$

with initial condition $\hat{x}_{KB}(0) = \bar{x}_0$ (the same as in (1)) and the gain matrix $K(t)$ given by $K(t) = P(t)C_{\theta(t)}(t)' \times (D_{\theta(t)}(t)D_{\theta(t)}(t))^{-1}$, where for each (given) realization of the jump process, $P(t)$ satisfies the matrix differential equation,

$$\dot{P}(t) = A_{\theta(t)}(t)P(t) + P(t)A_{\theta(t)}(t)' + E_{\theta(t)}(t)E_{\theta(t)}(t)' - P(t)C_{\theta(t)}(t)'(D_{\theta(t)}(t)D_{\theta(t)}(t)')^{-1}C_{\theta(t)}(t)P(t), \tag{9}$$

with initial condition $P(0) = \Sigma - \bar{x}_0\bar{x}_0'$. $P(t)$ provides the conditional second moment given as $\mathcal{E}\{e_{KB}(t)e_{KB}(t)'|\mathcal{G}_t\}$ where $e_{KB}(t)$ is the estimation error associated to the KBF. It is easy to see and important to keep in mind that the optimal value of the mean quadratic estimation error $||e(t)||^2$ is given by the trace of the expected value of $P(t)$, that is

$$\min_{\hat{x}} \mathcal{E}\{||e(t)||^2\} = \mathcal{E}\{||e_{KB}(t)||^2\} = \operatorname{tr}(\mathcal{E}\{P(t)\}), \tag{10}$$

where $e(t)$ is a function depending on estimators $\hat{x}$ belonging to a class of linear estimators. As $K(t)$ and $P(t)$ depend on the realizations of the jump process $\theta(t)$ (they both constitute stochastic processes), offline computation of the filtering gain is usually unadvisable. There are some "partially offline" schemes that can be explored, for instance, at time instant t and given $P(t)$ and $\theta(t) = i$, one can solve (9) as if $\theta(s) = i$ within an interval $t \leq s \leq t+T$, and then implement the obtained gain while the Markov state remains at i. Finally, it is worth noting that when $x(0)$ is assumed Gaussian the KB filter coincides with the conditional expectation $\mathcal{E}\{x(t)|\mathcal{F}_t, \mathcal{G}_t\}$ which is not only the best linear estimator but the best one among all the estimators that can be obtained.

3.2 Fragoso-Costa Filter

We now present a linear estimator which features a computational advantage over the KBF in the sense that its gain matrices can be calculated off-line. The optimal filtering problem formulation here introduces two important, relatively strong constraints - the filter has to be linear and Markovian, presenting a certain prescribed structure. It is obtained a set of filtering gain matrices from which the i-th gain is selected at time t when $\theta(t) = i$, while the system operation is in progress. This lead to require that $\pi_i(t) > 0$, for all $t \geq 0$ and each $i \in \mathcal{S}$, which turns to be a quite restrictive condition to the problem. Thus, the optimal FCF behaves due to a linear Markovian system given by

$$\begin{aligned} d\hat{x}_{FC}(t) &= (A_{\theta(t)}(t)\hat{x}_{FC}(t) + B_{\theta(t)}(t)u(t))dt + K_{\theta(t)}(t) \\ &\times (dy(t) - C_{\theta(t)}(t)\hat{x}_{FC}(t)dt), \end{aligned} \tag{11}$$

with initial condition $\hat{x}_{FC}(0) = \bar{x}_0$ and the gain matrix $K_i(t)$ given by

$$K_i(t) = M_i(t)C_i(t)'(D_i(t)D_i(t)\pi_i(t))^{-1}, \tag{12}$$

where, for each $i \in \mathcal{S}$, $M_i(t)$ satisfies the coupled matrix differential equation,

$$\begin{aligned} \dot{M}_i(t) &= A_i(t)M_i(t) + M_i(t)A_i(t)' + \sum_{j=1}^{N} \lambda_{ji} M_j(t) \\ &\quad + E_i(t)E_i(t)'\pi_i(t) - M_i(t)C_i(t)'(D_i(t)D_i(t)' \\ &\quad \times \pi_i(t))^{-1}C_i(t)M_i(t), \\ M_i(0) &= (\Sigma - \bar{x}_0\,\bar{x}_0{}')\pi_i(0). \end{aligned} \tag{13}$$

