

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Design of Estimation Algorithms from an Innovation Approach in Linear Discrete-Time Stochastic Systems with Uncertain Observations

J. Linares-Pérez, R. Caballero-Águila and I. García-Garrido

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/45777>

1. Introduction

The least-squares estimation problem in linear discrete-time stochastic systems in which the signal to be estimated is always present in the observations has been widely treated; as is well known, the Kalman filter [12] provides the least-squares estimator when the additive noises and the initial state are Gaussian and mutually independent.

Nevertheless, in many real situations, usually the measurement device or the transmission mechanism can be subject to random failures, generating observations in which the state appears randomly or which may consist of noise only due, for example, to component or interconnection failures, intermittent failures in the observation mechanism, fading phenomena in propagation channels, accidental loss of some measurements or data inaccessibility at certain times. In these situations where it is possible that information concerning the system state vector may or may not be contained in the observations, at each sampling time, there is a positive probability (called *false alarm probability*) that only noise is observed and, hence, that the observation does not contain the transmitted signal, but it is not generally known whether the observation used for estimation contains the signal or it is only noise. To describe this interrupted observation mechanism (*uncertain observations*), the observation equation, with the usual additive measurement noise, is formulated by multiplying the signal function at each sampling time by a binary random variable taking the values one and zero (Bernoulli random variable); the value one indicates that the measurement at that time contains the signal, whereas the value zero reflects the fact that the signal is missing and, hence, the corresponding observation is only noise. So, the observation equation involves both an additive and a multiplicative noise, the latter modeling the uncertainty about the signal being present or missing at each observation.

Linear discrete-time systems with uncertain observations have been widely used in estimation problems related to the above practical situations (which commonly appear, for example, in

Communication Theory). Due to the multiplicative noise component, even if the additive noises are Gaussian, systems with uncertain observations are always non-Gaussian and hence, as occurs in other kinds of non-Gaussian linear systems, the least-squares estimator is not a linear function of the observations and, generally, it is not easily obtainable by a recursive algorithm; for this reason, research in this kind of systems has focused special attention on the search of suboptimal estimators for the signal (mainly linear ones).

In some cases, the variables modeling the uncertainty in the observations can be assumed to be independent and, then, the distribution of the multiplicative noise is fully determined by the probability that each particular observation contains the signal. As it was shown by Nahi [17] (who was the first who analyzed the least-squares linear filtering problem in this kind of systems assuming that the state and observation additive noises are uncorrelated) the knowledge of the aforementioned probabilities allows to derive estimation algorithms with a recursive structure similar to the Kalman filter. Later on, Monzingo [16] completed these results by analyzing the least-squares smoothing problem and, subsequently, [3] and [4] generalized the least-squares linear filtering and smoothing algorithms considering that the additive noises of the state and the observation are correlated.

However, there exist many real situations where this independence assumption of the Bernoulli variables modeling the uncertainty is not satisfied; for example, in signal transmission models with stand-by sensors in which any failure in the transmission is detected immediately and the old sensor is then replaced, thus avoiding the possibility of the signal being missing in two successive observations. This different situation was considered by [9] by assuming that the variables modeling the uncertainty are correlated at consecutive time instants, and the proposed least-squares linear filtering algorithm provides the signal estimator at any time from those in the two previous instants. Later on, the state estimation problem in discrete-time systems with uncertain observations, has been widely studied under different hypotheses on the additive noises involved in the state and observation equations and, also, under several hypotheses on the multiplicative noise modeling the uncertainty in the observations (see e.g. [22] - [13], among others).

On the other hand, there are many engineering application fields (for example, in communication systems) where sensor networks are used to obtain all the available information on the system state and its estimation must be carried out from the observations provided by all the sensors (see [6] and references therein). Most papers concerning systems with uncertain observations transmitted by multiple sensors assume that all the sensors have the same uncertainty characteristics. In the last years, this situation has been generalized by several authors considering uncertain observations whose statistical properties are assumed not to be the same for all the sensors. This is a realistic assumption in several application fields, for instance, in networked communication systems involving heterogeneous measurement devices (see e.g. [14] and [8], among others). In [7] it is assumed that the uncertainty in each sensor is modeled by a sequence of independent Bernoulli random variables, whose statistical properties are not necessarily the same for all the sensors. Later on, in [10] and [1] the independence restriction is weakened; specifically, different sequences of Bernoulli random variables correlated at consecutive sampling times are considered to model the uncertainty at each sensor. This form of correlation covers practical situations where the signal cannot be missing in two successive observations. In [2] the least-squares linear and quadratic problems are addressed when the Bernoulli variables describing the uncertainty in

the observations are correlated at instants that differ two units of time. This study covers more general practical situations, for example, in sensor networks where sensor failures may happen and a failed sensor is replaced not immediately, but two sampling times after having failed. However, even if it is assumed that any failure in the transmission results from sensor failures, usually the failed sensor may not be replaced immediately but after m instants of time; in such situations, correlation among the random variables modeling the uncertainty in the observations at times k and $k + m$ must be considered and new algorithms must be deduced.

The current chapter is concerned with the state estimation problem for linear discrete-time systems with uncertain observations when the uncertainty at any sampling time k depends only on the uncertainty at the previous time $k - m$; this form of correlation allows us to consider certain models in which the signal cannot be missing in $m + 1$ consecutive observations.

The random interruptions in the observation process are modeled by a sequence of Bernoulli variables (at each time, the value one of the variable indicates that the measurement is the current system output, whereas the value zero reflects that only noise is available), which are correlated only at the sampling times $k - m$ and k . Recursive algorithms for the filtering and fixed-point smoothing problems are proposed by using an innovation approach; this approach, based on the fact that the innovation process can be obtained by a causal and invertible operation on the observation process, consists of obtaining the estimators as a linear combination of the innovations and simplifies considerably the derivation of the estimators due to the fact that the innovations constitute a white process.

The chapter is organized as follows: in Section 2 the system model is described; more specifically, we introduce the linear state transition model perturbed by a white noise, and the observation model affected by an additive white noise and a multiplicative noise describing the uncertainty. Also, the pertinent hypotheses to address the least-squares linear estimation problem are established. In Section 3 this estimation problem is formulated using an innovation approach. Next, in Section 4, recursive algorithms for the filter and fixed-point smoother are derived, including recursive formulas for the estimation error covariance matrices. Finally, the performance of the proposed estimators is illustrated in Section 5 by a numerical simulation example, where a two-dimensional signal is estimated and the estimation accuracy is analyzed for different values of the uncertainty probability and several values of the time period m .

2. Model description

Consider linear discrete-time stochastic systems with uncertain observations coming from multiple sensors, whose mathematical modeling is accomplished by the following equations.

The state equation is given by

$$x_k = F_{k-1}x_{k-1} + w_{k-1}, \quad k \geq 1, \quad (1)$$

where $\{x_k; k \geq 0\}$ is an n -dimensional stochastic process representing the system state, $\{w_k; k \geq 0\}$ is a white noise process and F_k , for $k \geq 0$, are known deterministic matrices.

We consider scalar uncertain observations $\{y_k^i; k \geq 1\}$, $i = 1, \dots, r$, coming from r sensors and perturbed by noises whose statistical properties are not necessarily the same for all the sensors. Specifically, we assume that, in each sensor and at any time k , the observation y_k^i , perturbed by an additive noise, can have no information about the state (thus being only noise) with a known probability. That is,

$$y_k^i = \begin{cases} H_k^i x_k + v_k^i, & \text{with probability } \bar{\theta}_k^i \\ v_k^i, & \text{with probability } 1 - \bar{\theta}_k^i \end{cases}$$

where, for $i = 1, \dots, r$, $\{v_k^i; k \geq 1\}$ is the observation additive noise process of the i -th sensor and H_k^i , for $k \geq 1$, are known deterministic matrices of compatible dimensions. If we introduce $\{\theta_k^i; k \geq 1\}$, $i = 1, \dots, r$, sequences of Bernoulli random variables with $P[\theta_k^i = 1] = \bar{\theta}_k^i$, the observations of the state can be rewritten as

$$y_k^i = \theta_k^i H_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, \dots, r. \quad (2)$$

Remark 1. If $\theta_k^i = 1$, which occurs with known probability $\bar{\theta}_k^i$, the state x_k is present in the observation y_k^i coming from the i -th sensor at time k , whereas if $\theta_k^i = 0$ such observation only contains additive noise, v_k^i , with probability $1 - \bar{\theta}_k^i$. This probability is called *false alarm probability* and it represents the probability that only noise is observed or, equivalently, that y_k^i does not contain the state.

