

Quadratic estimation for stochastic systems in the presence of random parameter matrices, time-correlated additive noise and deception attacks

R. Caballero-Águila^a, and J. Linares-Pérez^b

^aDepartamento de Estadística. Universidad de Jaén. Paraje Las Lagunillas. 23071. Jaén. Spain;

^bDepartamento de Estadística. Universidad de Granada. Avenida Fuentenueva. 18071. Granada. Spain

ARTICLE HISTORY

Compiled August 23, 2024

ABSTRACT

Networked systems usually face different random uncertainties that make the performance of the least-squares (LS) linear filter decline significantly. For this reason, great attention has been paid to the search for other kinds of suboptimal estimators. Among them, the LS quadratic estimation approach has attracted considerable interest in the scientific community for its balance between computational complexity and estimation accuracy. When it comes to stochastic systems subject to different random uncertainties and deception attacks, the quadratic estimator design has not been deeply studied. In this paper, using covariance information, the LS quadratic filtering and fixed-point smoothing problems are addressed under the assumption that the measurements are perturbed by a time-correlated additive noise, as well as affected by random parameter matrices and exposed to random deception attacks. The use of random parameter matrices covers a wide range of common uncertainties and random failures, thus better reflecting the engineering reality. The signal and observation vectors are augmented by stacking the original vectors with their second-order Kronecker powers; then, the linear estimator of the original signal based on the augmented observations provides the required quadratic estimator. A simulation example illustrates the superiority of the proposed quadratic estimators over the conventional linear ones and the effect of the deception attacks on the estimation performance.

KEYWORDS

Least-squares quadratic estimation; random parameter matrices; time-correlated additive noise; stochastic deception attacks.

1. Introduction

Sensor networks are becoming increasingly popular in a broad range of application fields, including health care, military, transportation, mining, agriculture, intelligent buildings and smart cities, among others [1]. As a result, the fusion estimation problem in networked systems is attracting tremendous research interest.

However, it must be mentioned that communication networks usually suffer resource constraints and, consequently, some network-induced phenomena will inevitably emerge during signal measurement or transmission [2]. Some of the most common networked-induced phenomena occurring in different application disciplines –presence of multiplicative noise, missing observations, or fading measurements, among others– can be globally characterized

by incorporating stochastic parameters in the measurement equations. Thus, the use of random parameter matrices in the measurement equation allows us to model the randomness in the measurements and to account for different uncertainties in many real-life scenarios – e.g., radar systems, wireless communication, sensor networks or environmental monitoring—where such uncertainties can occur. As a result, research on the estimation problem in systems with random parameter matrices has grown in popularity during the past few years. See, for example, [3]-[9] and the references therein for some sample contributions.

The traditional Kalman-type filtering problem is usually based on the assumption that the measurement additive noise is either white or finite-step correlated. In practice, however, infinite-step correlated measurement noises can be found in a wide variety of engineering applications, where the sampling frequency is typically high enough to cause measurement noises to be significantly correlated over two or more consecutive sampling periods. Over the last decade, numerous papers have addressed the estimation problem under the assumption that the infinite-step time-correlated channel noise is the output of a linear system model with white noise. The state augmentation method—which is simple and direct but computationally expensive—and the measurement differencing method—which avoids increasing dimensions, but requires two consecutive measurements to compute the difference—are the most popular methods for dealing with this type of noise correlation (see, e.g., [10]-[14]). More recently, alternative non-augmentation methods that do not require the availability of consecutive measurements have been proposed to address the state estimation problem in linear systems with time-correlated additive noises and random packet dropouts in [15] and [16].

When dealing with the estimation problem in networked systems, security is an important topic that should not be overlooked. The possibility of suffering cyber attacks is one of the most common weaknesses (see, e.g., [17] and [18]) and, particularly, the estimation problem in networked systems subject to deception attacks have inspired many significant research studies. Generally speaking, the major goal of deception attackers is altering the data integrity by maliciously falsifying their information in a random way. The centralized security-guaranteed filtering problem is studied for linear time-invariant stochastic systems with multirate-sensor fusion under deception attacks in [19]. The H_∞ -consensus filtering problem for discrete-time systems with multiplicative noises and deception attacks is investigated in [20] and the chance-constrained H_∞ state estimation problem is investigated for a class of time-varying neural networks subject to measurements degradation and randomly occurring deception attacks in [21]. The distributed estimation problem in sensor networks with a specific topology structure has been studied in [22]—under false data injection attacks—and in [23]-[26]—under deception attacks.

It is unquestionable that the LS estimation problem of random signals from noisy measurements has played a key role over the past decades. The well-known Kalman filter provides the LS signal estimator for linear systems subject to Gaussian, mutually independent initial signal and noise processes. However, in the presence of non-Gaussian disturbances, only the LS linear estimator is provided and the optimal estimator is, in general, computationally expensive. Besides, due to the network-induced random uncertainties described in the previous paragraphs, networked systems are generally non-Gaussian and the practical computation of the LS estimator usually involves a significant complexity. For this reason, a great deal of attention has been devoted to the design of simpler suboptimal estimators with satisfactory accuracy, being the design of LS linear estimation algorithms the most popular approach. A more effective scheme to address the estimation problem in non-Gaussian systems is the LS quadratic approach, due to its usual outperformance over the LS linear one and its adequate balance between computational burden and estimation accuracy. In linear discrete-time non-Gaussian systems, the input noise quadratic estimation problem is addressed in [27] and a recursive quadratic estimation algorithm for the system state is proposed in [28] under the

presence of random parameter matrices. A feedback quadratic filtering algorithm, that reduces the estimation error with respect to the plain quadratic filter, is proposed in [29]. The quadratic estimation problem has also been addressed in discrete-time systems with measurement delays and packet dropouts [30] and under the presence of multiplicative noises and quantization effects [31]. Recently, recursive quadratic estimation algorithms are proposed in [32] –for linear systems over time-correlated fading channels– and [33] –for nonlinear systems with energy-harvesting sensors. However, to the best of the authors’ knowledge, there have been scarce studies on the quadratic estimation problem in linear systems with random parameter matrices and time-correlated additive noise, let alone the scenario where random attacks are also involved.

Inspired by the discussion made so far, our aim is to address the LS quadratic filtering and fixed-point smoothing estimation problems for a class of stochastic systems in the presence of random parameter matrices, time-correlated additive noise and random deception attacks. The following are the key difficulties we are dealing with: (1) Analysis of the statistical properties of the augmented noises produced by the original time-correlated additive noises and their Kronecker products. (2) Development of an effective quadratic estimation method in the presence of random parameter matrices, time-correlated additive noise and random deception attacks. (3) Evaluation of the impact of deception attacks on estimation performance.

