Automatica 47 (2011) 466–476

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Linear minimum mean square filter for discrete-time linear systems with Markov jumps and multiplicative noises[☆]

Oswaldo L.V. Costa [*], Guilherme R.A.M. Benites

Departamento de Engenharia de Telecomunicações e Controle, Escola Politécnica da Universidade de São Paulo, 05508-970-São Paulo, SP, Brazil

ARTICLE INFO

Article history:
Available online 15 February 2011

Keywords:
Filtering theory
Kalman filters
Multiplicative noise
Jump process
Riccati equations

ABSTRACT

In this paper we obtain the linear minimum mean square estimator (LMMSE) for discrete-time linear systems subject to state and measurement multiplicative noises and Markov jumps on the parameters. It is assumed that the Markov chain is not available. By using geometric arguments we obtain a Kalman type filter conveniently implementable in a recurrence form. The stationary case is also studied and a proof for the convergence of the error covariance matrix of the LMMSE to a stationary value under the assumption of mean square stability of the system and ergodicity of the associated Markov chain is obtained. It is shown that there exists a unique positive semi-definite solution for the stationary Riccati-like filter equation and, moreover, this solution is the limit of the error covariance matrix of the LMMSE. The advantage of this scheme is that it is very easy to implement and all calculations can be performed offline.

1. Introduction

Linear systems subject to Markov jumps and multiplicative noises have been receiving a great deal of attention lately. This is due mainly to the fact that this kind of formulation has found many applications in engineering and finance. Some examples of such systems can be found in nuclear fission and heat transfer, population models and immunology, portfolio optimization, etc. (see, for instance, Costa & Kubrusly, 1996, Costa & de Paulo, 2007, Dragan & Morozan, 2002, Dragan & Morozan, 2006a,b, Geromel, Gonçalves, & Fioravanti, 2009 and Gershon & Shaked, 2006 and references therein for H_2 and H_∞ control problems, optimal filtering, robust stability and stabilizability conditions, predictive model-based control, etc.).

The filtering problem of this class of systems has also attracted a great deal of interest in the last past years under different hypothesis and performance criterions. For systems with only multiplicative noise we can mention (Chow & Birkemeier, 1990), in which it was considered that the influence of multiplicative noises affects only the measurements of the model, and a recursive structure was achieved by combining the previous estimate with a recursive innovation, which yields a linear combination of the most recent data samples and the previous estimate. The results in Chow and Birkemeier (1990) were somehow generalized in Zhang and Zhang (2007) to consider correlated additive noises. In Carravetta, Germani, and Raimondi (1997) the multiplicative noise affects only the state model and the theory developed covers linear systems with nonstationary and non-Gaussian noises. The authors were able to define a filter for systems with multiplicative state noises which is optimal in a class of polynomial transformations. In Yang, Wang, and Hung (2002) the authors considered a discrete time-varying system with both additive and multiplicative noises. The problem addressed is to design a linear system that yields an estimation error variance with an optimized guaranteed upper bound for all admissible uncertainties. The sufficient conditions for designing such a filter were derived in terms of two Riccati difference equations. The filtering and control problem for systems subject to multiplicative noises under the H_∞ criterion has been studied in Gershon, Shaked, and Yaesh (2001). For systems with only Markov jumps in the parameters and when only an output of the system is available, so that the values of the jump parameter are not known, the problem of optimal and sub-optimal filtering has been addressed in Ackerson and Fu (1970), Bar-Shalom and Li (1993), Blom and Bar-Shalom (1988), Chang and Athans (1978), Dufour and Elliott (1997) and Tugnait (1982) among other authors, under the hypothesis of Gaussian distribution for the disturbances, and by Zhang (1999, 2000) for the non-Gaussian case. Since the optimal estimator requires exponentially increasing memory and computation with time, sub-optimal algorithms are required. In the papers mentioned before the authors considered non-linear

[☆] The first author received financial support from CNPq (Brazilian National Research Council), grant 301067/2009-0. The material in this paper was partially presented at 49th IEEE Conference on Decision and Control, December 15–17, 2010, Atlanta, Georgia, USA. This paper was recommended for publication in revised form under the direction of Editor Berç Rüstem.

[*] Corresponding author. Tel.: +55 1130915771; fax: +55 1130915718.
E-mail addresses: oswaldo@lac.usp.br (O.L.V. Costa), guilherme@riskoffice.com.br (G.R.A.M. Benites).

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.01.015

sub-optimal estimators, which requires on-line calculations. In Costa (1994) it was obtained the linear minimum mean square error estimator (LMMSE) for MJLS, based on estimating $x(k)1_{\{\theta(k)=i\}}$ instead of estimating directly the state variable $x(k)$, where $1_{\{\cdot\}}$ stands for the Dirac measure. The advantage of this formulation is that it leads to a time-varying linear filter easy to implement and in which all the calculations (the filter gains) can be performed off-line. The stationary and robust linear filters based on an augmented algebraic Ricatti equation has been presented in Costa and Guerra (2002a,b), leading to a discrete-time invariant linear filter. A similar problem but using a linear matrix inequalities approach was considered in Fioravanti, Gonçalves, and Geromel (2008). A Kalman type filtering problem for linear systems subject to both state and measurement multiplicative noises and to Markov jumps was considered in Stoica, Dragan, and Yaesh (2009). The Markov parameter was assumed to be available for the filter and the state-dynamic and measurement noises were assumed to be independent of each other. It was obtained a mean-square filter in a stationary form, based on the solution of some coupled algebraic Riccati equations.

In this paper we obtain the LMMSE for discrete-time linear systems subject to state and measurement multiplicative noises, and Markov jumps on the parameters. As far as the authors are aware, this is the first time that the linear filtering problem under these general conditions (multiplicative noises and Markov jumps, and the Markov parameter not available) is considered in the literature. Indeed some previous papers dealt with the filtering problem (under different criterions and assumptions) for systems with either only Markov jumps (as for instance in Ackerson and Fu (1970), Bar-Shalom and Li (1993), Blom and Bar-Shalom (1988), Chang and Athans (1978), Costa (1994), Costa and Guerra (2002a), Costa and Guerra (2002b), Dufour and Elliott (1997), Fioravanti et al. (2008) and Tugnait (1982)) or only with multiplicative noises (as for instance in Chow and Birkemeier (1990), Gershon et al. (2001), Yang et al. (2002) and Zhang and Zhang (2007)). In Stoica et al. (2009) the linear filtering problem is studied for systems with both multiplicative noises and Makov jumps but, differently from here, in Stoica et al. (2009) the Markov parameter is assumed to be available. By using geometric arguments and estimating, as in Costa (1994), $x(k)1_{\{\theta(k)=i\}}$ instead of estimating directly the state variable $x(k)$, we obtain a Kalman type filter conveniently implementable in a recurrence form from 2 matricial equations, one of a Lyapunov type, and the other one of a Kalman type. The convergence of these iterations to the associated Lyapunov and Riccati-like equations is also studied, generalizing the results in Costa (1994) and Costa and Guerra (2002b), which only considered Markov jump systems. For the case in which there are neither Markov jumps nor multiplicative noises our results recover those for the standard Kalman filter for linear systems (see Remark 3).

The paper is organized as follows. The notation, assumptions, definitions, the problem formulation and some preliminary results are presented in Sections 2 and 3. The main results are presented in Sections 4 and 5, which derive the LMMSE and the convergence of the associated Riccati-like equation. We illustrate the developed results with a numerical example in Section 6. Final comments are made in Section 7. Some auxiliary results are presented in the Appendix.

2. Notation, definitions, and auxiliary results

Throughout the paper the n-dimensional real Euclidean space will be denoted by $\mathbb{R}^n$. For $\mathbb{X}$ and $\mathbb{Y}$ Banach spaces we set $\mathbb{B}(\mathbb{X},\mathbb{Y})$ for the Banach space of all bounded linear operators of $\mathbb{X}$ into $\mathbb{Y}$, with the uniform induced norm represented by $\|.\|$. For simplicity we shall set $\mathbb{B}(\mathbb{X}) := \mathbb{B}(\mathbb{X},\mathbb{X})$. For $\mathcal{T} \in \mathbb{B}(\mathbb{X})$ we denote $r_\sigma(\mathcal{T})$ the spectral radius of $\mathcal{T}$. The superscript $'$ will denote the transpose of a vector or a matrix, and for a matrix $M \in \mathbb{B}(\mathbb{R}^n)$, $M \geq 0$ ($M > 0$ respectively) means that M is a positive semi-definite (positive definite) matrix. Define $\mathbb{H}^{n,m} = \{\mathbf{Q} = (Q_1,\ldots,Q_N);\ Q_i \in \mathbb{B}(\mathbb{R}^n,\mathbb{R}^m),\ i = 1,\ldots,N\}$, $\mathbb{H}^n = \mathbb{H}^{n,n}$, and $\mathbb{H}^{n+} = \{\mathbf{Q} = (Q_1,\ldots,Q_N) \in \mathbb{H}^n;\ Q_i \geq 0,\ i = 1,\ldots,N\}$. For $\mathbf{Q} = (Q_1,\ldots,Q_N) \in \mathbb{H}^{n+}$, $\mathbf{V} = (V_1,\ldots,V_N) \in \mathbb{H}^{n+}$ we say that $\mathbf{Q} \geq \mathbf{V}$ if for each $i = 1,\ldots,N$ we have that $Q_i \geq V_i$. For $\mathbf{D} = (D_1,\ldots,D_N) \in \mathbb{H}^{n,m}$ we define $\text{diag}(D_j) \in \mathbb{B}(\mathbb{R}^{Nn},\mathbb{R}^{Nm})$ as the diagonal matrix formed by D_j in the diagonal and zero elsewhere, and define the operator $\text{Dg} : \mathbb{H}^{n,m} \to \mathbb{B}(\mathbb{R}^{Nn},\mathbb{R}^{Nm})$ as $\text{Dg}(\mathbf{D}) = \text{diag}(D_j)$. Finally $1_{\{\cdot\}}$ stands for the Dirac measure or equivalently, the indicator function of the event $\{\cdot\}$.