The aim is to address the state estimation problem considering all the available observations coming from the r sensors. For convenience, denoting $y_k = (y_k^1, \dots, y_k^r)^T$, $v_k = (v_k^1, \dots, v_k^r)^T$, $H_k = (H_k^{1T}, \dots, H_k^{rT})^T$ and $\Theta_k = \text{Diag}(\theta_k^1, \dots, \theta_k^r)$, Equation (2) is equivalent to the following stacked observation equation

$$y_k = \Theta_k H_k x_k + v_k, \quad k \geq 1. \quad (3)$$

2.1. Model hypotheses

In order to analyze the least-squares linear estimation problem of the state x_k from the observations y_1, \dots, y_L , with $L \geq k$, some considerations must be taken into account. On the one hand, it is known that the linear estimator of x_k , is the orthogonal projection of x_k onto the space of n -dimensional random variables obtained as linear transformations of the observations y_1, \dots, y_L , which requires the existence of the second-order moments of such observations. On the other hand, we consider that the variables describing the uncertainty in the observations are correlated in instants that differ m units of time to cover many practical situations where the independence assumption on such variables is not realistic. Specifically, the following hypotheses are assumed:

Hypothesis 1. The initial state x_0 is a random vector with $E[x_0] = \bar{x}_0$ and $\text{Cov}[x_0] = P_0$.

Hypothesis 2. The state noise $\{w_k; k \geq 0\}$ is a zero-mean white sequence with $\text{Cov}[w_k] = Q_k$, $\forall k \geq 0$.

Hypothesis 3. The observation additive noise $\{v_k; k \geq 1\}$ is a zero-mean white process with $Cov[v_k] = R_k, \forall k \geq 1$.

Hypothesis 4. For $i = 1, \dots, r, \{\theta_k^i; k \geq 1\}$ is a sequence of Bernoulli random variables with $P[\theta_k^i = 1] = \bar{\theta}_k^i$. For $i, j = 1, \dots, r$, the variables θ_k^i and θ_s^j are independent for $|k - s| \neq 0, m$ and $Cov[\theta_k^i, \theta_s^j]$ are known for $|k - s| = 0, m$. Defining $\theta_k = (\theta_k^1, \dots, \theta_k^r)^T$, the covariance matrices of θ_k and θ_s will be denoted by $K_{k,s}^\theta$.

Finally, we assume the following hypothesis on the independence of the initial state and noises:

Hypothesis 5. The initial state x_0 and the noise processes $\{w_k; k \geq 0\}$, $\{v_k; k \geq 1\}$ and $\{\theta_k; k \geq 1\}$ are mutually independent.

Remark 2. For the derivation of the estimation algorithms a matrix product, called *Hadamard product*, which is simpler than the conventional product, will be considered. Let $A, B \in M_{mn}$, the Hadamard product (denoted by \circ) of A and B is defined as $[A \circ B]_{ij} = A_{ij}B_{ij}$. From this definition it is easily deduced (see [7]) the next property that will be needed later.

For any random matrix $G_{m \times m}$ independent of $\{\Theta_k; k \geq 1\}$, the following equality is satisfied

$$E[\Theta_k G_{m \times m} \Theta_s] = E[\theta_k \theta_s^T] \circ E[G_{m \times m}].$$

Particularly, denoting $\bar{\Theta}_k = E[\Theta_k]$, it is immediately clear that

$$E[(\Theta_k - \bar{\Theta}_k) G_{m \times m} (\Theta_s - \bar{\Theta}_s)] = K_{k,s}^\theta \circ E[G_{m \times m}]. \quad (4)$$

Remark 3. Several authors assume that the observations available for the estimation come either from multiple sensors with identical uncertainty characteristics or from a single sensor (see [20] for the case when the uncertainty is modeled by independent variables, and [19] for the case when such variables are correlated at consecutive sampling times). Nevertheless, in the last years, this situation has been generalized by some authors considering multiple sensors featuring different uncertainty characteristics (see e.g. [7] for the case of independent uncertainty, and [1] for situations where the uncertainty in each sensor is modeled by variables correlated at consecutive sampling times). We analyze the state estimation problem for the class of linear discrete-time systems with uncertain observation (3), which, as established in Hypothesis 4, are characterized by the fact that the uncertainty at any sampling time k depends only on the uncertainty at the previous time $k - m$; this form of correlation allows us to consider certain models in which the signal cannot be absent in $m + 1$ consecutive observations.

3. Least-squares linear estimation problem

As mentioned above, our aim in this chapter is to obtain the least-squares linear estimator, $\hat{x}_{k/L}$, of the signal x_k based on the observations $\{y_1, \dots, y_L\}$, with $L \geq k$, by recursive formulas. Specifically, the problem is to derive recursive algorithms for the least-squares linear filter ($L = k$) and fixed-point smoother (fixed k and $L > k$) of the state using uncertain observations (3). For this purpose, we use an innovation approach as described in [11].

Since the observations are generally nonorthogonal vectors, we use the *Gram-Schmidt orthogonalization procedure* to transform the set of observations $\{y_1, \dots, y_L\}$ into an equivalent set of orthogonal vectors $\{v_1, \dots, v_L\}$; equivalent in the sense that they both generate the same linear subspace; that is,

$$\mathcal{L}(y_1, \dots, y_L) = \mathcal{L}(v_1, \dots, v_L) = \mathcal{L}_L.$$

Let $\{v_1, \dots, v_{k-1}\}$ be the set of orthogonal vectors satisfying $\mathcal{L}(v_1, \dots, v_{k-1}) = \mathcal{L}(y_1, \dots, y_{k-1})$, the next orthogonal vector, v_k , corresponding to the new observation y_k , is obtained by projecting y_k onto \mathcal{L}_{k-1} ; specifically

$$v_k = y_k - \text{Proj}\{y_k \text{ onto } \mathcal{L}_{k-1}\},$$

and, because of the orthogonality of $\{v_1, \dots, v_{k-1}\}$ the above projection can be found by projecting y_k along each of the previously found orthogonal vectors v_i , for $i \leq k-1$,

$$\text{Proj}\{y_k \text{ onto } \mathcal{L}_{k-1}\} = \sum_{i=1}^{k-1} \text{Proj}\{y_k \text{ along } v_i\} = \sum_{i=1}^{k-1} E[y_k v_i^T] \left(E[v_i v_i^T] \right)^{-1} v_i.$$

Since the projection of y_k onto \mathcal{L}_{k-1} is $\hat{y}_{k/k-1}$, the one-stage least-squares linear predictor of y_k , we have that

$$\hat{y}_{k/k-1} = \sum_{i=1}^{k-1} T_{k,i} \Pi_i^{-1} v_i, \quad k \geq 2 \quad (5)$$

where $T_{k,i} = E[y_k v_i^T]$ and $\Pi_i = E[v_i v_i^T]$ is the covariance of v_i .

Consequently, by starting with $v_1 = y_1 - E[y_1]$, the orthogonal vectors v_k are determined by $v_k = y_k - \hat{y}_{k/k-1}$, for $k \geq 2$. Hence, v_k can be considered as the "new information" or the "innovation" in y_k given $\{y_1, \dots, y_{k-1}\}$.

In summary, the observation process $\{y_k; k \geq 1\}$ has been transformed into an equivalent white noise $\{v_k; k \geq 1\}$ known as innovation process. Taking into account that both processes satisfy that

$$v_i \in \mathcal{L}(y_1, \dots, y_i) \quad \text{and} \quad y_i \in \mathcal{L}(v_1, \dots, v_i), \quad \forall i \geq 1,$$

we conclude that such processes are related to each other by a causal and causally invertible linear transformation, thus making the innovation process be uniquely determined by the observations.

This consideration allows us to state that the least-squares linear estimator of the state based on the observations, $\hat{x}_{k/L}$, is equal to the least-squares linear estimator of the state based on the innovations $\{v_1, \dots, v_L\}$. Thus, projecting x_k separately onto each v_i , $i \leq L$, the following general expression for the estimator $\hat{x}_{k/L}$ is obtained

$$\hat{x}_{k/L} = \sum_{i=1}^L S_{k,i} \Pi_i^{-1} v_i, \quad k \geq 1, \quad (6)$$

where $S_{k,i} = E[x_k v_i^T]$. This expression is the starting point to derive the recursive filtering and fixed-point smoothing algorithms in the next section.

4. Least-squares linear estimation recursive algorithms

In this section, using an innovation approach, recursive algorithms are proposed for the filter, $\hat{x}_{k/k}$, and the fixed-point smoother, $\hat{x}_{k/L}$, for fixed k and $L > k$.

4.1. Linear filtering algorithm

In view of the general expression (6) for $L = k$, it is clear that the state filter, $\hat{x}_{k/k}$, is obtained from the one-stage state predictor, $\hat{x}_{k/k-1}$, by

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + S_{k,k} \Pi_k^{-1} v_k, \quad k \geq 1; \quad \hat{x}_{0/0} = \bar{x}_0. \quad (7)$$

Hence, an equation for the predictor $\hat{x}_{k/k-1}$ in terms of the filter $\hat{x}_{k-1/k-1}$ and expressions for the innovation v_k , its covariance matrix Π_k and the matrix $S_{k,k}$ are required.