As we can see, the key point here is concerned with showing (see Fragoso and Costa (2005) for details) that the solution of the equation (13) is such that $M_i(t) = \mathcal{E}\{\epsilon_{FC}(t)\epsilon_{FC}(t)'1_{\{\theta(t)=i\}}\}$, where the inclusion of the indicator function expresses the intention of selecting the gain matrices at time t when $\theta(t) = i$. $\epsilon_{FC}(t)$ is the estimation error associated to the FCF and the minimal value of the quadratic estimation error is given by

$$\min_{\hat{x}} \mathcal{E}\{||\epsilon(t)||^2\} = \mathcal{E}\{||\epsilon_{FC}(t)||^2\} = \sum_{i=1}^{N} \mathrm{tr}(M_i(t)), \tag{14}$$

where $\epsilon(t)$ is a function depending on estimators $\hat{x}$ belonging to a class of linear Markovian estimators. Because the class of linear Markovian estimators is contained in the wider linear estimators class, we have that $\mathrm{tr}(\mathcal{E}\{P(t)\}) \leq \sum_{i=1}^{N} \mathrm{tr}(M_i(t))$, for all $t \geq 0$. The degree of suboptimality is strongly associated with the conjuncture that the FCF gains computation from (12) and (13) does not use all available information carried on $(\mathcal{F}_t, \mathcal{G}_t)$ but only the "non-filtered" distribution $\pi(t) = P(\theta(t) = i)$, which depends only on $\pi(0)$. This fact will be reflected in the example in Section 4.

We find it interesting to give some details of the analysis carried out in Fragoso and Costa (2005) to show that the linear Markovian estimator $\hat{x}_{FC}$ is optimal (in the sense of $\mathcal{E}\{||\epsilon(t)||^2\}$) among all the linear Markovian estimators. An estimator $\hat{x}_{op}$ with predetermined structure (linear, Markovian) is considered, with gain and covariance matrices as in (12) and (13) respectively. It is easy to see that the associated mean quadratic estimation error associated to $\hat{x}_{op}$ is given by $\mathcal{E}\{||\epsilon_{op}(t)||^2\} = \sum_{i=1}^{N} \mathrm{tr}(M_i(t))$. Then, we have that for any linear Markovian estimator $\hat{x}$,

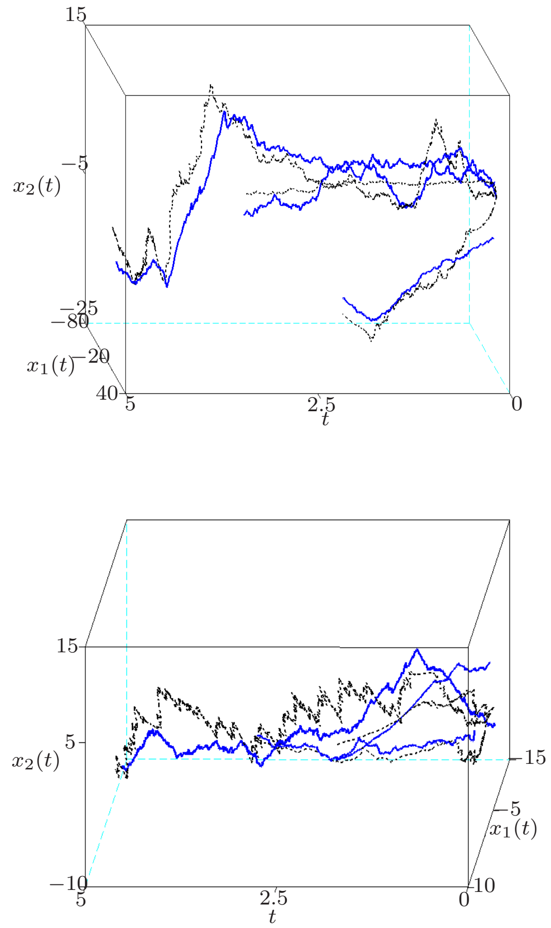

Fig. 1. *Phase plot of one realization for the state trajectory $x(t)$ (blue solid line on both graphics), for time horizons $T = 2, 3, 5$, being aproximated by the KBF $\hat{x}_{KB}(t)$ (black dotted line on top graphic) and the FCF $\hat{x}_{FC}(t)$ (black dotted line on lower graphic). The graphic on the top suggests that the KBF provides a better estimate of the trajectory as we shall see in subsequent examples.*