The main contributions of this paper are summarized as follows: (a) The class of systems investigated in this paper is quite comprehensive, as the use of random parameter matrices embraces a wide range of common uncertainties and random failures, thus better reflecting the technical reality. (b) The original observation vectors are augmented with their second-order Kronecker powers, so the quadratic estimation problem is reformulated as a linear estimation problem from the augmented observations and recursive formulas for the estimation error covariances are also proposed. (c) A covariance-based estimation approach is used, so the evolution model of the signal to be estimated does not need to be known. (d) The direct estimation of the time-correlated additive noise avoids the use of the differencing method. (e) The proposed LS quadratic filtering and fixed-point smoothing estimators outperform the conventional linear ones.

The paper is organized as follows. The characteristics of the observation model under consideration are described in Section 2. The LS quadratic estimation problem is formulated in Section 3, where the augmented vectors are defined. The study of the dynamics of the augmented vectors (subsection 3.1) and their second-order statistical properties (subsection 3.2) will be the key to obtain a new observation model, from which the recursive quadratic estimation algorithms are derived in Section 4. A simulation study, in Section 5, shows the effectiveness of the proposed filtering and fixed-point smoothing estimators, as well as their superiority over the conventional linear ones. Finally, some conclusions are given in Section 6, which is followed by three appendices that provide the mathematical proofs of the main theoretical results.

Notation. As far as possible, standard mathematical notation will be used throughout the paper. If not explicitly stated, the dimension of all vectors and matrices is assumed to be compatible with algebraic operations.

| | |
|----------------------------|---|
| \mathbb{R}^n | Set of n -dimensional real vectors |
| $\delta_{k,h}$ | Kronecker delta function |
| M^T and M^{-1} | Transpose and inverse of matrix M |
| $M^{(a)T}$ and $M^{(a)-1}$ | Shorthand for $(M^{(a)})^T$ and $(M^{(a)})^{-1}$ |
| $(M_1 \dots M_k)$ | Partitioned matrix whose blocks are the submatrices M_1, \dots, M_k |
| $Diag(N_1, \dots, N_m)$ | Block diagonal matrix with main-diagonal blocks N_1, \dots, N_m |
| 0 | Zero scalar or matrix of compatible dimension |
| \otimes | Kronecker product of matrices |
| K_{n^2} | $n^2 \times n^2$ matrix such that $K_{n^2}(z \otimes v) = v \otimes z + z \otimes v, \forall z, v \in \mathbb{R}^n$ |
| $M^{[2]} = M \otimes M$ | Second-order Kronecker power of vector or matrix M |
| $vec(\star)$ | vec operator |
| $E[a] = \bar{a}$ | Mathematical expectation of a random vector or matrix a |
| $G_k = G_{k,k}$ | Function $G_{k,h}$, depending on time instants k and h , when $h = k$ |
| $\Sigma_{k,s}^{ab}$ | Covariance of random vectors a_k and b_s ($\Sigma_{k,s}^a = \Sigma_{k,s}^{aa}$) |
| $\widehat{a}_{k/s}$ | $\Sigma_{k,s}^{ab} = Cov[a_k, b_s] = E[(a_k - \bar{a}_k)(b_s - \bar{b}_s)^T]$, $Cov[a_k] = \Sigma_k^a$ |
| | Optimal quadratic estimator of the vector a_k based on $\{y_1, \dots, y_s\}$ |

2. Observation model

Consider a random signal, $x_k \in \mathbb{R}^{n_x}$, to be estimated and assume that the actual measurements, $z_k \in \mathbb{R}^{n_z}$, are described by

$$z_k = H_k x_k + v_k, \quad k \geq 1, \quad (1)$$

where H_k are random parameter matrices and v_k is a time-correlated additive noise, satisfying

$$v_k = D_{k-1} v_{k-1} + u_{k-1}, \quad k \geq 1, \quad (2)$$

in which D_k are known, non-singular, time-varying matrices and u_k is a white noise.

Remark 1. Usually, in physical electronic systems, white noise becomes time-correlated when it passes through bandlimited channels. When the noise has a correlation time that is significantly shorter than the relevant time intervals of interest, it is generally regarded as white noise, and its colored nature is typically disregarded. However, if the correlation time is not negligible, the system may encounter significant interference from the colored noise. Model (2) is appropriate to simulate, for example, the signal strength colored noise in Global Positioning System (GPS) receivers or the GPS positioning noise due to carrier phase multipath errors (see [34] and references therein).

Let us assume that deception attacks are launched by an adversary, who injects a false signal modeled by

$$\check{z}_k = -z_k + w_k, \quad k \geq 1. \quad (3)$$

At each sampling time k , the attack can randomly succeed or fail and this fact is described by a Bernoulli random variable, λ_k , whose values –one or zero– represent a successful or failed attack, respectively. Therefore, the available observations, y_k , that will be used to estimate the signal, are given by $y_k = z_k + \lambda_k \check{z}_k$ or, equivalently, using (3),

$$y_k = (1 - \lambda_k) z_k + \lambda_k w_k, \quad k \geq 1. \quad (4)$$

Remark 2. According to (3), the false data injected by the attackers, $\check{z}_k = -z_k + w_k$, are assumed to be divided mathematically into two components: a neutralizing one, $-z_k$, that

will cancel the original measurement and a noise component, w_k , that represents the blurred deceptive information added by the attacker. Hence, at each instant of time, the compromised measurement (4) can be the actual measurement (if the attack fails) or only noise (if the attack succeeds). Another interesting kind of deception signal could be, for example, $\check{z}_k = w_k$, which involves adding random noise to the received measurement. This kind of deception attacks aims to degrade the measurement quality by introducing unwanted noise. Unlike (4), under this model, the resulting compromised measurements are given by $y_k = z_k + \lambda_k w_k$, thus always containing the actual measurement (with or without noise, depending on whether the attack is successful or not, respectively). A complete survey of different kinds of attacks can be seen, for example, in [17].

Remark 3. The mathematical model of the compromised measurements (4) looks similar to the packet loss model. The main difference between both models lies in the noise component, w_k , that represents the blurred deceptive information added by the attacker. In the random packet loss scenario, at each time instant, the processing center can receive the actual measurement (if there is no loss) or nothing (if the actual packet is lost), in which case it is usually compensated with either the most recently received packet or the prediction estimate of the lost measurement, thus always providing valuable information for the signal estimation. However, the possibility of random deception attacks means that, at each time instant, the processing center can receive the actual measurement (if the attack fails) or only noise (if the attack succeeds). Hence, in addition to neutralizing the original measurement, what could be mathematically similar to a packet dropout, the deception attack degrades the quality of the measurements by introducing unwanted noise, thus adding further difficulties in obtaining accurate estimation algorithms.