State predictor $\hat{x}_{k/k-1}$. From hypotheses 2 and 5, it is immediately clear that the filter of the noise w_{k-1} is $\hat{w}_{k-1/k-1} = E[w_{k-1}] = 0$ and hence, taking into account Equation (1), we have

$$\hat{x}_{k/k-1} = F_{k-1} \hat{x}_{k-1/k-1}, \quad k \geq 1. \quad (8)$$

Innovation v_k . We will now get an explicit formula for the innovation, $v_k = y_k - \hat{y}_{k/k-1}$, or equivalently for the one-stage predictor of the observation, $\hat{y}_{k/k-1}$. For this purpose, taking into account (5), we start by calculating $T_{k,i} = E[y_k v_i^T]$, for $i \leq k-1$.

From the observation equation (3) and hypotheses 3 and 5, it is clear that

$$T_{k,i} = E[\Theta_k H_k x_k v_i^T], \quad i \leq k-1. \quad (9)$$

Now, for $k \leq m$ or $k > m$ and $i < k-m$, hypotheses 4 and 5 guarantee that Θ_k is independent of the innovations v_i , and then we have that $T_{k,i} = \bar{\Theta}_k H_k E[x_k v_i^T] = \bar{\Theta}_k H_k S_{k,i}$. So, after some manipulations, we obtain

I. For $k \leq m$, it is satisfied that $\hat{y}_{k/k-1} = \bar{\Theta}_k H_k \sum_{i=1}^{k-1} S_{k,i} \Pi_i^{-1} v_i$, and using (6) for $L = k-1$ it is obvious that

$$\hat{y}_{k/k-1} = \bar{\Theta}_k H_k \hat{x}_{k/k-1}, \quad k \leq m. \quad (10)$$

II. For $k > m$, we have that $\hat{y}_{k/k-1} = \bar{\Theta}_k H_k \sum_{i=1}^{k-(m+1)} S_{k,i} \Pi_i^{-1} v_i + \sum_{i=1}^m T_{k,k-i} \Pi_{k-i}^{-1} v_{k-i}$ and adding

and subtracting $\sum_{i=1}^m \bar{\Theta}_k H_k S_{k,k-i} \Pi_{k-i}^{-1} v_{k-i}$, the following equality holds

$$\hat{y}_{k/k-1} = \bar{\Theta}_k H_k \sum_{i=1}^{k-1} S_{k,i} \Pi_i^{-1} v_i + \sum_{i=1}^m (T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i}) \Pi_{k-i}^{-1} v_{k-i}, \quad k > m. \quad (11)$$

Next, we determine an expression for $T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i}$, for $1 \leq i \leq m$.

Taking into account (9), it follows that

$$T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i} = E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k v_{k-i}^T \right], \quad 1 \leq i \leq m, \quad (12)$$

or equivalently,

$$T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i} = E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k y_{k-i}^T \right] - E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k \hat{y}_{k-i/k-(i+1)}^T \right].$$

To calculate the first expectation, we use again (3) for y_{k-i} and from hypotheses 3 and 5, we have that

$$E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k y_{k-i}^T \right] = E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k x_{k-i}^T H_{k-i}^T \Theta_{k-i} \right]$$

which, using Property (4), yields

$$E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k y_{k-i}^T \right] = K_{k,k-i}^\theta \circ \left(H_k E[x_k x_{k-i}^T] H_{k-i}^T \right).$$

Now, denoting $D_k = E[x_k x_k^T]$ and $\mathbb{F}_{k,i} = F_{k-1} \cdots F_i$, from Equation (1) it is clear that $E[x_k x_{k-i}^T] = \mathbb{F}_{k,k-i} D_{k-i}$, and hence

$$E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k y_{k-i}^T \right] = K_{k,k-i}^\theta \circ \left(H_k \mathbb{F}_{k,k-i} D_{k-i} H_{k-i}^T \right)$$

where D_k can be recursively obtained by

$$\begin{aligned} D_k &= F_{k-1} D_{k-1} F_{k-1}^T + Q_{k-1}, \quad k \geq 1; \\ D_0 &= P_0 + \bar{x}_0 \bar{x}_0^T. \end{aligned} \quad (13)$$

Summarizing, we have that

$$\begin{aligned} T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i} &= K_{k,k-i}^\theta \circ \left(H_k \mathbb{F}_{k,k-i} D_{k-i} H_{k-i}^T \right) \\ &\quad - E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k \hat{y}_{k-i/k-(i+1)}^T \right], \quad 1 \leq i \leq m. \end{aligned} \quad (14)$$

Taking into account the correlation hypothesis of the variables describing the uncertainty, the right-hand side of this equation is calculated differently for $i = m$ or $i < m$, as shown below.

(a) For $i = m$, since Θ_k is independent of the innovations v_i , for $i < k - m$, we have that

$$E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k \hat{y}_{k-m/k-(m+1)}^T \right] = 0, \text{ and from (14)}$$

$$T_{k,k-m} - \bar{\Theta}_k H_k S_{k,k-m} = K_{k,k-m}^\theta \circ \left(H_k \mathbb{F}_{k,k-m} D_{k-m} H_{k-m}^T \right). \quad (15)$$

(b) For $i < m$, from Hypothesis 4, $K_{k,k-i}^\theta = 0$ and, hence, from (14)

$$T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i} = -E \left[(\Theta_k - \bar{\Theta}_k) H_k x_k \hat{y}_{k-i/k-(i+1)}^T \right].$$

Now, from expression (5),

$$\hat{y}_{k-i/k-(i+1)} = \sum_{j=1}^{k-(i+1)} T_{k-i,j} \Pi_j^{-1} v_j,$$

and using again that Θ_k is independent of v_i , for $i \neq k-m$, it is deduced that

$$T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i} = -E[(\Theta_k - \bar{\Theta}_k) H_k x_k v_{k-m}^T] \Pi_{k-m}^{-1} T_{k-i,k-m}^T$$

or, equivalently, from (12) for $i = m$, (15) and noting

$$\Psi_{k,k-m} = K_{k,k-m}^\theta \circ \left(H_k \mathbb{F}_{k,k-m} D_{k-m} H_{k-m}^T \right) \Pi_{k-m}^{-1},$$

we have that

$$T_{k,k-i} - \bar{\Theta}_k H_k S_{k,k-i} = -\Psi_{k,k-m} T_{k-i,k-m}^T, \quad i < m. \quad (16)$$

Next, substituting (15) and (16) into (11) and using (6) for $\hat{x}_{k/k-1}$, it is concluded that

$$\hat{y}_{k/k-1} = \bar{\Theta}_k H_k \hat{x}_{k/k-1} + \Psi_{k,k-m} \left[v_{k-m} - \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} v_{k-i} \right], \quad k > m. \quad (17)$$

Finally, using (3) and (16) and taking into account that, from (1), $S_{k,k-i} = \mathbb{F}_{k,k-i} S_{k-i,k-i}$, the matrices $T_{k,k-i}$ in (17) are obtained by

$$T_{k,k-i} = \bar{\Theta}_k H_k \mathbb{F}_{k,k-i} S_{k-i,k-i}, \quad 2 \leq k \leq m, \quad 1 \leq i \leq k-1,$$

$$T_{k,k-i} = \bar{\Theta}_k H_k \mathbb{F}_{k,k-i} S_{k-i,k-i} - \Psi_{k,k-m} T_{k-i,k-m}^T, \quad k > m, \quad 1 \leq i \leq m-1.$$

Matrix $S_{k,k}$. Since $v_k = y_k - \hat{y}_{k/k-1}$, we have that $S_{k,k} = E[x_k v_k^T] = E[x_k y_k^T] - E[x_k \hat{y}_{k/k-1}^T]$. Next, we calculate these expectations.