$$\begin{aligned} \mathcal{E}\{||\epsilon(t)||^2\} &= \mathcal{E}\{||x - \hat{x}||^2\} \\ &= \mathcal{E}\{||\hat{x} - \hat{x}_{op}||^2\} + \sum_{i=1}^{N} \mathrm{tr}(M_i(t)) \\ &\geq \sum_{i=1}^{N} \mathrm{tr}(M_i(t)), \end{aligned} \tag{15}$$

which leads to the surprising fact that $(M_1(t), \ldots, M_N(t))$ generates a lower bound for $\mathcal{E}\{||\epsilon(t)||^2\}$. Thus, when $\hat{x} = \hat{x}_{op}$, (15) is minimized and, therefore, $\hat{x}_{op}$ is the FCF, that is, $\hat{x}_{FC} := \hat{x}_{op}$.

Remark 1. *As mentioned before, the most significant dissimilarity between the KBF and the FCF is that the later is Markovian. To achieve this, the computation of the gains $K_i(t)$ disregards the information on $\theta(s), 0 \leq s \leq t$ that is available at time t, which is the basic "source of suboptimality" in perspective with the KBF; note in this connection that the distribution $\pi(t)$ appearing in the Riccati equation (13) is irrespective of the observations. On*

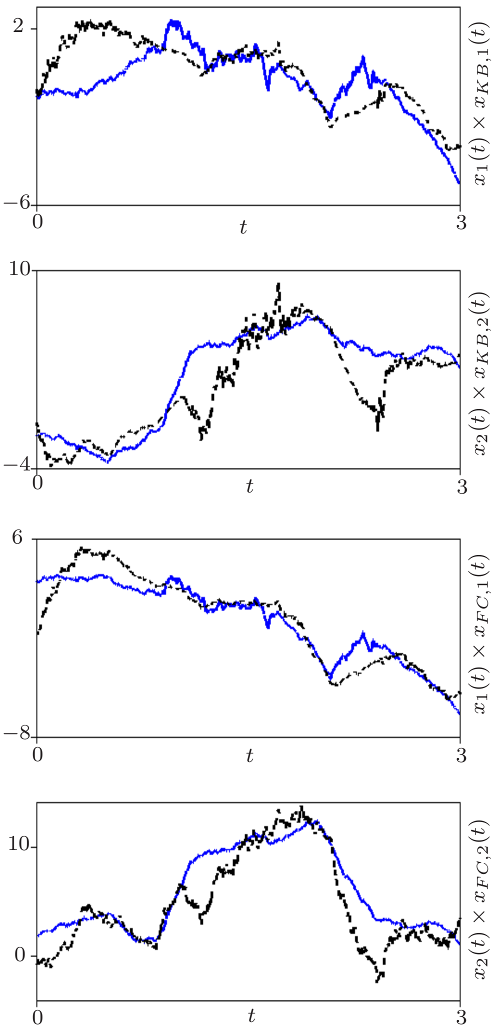

Fig. 2. *Phase plot of one realization for the components of* $x(t)$ *(blue solid line in the graphics), being aproximated by the components of the KBF* $\hat{x}_{KB}(t)$ *(black dotted line on the two top graphics) and the components of the FCF* $\hat{x}_{FC}(t)$ *(black dotted line on the two lower graphics).*

the bright side, this allows for pre-computation of gains, a much valuable feature for the filter implementation. Roughly speaking, the FCF trades information for simplicity. Then, $\pi(t)$ *is the key to understand the main differences in the fitlers performances. The more* $\pi(t)$ *is concentrated in a state* i *during a time interval, the less the actual state trajectory* $\theta(t)$ *leaves* i *in the same interval, and closer the performances are. At the other extreme, when the distribution is "balanced" (close to equiprobable) then the realization of the Markov chain is less predictible and the gap between the performances of the filters is high.*

4. NUMERICAL TESTS

In this section we present case studies involving situations where the difference in performance of the filters is large/small. The costs incurred by the combined controls and filters are estimated using Monte Carlo simulation based on 10.000 realizations. The average integral of the (square) estimation error is also presented, to illustrate the effectiveness of each filter.