Taking into account the difficulty in obtaining the least-squares (LS) optimal estimator in the presence of random uncertainties in the measurements, most studies are focused on the LS linear estimation problem. As it is well-known, when the random processes involved in the observation model have finite second-order moments, the LS linear estimator of x_k given the observations y_1, \dots, y_L is the orthogonal projection of x_k on the space of n_x -dimensional random vectors produced as linear transformations of such observations. Nonetheless, some efforts have also been directed towards the search for estimation algorithms that keep the advantages of the linear ones –recursivity and computational simplicity– and provide more precise estimates. Specifically, our aim is to design recursive algorithms to obtain LS quadratic estimators of the signal x_k given the observations y_1, \dots, y_L ; such estimators are given by the orthogonal projection of the vector x_k onto the linear space generated by the observations y_1, \dots, y_L and their second-order Kronecker powers $y_1^{[2]}, \dots, y_L^{[2]}$. Therefore, to address the LS quadratic estimation problem, it is necessary that the second-order moments of the Kronecker powers $y_1^{[2]}, \dots, y_L^{[2]}$ exist, for which all the random processes involved must be fourth-order processes and we will assume them to have known finite second, third and fourth moments.

To simplify the statement of the subsequent assumptions and properties of the processes involved, we introduce the following notations for the covariance and cross-covariance functions of a stochastic process $\{a_k\}_{k \geq 1}$ and its second-order Kronecker powers:

$$\begin{aligned} \Sigma_{k,s}^a &= \text{Cov}[a_k, a_s], \quad \Sigma_{k,s}^{a^{(2)}} = \text{Cov}[a_k^{[2]}, a_s^{[2]}], \quad \Sigma_{k,s}^{a^{(12)}} = \text{Cov}[a_k, a_s^{[2]}], \quad \Sigma_{k,s}^{a^{(2,1)}} = \Sigma_{s,k}^{a^{(12)T}} \\ \Sigma_k^a &= \text{Cov}[a_k], \quad \Sigma_k^{a^{(2)}} = \text{Cov}[a_k^{[2]}], \quad \Sigma_k^{a^{(12)}} = \text{Cov}[a_k, a_k^{[2]}], \quad \Sigma_k^{a^{(2,1)}} = \Sigma_k^{a^{(12)T}} \end{aligned}$$

The following assumptions are made:

(H1) The signal process $\{x_k\}_{k \geq 1}$ has zero mean and its covariance function, $\Sigma_{k,s}^x$, as well as the

covariance function of its second-order powers, $\Sigma_{k,s}^{x^{(2)}}$, can be factorized as follows:

$$\Sigma_{k,s}^x = A_k B_s^T, \quad \Sigma_{k,s}^{x^{(2)}} = \check{A}_k \check{B}_s^T, \quad s \leq k,$$

where the $n_x \times M_1$ matrices A_k, B_s and the $n_x^2 \times M_2$ matrices \check{A}_k, \check{B}_s are known for all $k, s \geq 1$. Also, the cross-covariance function of the signal and its second-order powers, $\Sigma_{k,s}^{x^{(12)}}$, can be expressed as:

$$\Sigma_{k,s}^{x^{(12)}} = \begin{cases} A_{1,k} B_{1,s}^T, & s \leq k, \\ B_{2,k} A_{2,s}^T, & k \leq s, \end{cases}$$

where, for all $k, s \geq 1$, $A_{1,k}, B_{1,s}, B_{2,k}$ and $A_{2,k}$ are $n_x \times P_1, n_x^2 \times P_1, n_x \times P_2$ and $n_x^2 \times P_2$ known matrices, respectively.

- (H2) $\{H_k\}_{k \geq 1}$ is a sequence of independent random parameter matrices with known mean matrices \bar{H}_k . The covariances and cross-covariances between the entries of the matrices H_k and $H_k^{[2]}$, are also assumed to be known.
- (H3) v_0 is a zero-mean random vector whose moments up to the fourth-order one are known.
- (H4) The noise processes $\{u_k\}_{k \geq 0}$ and $\{w_k\}_{k \geq 1}$ are zero-mean white sequences with known moments, up to the fourth-order ones.
- (H5) $\{\lambda_k\}_{k \geq 1}$, is a sequence of independent Bernoulli random variables with known probabilities $P(\lambda_k = 1) = \bar{\lambda}_k, k \geq 1$.
- (H6) The signal process $\{x_k\}_{k \geq 1}$, the vector v_0 and the processes $\{H_k\}_{k \geq 1}, \{u_k\}_{k \geq 0}, \{w_k\}_{k \geq 1}$ and $\{\lambda_k\}_{k \geq 1}$ are mutually independent.

Remark 4. The derivation of the proposed quadratic estimation algorithms will not require the evolution model of the signal; instead, we will use a covariance-based estimation approach. In this approach, although the signal evolution model is not necessary, a zero-mean signal is required and the covariance and cross-covariance functions of the signal and its second-order powers are to be expressed in a separable form. It should be noted that these assumptions, imposed in (H1), are met under the most commonly used signal evolution models. For instance, let us consider a zero-mean non-stationary signal obeying a linear evolution model $x_k = \phi_{k-1} x_{k-1} + \varepsilon_{k-1}, k \geq 1$, with non-singular transition matrices, ϕ_k , and the following assumptions:

- x_0 is a zero-mean random vector whose moments, up to the fourth-order one, are known.
- $\{\varepsilon_k\}_{k \geq 0}$ is a zero-mean white process with known moments, up to the fourth-order ones.

The covariance and cross-covariance functions of the signal and their second-order powers are given by

$$\Sigma_{k,s}^x = \phi_{k,s} \Sigma_s^x, \quad \Sigma_{k,s}^{x^{(2)}} = \phi_{k,s}^{[2]} \Sigma_s^{x^{(2)}}, \quad s \leq k; \quad \Sigma_{k,s}^{x^{(12)}} = \begin{cases} \phi_{k,s} \Sigma_s^{x^{(12)}}, & s \leq k, \\ \Sigma_k^{x^{(12)}} \phi_{s,k}^{[2]T}, & k \leq s, \end{cases}$$

where $\phi_{k,s} = \phi_{k-1} \cdots \phi_s, \phi_{k,s}^{[2]} = \phi_{k-1}^{[2]} \cdots \phi_s^{[2]}$ and the functions $\Sigma_s^x, \Sigma_s^{x^{(12)}}$ and $\Sigma_s^{x^{(2)}}$ are recursively obtained by

$$\begin{aligned} \Sigma_s^x &= \phi_{s-1} \Sigma_{s-1}^x \phi_{s-1}^T + \Sigma_{s-1}^\varepsilon, \quad s \geq 1, \\ \Sigma_s^{x^{(12)}} &= \phi_{s-1} \Sigma_{s-1}^{x^{(12)}} \phi_{s-1}^{[2]T} + \Sigma_{s-1}^{\varepsilon^{(12)}}, \quad s \geq 1, \\ \Sigma_s^{x^{(2)}} &= \phi_{s-1}^{[2]} \Sigma_{s-1}^{x^{(2)}} \phi_{s-1}^{[2]T} + K_{n_x^2} (\phi_{s-1} \Sigma_{s-1}^x \phi_{s-1}^T \otimes \Sigma_{s-1}^\varepsilon) K_{n_x^2} + \Sigma_{s-1}^{\varepsilon^{(2)}}, \quad s \geq 1. \end{aligned}$$

Clearly, according to assumption (H1), these covariance functions can be expressed in a sep-

arable form taking, for example, the following functions:

$$A_k = \phi_{k,0}, B_s^T = \phi_{s,0}^{-1} \Sigma_s^x; \quad \check{A}_k = \phi_{k,0}^{[2]}, \check{B}_s^T = (\phi_{s,0}^{[2]})^{-1} \Sigma_k^{x^{(2)}};$$

$$A_{1,k} = \phi_{k,0}, B_{1,s}^T = \phi_{s,0}^{-1} \Sigma_s^{x^{(12)}}; \quad B_{2,k} = \Sigma_k^{x^{(12)}} (\phi_{k,0}^{[2]T})^{-1}, A_{2,s}^T = \phi_{s,0}^{[2]T}.$$

An analogous reasoning can be carried out for a stationary signal, as it will be shown in Section 5; hence, the separability assumptions on the signal required in (H1) covers different types of stationary and non-stationary signals and the estimation based on such hypotheses, instead of the state-space model, provides a unifying context to obtain general algorithms which are applicable to a large number of practical situations.