- I. From Equation (3) and the independence hypothesis, it is clear that $E[x_k y_k^T] = D_k H_k^T \bar{\Theta}_k$, $\forall k \geq 1$, where $D_k = E[x_k x_k^T]$ is given by (13).
- II. To calculate $E[x_k \hat{y}_{k/k-1}^T]$, the correlation hypothesis of the random variables θ_k must be taken into account and two cases must be considered:
 - (a) For $k \leq m$, from (10) we obtain

$$E[x_k \hat{y}_{k/k-1}^T] = E[x_k \hat{x}_{k/k-1}^T] H_k^T \bar{\Theta}_k.$$

By using the orthogonal projection lemma, which assures that $E[x_k \hat{x}_{k/k-1}^T] = D_k - P_{k/k-1}$, where $P_{k/k-1} = E[(x_k - \hat{x}_{k/k-1})(x_k - \hat{x}_{k/k-1})^T]$ is the prediction error covariance matrix, we get

$$E[x_k \hat{y}_{k/k-1}^T] = (D_k - P_{k/k-1}) H_k^T \bar{\Theta}_k, \quad k \leq m.$$

(b) For $k > m$, from (17) it follows that

$$\begin{aligned} E[x_k \hat{y}_{k/k-1}^T] &= E[x_k \hat{x}_{k/k-1}^T] H_k^T \bar{\Theta}_k + E[x_k v_{k-m}^T] \Psi_{k,k-m}^T \\ &\quad - E \left[x_k \left(\sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} v_{k-i} \right)^T \right] \Psi_{k,k-m}^T, \end{aligned}$$

hence, using again the orthogonal projection lemma and taking into account that $S_{k,k-i} = E[x_k v_{k-i}^T]$, for $1 \leq i \leq m$, it follows that

$$\begin{aligned} E[x_k \hat{y}_{k/k-1}^T] &= (D_k - P_{k/k-1}) H_k^T \bar{\Theta}_k + S_{k,k-m} \Psi_{k,k-m}^T \\ &\quad - \sum_{i=1}^{m-1} S_{k,k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T, \quad k > m. \end{aligned}$$

Then, substituting these expectations in the expression of $S_{k,k}$ and simplifying, it is clear that

$$\begin{aligned} S_{k,k} &= P_{k/k-1} H_k^T \bar{\Theta}_k, \quad 1 \leq k \leq m, \\ S_{k,k} &= P_{k/k-1} H_k^T \bar{\Theta}_k - \left(S_{k,k-m} - \sum_{i=1}^{m-1} S_{k,k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \right) \Psi_{k,k-m}^T, \quad k > m. \end{aligned} \quad (18)$$

Now, an expression for the prediction error covariance matrix, $P_{k/k-1}$, is necessary. From Equation (1), it is immediately clear that

$$P_{k/k-1} = F_{k-1} P_{k-1/k-1} F_{k-1}^T + Q_{k-1}, \quad k \geq 1,$$

where $P_{k/k} = E[(x_k - \hat{x}_{k/k})(x_k - \hat{x}_{k/k})^T]$ is the filtering error covariance matrix. From Equation (7), it is concluded that

$$P_{k/k} = P_{k/k-1} - S_{k,k} \Pi_k^{-1} S_{k,k}^T, \quad k \geq 1; \quad P_{0/0} = P_0.$$

Covariance matrix of the innovation $\Pi_k = E[\nu_k \nu_k^T]$. From the orthogonal projection lemma, the covariance matrix of the innovation is obtained as $\Pi_k = E[y_k y_k^T] - E[\hat{y}_{k/k-1} \hat{y}_{k/k-1}^T]$.

From (3) and using Property (4), we have that

$$E[y_k y_k^T] = E[\theta_k \theta_k^T] \circ (H_k D_k H_k^T) + R_k, \quad k \geq 1.$$

To obtain $E[\hat{y}_{k/k-1} \hat{y}_{k/k-1}^T]$ two cases must be distinguished again, due to the correlation hypothesis of the Bernoulli variables θ_k :

I. For $k \leq m$, Equation (10) and Property (4) yield

$$E[\hat{y}_{k/k-1} \hat{y}_{k/k-1}^T] = (\bar{\theta}_k \bar{\theta}_k^T) \circ (H_k E[\hat{x}_{k/k-1} \hat{x}_{k/k-1}^T] H_k^T)$$

and in view of the orthogonal projection lemma,

$$E[\hat{y}_{k/k-1} \hat{y}_{k/k-1}^T] = (\bar{\theta}_k \bar{\theta}_k^T) \circ (H_k (D_k - P_{k/k-1}) H_k^T), \quad k \leq m.$$

II. For $k > m$, an analogous reasoning, but using now Equation (17), yields

$$\begin{aligned}
 E[\hat{y}_{k/k-1}\hat{y}_{k/k-1}^T] &= (\bar{\theta}_k\bar{\theta}_k^T) \circ (H_k(D_k - P_{k/k-1})H_k^T) + \Psi_{k,k-m}\Pi_{k-m}\Psi_{k,k-m}^T \\
 &+ \Psi_{k,k-m} \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} \Pi_{k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T \\
 &+ \bar{\Theta}_k H_k E[\hat{x}_{k/k-1} \nu_{k-m}^T] \Psi_{k,k-m}^T + \Psi_{k,k-m} E[\nu_{k-m} \hat{x}_{k/k-1}^T] H_k^T \bar{\Theta}_k \\
 &- \bar{\Theta}_k H_k E[\hat{x}_{k/k-1} \sum_{i=1}^{m-1} \nu_{k-i}^T \Pi_{k-i}^{-1} T_{k-i,k-m}] \Psi_{k,k-m}^T \\
 &- \Psi_{k,k-m} \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} E[\nu_{k-i} \hat{x}_{k/k-1}^T] H_k^T \bar{\Theta}_k.
 \end{aligned}$$

Next, again from the orthogonal projection lemma, $E[\hat{x}_{k/k-1} \nu_{k-i}^T] = E[x_k \nu_{k-i}^T] = S_{k,k-i}$, for $1 \leq i \leq m$, and therefore

$$\begin{aligned}
 E[\hat{y}_{k/k-1}\hat{y}_{k/k-1}^T] &= (\bar{\theta}_k\bar{\theta}_k^T) \circ (H_k(D_k - P_{k/k-1})H_k^T) + \Psi_{k,k-m}\Pi_{k-m}\Psi_{k,k-m}^T \\
 &+ \Psi_{k,k-m} \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T \\
 &+ \bar{\Theta}_k H_k \left(S_{k,k-m} \Psi_{k,k-m}^T - \sum_{i=1}^{m-1} S_{k,k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T \right) \\
 &+ \left(\Psi_{k,k-m} S_{k,k-m}^T - \Psi_{k,k-m} \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} S_{k,k-i}^T \right) H_k^T \bar{\Theta}_k.
 \end{aligned}$$

Finally, from Equation (18), we have

$$S_{k,k-m} \Psi_{k,k-m}^T - \sum_{i=1}^{m-1} S_{k,k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T = -(S_{k,k} - P_{k/k-1} H_k^T \bar{\Theta}_k),$$

and hence,

$$\begin{aligned}
 E[\hat{y}_{k/k-1}\hat{y}_{k/k-1}^T] &= (\bar{\theta}_k\bar{\theta}_k^T) \circ (H_k(D_k - P_{k/k-1})H_k^T) + \Psi_{k,k-m}\Pi_{k-m}\Psi_{k,k-m}^T \\
 &+ \Psi_{k,k-m} \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T \\
 &- \bar{\Theta}_k H_k \left(S_{k,k} - P_{k/k-1} H_k^T \bar{\Theta}_k \right) - \left(S_{k,k}^T - \bar{\Theta}_k H_k P_{k/k-1} \right) H_k^T \bar{\Theta}_k, \quad k > m.
 \end{aligned}$$

Finally, since $K_{k,k}^\theta = E[\theta_k \theta_k^T] - \bar{\theta}_k \bar{\theta}_k^T$, the above expectations lead to the following expression for the innovation covariance matrices

$$\Pi_k = K_{k,k}^\theta \circ (H_k D_k H_k^T) + R_k + \bar{\Theta}_k H_k S_{k,k}, \quad k \leq m,$$

$$\begin{aligned} \Pi_k &= K_{k,k}^\theta \circ \left(H_k D_k H_k^T \right) + R_k - \Psi_{k,k-m} \left(\Pi_{k-m} + \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} T_{k-i,k-m} \right) \Psi_{k,k-m}^T \\ &\quad + \bar{\Theta}_k H_k S_{k,k} + S_{k,k}^T H_k^T \bar{\Theta}_k - \bar{\Theta}_k H_k P_{k/k-1} H_k^T \bar{\Theta}_k, \quad k > m. \end{aligned}$$

All these results are summarized in the following theorem.

