Research Article

Unscented Filtering from Delayed Observations with Correlated Noises

A. Hermoso-Carazo and J. Linares-Pérez

Departamento de Estadística, Universidad de Granada, Avda. Fuentenueva, 18071 Granada, Spain

Correspondence should be addressed to J. Linares-Pérez, jlinares@ugr.es

A filtering algorithm based on the unscented transformation is proposed to estimate the state of a nonlinear system from noisy measurements which can be randomly delayed by one sampling time. The state and observation noises are perturbed by correlated nonadditive noises, and the delay is modeled by independent Bernoulli random variables.

1. Introduction

The signal estimation problem in time-delay stochastic systems plays an important role in different application fields. For example, in engineering applications involving communication networks with a heavy network traffic, the measurements available may not be uptodate. Although the delay can sometimes be interpreted as a known deterministic function of the time, the numerous sources of uncertainty make it preferable to interpret it as a stochastic process, including its statistical properties in the system model. This fact must be considered in the study of the signal estimation problem since the conventional algorithms are then not applicable.

In the past few years, attention has been focused on investigating estimation problems from measurements subject to a random delay which does not exceed one sampling time, modeling the delay values by a zero-one white noise with known probabilities indicating that the measurements either arrive on time or are delayed. In linear systems, Ray et al. [1] first modified the conventional algorithms to fit this observation model and since then many results based on this model have been reported (see, among others, [2–5] and references therein). Literature on nonlinear filtering from randomly delayed observations is less extensive. Recently, generalizations of extended and unscented Kalman filters, using one- and two-step randomly delayed observations, have been proposed and compared in [6, 7], respectively, for a class of nonlinear discrete-time systems with independent additive noises.

In this paper, we address the problem of estimating from nonlinear measurements subject to a random delay which does not exceed one sampling time, and when the last available measurement is used for the estimation at any time. This situation is modelled by considering Bernoulli random variables whose value of one indicates that the corresponding observation is not updated. Concretely, we propose an extension of the unscented filter in [6] to the case of correlated and nonadditive signal and measurement noises.

2. State and Observation Models

In this section, we present the nonlinear systems with one-step randomly delayed observations to be considered and we describe the assumption about the underlying processes.

The considered nonlinear discrete-time model is represented by the equations:

$$x_{k+1} = f_k(x_k, w_k), \quad k \geq 0,$$
$$\widetilde{y}_k = h_k(x_k, v_k), \quad k \geq 1, \tag{2.1}$$

where x_k and $\widetilde{y}_k$ are random vectors which describe the system state and output at time k, respectively. The process $\{w_k; \ k \geq 0\}$ is the state noise, $\{v_k; \ k \geq 1\}$ is the measurement noise, and, for all k, f_k and h_k are known analytic (not necessarily linear) functions.

We assume that at time $k = 1$ the real observation $\widetilde{y}_1$ is always available for the estimation but, as indicated previously, we go to consider the possibility that the current observation at any time $k > 1$, y_k, is either the current system output, $\widetilde{y}_k$, with probability of $1 - p_k$, or the previous one, $\widetilde{y}_{k-1}$, with probability of p_k (delay probability). Thus, the available observations for $k > 1$ are

$$y_k = \begin{cases} \widetilde{y}_{k-1}, & \text{with probability } p_k \\ \widetilde{y}_k, & \text{with probability } 1 - p_k, \end{cases} \tag{2.2}$$

and the delayed observation model can be described as [6]

$$y_k = (1 - \gamma_k)\widetilde{y}_k + \gamma_k \widetilde{y}_{k-1}, \quad k > 1; \quad y_1 = \widetilde{y}_1, \tag{2.3}$$

where $\{\gamma_k; \ k > 1\}$ are Bernoulli random variables (binary switching sequence taking the values 0 or 1) with $P(\gamma_k = 1) = p_k$, which model the delays in the observations. Indeed, if $\gamma_k = 1$ (which occurs with probability p_k), then $y_k = \widetilde{y}_{k-1}$ and the measurement is delayed by one sampling period; otherwise, $\gamma_k = 0$ implies that $y_k = \widetilde{y}_k$ or, equivalently, that the measurement is updated (which occurs with probability $1 - p_k$).