In what follows it will be convenient to introduce the following notation. For any sequence of second order random vectors $\{r(k)\}$ we define the "centered" random vector $r^c(k)$ as $r(k) - E(r(k))$, $\widehat{r}(k|t)$ the best affine estimator of $r(k)$ given $\{y(0),\ldots,y(t)\}$, and $\widetilde{r}(k|t) = r(k) - \widehat{r}(k|t)$. Similarly $\widehat{r}^c(k|t)$ is the best linear estimator of $r^c(k)$ given $\{y^c(0),\ldots,y^c(t)\}$ and $\widetilde{r}^c(k|t) = r^c(k) - \widehat{r}^c(k|t)$. It is well known (cf. Davis & Vinter, 1985, p. 109) that

$$\widehat{r}(k|t) = \widehat{r}^c(k|t) + E(r(k)) \tag{1}$$

and, in particular, $\widetilde{r}^c(k|t) = \widetilde{r}(k|t)$. We shall denote by $\mathfrak{L}(y^k)$ the linear subspace spanned by

$$(y)^k = \begin{pmatrix} y(0) \\ \vdots \\ y(k) \end{pmatrix},$$

(see Davis & Vinter, 1985), that is, a random variable $r \in \mathfrak{L}(y^k)$ if $r = \sum_{i=0}^{k} \alpha(i)' y(i)$ for some $\alpha(i) \in \mathbb{R}^m$, $i = 0,\ldots,k$. We recall that for the second order random vectors r and s taking values in $\mathbb{R}^n$, the inner product $\langle .; .\rangle$ is defined as $\langle r; s\rangle = E(s'r)$ and therefore we say that r and s are orthogonal if $\langle r; s\rangle = E(s'r) = 0$. For $t \leq k$, the best linear estimator $\widehat{r}^c(k|t)$ of the random vector $r^c(k)$, represented by

$$\widehat{r}^c(k|t) = \begin{pmatrix} \widehat{r}^c_1(k|t) \\ \vdots \\ \widehat{r}^c_n(k|t) \end{pmatrix}, \qquad r^c(k) = \begin{pmatrix} r^c_1(k) \\ \vdots \\ r^c_n(k) \end{pmatrix},$$

is the projection of $r^c(k)$ onto the subspace $\mathfrak{L}((y^c)^t)$ and satisfies the following properties (cf. Davis & Vinter, 1985, p. 108 and 113):

(1) $\widehat{r}^c_j(k|t) \in \mathfrak{L}((y^c)^t)$, $j = 1,\ldots,n$
(2) $\widetilde{r}_j(k|t)$ is orthogonal to $\mathfrak{L}((y^c)^t)$, $j = 1,\ldots,n$
(3) if $\text{cov}((y)^t)$ is non-singular then

$$\widehat{r}^c(k|t) = E(r^c(k)(y^c)^{t\prime})\text{cov}((y)^t)^{-1}(y^c)^t, \tag{2}$$

$$\widehat{r}^c(k|k) = \widehat{r}^c(k|k-1) + E(\widehat{r}^c(k)\widetilde{y}(k|k-1)') \times E(\widetilde{y}(k|k-1)\widetilde{y}(k|k-1)')^{-1}(y^c(k) - \widehat{y}^c(k|k-1)). \tag{3}$$

3. Problem formulation, assumptions and definitions

We consider the following discrete-time linear system with Markov jumps and multiplicative noise, on a probabilistic space $(\Omega, \mathbf{P}, \mathcal{F})$:

$$\begin{aligned} x(k+1) &= \left(\bar{A}_{\theta(k)}(k) + \sum_{s=1}^{\varepsilon^x} \widetilde{A}_{\theta(k),s}(k) w^x_s(k)\right) x(k) \\ &\quad + C_{\theta(k)}(k) w(k), \\ y(k) &= \left(\bar{H}_{\theta(k)}(k) + \sum_{s=1}^{\varepsilon^y} \widetilde{H}_{\theta(k),s}(k) w^y_s(k)\right) x(k) + G_{\theta(k)}(k) w(k), \\ x(0) &= x_0. \end{aligned} \tag{4}$$

Here the state variable is given by the n-dimensional vector $x(k)$, the measurement variable by $y(k)$, the additive noise by $w(k)$, the Markov chain by $\theta(k)$, the state multiplicative-noise by $w_s^x(k)$ and the measurement multiplicative noise by $w_s^y(k)$. The goal here is to obtain the linear minimum mean square estimator for $x(k)$ given the observations $\{y(0),\ldots,y(k)\}$ (we assume that $\theta(k)$ is not available). We assume that:

(i) the additive noise $\{w(k)\}$ is given by a sequence of null-mean independent random vectors with covariance matrix equal to the identity;
(ii) the multiplicative noises $\{w_s^x(k);\ s=1,\ldots,\varepsilon^x\}$ and $\{w_s^y(k);\ s=1,\ldots,\varepsilon^y\}$ are both zero-mean independent sequence of random variables with variance equal to 1 and $E(w_i^x(k)w_j^x(k))=0$, $E(w_i^y(k)w_j^y(k))=0$, for all k and $i\neq j$;
(iii) the mutual correlation between $w_{s_1}^x(k)$ and $w_{s_2}^y(k)$ is given by $E(w_{s_1}^x(k)w_{s_2}^y(k))=\rho_{s_1,s_2}$;
(iv) $\{w_s^x(k);\ s=1,\ldots,\varepsilon^x\}$ and $\{w_s^y(k);\ s=1,\ldots,\varepsilon^y\}$ are independent of $\{w(k)\}$ and x_0, as well as $\{w(k)\}$ is independent of x_0.
(v) the Markov chain $\{\theta(k)\}$ takes values in $\{1,\ldots,N\}$ and has transition probability matrix $P_M=[p_{ij}]$. We set $\pi_j(k):=\mathbf{P}(\theta(k)=j)$.
(vi) the Markov chain $\{\theta(k)\}$ is independent of $\{w_s^x(k);\ s=1,\ldots,\varepsilon^x\}$, $\{w_s^y(k);\ s=1,\ldots,\varepsilon^y\}$, $\{w(k)\}$ and x_0.
(vii) the initial condition x_0 is an n-dimensional random vector with $E(x_0)=\mu_0$ and $Q_0=E(x_0x_0')$.

Remark 1. It should be noticed that if the "gain" matrix $C_{\theta(k)}(k)$ depended linearly on the state variable $x(k)$ then the noise term associated to this expression could also be regarded as a state multiplicative noise and our results would remain valid provided assumption (iv) is satisfied.

We consider the following augmented matrices

$$\bar{A}(k):=\begin{pmatrix} p_{11}\bar{A}_1(k) & \cdots & p_{N1}\bar{A}_N(k)\\ \vdots & \ddots & \vdots\\ p_{1N}\bar{A}_1(k) & \cdots & p_{NN}\bar{A}_N(k)\end{pmatrix},$$
$$\widetilde{A}_s(k):=\begin{pmatrix} p_{11}\widetilde{A}_{1,s}(k) & \cdots & p_{N1}\widetilde{A}_{N,s}(k)\\ \vdots & \ddots & \vdots\\ p_{1N}\widetilde{A}_{1,s}(k) & \cdots & p_{NN}\widetilde{A}_{N,s}(k)\end{pmatrix},$$
$$\bar{H}(k):=\left(H_1(k)\cdots H_N(k)\right),$$
$$\widetilde{H}_s(k):=\left(\widetilde{H}_{1,s}(k)\cdots\widetilde{H}_{N,s}(k)\right),$$
$$G(k):=\left(G_1(k)\pi_1(k)^{1/2}\cdots G_N(k)\pi_N(k)^{1/2}\right),$$
$$C(k):=\begin{pmatrix} \pi_1(k)^{1/2}p_{11}C_1(k) & \cdots & \pi_N(k)^{1/2}p_{N1}C_N(k)\\ \vdots & \ddots & \vdots\\ \pi_1(k)^{1/2}p_{1N}C_1(k) & \cdots & \pi_N(k)^{1/2}p_{NN}C_N(k)\end{pmatrix}.$$

We define the linear operators $\mathcal{B}_1(k,.)$, $\mathcal{B}_2(k,.)$, $\mathcal{B}(k,.):\mathbb{H}^n\to\mathbb{B}(\mathbb{R}^{Nn})$ as follows: for $\boldsymbol{\Upsilon}=(\Upsilon_1,\ldots,\Upsilon_N)\in\mathbb{H}^n$,

$$\mathcal{B}_1(k,\boldsymbol{\Upsilon}):=\operatorname{diag}\left[\sum_{i=1}^N p_{ij}\bar{A}_i(k)\Upsilon_i\bar{A}_i(k)'\right]-\bar{A}(k)(\operatorname{diag}(\Upsilon_i))\bar{A}(k)', \quad (5)$$

$$\mathcal{B}_2(k,\boldsymbol{\Upsilon}):=\operatorname{diag}\left[\sum_{i=1}^N p_{ij}\sum_{s=1}^{\varepsilon^x}\widetilde{A}_{i,s}(k)\Upsilon_i\widetilde{A}_{i,s}(k)'\right], \quad (6)$$

$$\mathcal{B}(k,\boldsymbol{\Upsilon}):=\mathcal{B}_1(k,\boldsymbol{\Upsilon})+\mathcal{B}_2(k,\boldsymbol{\Upsilon}). \quad (7)$$

Remark 2. Notice that, as shown in Costa and Guerra (2002b), if $\boldsymbol{\Upsilon}\in\mathbb{H}^{n+}$ then $\mathcal{B}_1(\boldsymbol{\Upsilon})\geq 0$. It is easy to see that in this case we also have $\mathcal{B}_2(\boldsymbol{\Upsilon})\geq 0$ and thus $\mathcal{B}(\boldsymbol{\Upsilon})\geq 0$ whenever $\boldsymbol{\Upsilon}\in\mathbb{H}^{n+}$.

We present next the second moment matrices associated to the state variable in (4). For this we define $z(k,i)=x(k)1_{\{\theta(k)=i\}}$, $i=1,\ldots,N$,

$$z(k)=\begin{pmatrix} z(k,1)\\ \vdots\\ z(k,N)\end{pmatrix},$$
$$Q_i(k)=E(z(k,i)z(k,i)')=E(x(k)x(k)'1_{\{\theta(k)=i\}}),$$

and the operator $\mathcal{T}(k,.)\in\mathbb{B}(\mathbb{H}^n)$ as follows. For $\boldsymbol{\Upsilon}=(\Upsilon_1,\ldots,\Upsilon_N)\in\mathbb{H}^n$, $\mathcal{T}(k,\boldsymbol{\Upsilon})=(\mathcal{T}_1(k,\boldsymbol{\Upsilon}),\ldots,\mathcal{T}_N(k,\boldsymbol{\Upsilon}))$ is given, for $j=1,\ldots,N$, as

$$\mathcal{T}_j(k,\boldsymbol{\Upsilon})=\sum_{i=1}^N p_{ij}\left(\bar{A}_i(k)\Upsilon_i\bar{A}_i(k)'+\sum_{s=1}^{\varepsilon^x}\widetilde{A}_{i,s}(k)\Upsilon_i\widetilde{A}_{i,s}(k)'\right). \quad (8)$$

Define $\mathbf{D}(k)=(D_1(k),\ldots,D_N(k))\in\mathbb{H}^{n+}$ as

$$D_j(k)=\sum_{i=1}^N p_{ij}\pi_i(k)C_i(k)C_i(k)'.$$

Set $\mathbf{Q}(k)=(Q_1(k),\ldots,Q_N(k))$ and $Q(k)=\operatorname{Dg}(\mathbf{Q}(k))=\operatorname{diag}(Q_i(k))=E(z(k)z(k)')$. Following the same reasoning as in Proposition 3.35 of Costa, Fragoso, and Marques (2005), the Lyapunov-like recurrent equation for the second moment matrices $Q_i(k)$ is given by:

$$\mathbf{Q}(k+1)=\mathcal{T}(k,\mathbf{Q}(k))+\mathbf{D}(k),\qquad Q_i(0)=\pi_i(0)Q_0. \quad (9)$$

We set now

$$\mu(0)=E(z(0))=\begin{pmatrix}\mu_0\pi_1(0)\\ \vdots\\ \mu_0\pi_N(0)\end{pmatrix},$$
$$\begin{aligned}P(0)&=E((z(0)-\mu(0))(z(0)-\mu(0)'))\\ &=Q(0)-\mu(0)\mu(0)'.\end{aligned}$$

We make the following assumptions (A1) and (A2) to guarantee the existence of the inverse of some matrices that will be needed in the proof of Theorem 1:

(A1) We assume that

$$\bar{H}(0)P(0)\bar{H}(0)'+\sum_{s=1}^{\varepsilon^y}\widetilde{H}_s(0)Q(0)\widetilde{H}_s(0)'+G(0)G(0)'>0.$$

(A2) We assume that for $k=1,2,\ldots,$

$$\sum_{s=1}^{\varepsilon^y}\widetilde{H}_s(k)Q(k)\widetilde{H}_s(k)'+G(k)G(k)'>0.$$

It is easy to check that if we have the usual kind of assumption for Kalman filtering problems, which is in our case $G_i(k)G_i(k)'>0$ for each $k=0,1,2,\ldots$ and $i=1,\ldots,N$, then (A1) and (A2) will hold.