Theorem 1. The linear filter, $\hat{x}_{k/k}$, of the state x_k is obtained as

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + S_{k,k} \Pi_k^{-1} v_k, \quad k \geq 1; \quad \hat{x}_{0/0} = \bar{x}_0,$$

where the state predictor, $\hat{x}_{k/k-1}$, is given by

$$\hat{x}_{k/k-1} = F_{k-1} \hat{x}_{k-1/k-1}, \quad k \geq 1.$$

The innovation process satisfies

$$\begin{aligned} v_k &= y_k - \bar{\Theta}_k H_k \hat{x}_{k/k-1}, \quad k \leq m, \\ v_k &= y_k - \bar{\Theta}_k H_k \hat{x}_{k/k-1} + \Psi_{k,k-m} \left[v_{k-m} - \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} v_{k-i} \right], \quad k > m, \end{aligned}$$

where $\bar{\Theta}_k = E[\Theta_k]$ and $\Psi_{k,k-m} = K_{k,k-m}^\theta \circ \left(H_k \mathbb{F}_{k,k-m} D_{k-m} H_{k-m}^T \Pi_{k-m}^{-1} \right)$, with \circ the Hadamard product, $\mathbb{F}_{k,i} = F_{k-1} \cdots F_i$ and $D_k = E[x_k x_k^T]$ recursively obtained by

$$D_k = F_{k-1} D_{k-1} F_{k-1}^T + Q_{k-1}, \quad k \geq 1; \quad D_0 = P_0 + \bar{x}_0 \bar{x}_0^T.$$

The matrices $T_{k,k-i}$ are given by

$$\begin{aligned} T_{k,k-i} &= \bar{\Theta}_k H_k \mathbb{F}_{k,k-i} S_{k-i,k-i}, \quad 2 \leq k \leq m, \quad 1 \leq i \leq k-1, \\ T_{k,k-i} &= \bar{\Theta}_k H_k \mathbb{F}_{k,k-i} S_{k-i,k-i} - \Psi_{k,k-m} T_{k-i,k-m}^T, \quad k > m, \quad 1 \leq i \leq m-1. \end{aligned}$$

The covariance matrix of the innovation, $\Pi_k = E[v_k v_k^T]$, satisfies

$$\begin{aligned} \Pi_k &= K_{k,k}^\theta \circ \left(H_k D_k H_k^T \right) + R_k + \bar{\Theta}_k H_k S_{k,k}, \quad k \leq m, \\ \Pi_k &= K_{k,k}^\theta \circ \left(H_k D_k H_k^T \right) + R_k - \Psi_{k,k-m} \left(\Pi_{k-m} + \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} T_{k-i,k-m} \right) \Psi_{k,k-m}^T \\ &\quad + \bar{\Theta}_k H_k S_{k,k} + S_{k,k}^T H_k^T \bar{\Theta}_k - \bar{\Theta}_k H_k P_{k/k-1} H_k^T \bar{\Theta}_k, \quad k > m. \end{aligned}$$

The matrix $S_{k,k}$ is determined by the following expression

$$\begin{aligned} S_{k,k} &= P_{k/k-1} H_k^T \bar{\Theta}_k, \quad k \leq m, \\ S_{k,k} &= P_{k/k-1} H_k^T \bar{\Theta}_k - \left(S_{k,k-m} - \sum_{i=1}^{m-1} S_{k,k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \right) \Psi_{k,k-m}^T, \quad k > m, \end{aligned}$$

where $P_{k/k-1}$, the prediction error covariance matrix, is obtained by

$$P_{k/k-1} = F_{k-1}P_{k-1/k-1}F_{k-1}^T + Q_{k-1}, \quad k \geq 1,$$

with $P_{k/k}$, the filtering error covariance matrix, satisfying

$$P_{k/k} = P_{k/k-1} - S_{k,k}\Pi_k^{-1}S_{k,k}^T, \quad k \geq 1; \quad P_{0/0} = P_0. \quad (19)$$

4.2. Linear fixed-point smoothing algorithm

The following theorem provides a recursive fixed-point smoothing algorithm to obtain the least-squares linear estimator, $\hat{x}_{k/k+N}$, of the state x_k based on the observations $\{y_1, \dots, y_{k+N}\}$, for $k \geq 1$ fixed and $N \geq 1$. Moreover, to measure of the estimation accuracy, a recursive formula for the error covariance matrices, $P_{k/k+N} = E[(x_k - \hat{x}_{k/k+N})(x_k - \hat{x}_{k/k+N})^T]$, is derived.

Theorem 2. For each fixed $k \geq 1$, the fixed-point smoothers, $\hat{x}_{k/k+N}$, $N \geq 1$ are calculated by

$$\hat{x}_{k/k+N} = \hat{x}_{k/k+N-1} + S_{k,k+N}\Pi_{k+N}^{-1}v_{k+N}, \quad N \geq 1, \quad (20)$$

whose initial condition is the filter, $\hat{x}_{k/k}$, given in (7).

The matrices $S_{k,k+N}$ are calculated from

$$\begin{aligned} S_{k,k+N} &= \left(D_k \mathbb{F}_{k+N,k}^T - M_{k,k+N-1} F_{k+N-1}^T \right) H_{k+N}^T \bar{\Theta}_{k+N}, \quad k \leq m - N, \quad N \geq 1, \\ S_{k,k+N} &= \left(D_k \mathbb{F}_{k+N,k}^T - M_{k,k+N-1} F_{k+N-1}^T \right) H_{k+N}^T \bar{\Theta}_{k+N} \\ &\quad - \left(S_{k,k+N-m} - \sum_{i=1}^{m-1} S_{k,k+N-i} \Pi_{k+N-i}^{-1} T_{k+N-i,k+N-m} \right) \\ &\quad \times \Psi_{k+N,k+N-m}^T, \quad k > m - N, \quad N \geq 1. \end{aligned} \quad (21)$$

where the matrices $M_{k,k+N}$ satisfy the following recursive formula:

$$\begin{aligned} M_{k,k+N} &= M_{k,k+N-1} F_{k+N-1}^T + S_{k,k+N} \Pi_{k+N}^{-1} S_{k+N,k+N}^T, \quad N \geq 1, \\ M_{k,k} &= D_k - P_{k/k}. \end{aligned} \quad (22)$$

The innovations v_{k+N} , their covariance matrices Π_{k+N} , the matrices $T_{k+N,k+N-i}$, $\Psi_{k+N,k+N-m}$, D_k and $P_{k/k}$ are given in Theorem 1.

Finally, the fixed-point smoothing error covariance matrix, $P_{k/k+N}$, verifies

$$P_{k/k+N} = P_{k/k+N-1} - S_{k,k+N} \Pi_{k+N}^{-1} S_{k,k+N}^T, \quad N \geq 1, \quad (23)$$

with initial condition the filtering error covariance matrix, $P_{k/k}$, given by (19).

Proof. From the general expression (6), for each fixed $k \geq 1$, the recursive relation (20) is immediately clear.

Now, we need to prove (21) for $S_{k,k+N} = E[x_k v_{k+N}^T] = E[x_k y_{k+N}^T] - E[x_k \hat{y}_{k+N/k+N-1}^T]$, thus being necessary to calculate both expectations.

- I. From Equation (3), taking into account that $E[x_k v_{k+N}^T] = D_k \mathbb{F}_{k+N,k}^T$ and using that Θ_{k+N} and v_{k+N} are independent of x_k , we obtain

$$E[x_k y_{k+N}^T] = D_k \mathbb{F}_{k+N,k}^T H_{k+N}^T \bar{\Theta}_{k+N}, \quad N \geq 1.$$

- II. Based on expressions (10) and (17) for $\hat{y}_{k+N/k+N-1}$, which are different depending on $k+N \leq m$ or $k+N > m$, two options must be considered:

- (a) For $k \leq m - N$, using (10) for $\hat{y}_{k+N/k+N-1}$ with (8) for $\hat{x}_{k+N/k+N-1}$, we have that

$$E[x_k \hat{y}_{k+N/k+N-1}^T] = M_{k,k+N-1} F_{k+N-1}^T H_{k+N}^T \bar{\Theta}_{k+N},$$

where $M_{k,k+N-1} = E[x_k \hat{x}_{k+N-1/k+N-1}^T]$.

- (b) For $k > m - N$, by following a similar reasoning to the previous one but starting from (17), we get

$$\begin{aligned} E[x_k \hat{y}_{k+N/k+N-1}^T] &= M_{k,k+N-1} F_{k+N-1}^T H_{k+N}^T \bar{\Theta}_{k+N} + (S_{k,k+N-m} \\ &\quad - \sum_{i=1}^{m-1} S_{k,k+N-i} \Pi_{k+N-i}^{-1} T_{k+N-i,k+N-m}) \Psi_{k+N,k+N-m}^T. \end{aligned}$$

Then, the replacement of the above expectations in $S_{k,k+N}$ leads to expression (21).

The recursive relation (22) for $M_{k,k+N} = E[x_k \hat{x}_{k+N/k+N}^T]$ is immediately clear from (7) for $\hat{x}_{k+N/k+N}$ and its initial condition $M_{k,k} = E[x_k \hat{x}_{k/k}^T]$ is calculated taking into account that, from the orthogonality, $E[x_k \hat{x}_{k/k}^T] = E[\hat{x}_{k/k} \hat{x}_{k/k}^T] = D_k - P_{k/k}$.