In applications of communication networks, the noise $\{\gamma_k; \ k > 1\}$ usually represents the random delay from sensor to controller and the assumption of one-step sensor delay is based on the reasonable supposition that the induced data latency from the sensor to the controller is restricted so as not to exceed the sampling period.

To deal with the state estimation problem, the following assumptions about the processes involved in (2.1) and (2.3) are considered.

Assumption 1. The initial state in (2.1), x_0, is a random vector with $E[x_0] = \overline{x}_0$ and $E[(x_0 - \overline{x}_0)(x_0 - \overline{x}_0)^T] = P_0$.

Assumption 2. The noises $\{w_k; \ k \geq 0\}$ and $\{v_k; \ k \geq 1\}$ are correlated zero-mean white processes with $E[w_k w_k^T] = Q_k$, $E[v_k v_k^T] = R_k$, and $E[w_j v_k] = S_k \delta_{j,k-1}$, with $\delta_{k-1,k-1} = 1$ and $\delta_{j,k-1} = 0$, $j \neq k - 1$.

Assumption 3. $\{\gamma_k; \ k > 1\}$ is a sequence of independent Bernoulli random variables with known probabilities, $P(\gamma_k = 1) = p_k$, for all $k > 1$.

Assumption 4. The initial state, x_0, and the processes $(\{w_k; \ k \geq 0\}, \{v_k; \ k \geq 1\})$ and $\{\gamma_k; \ k > 1\}$ are mutually independent.

3. Unscented Filtering Algorithm

The unscented transformation (see [8] for details) approximates the distribution of a N-dimensional random vector X by sample distributions with the same mean and covariance, $\widehat{X}$ and P^X. The distributions correspond to a set of $2N + 1$ sigma-points defined as

$$\chi_0 = \widehat{X},$$

$$\chi_i = \widehat{X} + \left(\sqrt{(N + \lambda)P^X}\right)_i, \quad i = 1, \ldots, N,$$

$$\chi_i = \widehat{X} - \left(\sqrt{(N + \lambda)P^X}\right)_{i-N}, \quad i = N + 1, \ldots, 2N,$$

(3.1)

(expression $(P)_i$ denotes the i-th column of the matrix P) whose mean and covariance are $\widehat{X}$ and P^X, respectively, when the following weights, W_i^m for the mean and W_i^c for the covariance, are used:

$$W_0^m = \frac{\lambda}{(N + \lambda)}, \qquad W_0^c = \frac{\lambda}{N + \lambda} + \left(1 - \alpha^2 + \beta\right),$$

$$W_i^m = W_i^c = \frac{1}{2(N + \lambda)}, \quad i = 1, \ldots, 2N.$$

(3.2)

The parameters in (3.1) and (3.2) are $\lambda = \alpha^2(N + \kappa) - N$, where α is a scaling parameter determining the spread of the sigma-points, and κ, β are tuning parameters. When the mean and covariance of a nonlinear transformation $g(X)$ are approximated by the sample mean and covariance of the transformed sigma-points, $g(\chi_i)$, $i = 0, \ldots, 2N$, weighted with W_i^m and W_i^c, respectively, these approximations are accurate up to the second and first term of their Taylor expansion series, respectively.

On the basis of this procedure, unscented filtering uses the state equation (which provides x_k as a nonlinear function of x_{k-1} and w_{k-1}) to approximate the conditional mean and covariance of x_k given $Y^{k-1} = \mathrm{Col}(y_1, \ldots, y_{k-1})$ from those of x_{k-1} and w_{k-1}. These statistics are then updated with the observation y_k using the Kalman equations to obtain approximations of the conditional mean and covariance of x_k given Y^k.

For the update, the mean and covariance of y_k given Y^{k-1}, and hence those of $\widetilde{y}_{k-1} = h_{k-1}(x_{k-1}, v_{k-1})$ and $\widetilde{y}_k = h_k(x_k, v_k)$, are approximated using again the unscented procedure and, for this purpose, the conditional statistics of the vectors x_{k-1}, v_{k-1}, x_k, and v_k must be known. Therefore, in view of the requirements in the prediction and update steps, the starting point for obtaining the filter of x_k is the knowledge of the approximated conditional mean and covariance of the vector $X_{k-1} = \mathrm{Col}(x_{k-1}, v_{k-1}, w_{k-1}, v_k)$ given Y^{k-1}; these statistics, which are denoted by $\widehat{X}_{k-1/k-1}$ and $P^{XX}_{k-1,k-1/k-1}$, respectively, provide the approximations of the conditional statistics of X_k given Y^k which, in turn, provide those of x_k. The procedure is now detailed in the following two steps.