4. The LMMSE filter

We present now the LMMSE filter for model (4), derived from geometric arguments as in Davis and Vinter (1985).

Theorem 1. *Consider the system represented by (4). Then for $k = 0, 1, \ldots$, the LMMSE $\widehat{x}(k|k)$ is given by*

$$\widehat{x}(k|k) = \sum_{i=1}^{N} \widehat{z}(k, i|k) \tag{10}$$

where $\widehat{z}(k|k)$ satisfies the recursive equation

$$\widehat{z}(k|k) = \widehat{z}(k|k-1) + P(k)\bar{H}(k)'M(k)^{-1}(y(k) - \bar{H}(k)\widehat{z}(k|k-1)), \tag{11}$$

$$\widehat{z}(k+1|k) = \bar{A}(k)\widehat{z}(k|k-1) + V(k)(y(k) - \bar{H}(k)\widehat{z}(k|k-1)), \tag{12}$$

$$\widehat{z}(0|-1) = \mu(0) \tag{13}$$

and the matrices $M(k) > 0$ and $V(k)$ are given by

$$M(k) = \bar{H}(k)P(k)\bar{H}(k)' + \sum_{\ell=1}^{\varepsilon^y} \widetilde{H}_\ell(k)Q(k)\widetilde{H}_\ell(k)' + G(k)G(k)', \tag{14}$$

$$V(k) = \left(\bar{A}(k)P(k)\bar{H}(k)' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y} \rho_{s,\ell}\widetilde{A}_s(k)Q(k)\widetilde{H}_\ell(k)' + C(k)G(k)'\right)M(k)^{-1}. \tag{15}$$

The matrices $P(k) = E(\widetilde{z}(k|k-1)\widetilde{z}(k|k-1)') \geq 0$ satisfy the Riccati-like recurrent equation

$$\begin{aligned} P(k+1) = {} & \bar{A}(k)P(k)\bar{A}(k)' + \mathcal{B}(k, \mathbf{Q}(k)) + \mathrm{Dg}(\mathbf{D}(k)) \\ & - \left(\bar{A}(k)P(k)\bar{H}(k)' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y} \rho_{s,\ell}\widetilde{A}_s(k)Q(k)\widetilde{H}_\ell(k)' + C(k)G(k)'\right)\left(\bar{H}(k)P(k)\bar{H}(k)' + \sum_{\ell=1}^{\varepsilon^y} \widetilde{H}_\ell(k)Q(k)\widetilde{H}_\ell(k)' + G(k)G(k)'\right)^{-1} \\ & \times \left(\bar{A}(k)P(k)\bar{H}(k)' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y} \rho_{s,\ell}\widetilde{A}_s(k)Q(k)\widetilde{H}_\ell(k)' + C(k)G(k)'\right)'. \end{aligned} \tag{16}$$

Proof. Setting $\mu(k) = E(z(k))$ it follows that $\mu(k+1) = \bar{A}(k)\mu(k)$. Define

$$\varphi_{ij}(k) = 1_{\{\theta(k+1)=j\}} - p_{ij},$$

$$\ell(k) = \sum_{s=1}^{\varepsilon^x}\left(\sum_{i=1}^{N} \widetilde{A}_{i,s}z(k, i)\right) w_s^x(k),$$

$$c(k) = \sum_{i=1}^{N} C_i(k)w(k)1_{\{\theta(k)=i\}},$$

and

$$\Upsilon^1(k) = \begin{pmatrix} \sum_{i=1}^{N} \bar{A}_i z_i(k)\varphi_{i1}(k) \\ \vdots \\ \sum_{i=1}^{N} \bar{A}_i z_i(k)\varphi_{iN}(k) \end{pmatrix}, \tag{17}$$

$$\Upsilon^2(k) = \begin{pmatrix} \ell(k)1_{\{\theta(k+1)=1\}} \\ \vdots \\ \ell(k)1_{\{\theta(k+1)=N\}} \end{pmatrix}, \tag{18}$$

$$\Upsilon^3(k) = \begin{pmatrix} c(k)1_{\{\theta(k+1)=1\}} \\ \vdots \\ c(k)1_{\{\theta(k+1)=N\}} \end{pmatrix}. \tag{19}$$

From (4) we have that

$$z^c(k+1) = \bar{A}(k)z^c(k) + \Upsilon^1(k) + \Upsilon^2(k) + \Upsilon^3(k), \tag{20}$$

$$y^c(k) = \bar{H}(k)z^c(k) + \sum_{s=1}^{\varepsilon^y} \widetilde{H}_s(k)\xi_s^y(k) + G_{\theta(k)}(k)w(k), \tag{21}$$

where $\xi_s^y(k) = w_s^y(k)z(k)$. Let us denote by $\mathfrak{L}((y^c)^t)$ the linear subspace spanned by the random vector

$$(y^c)^t = \begin{pmatrix} y^c(0) \\ \vdots \\ y^c(t) \end{pmatrix},$$

and $\mathcal{P}^t$ the orthogonal projection onto $\mathfrak{L}((y^c)^t)$. In what follows we apply induction on k to show that from assumptions (A1) and (A2), $\mathrm{cov}((y^c)^k) > 0$. From (21) with $k = 0$ it is easy to see that $\mathrm{cov}(y^c(0)) = \bar{H}(0)P(0)\bar{H}(0)' + \sum_{s=1}^{\varepsilon^y} \widetilde{H}_s(0)Q(0)\widetilde{H}_s(0)' + G(0)G(0)' > 0$ from assumption (A1). Suppose that $\mathrm{cov}((y^c)^{k-1}) > 0$. From (2), for any null mean random vector z,

$$\mathcal{P}^{k-1}(z) = E(z(y^c)^{k-1\prime})\mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1}. \tag{22}$$

It follows from (22) and the independence hypothesis made in Section 3 that

$$\begin{aligned} \mathcal{P}^{k-1}(\xi_s^y(k)) &= E((w_s^y(k)z(k))(y^c)^{k-1\prime})\mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1} \\ &= E(w_s^y(k))E(z(k)(y^c)^{k-1\prime})\mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1} \\ &= 0. \end{aligned} \tag{23}$$

Similarly,

$$\begin{aligned} & \mathcal{P}^{k-1}(G_{\theta(k)}(k)w(k)) \\ & = \sum_{i=1}^{N} E(1_{\{\theta(k)=i\}}G_i(k)w(k)(y^c)^{k-1\prime})\mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1} \\ & = \sum_{i=1}^{N} G_i(k)E(w(k))E(1_{\{\theta(k)=i\}}(y^c)^{k-1\prime})\mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1} \\ & = 0. \end{aligned} \tag{24}$$

From (21), (23) and (24) it is immediate to see that

$$\widehat{y}^c(k|k-1) = \mathcal{P}^{k-1}(y^c(k)) = \bar{H}(k)\widehat{z}^c(k|k-1). \tag{25}$$

Thus, recalling that $\widetilde{y}(k|k-1) = y^c(k) - \widehat{y}^c(k|k-1)$, and $\widetilde{z}(k|k-1) = z^c(k) - \widehat{z}^c(k|k-1)$, it follows from (21) and (25) that

$$\widetilde{y}(k|k-1) = \bar{H}(k)\widetilde{z}(k|k-1) + \sum_{s=1}^{\varepsilon^y} \widetilde{H}_s(k)\xi_s^y(k) + G_{\theta(k)}(k)w(k). \tag{26}$$

From (26) and setting $P(k) = E(\widetilde{z}(k|k-1)\widetilde{z}(k|k-1)')$, it follows from the independence hypothesis made in Section 3 that

$$\begin{aligned} & \mathrm{cov}(\widetilde{y}(k|k-1)) \\ & = \bar{H}(k)P(k)\bar{H}(k)' + \sum_{s=1}^{\varepsilon^y} \widetilde{H}_s(k)Q(k)\widetilde{H}_s(k)' + G(k)G(k)' \\ & = M(k) \geq \sum_{s=1}^{\varepsilon^y} \widetilde{H}_s(k)Q(k)\widetilde{H}_s(k)' + G(k)G(k)' > 0 \end{aligned} \tag{27}$$

from assumption (A2). From Proposition 2 in the Appendix we get that $\mathrm{cov}((y)^k) > 0$. Since, from the independence hypothesis made in Section 3, we have

$$E(z^c(k)\xi_s^y(k)') = E(w_s^y(k))E(z^c(k)z(k)') = 0, \tag{28}$$

$$\begin{aligned} &E(z^c(k)(G_{\theta(k)}(k)w(k))') \\ &\quad= \sum_{i=1}^N E(z^c(k)(G_i(k)w(k))'1_{\{\theta(k)=i\}}) \\ &\quad= \sum_{i=1}^N E(1_{\{\theta(k)=i\}}z^c(k))E(w(k))'G_i(k)' = 0 \end{aligned} \tag{29}$$

it follows from (26), (28), (29), $z^c(k) = \widehat{z}^c(k|k-1) + \widetilde{z}(k|k-1)$ and orthogonality between $\widehat{z}^c(k|k-1)$ and $\widetilde{z}(k|k-1)$, that

$$\begin{aligned} &E(z^c(k)\widetilde{y}(k|k-1)') \\ &\quad= E(z^c(k)\widetilde{z}(k|k-1)')\bar{H}(k)' \\ &\quad= E((\widehat{z}^c(k|k-1)+\widetilde{z}(k|k-1))\widetilde{z}(k|k-1)')\bar{H}(k)' \\ &\quad= E(\widehat{z}^c(k|k-1)\widetilde{z}(k|k-1)')\bar{H}(k)' \\ &\qquad+ E(\widetilde{z}(k|k-1)\widetilde{z}(k|k-1)')\bar{H}(k)' \\ &\quad= P(k)\bar{H}(k)'. \end{aligned} \tag{30}$$

From (3), (25), (27) and (30) we get that

$$\begin{aligned} \widehat{z}^c(k|k) = \widehat{z}^c(k|k-1) + P(k)\bar{H}(k)'M(k)^{-1}(y^c(k) \\ - \bar{H}(k)\widehat{z}^c(k|k-1)). \end{aligned} \tag{31}$$

From (1) and noticing that $\widetilde{y}(k|k-1) = y^c(k) - \bar{H}(k)\widehat{z}^c(k|k-1) = y(k) - \bar{H}(k)\widehat{z}(k|k-1)$ we obtain (11). Let us derive now (12). From (20) it follows that

$$\begin{aligned} \widehat{z}^c(k+1|k) = \bar{A}(k)\widehat{z}^c(k|k) + \mathcal{P}^k(\Upsilon^1(k)) + \mathcal{P}^k(\Upsilon^2(k)) \\ + \mathcal{P}^k(\Upsilon^3(k)). \end{aligned} \tag{32}$$

From (2) we have that

$$\mathcal{P}^k(\Upsilon^1(k)) = E(\Upsilon^1(k)(y^c)^{k\prime})\mathrm{cov}((y)^k)^{-1}(y^c)^k.$$