Remark 5. As already indicated in the Introduction, in order to deal with the time-correlated noise, we will not use the measurement differencing method, but we will instead address the direct estimation of the noise. Since the derivation of the estimation algorithms will be carried out under a covariance-based approach, it is necessary to express the noise covariance function in a separable form.

From (2) and assumptions (H3) and (H4), we have that the covariance function $\Sigma_{k,s}^v$ of the time-correlated noise, v_k , is factorized in a separable form $\Sigma_{k,s}^v = \mathbf{D}_k \mathbf{F}_s^T$, in which $\mathbf{D}_k = D_{k,0}$, $\mathbf{F}_s^T = D_{s,0}^{-1} \Sigma_s^v$, where $D_{k,0} = D_{k-1} \cdots D_0$ and Σ_s^v is recursively computed by

$$\Sigma_s^v = D_{s-1} \Sigma_{s-1}^v D_{s-1}^T + \Sigma_{s-1}^u, \quad s \geq 1.$$

3. Least-squares quadratic estimation

Given the observation model (1)-(4), under hypotheses (H1)-(H6), our aim is finding the LS quadratic estimator, $\widehat{x}_{k/L}$, of the signal, x_k , with knowledge of the observation history up to the L th sampling time, y_1, \dots, y_L . More specifically, our goal is to construct recursive algorithms for the filter $\widehat{x}_{k/k}$ and the smoother $\widehat{x}_{k/L}$, at the arbitrary fixed point k , for any $L > k$.

The filtering algorithm provides estimators of the current signal, x_k , based on the measurements up to the present time, y_1, \dots, y_k . At each time step k , the filtering estimator, $\widehat{x}_{k/k}$, is updated based on the new measurement, y_k , and the previous estimator, $\widehat{x}_{k-1/k-1}$.

The fixed-point smoothing algorithm allows us to obtain the estimator of the signal at a fixed time, given an increasing number of posterior available measurements; that is, the fixed-point smoother is used to estimate the signal, x_k , at the fixed point k , not only based on measurements up to that time, but also using measurements taken beyond it, $k+1, k+2, \dots$.

As already indicated, the quadratic estimator is the orthogonal projection of the vector x_k onto the linear space generated by the observations y_1, \dots, y_L and the Kronecker powers $y_1^{[2]}, \dots, y_L^{[2]}$. The model hypotheses ensure the existence of the second-order moments of these Kronecker powers and, consequently, the existence of the quadratic estimators $\widehat{x}_{k/L}$ is guaranteed. By combining the original vectors and their second-order powers, the following augmented vectors are defined to derive such estimators:

$$\mathcal{X}_k = \begin{pmatrix} x_k \\ x_k^{[2]} \end{pmatrix}, \quad \mathcal{Z}_k = \begin{pmatrix} z_k \\ z_k^{[2]} \end{pmatrix}, \quad \mathcal{V}_k = \begin{pmatrix} v_k \\ v_k^{[2]} \end{pmatrix}, \quad \mathcal{Y}_k = \begin{pmatrix} y_k \\ y_k^{[2]} \end{pmatrix}, \quad \mathcal{W}_k = \begin{pmatrix} w_k \\ w_k^{[2]} \end{pmatrix}.$$

Noting that the space of n_x -dimensional linear transformations of y_1, \dots, y_L and $y_1^{[2]}, \dots, y_L^{[2]}$ is identical to that of linear transformations of $\mathcal{Y}_1, \dots, \mathcal{Y}_L$, it is obvious that the LS quadratic estimator of x_k based on y_1, \dots, y_L is the LS linear estimator of x_k based on

$\mathcal{Y}_1, \dots, \mathcal{Y}_L$. To obtain this linear estimator, firstly, an observation model for the vectors \mathcal{Y}_L will be created by studying the dynamics of the augmented vectors and, secondly, the second-order statistical characteristics of the processes included in this new model are examined.

3.1. Dynamics of the augmented vectors

By using (1) and the Kronecker product properties, the following expression for $z_k^{[2]}$ is obtained:

$$z_k^{[2]} = H_k^{[2]} x_k^{[2]} + K_{n_z^2}(H_k x_k \otimes v_k) + v_k^{[2]}, \quad k \geq 1.$$

Hence, the augmented measurement vectors \mathcal{Z}_k clearly satisfy

$$\mathcal{Z}_k = \mathcal{H}_k \mathcal{X}_k + \mathcal{V}_k + \mathcal{V}_k^*, \quad k \geq 1, \quad (5)$$

in which

$$\mathcal{H}_k = \begin{pmatrix} H_k & 0 \\ 0 & H_k^{[2]} \end{pmatrix}, \quad \mathcal{V}_k^* = \begin{pmatrix} 0 \\ K_{n_z^2}(H_k x_k \otimes v_k) \end{pmatrix}.$$

In order to obtain the evolution model for the augmented noise \mathcal{V}_k in this equation, let us observe that, from (2), it is clear that

$$v_k^{[2]} = D_{k-1}^{[2]} v_{k-1}^{[2]} + K_{n_z^2}(D_{k-1} v_{k-1} \otimes u_{k-1}) + u_{k-1}^{[2]}, \quad k \geq 1.$$

Hence, the augmented noise \mathcal{V}_k is time-correlated and the following equation holds:

$$\mathcal{V}_k = \mathcal{D}_{k-1} \mathcal{V}_{k-1} + \mathcal{U}_{k-1}, \quad k \geq 1. \quad (6)$$

in which

$$\mathcal{D}_k = \begin{pmatrix} D_k & 0 \\ 0 & D_k^{[2]} \end{pmatrix}, \quad \mathcal{U}_k = \begin{pmatrix} u_k \\ K_{n_z^2}(D_k v_k \otimes u_k) + u_k^{[2]} \end{pmatrix}.$$