Prediction step

From the independence of the vectors $\mathrm{Col}(v_k, w_k, v_{k+1})$ and Y^{k-1} and the conditional independence of $\mathrm{Col}(x_k, v_k)$ and $\mathrm{Col}(w_k, v_{k+1})$, the conditional mean and covariance of $X_k = \mathrm{Col}(x_k, v_k, w_k, v_{k+1})$ given Y^{k-1} are given by

$$
\widehat{X}_{k/k-1} = \begin{pmatrix} \widehat{x}_{k/k-1} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad
P^{XX}_{k,k/k-1} = \begin{pmatrix} P^{xx}_{k,k/k-1} & P^{xv}_{k,k/k-1} & 0 & 0 \\ P^{vx}_{k,k/k-1} & R_k & 0 & 0 \\ 0 & 0 & Q_k & S_{k+1} \\ 0 & 0 & S^T_{k+1} & R_{k+1} \end{pmatrix}, \tag{3.3}
$$

and then, the problem is to obtain approximations for the conditional mean and covariance of x_k, $\widehat{x}_{k/k-1}$ and $P^{xx}_{k,k/k-1}$, as well as for the conditional cross-covariance of x_k and v_k, $P^{xv}_{k,k/k-1}$.

Since $x_k = f_{k-1}(x_{k-1}, w_{k-1})$ and v_k are both functions of the vector X_{k-1}, in order to approximate their conditional statistics we use the sigma-points $\chi_{i,k-1} = \mathrm{Col}(\chi^x_{i,k-1}, \chi^v_{i,k-1}, \chi^w_{i,k-1}, \chi^v_{i,k})$, $i = 0,\ldots,2N$, defined from $\widehat{X}_{k-1/k-1}$ and $P^{XX}_{k-1,k-1/k-1}$ as in (3.1) (here N is the dimension of the augmented vector X_{k-1}), and the required statistics are approximated by those corresponding to the transformed sigma-points, $f^a_{k-1}(\chi_{i,k-1}) = f_{k-1}(\chi^x_{i,k-1}, \chi^w_{i,k-1})$, by using the weights defined in (3.2):

$$
\widehat{x}_{k/k-1} = \sum_{i=0}^{2N} W^m_i f^a_{k-1}(\chi_{i,k-1}),
$$

$$
P^{xx}_{k,k/k-1} = \sum_{i=0}^{2N} W^c_i \left(f^a_{k-1}(\chi_{i,k-1}) - \widehat{x}_{k/k-1} \right) \left(f^a_{k-1}(\chi_{i,k-1}) - \widehat{x}_{k/k-1} \right)^T, \tag{3.4}
$$

$$
P^{xv}_{k,k/k-1} = \sum_{i=0}^{2N} W^c_i f^a_{k-1}(\chi_{i,k-1}) \chi^{vT}_{i,k}.
$$

Update step

As previously commented, the obtaining of $\widehat{X}_{k/k}$ and $P^{XX}_{k,k/k}$ from $\widehat{X}_{k/k-1}$ and $P^{XX}_{k,k/k-1}$ is carried out by using the Kalman filter equations and hence, the mean and covariance of y_k given Y^{k-1}, as well as the conditional cross-covariance of y_k and X_k, need to be approximated; next, we describe the approximation procedure.