Let us denote by $\mathcal{F}_k$ the σ-field generated by the random variable and vectors $\theta(k)$, $z(k)$ and $(y^c)^k$. We have that

$$\begin{aligned} &\mathcal{P}^k\left(\sum_{i=1}^N \bar{A}_i z_i(k)\varphi_{ij}(k)\right) \\ &\quad= \sum_{i=1}^N E(\bar{A}_i(k)z_i(k)\varphi_{ij}(k)(y^c)^{k\prime})\mathrm{cov}((y)^k)^{-1}(y^c)^k \\ &\quad= \sum_{i=1}^N E(E(\bar{A}_i(k)z_i(k)\varphi_{ij}(k)(y^c)^{k\prime}|\mathcal{F}_k))\mathrm{cov}((y)^k)^{-1}(y^c)^k \\ &\quad= \sum_{i=1}^N E(\bar{A}_i(k)z_i(k)E(\varphi_{ij}(k)|\mathcal{F}_k)1_{\{\theta(k)=i\}}(y^c)^{k\prime}) \\ &\qquad\times \mathrm{cov}((y)^k)^{-1}(y^c)^k = 0 \end{aligned}$$

since

$$E(\varphi_{ij}(k)|\mathcal{F}_k)1_{\{\theta(k)=i\}} = (P(\theta(k+1)=j|\mathcal{F}_k) - p_{ij})1_{\{\theta(k)=i\}} = 0.$$

This shows that $\mathcal{P}^k(\Upsilon^1(k)) = 0$. From (3), we have for $\kappa = 2, 3$,

$$\begin{aligned} \mathcal{P}^k(\Upsilon^\kappa(k)) = \mathcal{P}^{k-1}(\Upsilon^\kappa(k)) \\ + E(\Upsilon^\kappa(k)\widetilde{y}(k|k-1)')M(k)^{-1}\widetilde{y}(k|k-1). \end{aligned}$$

From the independence hypothesis made in Section 3 and (2) we have that

$$\begin{aligned} &\mathcal{P}^{k-1}(\ell(k)1_{\{\theta(k+1)=j\}}) \\ &\quad= \sum_{s=1}^{\varepsilon^x}\left(\sum_{i=1}^N \widetilde{A}_{i,s}(k)E(z(k,i)w_s^x(k)(y^c)^{k-1\prime}1_{\{\theta(k+1)=j\}})\right) \\ &\qquad\times \mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1} \\ &\quad= \sum_{s=1}^{\varepsilon^x}\left(\sum_{i=1}^N \widetilde{A}_{i,s}(k)E(w_s^x(k))E(z(k,i)(y^c)^{k-1\prime}1_{\{\theta(k+1)=j\}})\right) \\ &\qquad\times \mathrm{cov}((y)^{k-1})^{-1}(y^c)^{k-1} = 0 \end{aligned}$$

and from (26), that

$$\begin{aligned} &E(1_{\{\theta(k+1)=j\}}z(k,i)w_s^x(k)\widetilde{y}(k|k-1)') \\ &\quad= E\Bigg(1_{\{\theta(k+1)=j\}}z(k,i)w_s^x(k)\Bigg(\bar{H}(k)\widetilde{z}(k|k-1) \\ &\qquad+ \sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell(k)w_\ell^y(k)z(k) + G_{\theta(k)}(k)w(k)\Bigg)'\Bigg) \\ &\quad= E(w_s^x(k))E\Big(1_{\{\theta(k+1)=j\}}z(k,i)\Big(\bar{H}(k)\widetilde{z}(k|k-1) \\ &\qquad+ G_{\theta(k)}(k)w(k)\Big)'\Big) \\ &\qquad+ \sum_{\ell=1}^{\varepsilon^y}E(w_s^x(k)w_\ell^y(k))E\Big(1_{\{\theta(k+1)=j\}}z(k,i)z(k)'\widetilde{H}_\ell(k)'\Big) \\ &\quad= \sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}Q_i(k)\widetilde{H}_{i,\ell}(k)'\pi_i(k)p_{ij} \end{aligned}$$

so that

$$\begin{aligned} &E(\ell(k)1_{\{\theta(k+1)=j\}}\widetilde{y}(k|k-1)') \\ &\quad= \sum_{s=1}^{\varepsilon^x}\left(\sum_{i=1}^N \widetilde{A}_{i,s}(k)E(1_{\{\theta(k+1)=j\}}z(k,i)w_s^x(k)\widetilde{y}(k|k-1)')\right) \\ &\quad= \sum_{i=1}^N \pi_i(k)p_{ij}\sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_{i,s}(k)Q_i(k)\widetilde{H}_{i,\ell}(k)', \end{aligned}$$

and thus

$$\mathcal{P}^k(\Upsilon^2(k)) = \left(\sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_s(k)Q(k)\widetilde{H}_\ell(k)'\right)M(k)^{-1}\widetilde{y}(k|k-1).$$

Similar reasoning shows that

$$\mathcal{P}^k(\Upsilon^3(k)) = C(k)G(k)'M(k)^{-1}\widetilde{y}(k|k-1).$$

From (31), (32) and the above results we get that

$$\widehat{z}^c(k+1|k) = \bar{A}(k)\widehat{z}^c(k|k) + V(k)\widetilde{y}(k|k-1)$$

and (12) follows from (1) and the facts that $\mu(k+1) = \bar{A}(k)\mu(k)$ and $\widetilde{y}(k|k-1) = y(k) - \bar{H}(k)\widehat{z}(k|k-1)$. Finally, from (12) and (17)–(21), we have that

$$\begin{aligned} \widetilde{z}(k+1|k) = (\bar{A}(k) - V(k)\bar{H}(k))\widetilde{z}(k|k-1) + \Upsilon^1(k) \\ + \Upsilon^2(k) + \Upsilon^3(k) - V(k)\sum_{s=1}^{\varepsilon^y}\widetilde{H}_s(k)w_s^y(k)z(k) \\ - V(k)G_{\theta(k)}(k)w(k). \end{aligned} \tag{33}$$

Set

$$\Upsilon^0(k) = (\bar{A}(k) - V(k)\bar{H}(k))\widetilde{z}(k|k-1),$$

$$\Upsilon^4(k) = -V(k)\sum_{s=1}^{\varepsilon^y}\widetilde{H}_s(k)w_s^y(k)z(k),$$

$\Upsilon^5(k) = -V(k)G_{\theta(k)}(k)w(k),$

so that from (33), $\widetilde{z}(k+1|k) = \sum_{\kappa=0}^{5}\Upsilon^\kappa(k)$. From the independence hypothesis made in Section 3 we get that $E(\Upsilon^0(k)\Upsilon^\kappa(k)') = 0$, for $\kappa = 1,2,3,4,5$, $E(\Upsilon^1(k)\Upsilon^\kappa(k)') = 0$, for $\kappa = 2,3,4,5$, $E(\Upsilon^2(k)\Upsilon^\kappa(k)') = 0$, for $\kappa = 3,5$, $E(\Upsilon^3(k)\Upsilon^4(k)') = 0$ and $E(\Upsilon^4(k)\Upsilon^5(k)') = 0$. Moreover we have that

$$E(\Upsilon^2(k)\Upsilon^4(k)') = -\sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_s(k)Q(k)\widetilde{H}_\ell(k)'V(k)'),$$

$$E(\Upsilon^3(k)\Upsilon^5(k)') = -C(k)G(k)'V(k)',$$

$$E(\Upsilon^0(k)\Upsilon^0(k)') = (\bar{A}(k) - V(k)\bar{H}(k))P(k)(\bar{A}(k) - V(k)\bar{H}(k))',$$

$$\begin{aligned} E(\Upsilon^1(k)\Upsilon^1(k)') &= \operatorname{diag}\left[\sum_{i=1}^{N} p_{ij}\bar{A}_i(k)Q_i(k)\bar{A}_i(k)'\right] \\ &\quad -\bar{A}(k)\operatorname{diag}(Q_i(k))\bar{A}(k)' = \mathscr{B}_1(k,\mathbf{Q}(k)), \end{aligned}$$

$$\begin{aligned} E(\Upsilon^2(k)\Upsilon^2(k)') &= \operatorname{diag}\left[\sum_{i=1}^{N} p_{ij}\sum_{s=1}^{\varepsilon^x}\widetilde{A}_{i,s}(k)Q_i(k)\widetilde{A}_{i,s}(k)'\right] \\ &= \mathscr{B}_2(k,\mathbf{Q}(k)), \end{aligned}$$

$$\begin{aligned} E(\Upsilon^3(k)\Upsilon^3(k)') &= \operatorname{diag}\left[\sum_{i=1}^{N} p_{ij}\pi_i(k)C_i(k)C_i(k)'\right] \\ &= \operatorname{Dg}(\mathbf{D}(k)), \end{aligned}$$

$$E(\Upsilon^4(k)\Upsilon^4(k)') = V(k)\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell(k)Q(k)\widetilde{H}_\ell(k)'\right)V(k)',$$

$$E(\Upsilon^5(k)\Upsilon^5(k)') = V(k)G(k)G(k)'V(k)'.$$

Putting all these results together we get that

$$\begin{aligned} P(k+1) &= (\bar{A}(k) - V(k)\bar{H}(k))P(k)(\bar{A}(k) - V(k)\bar{H}(k))' \\ &\quad + \mathscr{B}_1(k,\mathbf{Q}(k)) + \mathscr{B}_2(k,\mathbf{Q}(k)) + \operatorname{Dg}(\mathbf{D}(k)) \\ &\quad - C(k)G(k)'V(k)' - V(k)G(k)C(k)' \\ &\quad + V(k)G(k)G(k)'V(k)' \\ &\quad + V(k)\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell(k)Q(k)\widetilde{H}_\ell(k)'\right)V(k)' \\ &\quad - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_s(k)Q(k)\widetilde{H}_\ell(k)'V(k)' \\ &\quad + V(k)\widetilde{H}_\ell(k)Q(k)A_s(k)) \end{aligned} \tag{34}$$

and after some algebraic manipulations we obtain (16). □

Remark 3. Notice that for the case in which there are neither jumps ($N = 1$) nor multiplicative noises we have that $\mathscr{B}(k,\mathbf{Q}(k)) = 0$ and all the terms with $\widetilde{A}_{i,s}$ and $\widetilde{H}_{i,s}$ will disappear, so that (16) becomes the usual recursive Riccati equation (see Davis & Vinter, 1985, page 118) and the filter derived in Theorem 1 reduces to the standard Kalman filter (see, for instance, Davis & Vinter, 1985, Theorem 3.3.1).

Remark 4. It should be noticed that the filter described in (11)–(16) does not depend on the knowledge of the Markov chain $\theta(k)$. However it is necessary to know the initial distribution $\{\pi_i(0)\}$ of $\theta(0)$.

5. Stationary solution

In this section we consider that all the matrices in (4) are time invariant. In particular we can suppress the dependence on k of the operators $\mathscr{B}_1$, $\mathscr{B}_2$, $\mathscr{B}$ and $\mathscr{T}$ defined respectively in (5)–(8). The goal of this section is to show that, under the hypothesis of mean square stability of the system (4) (see Definition 1), there exists a stationary solution $\mathbf{Q} \geq 0$ and $P \geq 0$ for the Lyapunov and Riccati-like equations related to (9) and (16) respectively and, moreover $\mathbf{Q}(k)$ and $P(k)$ given by (9) and (16) respectively converge to $\mathbf{Q}$ and P whatever the initial condition $\mathbf{Q}(0) \geq 0$ and $P(0) \geq 0$ are. From now on we shall make the following assumption.

(A3) The Markov chain $\{\theta(k)\}$ is ergodic.