Finally, using (4) and taking into account that $\lambda_k(1 - \lambda_k) = 0$, the augmented observation vectors \mathcal{Y}_k can be expressed as

$$\mathcal{Y}_k = (1 - \lambda_k) \mathcal{Z}_k + \lambda_k \mathcal{W}_k, \quad k \geq 1. \quad (7)$$

Note that the augmented vectors in this new model have non-zero mean. For simplicity in both the study of the statistical properties of the augmented vectors and the derivation of estimation algorithms, the quadratic estimation problem will be addressed by defining the centered vectors:

$$\begin{aligned} \mathbf{x}_k &= \mathcal{X}_k - \overline{\mathcal{X}}_k, & \mathbf{z}_k &= \mathcal{Z}_k - \overline{\mathcal{Z}}_k, & \mathbf{v}_k &= \mathcal{V}_k - \overline{\mathcal{V}}_k, \\ \mathbf{y}_k &= \mathcal{Y}_k - \overline{\mathcal{Y}}_k, & \mathbf{w}_k &= \mathcal{W}_k - \overline{\mathcal{W}}_k, & \mathbf{u}_k &= \mathcal{U}_k - \overline{\mathcal{U}}_k. \end{aligned}$$

Clearly, the LS linear estimator of x_k based on $\mathcal{Y}_k, \dots, \mathcal{Y}_L$ is equal to the one based on

$\mathbf{y}_1, \dots, \mathbf{y}_L$. Theorem 3.1 summarizes the dynamics of the centered augmented vectors (hereafter simply referred to as augmented vectors).

Theorem 3.1. *The dynamics of the augmented vectors are specified as follows:*

(a) *The augmented measurements $\{\mathbf{z}_k\}_{k \geq 1}$ obey the following equation*

$$\mathbf{z}_k = \mathcal{H}_k \mathbf{x}_k + \mathbf{v}_k^* + \mathbf{v}_k^{**}, \quad k \geq 1, \quad (8)$$

where

$$\mathbf{v}_k^* = \mathbf{v}_k + \begin{pmatrix} 0 \\ K_{n_z^2}(\bar{H}_k \mathbf{x}_k \otimes v_k) \end{pmatrix},$$

$$\mathbf{v}_k^{**} = \begin{pmatrix} 0 \\ \bar{H}_k^{[2]} \text{vec}(A_k B_k^T) + K_{n_z^2}(\tilde{H}_k \mathbf{x}_k \otimes v_k) \end{pmatrix},$$

in which

$$\tilde{H}_k = H_k - \bar{H}_k, \quad \bar{H}_k^{[2]} = H_k^{[2]} - E[H_k^{[2]}]$$

and $\{\mathbf{v}_k\}_{k \geq 0}$ is a time-correlated noise, satisfying

$$\mathbf{v}_k = \mathcal{D}_{k-1} \mathbf{v}_{k-1} + \mathbf{u}_{k-1}, \quad k \geq 1. \quad (9)$$

(b) *The following equation holds for the augmented observation process $\{\mathbf{y}_k\}_{k \geq 1}$:*

$$\mathbf{y}_k = (1 - \lambda_k) \mathbf{z}_k + \lambda_k \mathbf{w}_k - (\lambda_k - \bar{\lambda}_k) \mathbf{g}_k, \quad k \geq 1, \quad (10)$$

where $\mathbf{g}_k = \begin{pmatrix} 0 \\ \bar{H}_k^{[2]} \text{vec}(A_k B_k^T) + \text{vec}(\mathbf{D}_k \mathbf{F}_k^T) - \text{vec}(\Sigma_k^w) \end{pmatrix}$.

Proof. See Appendix A. □

3.2. Second-order statistical properties of the augmented processes

The second-order statistical properties of the augmented processes involved in equations (8)-(10) are set out in the following propositions.

Proposition 3.2. *Under hypotheses (H1)-(H6), the processes involved in (8) satisfy the following properties:*

(a) *The augmented signal process $\{\mathbf{x}_k\}_{k \geq 1}$ has zero mean and its covariance function admits the following factorization:*

$$\Sigma_{k,s}^{\mathbf{x}} = \mathbf{A}_k \mathbf{B}_s^T, \quad s \leq k, \quad (11)$$

where

$$\mathring{\mathbf{A}}_k = \begin{pmatrix} A_k & A_{1,k} & 0 & 0 \\ 0 & 0 & A_{2,k} & \mathring{A}_k \end{pmatrix}, \quad \mathring{\mathbf{B}}_k = \begin{pmatrix} B_k & 0 & B_{2,k} & 0 \\ 0 & B_{1,k} & 0 & \mathring{B}_k \end{pmatrix}.$$

Also, the expectations $E[x_k \mathbf{x}_s^T]$ are factorized as follows:

$$E[x_k \mathbf{x}_s^T] = \begin{cases} \mathring{\mathbf{A}}_k \mathring{\mathbf{B}}_s^T, & s \leq k, \\ \mathring{\mathbf{B}}_k \mathring{\mathbf{A}}_s^T, & k \leq s, \end{cases} \quad (12)$$

where

$$\mathring{\mathbf{A}}_k = (A_k | A_{1,k} | 0 | 0), \quad \mathring{\mathbf{B}}_k = (B_k | 0 | B_{2,k} | 0).$$

(b) The noise process $\{\mathbf{v}_k^*\}_{k \geq 1}$ is a sequence of zero-mean random vectors and its covariance function $\Sigma_{k,s}^{\mathbf{v}^*}$ can be factorized as follows:

$$\Sigma_{k,s}^{\mathbf{v}^*} = \Delta_k \Upsilon_s^T, \quad s \leq k, \quad (13)$$

with $\Delta_k = (\mathbb{C}_k | \mathbb{D}_k)$, $\Upsilon_k = (\mathbb{E}_k | \mathbb{F}_k)$, $k \geq 1$, in which

$$\mathbb{C}_k = \begin{pmatrix} 0 \\ K_{n_z^2} (\overline{H}_k A_k \otimes \mathbf{D}_k) \end{pmatrix}, \quad \mathbb{D}_k = \mathcal{D}_{k,0},$$

$$\mathbb{E}_k = \begin{pmatrix} 0 \\ K_{n_z^2} (\overline{H}_k B_k \otimes \mathbf{F}_k) \end{pmatrix}, \quad \mathbb{F}_k = \Sigma_k^{\mathbf{v}} \mathcal{D}_{k,0}^{-1T},$$

where $\mathcal{D}_{k,0} = \mathcal{D}_{k-1} \cdots \mathcal{D}_0$ and $\Sigma_s^{\mathbf{v}}$ is recursively computed as follows:

$$\Sigma_s^{\mathbf{v}} = \mathcal{D}_{s-1} \Sigma_{s-1}^{\mathbf{v}} \mathcal{D}_{s-1}^T + \Sigma_{s-1}^{\mathbf{u}}, \quad s \geq 1; \quad \Sigma_0^{\mathbf{v}} = \begin{pmatrix} \Sigma_0^{\mathbf{v}} & \Sigma_0^{\mathbf{v}(12)} \\ \Sigma_0^{\mathbf{v}(21)} & \Sigma_0^{\mathbf{v}(22)} \end{pmatrix}.$$