Taking into account (2.3) and since, from the independence, $P(\gamma_k = 1/Y^{k-1}) = p_k$, the conditional statistics of y_k given Y^{k-1} are expressed in terms of those corresponding to $\widetilde{y}_{k-1}$ and $\widetilde{y}_k$ as follows:

$$\widehat{y}_{k/k-1} = (1 - p_k)\widehat{\widetilde{y}}_{k/k-1} + p_k\widehat{\widetilde{y}}_{k-1/k-1},$$

$$P^{yy}_{k,k/k-1} = (1 - p_k)P^{\widetilde{y}\widetilde{y}}_{k,k/k-1} + p_k P^{\widetilde{y}\widetilde{y}}_{k-1,k-1/k-1}$$

$$+ p_k(1 - p_k)\left(\widehat{\widetilde{y}}_{k/k-1} - \widehat{\widetilde{y}}_{k-1/k-1}\right)\left(\widehat{\widetilde{y}}_{k/k-1} - \widehat{\widetilde{y}}_{k-1/k-1}\right)^T,$$

$$P^{Xy}_{k,k/k-1} = (1 - p_k)P^{X\widetilde{y}}_{k,k/k-1} + p_k P^{X\widetilde{y}}_{k,k-1/k-1},$$

(3.5)

where again applying the independence hypotheses of the model,

$$P^{X\widetilde{y}}_{k,k/k-1} = \begin{pmatrix} P^{x\widetilde{y}}_{k,k/k-1} \\ P^{v\widetilde{y}}_{k,k/k-1} \\ 0 \\ 0 \end{pmatrix}, \qquad P^{X\widetilde{y}}_{k,k-1/k-1} = \begin{pmatrix} P^{x\widetilde{y}}_{k,k-1/k-1} \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

(3.6)

As in the prediction step, the conditional statistics of $\widetilde{y}_{k-1} = h_{k-1}(x_{k-1}, v_{k-1})$ are approximated from the sigma-points $\chi_{i,k-1}$ associated with $\widehat{X}_{k-1/k-1}$ and $P^{XX}_{k-1,k-1/k-1}$, by defining $h^a_{k-1}(\chi_{i,k-1}) = h_{k-1}(\chi^x_{i,k-1}, \chi^v_{i,k-1})$:

$$\widehat{\widetilde{y}}_{k-1/k-1} = \sum_{i=0}^{2N} W^m_i h^a_{k-1}(\chi_{i,k-1}),$$

$$P^{\widetilde{y}\widetilde{y}}_{k-1,k-1/k-1} = \sum_{i=0}^{2N} W^c_i \left(h^a_{k-1}(\chi_{i,k-1}) - \widehat{\widetilde{y}}_{k-1/k-1}\right)\left(h^a_{k-1}(\chi_{i,k-1}) - \widehat{\widetilde{y}}_{k-1/k-1}\right)^T,$$

$$P^{x\widetilde{y}}_{k,k-1/k-1} = \sum_{i=0}^{2N} W^c_i \left(f^a_{k-1}(\chi_{i,k-1}) - \widehat{x}_{k/k-1}\right)\left(h^a_{k-1}(\chi_{i,k-1}) - \widehat{\widetilde{y}}_{k-1/k-1}\right)^T.$$

(3.7)

However, to approximate the statistics of $\widetilde{y}_k = h_k(x_k, v_k)$, which is a function of the vector $\mathrm{Col}(x_k, v_k)$, we use the information given in (3.3) and (3.4) about their conditional statistics. Thus, we consider a set of sigma-points, $\mathrm{Col}(\xi^x_{i,k}, \xi^v_{i,k})$, $i = 0, \ldots, 2M$ (M being the dimension of the vector $\mathrm{Col}(x_k, v_k)$), defined in a similar way to those in (3.1) for the two first block components of $\widehat{X}_{k/k-1}$ and $P^{XX}_{k,k/k-1}$, with weights w^m_i and w^c_i defined as in (3.2), and the

following approximations are used:

$$\widehat{\widetilde{y}}_{k/k-1} = \sum_{i=0}^{2M} w_i^m h_k\left(\xi_{i,k}^x, \xi_{i,k}^v\right),$$

$$P_{k,k/k-1}^{\widetilde{y}\widetilde{y}} = \sum_{i=0}^{2M} w_i^c \left(h_k\left(\xi_{i,k}^x, \xi_{i,k}^v\right) - \widehat{\widetilde{y}}_{k/k-1}\right)\left(h_k(\xi_{i,k}^x, \xi_{i,k}^v) - \widehat{\widetilde{y}}_{k/k-1}\right)^T,$$

$$P_{k,k/k-1}^{x\widetilde{y}} = \sum_{i=0}^{2M} w_i^c \left(\xi_{i,k}^x - \widehat{x}_{k/k-1}\right)\left(h_k(\xi_{i,k}^x, \xi_{i,k}^v) - \widehat{\widetilde{y}}_{k/k-1}\right)^T,$$

$$P_{k,k/k-1}^{v\widetilde{y}} = \sum_{i=0}^{2M} w_i^c \xi_{i,k}^v \left(h_k(\xi_{i,k}^x, \xi_{i,k}^v)\right)^T.$$