From hypothesis (A3) it follows that there exist $\pi_i > 0$, $i = 1,\ldots,N$, $\sum_{i=1}^{N}\pi_i = 1$ such that $\lim_{k\to\infty}\pi_i(k) = \pi_i$ exponentially fast and independent from $\theta(0)$. We set $\pi(k) = (\pi_1(k),\ldots,\pi_N(k))$ and $\pi = (\pi_1,\ldots,\pi_N)$. For $\alpha = (\alpha_1,\ldots,\alpha_N)$, $\alpha_i \geq 0$, we define $\mathbf{D}(\alpha) := (D_1(\alpha),\ldots,D_N(\alpha)) \in \mathbb{H}^{n+}$ as $D_j(\alpha) := \sum_{i=1}^{N} p_{ij}\alpha_i C_i C_i'$, and

$$G(\alpha) := \left(G_1\alpha_1^{1/2}\cdots G_N\alpha_N^{1/2}\right),$$

$$C(\alpha) := \begin{pmatrix} \alpha_1^{1/2}p_{11}C_1 & \cdots & \alpha_N^{1/2}p_{N1}C_N \\ \vdots & \ddots & \vdots \\ \alpha_1^{1/2}p_{1N}C_1 & \cdots & \alpha_N^{1/2}p_{NN}C_N \end{pmatrix},$$

and $\mathbf{D} := \mathbf{D}(\pi)$, $G := G(\pi)$, $C := C(\pi)$.

We have the following definition for mean square stability of (4).

Definition 1. We say that system (4) is mean square stable (MSS) if, with $w(k) = 0$, we have that $E(\|x(k)\|^2) \to 0$ as $k \to \infty$ whatever the initial condition x_0 satisfying $E(\|x_0\|^2) < \infty$ is.

The next result follows from Propositions 2.5, 2.6, and 2.9 in Costa et al. (2005) (see also Dragan & Morozan, 2006a).

Proposition 1. *The following affirmatives are equivalent:*

(a) *System* (4) *is MSS.*
(b) $r_\sigma(\mathscr{T}) < 1$.
(c) *There exists* $\mathbf{W} > 0$ *such that* $\mathbf{W} - \mathscr{T}(\mathbf{W}) > 0$.

Moreover if the above holds then there exists a unique solution $\mathbf{Q}$ *satisfying*

$$\mathbf{Q} = \mathscr{T}(\mathbf{Q}) + \mathbf{D} \tag{35}$$

and $\mathbf{Q}(k) \to \mathbf{Q}$ *as* $\to \infty$ *exponentially fast, where* $\mathbf{Q}(k)$ *satisfies the recursive equation* (9).

It follows that if system (4) is MSS then from Proposition 1(c) we have that for each $j = 1,\ldots,N$,

$$\begin{aligned} 0 &< W_j - \sum_{i=1}^{N} p_{ij}\left(\bar{A}_i W_i \bar{A}_i' + \sum_{s=0}^{\varepsilon^y}\widetilde{A}_{i,s}W_i\widetilde{A}_{i,s}'\right) \\ &\leq W_j - \sum_{i=1}^{N} p_{ij}\bar{A}_i W_i \bar{A}_i' \end{aligned} \tag{36}$$

and from Theorem 3.9 and Proposition 3.6. in Costa et al. (2005) we have that (36) implies that $r_\sigma(\bar{A}) < 1$. In what follows we set for matrices $K, Z \geq 0$, of appropriate dimensions and $\mathbf{U} \in \mathbb{H}^{n+}$,

$$M(Z,\mathbf{U},\alpha) := \bar{H}Z\bar{H}' + \sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell\operatorname{Dg}(\mathbf{U})\widetilde{H}_\ell' + G(\alpha)G(\alpha)', \tag{37}$$

$$\begin{aligned} V(Z,\mathbf{U},\alpha) &:= \Bigg(\bar{A}Z\bar{H}' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_s\operatorname{Dg}(\mathbf{U})\widetilde{H}_\ell' \\ &\quad + C(\alpha)G(\alpha)'\Bigg)M(Z,\mathbf{U},\alpha)^{-1}, \end{aligned} \tag{38}$$

$$\Phi(K,\alpha) := \mathrm{Dg}(\mathbf{D}(\alpha)) - C(\alpha)G(\alpha)'K' - KG(\alpha)C(\alpha)' + KG(\alpha)G(\alpha)'K', \tag{39}$$

$$\Psi(K,\mathbf{U}) := \mathcal{B}_2(\mathbf{U}) + K\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell \mathrm{Dg}(\mathbf{U})\widetilde{H}_\ell'\right)K' - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_s\mathrm{Dg}(\mathbf{U})\widetilde{H}_\ell'K' + K\widetilde{H}_\ell \mathrm{Dg}(\mathbf{U})A_s') \tag{40}$$

and $M(Z,\mathbf{U}) := M(Z,\mathbf{U},\pi)$, $V(Z,\mathbf{U}) := V(Z,\mathbf{U},\pi)$, $\Phi(K) := \Phi(K,\pi)$.

Let κ be such that $\inf_{\ell\geq\kappa}\pi_i(\ell) > 0$ for all $i = 1,\ldots,N$ (since $\pi_i(k) \overset{k\to\infty}{\to} \pi_i > 0$ we have that this number exists). Define $\alpha_i(k) = \inf_{\ell\geq k}\pi_i(\ell+\kappa)$. Obviously for $k = 1,2,\ldots,\ i = 1,\ldots,N$,

$$\pi_i(k+\kappa) \geq \alpha_i(k) \geq \alpha_i(k-1) > 0, \tag{41}$$

and $\alpha_i(k) \overset{k\to\infty}{\to} \pi_i$ exponentially fast. Set $\alpha(k) = (\alpha_1(k),\ldots,\alpha_N(k))$, and the sequence $\mathbf{Q}_0(k) \in \mathbb{H}^{n+}$ as:

$$\mathbf{Q}_0(k+1) = \mathcal{T}(\mathbf{Q}_0(k)) + \mathbf{D}(\alpha(k)), \qquad \mathbf{Q}_0(0) = 0. \tag{42}$$

From Proposition 4 in the Appendix, $\mathbf{Q}_0(k) \overset{k\to\infty}{\to} \mathbf{Q}$ and for each $k = 0,1,2,\ldots$, $\mathbf{Q}(k+\kappa) \geq \mathbf{Q}_0(k) \geq \mathbf{Q}_0(k-1)$. Define $Q_0(k) = \mathrm{Dg}(\mathbf{Q}_0(k))$, so that $Q_0(k) \leq Q_0(k+1) \leq Q$. We need the following assumption:

(A4) $G(\alpha(0))G(\alpha(0))' > 0$.

As before, if for each $i = 1,\ldots,N$ we have $G_iG_i' > 0$ then clearly (A4) is satisfied. From (41) it is easy to see that

$$G(\alpha(0))G(\alpha(0))' = \sum_{i=1}^N \alpha_i(0)G_iG_i' \leq \sum_{i=1}^N \alpha_i(k)G_iG_i' = G(\alpha(k))G(\alpha(k))' \leq \sum_{i=1}^N \pi_iG_iG_i' = GG'$$

and since $0 \leq Q_0(k) \leq Q$ we get from (A4) that

$$0 < \sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q_0(k)\widetilde{H}_\ell' + G(\alpha(k))G(\alpha(k))' \leq \sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q\widetilde{H}_\ell' + GG'. \tag{43}$$

From (43) it follows that for any $Z \geq 0$ and $k = 0,1,\ldots$, $M(Z,\mathbf{Q}_0(k),\alpha(k)) > 0$ and $M(Z,\mathbf{Q}) > 0$. We have the following theorem.

Theorem 2. *Suppose that system (4) is MSS. Consider the algebraic Riccati-like equation in Z,*

$$Z = \bar{A}Z\bar{A}' + \mathcal{B}(\mathbf{Q}) + \mathrm{Dg}(\mathbf{D}) - \left(\bar{A}Z\bar{H}' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_sQ\widetilde{H}_\ell' + CG'\right) \times\left(\bar{H}Z\bar{H}' + \sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q\widetilde{H}_\ell' + GG'\right)^{-1} \times\left(\bar{A}Z\bar{H}' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_sQ\widetilde{H}_\ell' + CG'\right)', \tag{44}$$

where $\mathbf{Q}$ is the unique solution of (35) and $Q = \mathrm{Dg}(\mathbf{Q})$. Then there exists a unique positive semi-definite solution P to (44). Moreover $r_\sigma(\bar{A} - V(P,\mathbf{Q})\bar{H}) < 1$ and for any $\mathbf{Q}(0) \geq 0$, $P(0) \geq 0$ we have that $\mathbf{Q}(k)$ and $P(k)$ given by (9) and (16) satisfy $\mathbf{Q}(k) \to \mathbf{Q}$ and $P(k) \to P$ as $k \to \infty$.

Proof. According to Proposition 3 in the Appendix, we can find matrices $\mathcal{C}$ and $\mathcal{G}$ such that (44) can be re-written as

$$Z = \bar{A}Z\bar{A}' + \mathcal{C}\mathcal{C}' - \left(\bar{A}Z\bar{H}' + \mathcal{C}\mathcal{G}'\right)\left(\bar{H}Z\bar{H}' + \mathcal{G}\mathcal{G}'\right)^{-1} \times\left(\bar{A}Z\bar{H}' + \mathcal{C}\mathcal{G}'\right)'. \tag{45}$$

Since $r_\sigma(\bar{A}) < 1$ we have from Theorem 3.3.3 in Davis and Vinter (1985) that there exists a unique positive semi-definite solution P to (45) and moreover the matrix

$$\bar{A} - (\bar{A}P\bar{H} + \mathcal{C}\mathcal{G}')(\bar{H}P\bar{H}' + \mathcal{G}\mathcal{G}')^{-1}\bar{H} = \bar{A} - V(P,\mathbf{Q})\bar{H}$$

is stable, showing the first part of the theorem. Next we want to show the convergence $P(k) \to P$. Define $V = V(P,\mathbf{Q})$, $M(k) = V(P(k),\mathbf{Q}(k))$, $V(k) = V(P(k),\mathbf{Q}(k))$, and the linear recurrence sequence

$$J(k+1) = (\bar{A} - V\bar{H})J(k)(\bar{A} - V\bar{H})' + \mathcal{B}(\mathbf{Q}(k)) + \Phi(V) + V\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q(k)\widetilde{H}_\ell'\right)V' - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_sQ(k)\widetilde{H}_\ell'V' + V\widetilde{H}_\ell Q(k)\widetilde{A}_s') \tag{46}$$

with $J(0) = P_0$. Let us show by induction that $J(k) \geq P(k)$ for all k. For $k = 0$ the result follows by definition. Suppose it holds for k. After some algebraic manipulation in (16), we get that

$$P(k+1) = (\bar{A} - V\bar{H})P(k)(\bar{A} - V\bar{H})' + \mathcal{B}(\mathbf{Q}(k)) + \Phi(V) + V\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q(k)\widetilde{H}_\ell'\right)V' - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_sQ(k)\widetilde{H}_\ell'V' + V\widetilde{H}_\ell Q(k)\widetilde{A}_s') - (V(k) - V)M(k)(V(k) - V)' \tag{47}$$

and subtracting (46) from (47) we get that

$$(J(k+1) - P(k+1)) = (\bar{A} - V\bar{H})(J(k) - P(k))(\bar{A} - V\bar{H})' + (V(k) - V)M(k)(V(k) - V)' \geq 0$$

showing the desired result. From the fact that $r_\sigma(\mathcal{T}) < 1$ and $r_\sigma(\bar{A} - V\bar{H}) < 1$ it follows that there exists a unique solution (see Proposition 2.6 in Costa et al. (2005)) $\mathbf{Q}, J$ satisfying (35) and

$$J = (\bar{A} - V\bar{H})J(\bar{A} - V\bar{H})' + \mathcal{B}(\mathbf{Q}) + \Phi(V) + V\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q\widetilde{H}_\ell'\right)V' - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_sQ\widetilde{H}_\ell'V' + V\widetilde{H}_\ell Q\widetilde{A}_s'), \tag{48}$$

and (see Propostion 2.9 in Costa et al. (2005)) $J(k) \to J$ as $k \to \infty$. But after some algebraic manipulations on (44) we get that P also satisfies (48) so that $J = P$, and thus we have shown that $P(k) \leq J(k)$ and $J(k) \to P$ as $k \to \infty$.