(c) The noise process $\{\mathbf{v}_k^{**}\}_{k \geq 1}$ is a sequence of zero-mean mutually uncorrelated random vectors with covariance

$$\Sigma_k^{\mathbf{v}^{**}} = \begin{pmatrix} 0 & 0 \\ 0 & (\Sigma_k^{\mathbf{v}^{**}})_{22} \end{pmatrix},$$

in which

$$(\Sigma_k^{\mathbf{v}^{**}})_{22} = E[\widetilde{H}_k^{[2]} \text{vec}(A_k B_k^T) \text{vec}^T(A_k B_k^T) \widetilde{H}_k^{[2]T}] + K_{n_z^2} (E[\widetilde{H}_k A_k B_k^T \widetilde{H}_k] \otimes \mathbf{D}_k \mathbf{F}_k^T) K_{n_z^2}. \quad (14)$$

Proof. See Appendix B. □

Proposition 3.3. Under hypotheses (H1)-(H6), the noise processes $\{\mathbf{u}_k\}_{k \geq 0}$ and $\{\mathbf{w}_k\}_{k \geq 1}$, involved in (9) and (10), respectively, are sequences of zero-mean mutually uncorrelated ran-

dom vectors with covariance matrices given by

$$\Sigma_k^w = \begin{pmatrix} \Sigma_k^w & \Sigma_k^{w(12)} \\ \Sigma_k^{w(21)} & \Sigma_k^{w(22)} \end{pmatrix}, \quad k \geq 1; \quad \Sigma_k^u = \begin{pmatrix} \Sigma_k^u & \Sigma_k^{u(12)} \\ \Sigma_k^{u(21)} & \Sigma_k^{u(22)} \end{pmatrix}, \quad k \geq 0,$$

in which $(\Sigma_k^u)_{22} = K_{n_z^2} (D_k \text{vec}(\mathbf{D}_k \mathbf{F}_k^T) D_k^T \otimes \Sigma_k^u) K_{n_z^2} + \Sigma_k^{u[2]}$.

Moreover, the stochastic processes $\{\mathbf{x}_k\}_{k \geq 1}$, $\{\mathbf{u}_k\}_{k \geq 0}$, $\{\mathbf{w}_k\}_{k \geq 1}$, $\{\mathbf{v}_k^*\}_{k \geq 1}$ and $\{\mathbf{v}_k^{**}\}_{k \geq 1}$ are uncorrelated to each other.

Proof. The proof is omitted, as it is easily deduced from the model hypotheses. \square

Remark 6. From equations (8) and (10), together with the properties established in Propositions 3.2 and 3.3, it is immediately deduced that

$$\begin{aligned} \Sigma_k^z &= E[\mathcal{H}_k \mathbf{A}_k \mathbf{B}_k^T \mathcal{H}_k^T] + \Delta_k \Upsilon_k^T + \Sigma_k^{\check{\mathbf{v}}}, \quad k \geq 1, \\ \Sigma_k^y &= (1 - \bar{\lambda}_k) \Sigma_k^z + \bar{\lambda}_k \Sigma_k^w + \bar{\lambda}_k (1 - \bar{\lambda}_k) \mathbf{g}_k \mathbf{g}_k^T, \quad k \geq 1. \end{aligned} \tag{15}$$

4. Quadratic estimation algorithms

The above statistical properties ensure that the augmented processes involved in equations (8)-(10) all have finite second-order moments. Consequently, the existence of the LS quadratic estimator of x_k based on the original observations (or, equivalently, the LS linear estimator of x_k based on the augmented observations $\mathbf{y}_1, \dots, \mathbf{y}_L$) is guaranteed. Using an innovation approach, the following recursive algorithm for the quadratic filtering ($L = k$) and fixed-point smoothing ($L = k + N$, $N \geq 1$) estimators is deduced.

Theorem 4.1. The LS quadratic filtering estimators, $\widehat{x}_{k/k}$, and the error covariance matrices, $\widehat{\Sigma}_{k/k} = E[(x_k - \widehat{x}_{k/k})(x_k - \widehat{x}_{k/k})^T]$, are recursively obtained by

$$\widehat{x}_{k/k} = (\check{\mathbf{A}}_k | 0) \mathbf{e}_k, \quad k \geq 1, \tag{16}$$

$$\widehat{\Sigma}_{k/k} = \mathbf{A}_k \mathbf{B}_k^T - (\check{\mathbf{A}}_k | 0) \mathbf{T}_k (\check{\mathbf{A}}_k | 0)^T, \quad k \geq 1, \tag{17}$$

in which the vectors \mathbf{e}_k and the matrices \mathbf{T}_k satisfy the following recurrence relations

$$\mathbf{e}_k = \mathbf{e}_{k-1} + \mathbf{\Psi}_k \Pi_k^{-1} \mu_k, \quad k \geq 1; \quad \mathbf{e}_0 = 0, \tag{18}$$

$$\mathbf{T}_k = \mathbf{T}_{k-1} + \mathbf{\Psi}_k \Pi_k^{-1} \mathbf{\Psi}_k^T, \quad k \geq 1; \quad \mathbf{T}_0 = 0, \tag{19}$$

where the innovation μ_k is calculated by

$$\mu_k = \mathbf{y}_k - (1 - \bar{\lambda}_k) (\overline{\mathcal{H}}_k \mathbf{A}_k | \Delta_k) \mathbf{e}_{k-1}, \quad k \geq 1. \tag{20}$$

The matrices Ψ_k are given by the following expression

$$\Psi_k = (1 - \bar{\lambda}_k) \left((\overline{\mathcal{H}}_k \mathbb{B}_k | \Upsilon_k) - (\overline{\mathcal{H}}_k \mathbb{A}_k | \Delta_k) \mathbf{T}_{k-1} \right)^T, \quad k \geq 1, \quad (21)$$

and the innovation covariance matrices, Π_k , are calculated by

$$\Pi_k = \Sigma_k^y - (1 - \bar{\lambda}_k)^2 (\overline{\mathcal{H}}_k \mathbb{A}_k | \Delta_k) \mathbf{T}_{k-1} (\overline{\mathcal{H}}_k \mathbb{A}_k | \Delta_k)^T, \quad k \geq 1, \quad (22)$$

where Σ_k^y is obtained by (15).