Finally, since $P_{k,k+N} = E[x_k x_k^T] - E[\hat{x}_{k/k+N} \hat{x}_{k/k+N}^T]$, using (20) and taking into account that $\hat{x}_{k/k+N-1}$ is uncorrelated with v_{k+N} , we have

$$P_{k,k+N} = E[x_k x_k^T] - E[\hat{x}_{k/k+N-1} \hat{x}_{k/k+N-1}^T] - S_{k,k+N} \Pi_{k+N}^{-1} S_{k,k+N}^T, \quad N \geq 1$$

and, consequently, expression (23) holds. □

5. Numerical simulation example

In this section, we present a numerical example to show the performance of the recursive algorithms proposed in this chapter. To illustrate the effectiveness of the proposed estimators, we ran a program in MATLAB which, at each iteration, simulates the state and the observed values and provides the filtering and fixed-point smoothing estimates, as well as the corresponding error covariance matrices, which provide a measure of the estimators accuracy.

Consider a two-dimensional state process, $\{x_k; k \geq 0\}$, generated by the following first-order autoregressive model

$$x_k = \left(1 + 0.2 \sin \left(\frac{(k-1)\pi}{50} \right)\right) \begin{pmatrix} 0.8 & 0 \\ 0.9 & 0.2 \end{pmatrix} x_{k-1} + w_{k-1}, \quad k \geq 1$$

with the following hypotheses:

- The initial state, x_0 , is a zero-mean Gaussian vector with covariance matrix given by $P_0 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$.
- The process $\{w_k; k \geq 0\}$ is a zero-mean white Gaussian noise with covariance matrices $Q_k = \begin{pmatrix} 0.36 & 0.3 \\ 0.3 & 0.25 \end{pmatrix}, \forall k \geq 0$.

Suppose that the scalar observations come from two sensors according to the following observation equations:

$$y_k^i = \theta_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2.$$

where $\{v_k^i; k \geq 1\}, i = 1, 2$, are zero-mean independent white Gaussian processes with variances $R_k^1 = 0.5$ and $R_k^2 = 0.9, \forall k \geq 1$, respectively.

According to our theoretical model, it is assumed that, for each sensor, the uncertainty at time k depends only on the uncertainty at the previous time $k - m$. The variables $\theta_k^i, i = 1, 2$, modeling this type of uncertainty correlation in the observation process are modeled by two independent sequences of independent Bernoulli random variables, $\{\gamma_k^i; k \geq 1\}, i = 1, 2$, with constant probabilities $P[\gamma_k^i = 1] = \gamma_i$. Specifically, the variables θ_k^i are defined as follows

$$\theta_k^i = 1 - \gamma_{k+m}^i (1 - \gamma_k^i), \quad i = 1, 2.$$

So, if $\theta_k^i = 0$, then $\gamma_{k+m}^i = 1$ and $\gamma_k^i = 0$, and hence, $\theta_{k+m}^i = 1$; this fact guarantees that, if the state is absent at time k , after $k + m$ instants of time the observation necessarily contains the state. Therefore, there cannot be more than m consecutive observations consisting of noise only.

Moreover, since the variables γ_k^i and γ_s^i are independent, θ_k^i and θ_s^i also are independent for $|k - s| \neq 0, m$. The common mean of these variables is $\bar{\theta}^i = 1 - \gamma_i(1 - \gamma_i)$ and its covariance function is given by

$$K_{k,s}^\theta = E[(\theta_k^i - \bar{\theta}^i)(\theta_s^i - \bar{\theta}^i)] = \begin{cases} 0 & \text{if } |k - s| \neq 0, m, \\ -(1 - \bar{\theta}^i)^2 & \text{if } |k - s| = m, \\ \bar{\theta}^i(1 - \bar{\theta}^i) & \text{if } |k - s| = 0. \end{cases}$$

To illustrate the effectiveness of the respective estimators, two hundred iterations of the proposed algorithms have been performed and the results obtained for different values of the uncertainty probability and several values of m have been analyzed.

Let us observe that the mean function of the variables θ_k^i , for $i = 1, 2$ are the same if $1 - \gamma_i$ is used instead of γ_i ; for this reason, only the case $\gamma_i \leq 0.5$ will be considered here. Note that, in such case, the false alarm probability at the i -th sensor, $1 - \bar{\theta}^i$, is an increasing function of γ_i .

Firstly, the values of the first component of a simulated state together with the filtering and the fixed-point smoothing estimates for $N = 2$, obtained from simulated observations of the state for $m = 3$ and $\gamma^1 = \gamma^2 = 0.5$ are displayed in Fig. 1. This graph shows that the fixed-point smoothing estimates follow the state evolution better than the filtering ones.

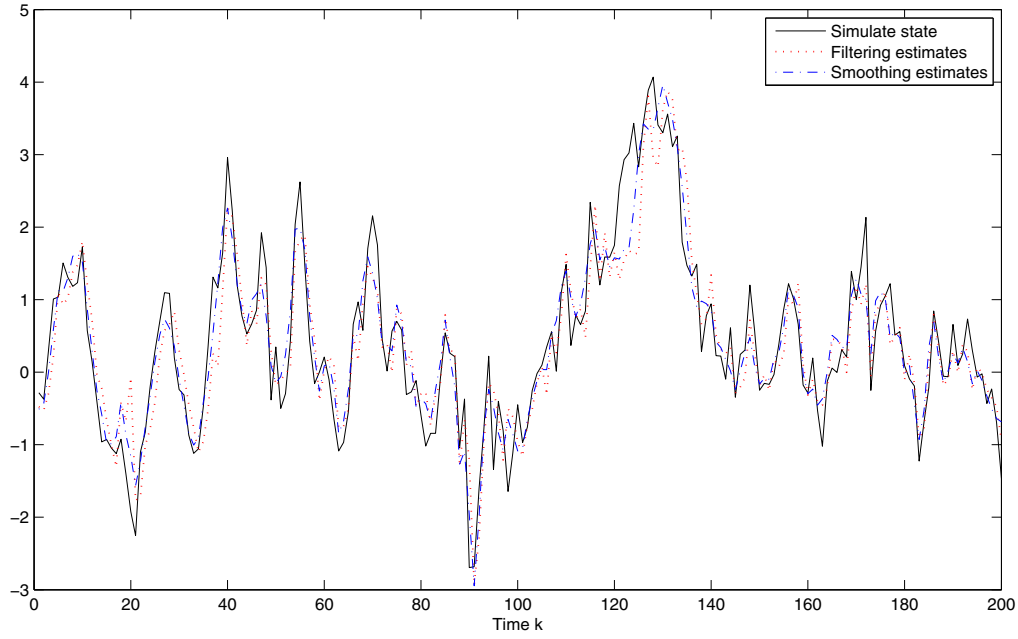


Figure 1. First component of the simulate state, filtering and fixed-point smoothing estimates for $N = 2$, when $m = 3$ and $\gamma^1 = \gamma^2 = 0.5$.

Next, assuming again that the Bernoulli variables of the observations are correlated at sampling times that differ three units of time ($m = 3$), we compare the effectiveness of the proposed filtering and fixed-point smoothing estimators considering different values of the probabilities γ_1 and γ_2 , which provides different values of the false alarm probabilities $1 - \bar{\theta}^i$, $i = 1, 2$; specifically, $\gamma_1 = 0.2$, $\gamma_2 = 0.4$ and $\gamma_1 = 0.1$, $\gamma_2 = 0.3$. For these values, Fig. 2 shows the filtering and fixed-point smoothing error variances, when $N = 2$ and $N = 5$, for the first state component. From this figure it is observed that:

- i) As both γ_1 and γ_2 decrease (which means that the false alarm probability decreases), the error variances are smaller and, consequently, better estimations are obtained.
- ii) The error variances corresponding to the fixed-point smoothers are less than those of the filters and, consequently, agreeing with the comments on the previous figure, the fixed-point smoothing estimates are more accurate.
- iii) The accuracy of the smoothers at each fixed-point k is better as the number of available observations increases.

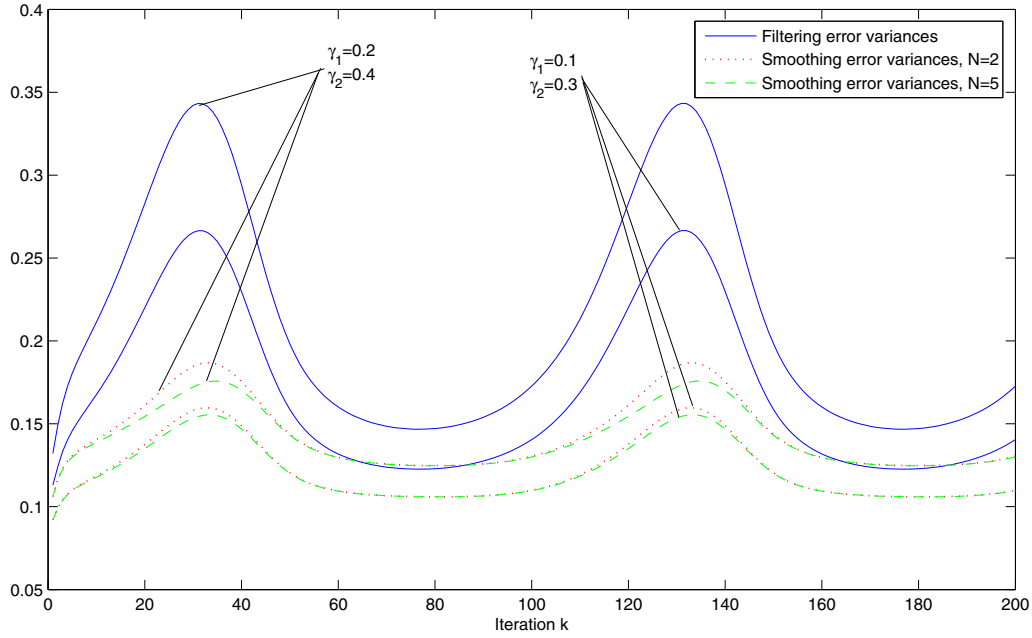


Figure 2. Filtering and smoothing error variances for the first state component for $\gamma_1 = 0.2$, $\gamma_2 = 0.4$ and $\gamma_1 = 0.1$, $\gamma_2 = 0.3$, when $m = 3$.