Example 1 Consider the MJLS given in (1) with matrices:

$$A_1 = \begin{bmatrix} -1 & 0.05 \\ 10 & -1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & -0.9 \\ 1.1 & 0.6 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 0 & -1.7 \\ 1.4 & -0.5 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 \\ -0.5 \end{bmatrix},$$

$$E_1 = \begin{bmatrix} 1 & 0.2 & -1.9 \\ -0.1 & 1.4 & -0.3 \end{bmatrix}, E_2 = \begin{bmatrix} 0.4 & 0.3 & 0.8 \\ 0.9 & 0 & 0.1 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 0 & -0.9 & 0 \\ 1.2 & -0.2 & 0.1 \end{bmatrix}, C_1 = \begin{bmatrix} -0.3 & 0.3 \\ 0.9 & -0.7 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 2.5 \end{bmatrix}, C_3 = \begin{bmatrix} -1 & 0.5 \\ 0 & -1.2 \end{bmatrix}, D_i = I, \ i = 1:3,$$

$$Q_1 = I, \ Q_2 = 2Q_1, \ Q_3 = 0, \text{ and,}$$

$R_1 = 10$, $R_2 = 0.5$, and $R_3 = 1$, with I being the identity matrix of appropriate dimension. Let $x(0)$ with a Gaussian distribution and assume also that each element of the initial distribution vector $\pi(0)$ is strictly positive. Finally, the transition probability matrix is,

$$\Lambda = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -3 & 2 \\ 1.5 & 0.5 & -2 \end{bmatrix}.$$

The costs incurred by the combined KBF/optimal control and the FCF/optimal control were estimated by Monte Carlo simulation, and displayed in the following table. The square estimation errors are also presented.

J^* values using KBF and FCF

	J^*-KBF	**J^*-FCF**
$T = 1$:	11,941.083	12,132.159
$T = 2$:	53,754.037	56,458.934
$T = 3$:	212,982.22	223,101.63

The next table clearly confirms the KBF as the best linear estimator for the state vector $x(t)$ in the sense that it minimizes the function $\ell = \int_0^T \mathcal{E}\{||e(t)||^2\}dt$. Observe that the costs and the quadratic error values attained by the KBF are lower than the ones by the FCF, confirming the suboptimality described in Section 3.2.

ℓ values using KBF and FCF

	ℓ-KBF	**ℓ-FCF**
$T = 1$:	2,794.7941	3,320.6918
$T = 2$:	5,260.8757	6,483.0149
$T = 3$:	8,847.1907	10,756.499

Example 2 Consider the MJLS of Example 1 with

$$\Lambda = \begin{bmatrix} -10 & 10 & 0 \\ 15 & -20 & 5 \\ 10^{-2} & 0 & -10^{-2} \end{bmatrix},$$

where we have a rather slow state $\theta = 3$, yielding $\lim_{t\to\infty} \pi_3(t) \approx 1$. Typically, in this situation and for large enough t, one has that $M_i(t) \approx 0$ for each $i \neq 3$ (see (13)) and, on the other hand, most realizations of the Markov chain provide $\theta(t) = 3$, for large t. This makes $P(t)$ (given by (9)) to approach $M_3(t)$ and the filter performances get similar after a time transient. In fact, we have obtained $J^\star$-KBF$= 16,653.748$ and $J^\star$-FCF$= 16,662.768$, ℓ-KBF$= 4,364.098$ and ℓ-FCF$= 4,585.139$ considering time horizon $T = 5$, and these figures are much closer to each other in this example than in Example 1. This is in perfect agreement with Remark 1.

Example 3 We consider 1000 MJLS with $n = 2$ and $N \in \{1, 2\}$. In each system we randomly generated (using the command "rand" in Scilab Scilab Enterprises (2012)) the collections of matrices A, B, C, D, E, Q. The transition rate matrix is fixed,

$$\Lambda = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

In order to make clear the differences in the performances, we consider the quantity $R = J^\star$-KBF$/J^\star$-FCF for each system and present in Table 4 the average $\bar{R}$ for different time horizons, which confirms that the KBF has a significantly better performance for large time horizons. In 90% of the systems we observed that the cost attained by using the KBF is at least 10% smaller when compared with the FCF. The limiting distribution is equiprobable, which, accordingly to Remark 1, is one of the most desfavourable situations for the FCF.