At any fixed sampling time $k \geq 1$, by starting from the filter and its error covariance matrix as initial conditions, the LS quadratic fixed-point smoothers, $\widehat{x}_{k/k+N}$, and their error covariances, $\widehat{\Sigma}_{k/k+N}$, admit the following recursive relations

$$\widehat{x}_{k/k+N} = \widehat{x}_{k/k+N-1} + \check{\mathbf{S}}_{k,k+N}^x \Pi_{k+N}^{-1} \mu_{k+N}, \quad N \geq 1, \quad (23)$$

$$\widehat{\Sigma}_{k/k+N} = \widehat{\Sigma}_{k/k+N-1} - \check{\mathbf{S}}_{k,k+N}^x \Pi_{k+N}^{-1} \check{\mathbf{S}}_{k,k+N}^{xT}, \quad N \geq 1, \quad (24)$$

where

$$\check{\mathbf{S}}_{k,k+N}^x = (1 - \bar{\lambda}_{k+N}) \left((\check{\mathbb{B}}_k | 0) - \mathbf{M}_{k,k+N-1} \right) (\overline{\mathcal{H}}_{k+N} \mathbb{A}_{k+N} | \Delta_{k+N})^T, \quad N \geq 1, \quad (25)$$

and

$$\begin{aligned} \mathbf{M}_{k,k+N} &= \mathbf{M}_{k,k+N-1} + \check{\mathbf{S}}_{k,k+N}^x \Pi_{k+N}^{-1} \Psi_{k+N}^T, \quad N \geq 1, \\ \mathbf{M}_k &= (\check{\mathbb{A}}_k | 0) \mathbf{T}_k, \quad k \geq 1. \end{aligned} \quad (26)$$

Proof. See Appendix C. □

4.1. Computational procedure

Next, the computational procedure of the proposed quadratic filtering and fixed-point smoothing algorithms is summarized.

1) Covariance matrices of the augmented processes.

- 1a) The covariance matrices $\Sigma_{k,s}^x$, $\Sigma_{k,s}^{y*}$, Σ_k^{y**} , Σ_k^w and Σ_k^u of the augmented signal and noise processes are obtained by using the expressions established in Propositions 1 and 2.
- 1b) From the matrices obtained in 1a), the covariance matrices Σ_k^z and Σ_k^y are computed by (15).

All these covariance matrices only depend on the system model information, so they can be calculated offline, before the observations are available.

2) LS quadratic filtering recursive algorithm.

At the sampling time k , starting with the prior knowledge of the $(k-1)$ -th iteration (consequently, Ψ_{k-1} , Π_{k-1} , \mathbf{T}_{k-1} , μ_{k-1} and \mathbf{e}_{k-1} are known), the proposed quadratic filtering algorithm operates as follows:

- 2a) *Filtering error covariance matrices.* Compute Ψ_k by (21) and, from it and with Σ_k^y obtained in 1a), the innovation covariance matrix Π_k is provided by (22). Then

\mathbf{T}_k is obtained by (19) and, from it, the filtering error covariance matrices, $\widehat{\Sigma}_{k/k}$, are obtained by (17). It should be noted that these matrices do not depend on the measurements, thus providing a measure of the filter performance even before we get any observed data.

2b) *Quadratic filtering estimators.* When the new measurement \mathbf{y}_k is available, the innovation μ_k is computed by expression (20), and, from it, \mathbf{e}_k is obtained by (18). Then, the quadratic filtering estimators, $\widehat{x}_{k/k}$ are computed by (16).

3) *LS quadratic fixed-point smoothing recursive algorithm.* At any fixed sampling time $k \geq 1$, once the filter, $\widehat{x}_{k/k}$, and the filtering error covariance matrix, $\widehat{\Sigma}_{k/k}$ are available, the proposed quadratic smoothing estimators and the corresponding error covariance matrix are obtained as follows:

For $N = k + 1, k + 2, \dots$, compute the matrices $\mathbf{M}_{k,k+N-1}$ using (26) and, from these matrices, $\check{\mathbf{S}}_{k,k+N}^x$ is derived by (25); then, the smoothers $\widehat{x}_{k/k+N}$ and their error covariance matrices $\widehat{\Sigma}_{k/k+N}$ are obtained from (23) and (24), respectively.

5. Simulation study

In this section, a simulation numerical example is considered to analyze the implementation and performance of the proposed quadratic filtering and fixed-point smoothing algorithm.

AR(1) scalar signal. Consider a scalar signal process $\{x_k\}_{k \geq 1}$ generated by the following first-order autoregressive model:

$$x_k = 0.95x_{k-1} + \varepsilon_{k-1}, \quad k \geq 1,$$

where the initial signal x_0 is a zero-mean Gaussian variable with variance $\Sigma_0^x = 0.1$, and $\{\varepsilon_k\}_{k \geq 0}$ is a zero-mean white Gaussian noise with variance $\Sigma_k^\varepsilon = 0.1$, $\forall k \geq 0$.

Assuming that x_0 and the sequence $\{\varepsilon_k\}_{k \geq 0}$ are mutually independent and taking into account that the third and fourth-order moments of a zero-mean Gaussian variable with variance σ^2 are 0 and $3\sigma^4$, respectively, the covariance and cross-covariance functions of this signal and their second-order powers are given by

$$\Sigma_{k,s}^x = 0.95^{k-s} \Sigma_s^x, \quad \Sigma_{k,s}^{x(2)} = 0.95^{2(k-s)} \Sigma_s^{x(2)}, \quad s \leq k; \quad \Sigma_{k,s}^{x(12)} = 0, \quad \forall k, s,$$

where the functions Σ_s^x and $\Sigma_s^{x(2)}$ are recursively obtained by

$$\Sigma_s^x = 0.9025 \Sigma_{s-1}^x + 0.1, \quad s \geq 1,$$

$$\Sigma_s^{x(2)} = 0.8145 \Sigma_{s-1}^{x(2)} + 0.361 \Sigma_{s-1}^x + 0.02, \quad s \geq 1.$$

According to assumption (H1), it is clear that these covariance functions can be expressed in a separable form defining, for example, the following functions:

$$A_k = 0.95^k, \quad B_k = 0.95^{-k} \Sigma_k^x; \quad \check{A}_k = 0.95^{2k}, \quad \check{B}_k = 0.95^{-2k} \Sigma_k^{x(2)}; \\ A_{1,k} = B_{1,k} = 0; \quad A_{2,k} = B_{2,k} = 0.$$

Actual measurements. Assume that the real measurements of the signal, z_k , are described by the model (1) with the following parameters:

- $H_k = 0.9\theta_k$, in which $\{\theta_k\}_{k \geq 1}$ is a sequence of independent identically distributed Bernoulli random variables with probability $P(\theta_k = 1) = \bar{\theta}$. These variables model whether the signal is present ($\theta_k = 1$) or not ($\theta_k = 0$) in the actual measurements and, therefore, $\bar{\theta}$ is the probability that the observations contain the signal to be estimated.
- The noise process $\{v_k\}_{k \geq 0}$ is generated by (2) where $D_k = 0.75$, $\{u_k\}_{k \geq 0}$ is a zero-mean white Gaussian noise with $\Sigma_k^u = 0.01$, $\forall k \geq 0$, and v_0 is a zero-mean Gaussian variable with $\Sigma_0^v = 0.1$; hence,

$$\Sigma_0^{(12)} = \Sigma_k^{u(12)} = 0, \quad \Sigma_0^{v(2)} = 0.02, \quad \Sigma_k^{u(2)} = 0.0002.$$

According to the theoretical model, let us suppose that the measurements are subject to deception attacks and the signal injected by the adversaries is given by (3). The false data injection attack noise $\{w_k\}_{k \geq 1}$ is a white non-Gaussian sequence with distribution

$$P(w_k = -8) = 1/8, \quad P(w_k = 8/7) = 7/8, \quad \forall k \geq 1;$$

hence,

$$E[w_k] = 0, \quad \Sigma_k^w = 9.1429, \quad \Sigma_k^{w(12)} = -62.6939, \quad \Sigma_k^{w(2)} = 429.9009.$$

Available observations. Finally, again in line with our theoretical study, we suppose that the available observations for the estimation are given by (4), where the white sequence of Bernoulli random variables $\{\lambda_k\}_{k \geq 1}$, modeling whether the deception attacks actually succeed or not, are identically distributed with probabilities $P(\lambda_k = 1) = \bar{\lambda}$.