On the other hand, in order to show more precisely the dependence of the error variance on the values γ_1 and γ_2 , Fig. 3 displays the filtering and fixed-point smoothing error variances of the first state component, at a fixed iteration (namely, $k = 200$) for $m = 3$, when both γ_1 and γ_2 are varied from 0.1 to 0.5, which provide different values of the probabilities $\bar{\theta}^1$ and $\bar{\theta}^2$. More specifically, we have considered the values $\gamma_i = 0.1, 0.2, 0.3, 0.4, 0.5$, which lead to the false alarm probabilities $1 - \bar{\theta}^i = 0.09, 0.16, 0.22, 0.24, 0.25$, respectively.

In this figure, both graphs (corresponding to the filtering and fixed-point smoothing error variances) corroborate the previous results, showing again that, as the false alarm probability increases, the filtering and fixed-point smoothing error variances ($N = 2$) become greater and consequently, worse estimations are obtained. Also, it is concluded that the smoothing error variances are better than the filtering ones.

Analogous results to those of Fig. 1-3 are obtained for the second component of the state. As example, Fig. 4 shows the filtering and fixed-point smoothing error variances of the second state component, at $k = 200$, versus γ_1 for constant values of γ_2 , when $m = 3$ and similar comments to those made from Fig. 3 are deduced.

Finally, for $\gamma_1 = 0.2$, $\gamma_2 = 0.4$ the performance of the estimators is compared for different values of m ; specifically, for $m = 1, 3, 6$, the filtering error variances of the first state component are displayed in Fig. 5. From this figure it is gathered that the estimators are more accurate as the values of m are lower. In other words, a greater distance between the instants at which the variables are correlated (which means that more consecutive observations may not contain state information) yields worse estimations.

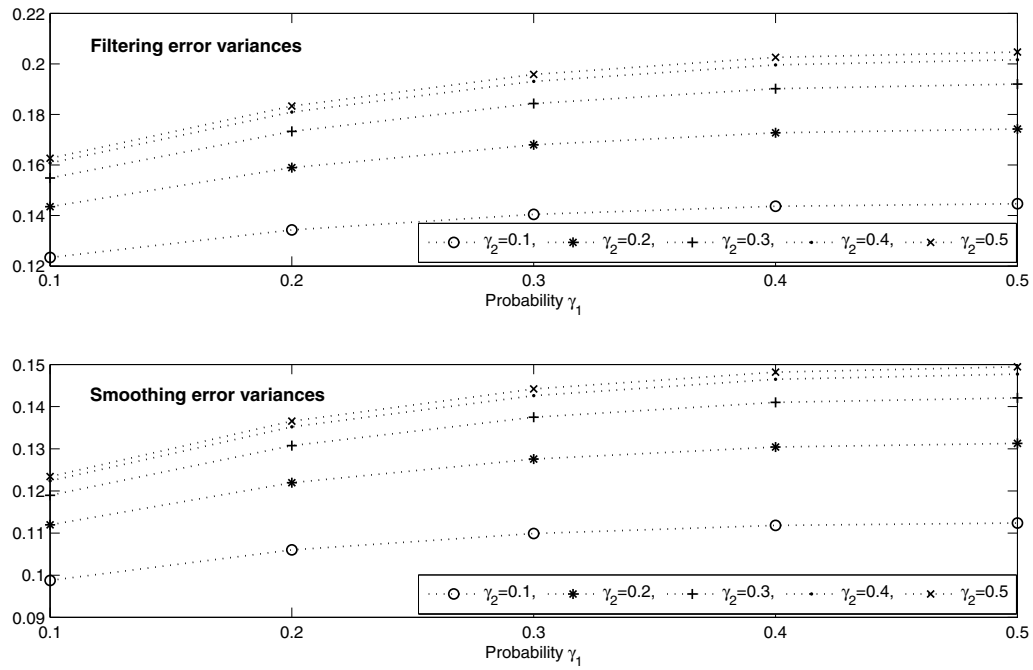


Figure 3. Filtering error variances and smoothing error variances for $N = 2$ of the first state component at $k = 200$ versus γ_1 with γ_2 varying from 0.1 to 0.5 when $m = 3$.

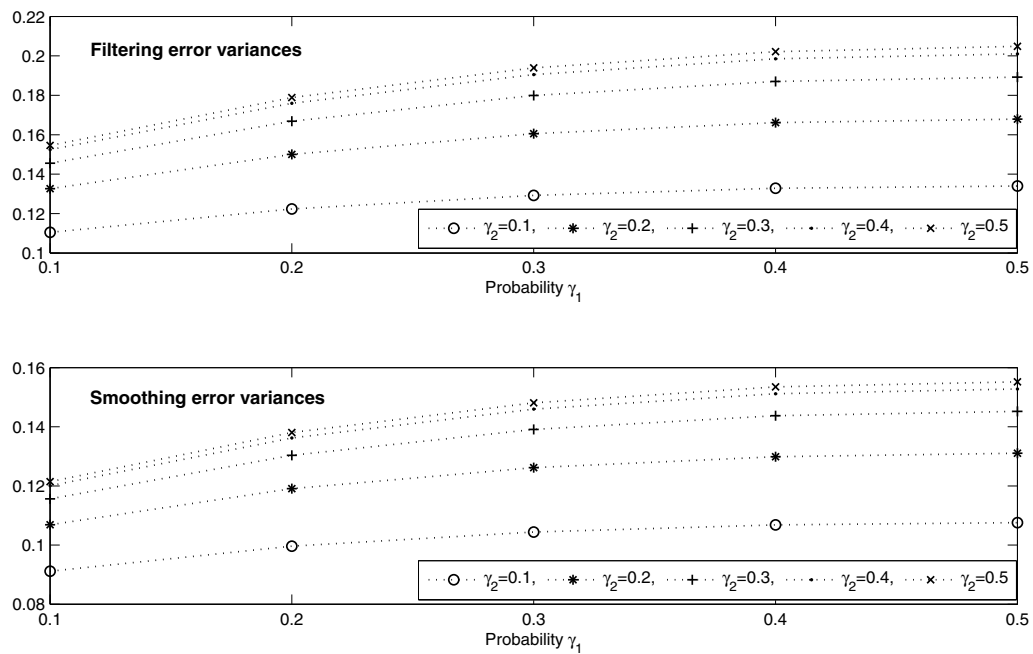


Figure 4. Filtering error variances and smoothing error variances for $N = 2$ of the second state component at $k = 200$ versus γ_1 with γ_2 varying from 0.1 to 0.5 when $m = 3$.

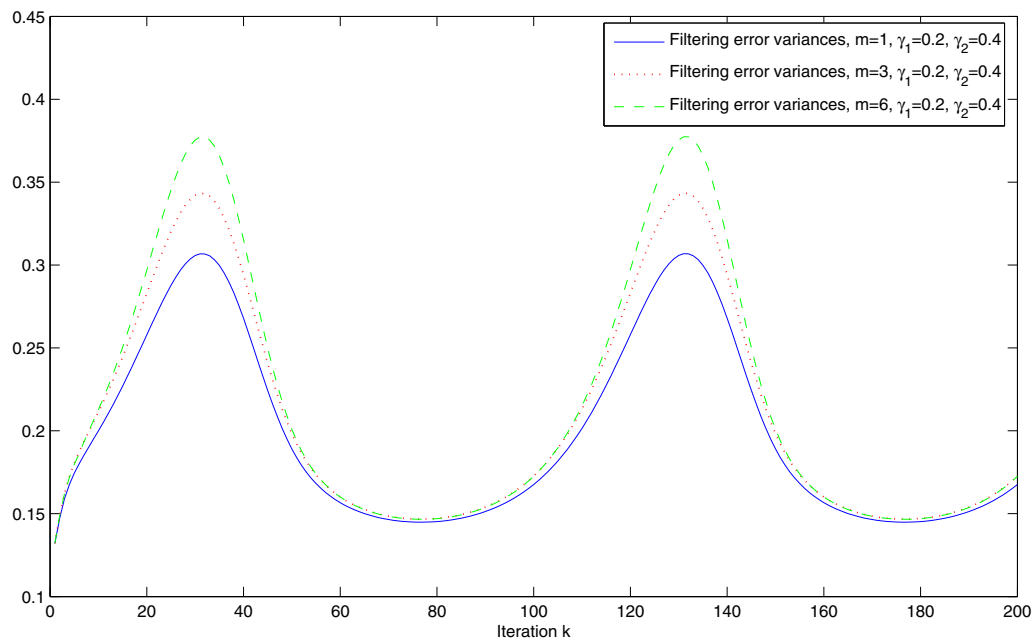


Figure 5. Filtering error variances for $\gamma_1 = 0.2$, $\gamma_2 = 0.4$ and $m = 1, 3, 6$.