Define the sequence $P_0(k)$ as follows:

$$P_0(k+1) = \bar{A}P_0(k)\bar{A}' + \mathcal{B}(\mathbf{Q}_0(k)) + \mathrm{Dg}(\mathbf{D}(\alpha(k))) - \left(\bar{A}P_0(k)\bar{H}' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_sQ_0(k)\widetilde{H}_\ell' + C(\alpha(k))G(\alpha(k))'\right) \times\left(\bar{H}P_0(k)\bar{H}' + \sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q_0(k)\widetilde{H}_\ell' + G(\alpha(k))G(\alpha(k))'\right)^{-1}$$

$$\times\left(\bar{A}P_0(k)\bar{H}' + \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}\widetilde{A}_s Q_0(k)\widetilde{H}'_\ell + C(\alpha(k))G(\alpha(k))'\right)', \qquad P_0(0) = 0. \tag{49}$$

Suppose for the moment that $P_0(k) \le P_0(k+1) \le P(k+1+\kappa)$. Then, from the fact that $P(k) \le J(k)$ and $J(k) \to P$ we have from the monotone convergence lemma for positive operators (see, for instance, Lemma 3.1 in Wonham (1968)) that $P_0(k) \uparrow P_\infty$ for some $P_\infty \ge 0$. Since $\mathbf{Q}_0(k) \to \mathbf{Q}$ as $k \to \infty$ we have, taking the limit as $k \to \infty$ in (49) that $P_\infty \ge 0$ is a solution of (44) and, from uniqueness, $P_\infty = P$. Combining the results we have that $P_0(k) \le P(k+\kappa) \le J(k+\kappa)$ and $P_0(k) \to P$, $J(k) \to P$ and thus $P(k) \to P$ as $k \to \infty$, showing the final part of the theorem. It remains to show that $P_0(k) \le P_0(k+1) \le P(k+1+\kappa)$. For this we set $M_0(k) = M(P_0(k), \mathbf{Q}_0(k), \alpha(k))$ and $V_0(k) = V(P_0(k), \mathbf{Q}_0(k), \alpha(k))$. Let us show first by induction that $P_0(k+1) \ge P_0(k)$. For $k = 0$ the result is immediate since $P_0(1) \ge 0 = P_0(0)$. Suppose it holds for $k-1$. After rearranging (49) we get that

$$\begin{aligned} P_0(k+1) = {} & (\bar{A} - V_0(k)\bar{H})P_0(k)(\bar{A} - V_0(k)\bar{H})' \\ & + \mathcal{B}_1(\mathbf{Q}_0(k)) + \Phi(V_0(k), \alpha(k)) + \mathcal{B}_2(\mathbf{Q}_0(k)) \\ & + V_0(k)\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q_0(k)\widetilde{H}'_\ell\right)V_0(k)' \\ & - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_s Q_0(k)\widetilde{H}'_\ell V_0(k)' \\ & + V_0(k)\widetilde{H}_\ell Q_0(k)A'_s) \end{aligned} \tag{50}$$

and

$$\begin{aligned} P_0(k) = {} & (\bar{A} - V_0(k)\bar{H})P_0(k-1)(\bar{A} - V_0(k)\bar{H})' \\ & + \mathcal{B}_1(\mathbf{Q}_0(k-1)) + \Phi(V_0(k), \alpha(k-1)) \\ & + \mathcal{B}_2(\mathbf{Q}_0(k-1)) + V_0(k)\left(\sum_{\ell=1}^{\varepsilon^y}\widetilde{H}_\ell Q_0(k-1)\widetilde{H}'_\ell\right)V_0(k)' \\ & - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y}\rho_{s,\ell}(\widetilde{A}_s Q_0(k-1)\widetilde{H}'_\ell V_0(k)' \\ & + V_0(k)\widetilde{H}_\ell Q_0(k-1)A'_s) \\ & - (V_0(k) - V_0(k-1))M_0(k-1)(V_0(k) - V_0(k-1))'. \end{aligned} \tag{51}$$

Recalling the definition of Ψ in (40) and subtracting (50) from (51) we get that

$$\begin{aligned} (P_0(k+1) - P_0(k)) = {} & (\bar{A} - V_0(k)\bar{H})(P_0(k) - P_0(k-1)) \\ & \times (\bar{A} - V_0(k)\bar{H})' + \mathcal{B}_1(\mathbf{Q}_0(k) - \mathbf{Q}_0(k-1)) + \Phi(V_0(k), \alpha(k)) \\ & - \Phi(V_0(k), \alpha(k-1)) + \Psi(V_0(k), \mathbf{Q}_0(k) - \mathbf{Q}_0(k-1)) \\ & + (V_0(k) - V_0(k-1))M_0(k-1)(V_0(k) - V_0(k-1))'. \end{aligned} \tag{52}$$

From the induction hypothesis, $P_0(k+1) - P_0(k) \ge 0$, and thus $(\bar{A} - V_0(k)\bar{H})(P_0(k) - P_0(k-1))(\bar{A}(k) - V_0(k)\bar{H})' \ge 0$. From Proposition 4, $\mathbf{Q}_0(k) - \mathbf{Q}_0(k-1) \ge 0$, and thus from Remark 2, $\mathcal{B}_1(\mathbf{Q}_0(k) - \mathbf{Q}_0(k-1)) \ge 0$. From Proposition 5 and (41), $\Phi(V_0(k), \alpha(k)) - \Phi(V_0(k), \alpha(k-1)) \ge 0$ and from Proposition 6, $\Psi(V_0(k), \mathbf{Q}(k) - \mathbf{Q}(k-1)) \ge 0$, so that from (52) we get that $P_0(k+1) \ge P_0(k)$. Let us show now by induction that $P_0(k) \le P(k+\kappa)$ for all $k \ge 0$. Again for $k = 0$ the result is immediate since $P_0(0) = 0 \le P(\kappa)$. Suppose it holds for k. Clearly from Proposition 4, $\mathbf{Q}(k+\kappa) - \mathbf{Q}_0(k) \ge 0$ and similarly as in (52) we have that

$$\begin{aligned} & (P(k+\kappa+1) - P_0(k+1)) \\ & = (\bar{A} - V(k+\kappa)\bar{H})(P(k+\kappa) - P_0(k))(\bar{A} - V(k+\kappa)\bar{H})' \\ & \quad + \mathcal{B}_1(\mathbf{Q}(k+\kappa) - \mathbf{Q}_0(k)) + \Phi(V(k+\kappa), \pi(k+\kappa)) \\ & \quad - \Phi(V(k+\kappa), \alpha(k)) + \Psi(V(k+\kappa), \mathbf{Q}(k+\kappa) - \mathbf{Q}_0(k)) \\ & \quad + (V(k+\kappa) - V_0(k))M_0(k)(V(k+\kappa) - V_0(k))'. \end{aligned} \tag{53}$$

From Proposition 5 and (41), $\Phi(V(k+\kappa), \pi(k+\kappa)) - \Phi(V(k+\kappa), \alpha(k)) \ge 0$. From Proposition 6, $\Psi(V(k+\kappa), \mathbf{Q}(k+\kappa) - \mathbf{Q}_0(k)) \ge 0$, so that (53) and the induction hypothesis yields that $P(k+\kappa+1) \ge P_0(k+1)$, completing the proof. $\square$

We conclude this section showing how the algebraic Lyapunov and Riccati-like Eqs. (35), (44) and the result proved in Theorem 2 could be used to design a stationary minimum mean square filter. The stationary minimum mean square filter problem is defined as follows. Find matrices $\widehat{A}, \widehat{B}, \widehat{L}$, with $r_\sigma(\widehat{A}) < 1$, such that minimizes $\lim_{k\to\infty} E(\|e(k)\|^2)$ where

$$\widehat{z}(k+1) = \widehat{A}\widehat{z}(k) + \widehat{B}y(k), \tag{54}$$

$$e(k) = x(k) - \widehat{L}\widehat{z}(k). \tag{55}$$

Let $\Xi(k) = E(e(k)e(k)')$. From $r_\sigma(\widehat{A}) < 1$ and $r_\sigma(\mathcal{T}) < 1$ it is easy to show that $\Xi(k) \to \Xi$ as $k \to \infty$ for some $\Xi \ge 0$ (see for instance Proposition 5 in Costa and Guerra (2002a)). From the optimality of the minimum mean square filter deduced in Section 3, we have that for all k, $P(k) \le \Xi(k)$, and taking the limit as $k \to \infty$ we have from Theorem 2 that $P = \lim_{k\to\infty} P(k) \le \lim_{k\to\infty} \Xi(k) = \Xi$, where P is the unique positive semi-definite solution to (44). As in the proof of Theorem 2, set $V = V(P, \mathbf{Q})$, where $\mathbf{Q}$ is the unique solution of (35). Consider now the following matrices for the filter (54), (55):

$$\widehat{A}_{\mathrm{op}} = \bar{A} - V\bar{H}, \qquad \widehat{B}_{\mathrm{op}} = V, \qquad \widehat{L}_{\mathrm{op}} = \begin{pmatrix} I & \cdots & I \end{pmatrix}$$

(recall from Theorem 2 that $r_\sigma(\bar{A} - V\bar{H}) < 1$), which yields the following filter equations

$$\widehat{z}_{\mathrm{op}}(k+1) = \widehat{A}_{\mathrm{op}}\widehat{z}_{\mathrm{op}}(k) + \widehat{B}_{\mathrm{op}}y(k),$$

$$e_{\mathrm{op}}(k) = x(k) - \widehat{L}_{\mathrm{op}}\widehat{z}_{\mathrm{op}}(k).$$

We have, as in the deduction of (33) and (34), that $E(e_{\mathrm{op}}(k)e_{\mathrm{op}}(k)') = J(k)$, where $J(k)$ satisfies the linear recursive equation (46). But as shown in the proof of Theorem 2, $\lim_{k\to\infty} J(k) = P$, and thus

$$\begin{aligned} \lim_{k\to\infty} E(\|e(k)\|^2) &= \lim_{k\to\infty} \mathrm{tr}(\Xi(k)) = \mathrm{tr}(\Xi) \ge \mathrm{tr}(P) \\ &= \lim_{k\to\infty} \mathrm{tr}(J(k)) = \lim_{k\to\infty} E(\|e_{\mathrm{op}}(k)\|^2) \end{aligned}$$

showing the optimality of the filter given by $\widehat{A}_{\mathrm{op}} = \bar{A} - V\bar{H}$ and $\widehat{B}_{\mathrm{op}} = V$.