(3.8)

The conditional statistics of $\widetilde{y}_{k-1}$ and $\widetilde{y}_k$ are substituted in (3.5) and (3.6) to obtain those of y_k, which are used in the following equations providing the filter of X_k and the error covariance:

$$\widehat{X}_{k/k} = \widehat{X}_{k/k-1} + P_{k,k/k-1}^{Xy}\left(P_{k,k/k-1}^{yy}\right)^{-1}(y_k - \widehat{y}_{k/k-1}),$$

$$P_{k,k/k}^{XX} = P_{k,k/k-1}^{XX} - P_{k,k/k-1}^{Xy}\left(P_{k,k/k-1}^{yy}\right)^{-1}\left(P_{k,k/k-1}^{Xy}\right)^T.$$

(3.9)

The initial conditions of the proposed algorithm are given by

$$\widehat{X}_{0/0} = \begin{pmatrix} \overline{x}_0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad P_{0,0/0}^{XX} = \begin{pmatrix} P_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & Q_0 & S_1 \\ 0 & 0 & S_1^T & R_1 \end{pmatrix},$$

(3.10)

which is easily obtained from the independence hypotheses and initial conditions of the model.

Summarizing, given $\widehat{X}_{k-1/k-1}$ and $P_{k-1,k-1/k-1}^{XX}$, the computation procedure of the proposed unscented filter is as follows:

Step 1. Compute the sigma-points defined from $\widehat{X}_{k-1/k-1}$ and $P_{k-1,k-1/k-1}^{XX}$ as in (3.1), and

  (i) compute $\widehat{x}_{k/k-1}$, $P_{k,k/k-1}^{xx}$, and $P_{k,k/k-1}^{xv}$ by (3.4), and compute $\widehat{X}_{k/k-1}$ and $P_{k,k/k-1}^{XX}$ by (3.3).

  (ii) compute $\widehat{\widetilde{y}}_{k-1/k-1}$, $P_{k-1,k-1/k-1}^{\widetilde{y}\widetilde{y}}$, and $P_{k,k-1/k-1}^{x\widetilde{y}}$ by (3.7).

Step 2. Compute the sigma-points defined from $\widehat{X}_{k/k-1}$ and $P_{k,k/k-1}^{XX}$ as in (3.1), and compute $\widehat{\widetilde{y}}_{k/k-1}$, $P_{k,k/k-1}^{\widetilde{y}\widetilde{y}}$, $P_{k,k/k-1}^{x\widetilde{y}}$, and $P_{k,k/k-1}^{v\widetilde{y}}$ by (3.8).

Step 3. Compute $P_{k,k/k-1}^{X\widetilde{y}}$ and $P_{k,k-1/k-1}^{X\widetilde{y}}$ by (3.6).
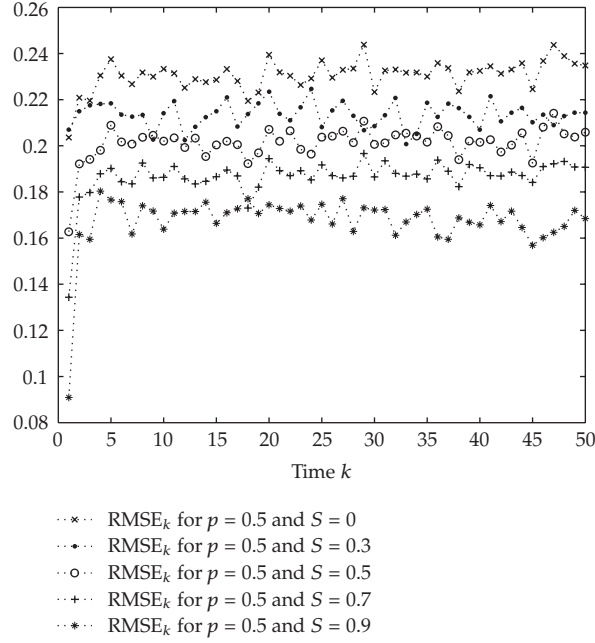
Figure 1: RMSE_k when $p = 0.5$ and $S = 0,\ 0.3,\ 0.5,\ 0.7,\ 0.9$.