6. Conclusions and future research

In this chapter, the least-squares linear filtering and fixed-point smoothing problems have been addressed for linear discrete-time stochastic systems with uncertain observations coming from multiple sensors. The uncertainty in the observations is modeled by a binary variable taking the values one or zero (Bernoulli variable), depending on whether the signal is present or absent in the corresponding observation, and it has been supposed that the uncertainty at any sampling time k depends only on the uncertainty at the previous time $k - m$. This situation covers, in particular, those signal transmission models in which any failure in the transmission is detected and the old sensor is replaced after m instants of time, thus avoiding the possibility of missing signal in $m + 1$ consecutive observations.

By applying an innovation technique, recursive algorithms for the linear filtering and fixed-point smoothing estimators have been obtained. This technique consists of obtaining the estimators as a linear combination of the innovations, simplifying the derivation of these estimators, due to the fact that the innovations constitute a white process.

Finally, the feasibility of the theoretical results has been illustrated by the estimation of a two-dimensional signal from uncertain observations coming from two sensors, for different uncertainty probabilities and different values of m . The results obtained confirm the greater effectiveness of the fixed-point smoothing estimators in contrast to the filtering ones and conclude that more accurate estimations are obtained as the values of m are lower.

In recent years, several problems of signal processing, such as signal prediction, detection and control, as well as image restoration problems, have been treated using quadratic estimators and, generally, polynomial estimators of arbitrary degree. Hence, it must be noticed that the current chapter can be extended by considering the least-squares polynomial estimation

problems of arbitrary degree for such linear systems with uncertain observations correlated in instants that differ m units of time. On the other hand, in practical engineering, some recent progress on the filtering and control problems for nonlinear stochastic systems with uncertain observations is being achieved. Nonlinearity and stochasticity are two important sources that are receiving special attention in research and, therefore, filtering and smoothing problems for nonlinear systems with uncertain observations would be relevant topics on which further investigation would be interesting.

Acknowledgements

This research is supported by Ministerio de Educación y Ciencia (Programa FPU and grant No. MTM2011-24718) and Junta de Andalucía (grant No. P07-FQM-02701).

Author details

J. Linares-Pérez

Dpto. de Estadística e I.O. Universidad de Granada. Avda. Fuentenueva. 18071. Granada, Spain

R. Caballero-Águila

Dpto. de Estadística e I.O. Universidad de Jaén. Paraje Las Lagunillas. 23071. Jaén, Spain

I. García-Garrido

Dpto. de Estadística e I.O. Universidad de Granada. Avda. Fuentenueva. 18071. Granada, Spain

7. References

- [1] Caballero-Águila, R., Hermoso-Carazo, A. & Linares-Pérez, J. [2011]. Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty, *Signal Processing* 91(No. 2): 330–337.
- [2] García-Garrido, I., Linares-Pérez, J., Caballero-Águila, R. & Hermoso-Carazo, A. [2012]. A solution to the filtering problem for stochastic systems with multi-sensor uncertain observations, *International Mathematical Forum* 7(No. 18): 887–903.
- [3] Hermoso-Carazo, A. & Linares-Pérez, J. [1994]. Linear estimation for discrete-time systems in the presence of time-correlated disturbances and uncertain observations, *IEEE Transactions on Automatic Control* 39(No. 8): 1636–1638.
- [4] Hermoso-Carazo, A. & Linares-Pérez, J. [1995]. Linear smoothing for discrete-time systems in the presence of correlated disturbances and uncertain observations, *IEEE Transactions on Automatic Control* 40(No. 8): 243–251.
- [5] Hermoso-Carazo, A., Linares-Pérez, J., Jiménez-López, J., Caballero-Águila, R. & Nakamori, S. [2008]. Recursive fixed-point smoothing algorithm from covariances based on uncertain observations with correlation in the uncertainty, *Applied Mathematics and Computation* 203(No. 1): 243–251.
- [6] Hespanha, J., Naghshtabrizi, P. & Xu, Y. [2007]. A survey of recent results in networked control systems, *Proceedings of the IEEE* 95(No. 1): 138–172.
- [7] Hounkpevi, F. & Yaz, E. [2007]. Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica* 43(No. 7): 1274–1280.

- [8] Huang, R. & Záruba, G. [2007]. Incorporating data from multiple sensors for localizing nodes in mobile ad hoc networks, *IEEE Transactions on Mobile Computing* 6(No. 9): 1090–1104.
- [9] Jackson, R. & Murthy, D. [1976]. Optimal linear estimation with uncertain observations, *IEEE Transactions on Information Theory* 22(No. 3): 376–378.
- [10] Jiménez-López, J., Linares-Pérez, J., Nakamori, S., Caballero-Águila, R. & Hermoso-Carazo, A. [2008]. Signal estimation based on covariance information from observations featuring correlated uncertainty and coming from multiple sensors, *Signal Processing* 88(No. 12): 2998–3006.
- [11] Kailath, T., Sayed, A. & Hassibi, B. [2000]. *Linear Estimation*, Prentice Hall, New Jersey.
- [12] Kalman, R. [1960]. A new approach to linear filtering and prediction problems, *Transactions of the ASME. Journal of Basic Engineering* (No. 82 (Series D)): 35–45.
- [13] Ma, J. & Sun, S. [2011]. Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations, *IEEE Transactions on Signal Processing* 59(No. 11): 5181–5192.
- [14] Malyavej, V., Manchester, I. & Savkin, A. [2006]. Precision missile guidance using radar/multiple-video sensor fusion via communication channels with bit-rate constraints, *Automatica* 42(No. 5): 763–769.
- [15] Moayed, M., Foo, Y. & Soh, Y. [2010]. Adaptive kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements, *IEEE Transactions on Signal Processing* 58(No. 3): 1577–1578.
- [16] Monzingo, R. [1981]. Discrete linear recursive smoothing for systems with uncertain observations, *IEEE Transactions on Automatic Control* AC-26(No. 3): 754–757.
- [17] Nahi, N. [1969]. Optimal recursive estimation with uncertain observation, *IEEE Transactions on Information Theory* 15(No. 4): 457–462.
- [18] Nakamori, S., Caballero-Águila, R., Hermoso-Carazo, A., Jiménez-López, J. & Linares-Pérez, J. [2006]. Least-squares vth-order polynomial estimation of signals from observations affected by non-independent uncertainty, *Applied Mathematics and Computation* Vol. 176(No. 2): 642–653.
- [19] Nakamori, S., Caballero-Águila, R., Hermoso-Carazo, A., Jiménez-López, J. & Linares-Pérez, J. [2007]. Signal polynomial smoothing from correlated interrupted observations based on covariances, *Mathematical Methods in the Applied Sciences* Vol. 30(No. 14): 1645–1665.
- [20] NaNacara, W. & Yaz, E. [1997]. Recursive estimator for linear and nonlinear systems with uncertain observations, *Signal Processing* Vol. 62(No. 2): 215–228.
- [21] Sahebsara, M., Chen, T. & Shah, S. [2007]. Optimal \mathcal{H}_2 filtering with random sensor delay, multiple packet dropout and uncertain observations, *International Journal of Control* 80(No. 2): 292–301.
- [22] Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. & Sastry, S. [2004]. Kalman filtering with intermittent observations, *IEEE Transactions on Automatic Control* 49(No. 9): 1453–1464.
- [23] Wang, Z., Yang, F., Ho, D. & Liu, X. [2005]. Robust finite-horizon filtering for stochastic systems with missing measurements, *IEEE Signal Processing Letters* 12(No. 6): 437–440.

- [24] Wang, Z., Yang, F., Ho, D. & Liu, X. [2006]. Robust H_∞ filtering for stochastic time-delay systems with missing measurements, *IEEE Transactions on Signal Processing* 54(No. 7): 2579–2587.

IntechOpen

IntechOpen