6. Numerical example

In this section we present a numerical comparison between the LMMSE filter and a standard time-varying Kalman filter based on the expected values of the stochastic parameters of the system, which we will call averaged Kalman filter. Recall that to implement the LMMSE filter we have first to iterate the system of linear equation (9) with solution $\mathbf{Q}(k) = (Q_1(k), \ldots, Q_N(k))$. After that we plug $\mathbf{Q}(k)$ and $Q(k) = \mathrm{Dg}(\mathbf{Q}(k))$ into the Riccati difference equation (16) and obtain $P(k)$ and, from (15), $V(k)$. The LMMSE estimator $\hat{x}(k|k)$ is given by (10) where $\hat{z}(k|k)$ satisfies the recursive equations (11)–(13). So the filter is very easy to implement and all calculations can be performed off-line. For the numerical example we considered a scalar problem with $x(0)$ Gaussian with mean 10

Table 1
Parameters considered in the simulations.

$\bar{a}_1$	$\bar{a}_2$	$\tilde{a}_1$	$\tilde{a}_2$	c_1	c_2
0.8	0.6	0.1	0.2	0.1	5.0

Table 2
Parameters considered in the simulations.

$\bar{h}_1$	$\bar{h}_2$	$\tilde{h}_1$	$\tilde{h}_2$	g_1	g_2	ρ
1.0	1.0	0.2	0.3	5.0	5.0	0.5

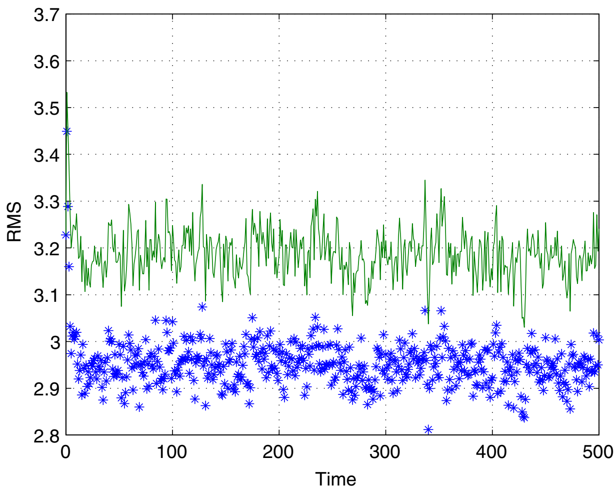


Fig. 1. Comparison between the averaged Kalman filter (solid line) and LMMSE (star line) filter.

and variance 10, $\theta(k) \in \{1, 2\}$, $\pi_1(0) = \pi_2(0) = 0.5$, $\varepsilon^x = \varepsilon^y = 1$ and $\bar{A}_i = \bar{a}_i$, $\tilde{A}_i = \tilde{a}_i$, $\bar{H}_i = \bar{h}_i$, $\tilde{H}_i = \tilde{h}_i$, $C_i = \begin{pmatrix} c_i & 0 \end{pmatrix}$, $G_i = \begin{pmatrix} 0 & g_i \end{pmatrix}$, $\rho_{1,1} = \rho$, as shown in Tables 1 and 2, and transition probability matrix given by $p_{11} = 0.8$, $p_{22} = 0.6$. All the noises were assumed to be Gaussian. For the averaged Kalman filter we considered the usual time varying Kalman filter with the expected value parameters $A(k) = \pi_1(k)\bar{A}_1 + \pi_2(k)\bar{A}_2$, $C(k) = \pi_1(k)C_1 + \pi_2(k)C_2$, $H(k) = \pi_1(k)\bar{H}_1 + \pi_2(k)\bar{H}_2$, $G(k) = \pi_1(k)G_1 + \pi_2(k)G_2$. We performed 5000 Monte Carlo simulations from $k = 0$ to 500, with the values of $\theta(k)$ generated randomly. Both filters were compared under the same conditions.

The results obtained are in Fig. 1, showing the square root of the mean square error (rms) through time. We can see from the simulations that the LMMSE had, as expected, a better performance than the averaged Kalman filter. It is easy to verify that for this example the system is MSS, and therefore from Theorem 2 we have the convergence of $P(k)$ to a stationary value P. Indeed we observe that for our problem $P = \begin{pmatrix} 7.5715 & -1.1005 \\ -1.1005 & 7.6593 \end{pmatrix}$. We believe that these simulations suggest that the LMMSE can be a good alternative in situations where computing power is at a premium. Besides being very simple to implement, all calculations for the filter can be performed off-line, so that the amount of on-line signal processing required is very modest.

7. Conclusions

In this paper we have obtained the linear minimum mean square filter for discrete-time linear systems subject to state and measurement multiplicative noises, and Markov jumps on the parameters. A Kalman-like filter is deduced in Theorem 1, based on Lyapunov and Riccati-like difference equations presented in (9) and (16). We have also shown in Theorem 2 that under the hypothesis that system (4) is MSS and ergodicity of the Markov chain there exists a unique positive semi-definite solution to the algebraic-like Riccati equation (44) and moreover the convergence of the error covariance matrix to this stationary value holds. Furthermore the filter error equation with the stationary gain will be stable. It is interesting to notice that due to the presence of the multiplicative noises the usual assumption $G_i(k)G_i(k) > 0$ and $G_iiG_i > 0$ can be replaced by the weaker conditions in assumptions (A1), (A2) and (A4). The filter proposed in this paper is very simple to implement, and all calculations can be performed offline, resulting in a discrete-time linear filter. Another advantage is that we can consider uncertainties in the parameters of the system through, for instance, an LMI approach. This formulation is being studied at the moment.

Acknowledgements

The authors are grateful to anonymous referees for their suggestions which have greatly improved the presentation of the paper.

Appendix

The following auxiliary result is useful in the proof of Theorem 1.

Proposition 2. *Suppose that* $\text{cov}(y^{t-1}) > 0$. *Then* $\text{cov}(y^t) > 0$ *if and only if* $\text{cov}(\tilde{y}(t|t-1)) > 0$.

Proof. For some matrix A we have that $\widehat{y}^c(t|t-1) = A(y^c)^{t-1}$ and $y^c(t) = \tilde{y}(t|t-1) + A(y^c)^{t-1}$. Thus from orthogonality between $(y^c)^{t-1}$ and $\tilde{y}(t|t-1)$, we have for any vectors a and b of appropriate dimensions,

$$\begin{pmatrix} a' & b' \end{pmatrix} \text{cov}(y^t) \begin{pmatrix} a \\ b \end{pmatrix} = (a + A'b)'\text{cov}(y^{t-1})(a + A'b) + b'\text{cov}(\tilde{y}(t|t-1))b$$

and the results follows. □

The following auxiliary results are needed to prove Theorem 2.

Proposition 3. *There are matrices* $\mathcal{C}$ *and* $\mathcal{G}$ *such that*

$$\begin{aligned} \mathcal{C}\mathcal{C}' &= \mathcal{B}(\mathbf{Q}) + \text{Dg}(\mathbf{D}) \\ &= \mathcal{B}_1(\mathbf{Q}) + \text{diag}\left(\sum_{i=1}^{N} p_{ij}\left(\sum_{s=1}^{\varepsilon^x} \tilde{A}_{i,s}Q_i\tilde{A}'_{i,s} + \pi_i C_i C'_i\right)\right), \end{aligned} \tag{56}$$

$$\mathcal{C}\mathcal{G}' = \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y} \rho_{s,\ell}\tilde{A}_s Q \tilde{H}'_\ell + CG', \tag{57}$$

$$\mathcal{G}\mathcal{G}' = \sum_{\ell=1}^{\varepsilon^y} \tilde{H}_\ell Q \tilde{H}'_\ell + GG'. \tag{58}$$

Proof. Set $\Pi_{11} = \text{diag}\left(\sum_{i=1}^{N} p_{ij}\left(\sum_{s=1}^{\varepsilon^x} \tilde{A}_{i,s}Q_i\tilde{A}'_{i,s} + \pi_i C_i C'_i\right)\right)$, $\Pi_{12} = \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y} \rho_{s,\ell}\tilde{A}_s Q\tilde{H}'_\ell + CG'$, $\Pi_{22} = \sum_{\ell=1}^{\varepsilon^y} \tilde{H}_\ell Q \tilde{H}'_\ell + GG'$, and

$$\hat{C} = \begin{pmatrix} \hat{C}_1 \\ \vdots \\ \hat{C}_N \end{pmatrix}$$

with $\hat{C}_j = \begin{pmatrix} C_{\theta(0)} & \frac{1}{\pi_{\theta(0)}^{1/2}} \sum_{s=1}^{\varepsilon^x} \widetilde{A}_{\theta(0),s} w_s^x(0) Q_{\theta(0)}^{1/2} \end{pmatrix} 1_{\{\theta(0)=j\}}$, and

$$\hat{G} = \begin{pmatrix} G_{\theta(0)} & \sum_{\ell=1}^{\varepsilon^y} \frac{1}{\pi_{\theta(0)}^{1/2}} \widetilde{H}_{\theta(0),\ell} w_\ell^y(0) Q_{\theta(0)}^{1/2} \end{pmatrix}.$$

Writing $E_\pi(.)$ the expected value operator when $\mathbf{P}(\theta(0) = i) = \pi_i$, we get that

$$0 \le E_\pi\left(\begin{pmatrix}\hat{C} \\ \hat{G}\end{pmatrix}\begin{pmatrix}\hat{C}' & \hat{G}'\end{pmatrix}\right) = \Pi := \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12}' & \Pi_{22}\end{pmatrix}.$$

We set $0 \le \Gamma = \Pi^{1/2}$, with $\Gamma := \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}' & \Gamma_{22}\end{pmatrix}$ and $\mathcal{C} := \begin{pmatrix}\Gamma_{11} & \Gamma_{12} & \mathcal{B}_1(\mathbf{Q})^{1/2}\end{pmatrix}$, $\mathcal{G} := \begin{pmatrix}\Gamma_{12}' & \Gamma_{22} & 0\end{pmatrix}$. Then (56)–(58) follows from $\Gamma^2 = \Pi$. □

Proposition 4. *Consider the sequence $\mathbf{Q}_0(k) \in \mathbb{H}^{n+}$ defined as in (42) and $\mathbf{Q} \in \mathbb{H}^{n+}$ the unique solution satisfying (35). Then $\mathbf{Q}_0(k) \overset{k\to\infty}{\to} \mathbf{Q}$ and for each $k = 1, 2, \dots$,*

$$\mathbf{Q}(k+\kappa) \ge \mathbf{Q}_0(k) \ge \mathbf{Q}_0(k-1) \ge 0 \tag{59}$$

where $\mathbf{Q}(k) \in \mathbb{H}^{n+}$ satisfies the recursive equation (9).