Step 4. Compute $\widehat{y}_{k/k-1}$, $P^{yy}_{k,k/k-1}$ and $P^{Xy}_{k,k/k-1}$ by (3.5).

Step 5. Compute $\widehat{X}_{k/k}$ and $P^{XX}_{k,k/k}$ by (3.9).

Finally, by extracting the first block components of $\widehat{X}_{k/k}$ and $P^{XX}_{k,k/k}$, the filter of the original vector x_k and the error covariance are obtained.

4. Simulation Example

To illustrate the performance of the proposed unscented filter, we consider the following logistic type of transition and measurement equations, used previously in [9] to compare the performance of various nonlinear filters in the case of mutually independent noises and nondelayed observations:

$$x_{k+1} = \frac{\exp(x_k)}{\exp(x_k) + \exp(w_k)}, \quad k \geq 0; \qquad \widetilde{y}_k = \frac{\exp(x_k)}{\exp(x_k) + \exp(v_k)}, \quad k \geq 1, \qquad (4.1)$$

where the initial state x_0 is a random variable with uniform distribution between zero and one; the state and observation noises are assumed to be zero-mean Gaussian joint processes with $Q_k = 1$, $R_k = 1$ and known $S_k = S$, for all k.

To apply the proposed algorithm, we assume that the observations available for the estimation can be randomly delayed by one sampling period; that is,

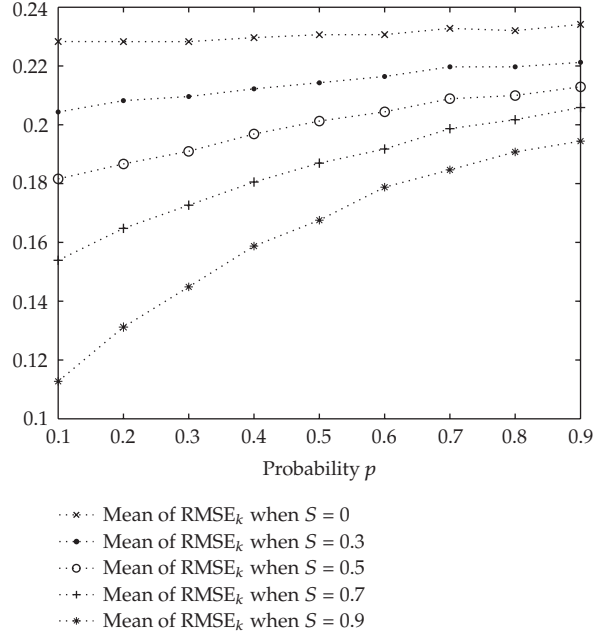$$y_k = (1 - \gamma_k)\widetilde{y}_k + \gamma_k \widetilde{y}_{k-1}, \quad k > 1; \quad y_1 = \widetilde{y}_1, \qquad (4.2)$$

Figure 2: Mean of RMSE_k when $S = 0,\ 0.3,\ 0.5,\ 0.7,\ 0.9$, versus p.

and that the noise $\{\gamma_k;\ k > 1\}$ modeling the delays is a sequence of independent Bernoulli variables with known delay probability $P(\gamma_k = 1) = p$, for all k.

We have implemented a MATLAB program which simulates the state, x_k, and the real, $\widetilde{y}_k$, and delayed measurements, y_k, for $k = 1, \ldots, 50$, for different values of S and p, and which provides the unscented filtering estimates of x_k. The root mean square error (RMSE) criterion has been used to quantify the performance of the estimates.