Average ratio between $J^\star$-KBF and $J^\star$-FCF values

	$\bar{R}$
$T = 1$:	0.9754584
$T = 2$:	0.7865339
$T = 3$:	0.5533218

5. CONCLUSIONS

In this paper we have presented the formulation of two different filters available in literature for Markov jump linear systems in the scenario where the jump variable is observed. We have explained that the filters are optimal in different senses, and that the KBF provides better estimates (in terms of mean square estimation error), however is considerably more difficult to implement as offline computation is not viable, whereas the FCF can be implemented offline in a relatively simple manner under the assumption that each element of the probability distribution vector $\pi(t)$ is positive. By combining these estimates with the optimal control for the complete state observation problem, in a typical "separated" implementation, we have obtained that the cost incurred by the KBF is smaller than the one by the FCF, in particular making clear that the FCF does not attain the standard separation principle for the partial observation control problem defined in (2). This is confirmed by numerical examples, which suggest that the gap between the performances is in direct proportion to the time horizon T, however that the difference is less marked when the Markov state distribution is concentrated for long periods of time.

Putting the filters in perspective has made clear the main reasons why they are different, which helps to choose which filter to use in real-world applications. Systems featuring absorbing or slow Markov states (along with fast ones) tend to present good performance with the FCF, while the KBF may be preferable for systems with "balanced" distributions, depending on the viability of its implementation. As for future works we plan to propose hybrid filters combining both KBF and FCF, taking into account some measure of how concentrated the distribution is at each time instant.

REFERENCES

Anderson, B.D.O. and Moore, J.B. (1979). *Optimal Filtering*. Prentice-Hall, London, first edition.

Athans, M., Castanon, D., Dunn, K.P., Greene, C.S., Lee, W.H., Sandell, N.R., and Willsky, A.S. (1977). The stochastic control of the F-8C aircraft using a multiple model adaptive control (MMAC) method - Part i: Equilibrium flight. *IEEE Transactions on Automatic Control*, 22, 768–780.

Boukas, E., Shi, P., and Andijani, A. (1999). Robust inventory-production control problem with stochastic demand. *Optimal Control Appl. & Methods*, 20(1), 1–20.

Costa, E.F. and Astolfi, A. (2009a). Characterization of exponential divergence of the kalman filter for time-varying systems. *SIAM Journal on Control and Optimization*.

Costa, E.F. and Astolfi, A. (2009b). Stochastic detectability and mean bounded error covariance of the recursive Kalman filter with Markov jump parameters. *Stochastic Analysis and Applications*.

do Val, J. and Basar, T. (1999). Receding horizon control of jump linear systems and a macroeconomic policy problem. *Journal of Economic Dynamics & Control*, 23, 1099–1131.

Fragoso, M.D. and Costa, O.L.V. (2005). A separation principle for the continuous-time lq-problem with markovian jump parameters. *IEEE Transactions on Automatic Control*, 55.

Gomes, M.J.F. and Costa, E.F. (2010). On the stability of the recursive kalman filter with markov jump parameters. In *Proc. ACC'10 American Control Conference*. Baltimore,USA.

Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press.

Ji, Y. and Chizeck, H.J. (1992). Jump linear quadratic gaussian conrol in continuous time. *IEEE Transactions on Automatic Control*, 37(12), 1884–1892.

Kumar, P.R. and Varaiya, P. (1986). *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice-Hall.

Kwakernaak, H. and Sivan, R. (1972). *Linear Optimal Control Systems*. Wiley-Interscience.

Miller, B.M. and Runggaldier, W.J. (1997). Kalman filtering for linear systems with coefficients driven by a hidden markov jump process. *Systems & Control Letters*, 31, 93–102.

Scilab Enterprises (2012). *Scilab: Free and Open Source software for numerical computation*. Scilab Enterprises, Orsay, France. URL `http://www.scilab.org`.

Sworder, D.D. and Rogers, R.O. (1983). An LQ-solution to a control problem associated with a solar thermal central receiver. *IEEE Transactions on Automatic Control*, 28(10), 971–978.