Proof. Since $r_\sigma(\mathcal{T}) < 1$ and $\alpha_i(k) \overset{k\to\infty}{\to} \pi_i$ exponentially fast we have from (42) and Proposition 2.9 in Costa et al. (2005) that

$$\begin{aligned}\lim_{k\to\infty} \mathbf{Q}_0(k) &= (\mathcal{I} - \mathcal{T})^{-1} \lim_{k\to\infty} \mathbf{D}(\alpha(k)) \\ &= (\mathcal{I} - \mathcal{T})^{-1}\mathbf{D}(\pi) = (\mathcal{I} - \mathcal{T})^{-1}\mathbf{D}\end{aligned}$$

where $\mathcal{I}$ is the identity operator in $\mathbb{B}(\mathbb{H}^n)$ and we recall that by definition $\mathbf{D} = \mathbf{D}(\pi)$. But from (35) it is clear that $\mathbf{Q} = (\mathcal{I} - \mathcal{T})^{-1}\mathbf{D}$ and thus we get that $\mathbf{Q}_0(k) \overset{k\to\infty}{\to} \mathbf{Q}$. Let us show now (59) by induction on k. For $k = 0$ the result is immediate, since from (9) and (42), $\mathbf{Q}(\kappa) \ge 0 = \mathbf{Q}_0(0)$ and $\mathbf{Q}_0(1) \ge 0 = \mathbf{Q}_0(0)$. Suppose that (59) holds for k. Then from (9), (41) and (59) we have that

$$\begin{aligned} &Q_j(k+1+\kappa) \\ &= \sum_{i=1}^N p_{ij}\left(\bar{A}_i Q_i(k+\kappa)\bar{A}_i' + \sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} Q_i \widetilde{A}_{i,s}' + \pi_i(k+\kappa) C_i C_i'\right) \\ &\ge \sum_{i=1}^N p_{ij}\left(\bar{A}_i Q_{0,i}(k)\bar{A}_i' + \sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} Q_{0,i} \widetilde{A}_{i,s}' + \alpha_i(k) C_i C_i'\right) \\ &= Q_{0,j}(k+1). \end{aligned}$$

Similarly,

$$\begin{aligned} &Q_{0,j}(k+1) \\ &= \sum_{i=1}^N p_{ij}\left(\bar{A}_i Q_{0,i}(k)\bar{A}_i' + \sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} Q_{0,i} \widetilde{A}_{i,s}' + \alpha_i(k) C_i C_i'\right) \\ &\ge \sum_{i=1}^N p_{ij}\left(\bar{A}_i Q_{0,i}(k-1)\bar{A}_i' + \sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} Q_{0,i}(k-1) \widetilde{A}_{i,s}' \right. \\ &\qquad \left. + \alpha_i(k-1) C_i C_i'\right) \\ &= Q_{0,j}(k) \end{aligned}$$

completing the induction argument in (59). □

For the next propositions recall the definitions of Φ and Ψ in (39) and (40) respectively.

Proposition 5. *Consider $\alpha = (\alpha_1, \dots, \alpha_N)$ and $\beta = (\beta_1, \dots, \beta_N)$ such that $\alpha_i \ge \beta_i \ge 0$ for $i = 1, \dots, N$. Then $\Phi(V, \alpha) \ge \Phi(V, \beta) \ge 0$.*

Proof. By definition $\Phi(V, \alpha)$ is a symmetric matrix. Consider an arbitrary vector x of appropriate dimension with

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \tag{60}$$

and set $y = V'x$. Writing $E_i(.)$ the expected value operator when $\mathbf{P}(\theta(0) = i) = 1$ we get after some algebraic manipulations that

$$\begin{aligned} &x'\Phi(V, \alpha)x \\ &= \sum_{i=1}^N \alpha_i\left(\sum_{j=1}^N p_{ij} x_j' C_i' C_i x_j - \left(\sum_{j=1}^N p_{ij} x_j\right)' C_i G_i' y \right. \\ &\qquad \left. - y' G_i C_i' \left(\sum_{j=1}^N p_{ij} x_j\right) + y' G_i G_i' y\right) \\ &= \sum_{i=1}^N \alpha_i E_i\left(\|C_i' x_{\theta(1)} - G_i' y\|^2\right) \\ &\ge \sum_{i=1}^N \beta_i E_i\left(\|C_i' x_{\theta(1)} - G_i' y\|^2\right) = x'\Phi(V, \beta)x \ge 0 \end{aligned}$$

showing that $\Phi(V, \alpha) \ge \Phi(V, \beta) \ge 0$. □

Proposition 6. *For any $\mathbf{U} \in \mathbb{H}^{n+}$ and V of appropriate dimension we have that $\Psi(V, \mathbf{U}) \ge 0$.*

Proof. Consider an arbitrary vector x as in (60) and $y = V'x$ and write again $E_i(.)$ as the expected value operator when $\mathbf{P}(\theta(0) = i) = 1$. The result follows after some algebraic manipulations yielding:

$$\begin{aligned} x'\Psi(V, \mathbf{U})x &= \sum_{i=1}^N \left(\sum_{j=1}^N p_{ij} x_j' \left(\sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} U_i \widetilde{A}_{i,s}'\right) x_j \right. \\ &\quad - \sum_{s=1}^{\varepsilon^x}\sum_{\ell=1}^{\varepsilon^y} \rho_{s,\ell}\left(\left(\sum_{j=1}^N p_{ij} x_j'\right)\widetilde{A}_{i,s} U_i \widetilde{H}_{i,\ell}' y \right. \\ &\quad \left.\left. + y'\widetilde{H}_{i,\ell} U_i \widetilde{A}_{i,s}' \left(\sum_{j=1}^N p_{ij} x_j\right)\right) + \sum_{\ell=1}^{\varepsilon^y} y' \widetilde{H}_{i,\ell} U_i \widetilde{H}_{i,\ell}' y\right) \\ &= \sum_{i=1}^N E_i\left(\left(\sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} w_s^x(0) x_{\theta(1)} - \sum_{\ell=1}^{\varepsilon^y} \widetilde{H}_{i,\ell} w_\ell^y(0) y\right) U_i \right. \\ &\quad \left. \times \left(\sum_{s=1}^{\varepsilon^x} \widetilde{A}_{i,s} w_s^x(0) x_{\theta(1)} - \sum_{\ell=1}^{\varepsilon^y} \widetilde{H}_{i,\ell} w_\ell^y(0) y\right)'\right) \ge 0. \quad \square \end{aligned}$$

References

Ackerson, G. A., & Fu, K. S. (1970). On the estate estimation in switching environments. *IEEE Transactions on Automatic Control*, *15*, 10–17.

Bar-Shalom, Y., & Li, X. R. (1993). *Estimation and tracking. principles, techniques, and software*. Artech House.

Blom, H. A. P., & Bar-Shalom, Y. (1988). The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, *33*(8), 780–783.

Carravetta, F., Germani, A., & Raimondi, M. (1997). Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise. *IEEE Transactions on Automatic Control*, *42*, 1106–1126.

Chang, C. G., & Athans, M. (1978). State estimation for discrete systems with switching parameters. *IEEE Transactions on Aerospace and Electronic Systems*, *14*, 418–424.

Chow, B. S., & Birkemeier, W. P. (1990). A new recursive filter for systems with multiplicative noise. *IEEE Transactions on Information Theory*, *36*, 1430–1435.

Costa, O. L. V. (1994). Linear minimum mean square error estimation for discrete-time Markovian jump linear systems. *IEEE Transactions on Automatic Control*, *39*, 1685–1689.

Costa, O. L. V., Fragoso, M. D., & Marques, R. P. (2005). *Discrete-time Markov jump linear systems*. Springer-Verlag.

Costa, O. L. V., & Guerra, S. (2002a). Robust linear filtering for discrete-time hybrid Markov linear systems. *International Journal of Control*, *75*, 712–727.

Costa, O. L. V., & Guerra, S. (2002b). Stationary filter for linear minimum mean square error estimator of discrete-time Markovian jump systems. *IEEE Transactions on Automatic Control*, *47*, 1351–1356.

Costa, O. L. V., & Kubrusly, C. S. (1996). State-feedback H_∞-control for discrete-time infinite-dimensional stochastic bilinear systems. *Journal of Mathematical Systems, Estimation, and Control*, *6*, 1–32.

Costa, O. L. V., & de Paulo, W. L. (2007). Indefinite quadratic with linear cost optimal control of Markovian jump with multiplicative noise systems. *Automatica*, *43*, 587–597.

Davis, M. H. A., & Vinter, R. B. (1985). *Stochastic modelling and control*. Chapman and Hall.

Dragan, V., & Morozan, T. (2002). Stability and robust stabilization to linear stochastic systems described by differential equations with Markovian jumping and multiplicative white noise. *Stochastic Analysis and Applications*, *20*, 33–92.

Dragan, V., & Morozan, T. (2006a). Mean square exponential stability for some stochastic linear discrete-time systems. *European Journal of Control*, *11*, 1–23.

Dragan, V., & Morozan, T. (2006b). Observability and detectability of a class of discrete-time stochastic linear systems. *IMA Journal of Mathematical Control and Information*, *23*, 371–394.

Dufour, F., & Elliott, R. J. (1997). Adaptive control of linear systems with Markov perturbations. *IEEE Transactions on Automatic Control*, *43*, 351–372.

Fioravanti, A. R., Gonçalves, A. P. C., & Geromel, J. C. (2008). H_2 filtering of discrete-time Markov jump linear systems through linear matrix inequalities. *International Journal of Control*, *81*, 1221–1231.

Geromel, J. C., Gonçalves, A. P., & Fioravanti, A. R. (2009). Dynamic output feedback control of discrete-time Markov jump linear systems through linear matrix inequalities. *SIAM Journal on Control and Optimization*, *48*, 573–593.

Gershon, E., & Shaked, U. (2006). Static H_2 and H_∞ output-feedback of discrete-time LTI systems with state multiplicative noise. *Systems & Control Letters*, *55*, 232–239.

Gershon, E., Shaked, U., & Yaesh, I. (2001). H_∞ control and filltering of discrete-time stochastic systems with multiplicative noise. *Automatica*, *37*, 409–417.

Stoica, A. M., Dragan, V., & Yaesh, I. (2009). Kalman-type filtering for stochastic systems with state-dependent noise and Markovian jumps. In *Proceedings of the 15th IFAC symposium on system identification. Saint-Maio, France. July* (pp. 1375–1380).

Tugnait, J. K. (1982). Detection and estimation for abruptly changing systems. *Automatica*, *18*, 607–615.

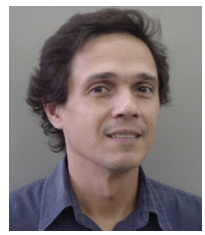
Wonham, W. M. (1968). On a matrix Riccati equation of stochastic control. *SIAM Journal on Control*, *6*, 681–697.

Yang, F., Wang, Z., & Hung, Y. S. (2002). Robust Kalman filtering for discrete time-varying uncertain systems with multiplicative noises. *IEEE Transactions on Automatic Control*, *47*, 1179–1183.

Zhang, Q. (1999). Optimal filtering of discrete-time hybrid systems. *Journal of Optimization Theory and Applications*, *100*, 123–144.

Zhang, Q. (2000). Hybrid filtering for linear systems with non-Gaussian disturbances. *IEEE Transactions on Automatic Control*, *45*, 50–61.

Zhang, L, & Zhang, X. D. (2007). An optimal filtering algorithm for systems with multiplicative/additive noises. *IEEE Signal Processing Letters*, *14*, 469–472.



Oswaldo L.V. Costa was born in 1959 in Rio de Janeiro, RJ, Brazil. He obtained his B.Sc. and M.Sc. degrees both in Electrical Engineering from the Catholic University of Rio de Janeiro, Brazil, in 1981 and 1983 respectively, and the Ph.D. degree in Electrical Engineering from Imperial College of Science and Technology in London in 1987. He is presently Professor in the Control Group of the Department of Telecommunications and Control Engineering of the Polytechnic School of the University of São Paulo, Brazil. His research interests include stochastic control, optimal control, and jump systems.

Guilherme R.A.M. Benites was born in 1982 in São Paulo, SP, Brazil. He obtained his B.S. and M.S. degrees from the University of São Paulo, São Paulo, Brazil, in 2003, and 2006, respectively. He is currently working toward the Ph.D. degree at the Department of Telecommunications and Control Engineering of the Polytechnic School of the University of São Paulo, Brazil. His research focuses mainly on stochastic control of systems with Markovian jumps.