Considering 1000 independent simulations and denoting by $\{x_k^{(s)},\ k = 1, \ldots, 50\}$ the sth set of the artificially simulated states and by $\widehat{x}_{k/k}^{(s)}$ the filtering estimate at time k in the sth simulation run, the RMSE of the filter at time k is calculated by

$$\text{RMSE}_k = \left(\frac{1}{1000} \sum_{s=1}^{1000} \left(x_k^{(s)} - \widehat{x}_{k/k}^{(s)} \right)^2 \right)^{1/2}. \tag{4.3}$$

Let us first examine the performance of the algorithm with respect to different values of S; Figure 1 illustrates the RMSE_k when the delay probability is $p = 0.5$ and different values of S are considered; specifically, $S = 0,\ 0.3,\ 0.5,\ 0.7$ and $S = 0.9$; this figure shows, as expected, that the higher the value of S (which means that the correlation between the state and the observations increases) the smaller that of RMSE_k and, consequently, the performance of the estimators is better. Analogous results are obtained for other values of p and S.

Moreover, in order to compare the performance of the estimators as a function of the delay probability p, the arithmetic average corresponding to the 50 iterations of RMSE_k was calculated for $p = 0.1, \ldots, 0.9$. The results of this are shown in Figure 2, from which it is apparent that the means increase as p increases, the increase being greater when S is greater,

Table 1: Mean of RMSE_k for the proposed unscented filter (UF) and the extended Kalman filter (EKF) when $p = 0.3,\ 0.5,\ 0.7,\ 0.9$ and $S = 0.7,\ 0.9$.

	$S = 0.7$		$S = 0.9$	
	EKF	*UF*	*EKF*	*UF*
$p = 0.3$	0.171999	0.171981	0.156515	0.146600
$p = 0.5$	0.192260	0.185108	0.186523	0.168968
$p = 0.7$	0.209041	0.194751	0.211315	0.183530
$p = 0.9$	0.224603	0.202314	0.233059	0.195062

and consequently, as expected, the performance of the estimators deteriorates as the delay probability p rises. From this figure, it is also inferred that, for each fixed value of p, the means decrease with increasing S, which extends the result in Figure 1 to different values of p.

Finally, to compare the performance of the proposed and the EKF algorithms, the latter was applied to the observation data of the simulation example for different values of p and S. The results show that the proposed algorithm outperforms the EKF algorithm and the improvement is greater when the delay probability p is greater and, also, when the correlation S increases. Table 1, showing the means of RMSE_k for both algorithms considering $p = 0.3,\ 0.5,\ 0.7,\ 0.9$ and $S = 0.7,\ 0.9$, illustrates this fact.

Acknowledgment

References

[1] A. Ray, L. W. Liou, and J. H. Shen, "State estimation using randomly delayed measurements," *Journal of Dynamic Systems, Measurement and Control*, vol. 115, no. 1, pp. 19–26, 1993.

[2] M. Sahebsara, T. Chen, and S. L. Shah, "Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations," *International Journal of Control*, vol. 80, no. 2, pp. 292–301, 2007.

[3] Z. Wang, D. W. C. Ho, and X. Liu, "Robust filtering under randomly varying sensor delay with variance constraints," *IEEE Transactions on Circuits and Systems II*, vol. 51, no. 6, pp. 320–326, 2004.

[4] F. O. Hounkpevi and E. E. Yaz, "Minimum variance generalized state estimators for multiple sensors with different delay rates," *Signal Processing*, vol. 87, no. 4, pp. 602–613, 2007.

[5] S. Zhou and G. Feng, "H_∞ filtering for discrete-time systems with randomly varying sensor delays," *Automatica*, vol. 44, no. 7, pp. 1918–1922, 2008.

[6] A. Hermoso-Carazo and J. Linares-Pérez, "Extended and unscented filtering algorithms using one-step randomly delayed observations," *Applied Mathematics and Computation*, vol. 190, no. 2, pp. 1375–1393, 2007.

[7] A. Hermoso-Carazo and J. Linares-Pérez, "Unscented filtering algorithm using two-step randomly delayed observations in nonlinear systems," *Applied Mathematical Modelling*, vol. 33, no. 9, pp. 3705–3717, 2009.

[8] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[9] H. Tanizaki, "Nonlinear and non-Gaussian state estimation: a quasi-optimal estimator," *Communications in Statistics—Theory and Methods*, vol. 29, no. 12, pp. 2805–2834, 2000.