Our goal with this example is threefold. First, we aim at showing the feasibility and effectiveness of the proposed quadratic estimators, illustrating their performance and the superiority of the quadratic estimators over the linear ones (the linear filtering and fixed point smoothing algorithms are given in Appendix D). Second, we intend to show how the probability $\bar{\theta}$ that the signal is present in the actual measurements influence the performance of the estimators. Third, we attempt to show the effect of the successful deception attack probability $\bar{\lambda}$ over the performance of the estimators.

For this purpose, a MATLAB program has been developed to obtain the linear and quadratic estimators and, in order to quantify the estimation accuracy, the corresponding estimation error variances were calculated for different values of the probabilities $\bar{\theta}$ and $\bar{\lambda}$.

Performance of the quadratic filtering and fixed-point smoothing estimators. Considering the same fixed value 0.5 for the probabilities $\bar{\theta}$ and $\bar{\lambda}$, the error variances of the linear and quadratic estimators are calculated to compare the performance of both filtering and fixed-point smoothing estimators. The results of this comparison are displayed in Figure 1, which shows, on the one hand, that the quadratic estimators present lower error variances than the linear ones, thus confirming the superiority of the former over the latter. On the other hand, it is gathered that, for both linear and quadratic estimators, the smoothing error variances are less than the corresponding filtering ones and, also, that as the number of available observations increases, the fixed-point smoothers become more accurate. Furthermore, it is observed that the values of the fixed-point smoothing error variance decrease with increasing N , although this decrease becomes almost negligible for $N \geq 9$.

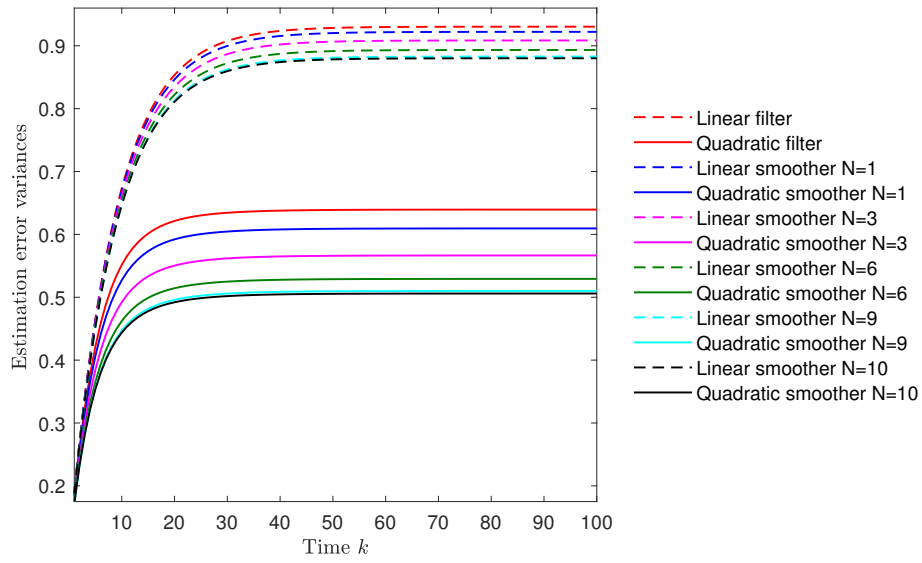


Figure 1. Estimation error variance comparison of the linear and quadratic filtering and smoothing estimators when $\bar{\theta} = \bar{\lambda} = 0.5$.

Figure 2 displays a simulated signal trajectory and their corresponding linear and quadratic filtering and smoothing estimates. Agreeing with the comments made about Figure 1, it is observed that the quadratic filtering and smoothing estimates track the signal evolution better than the linear ones. It is also noticed that the accuracy of the quadratic smoothing estimate is higher than that of the quadratic filtering estimate.

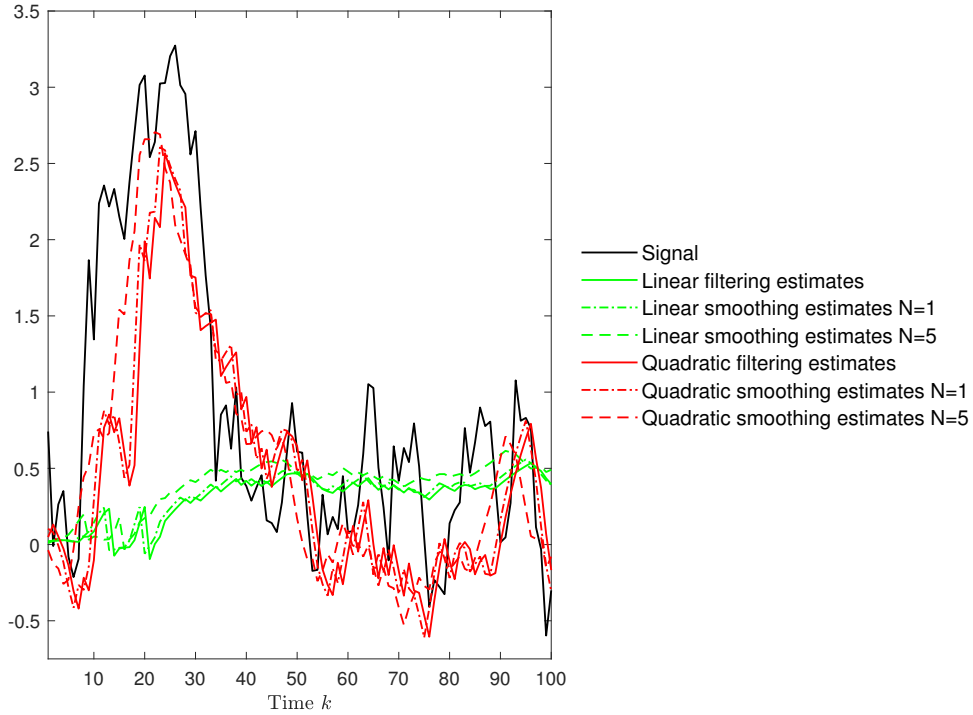


Figure 2. Simulated signal, linear and quadratic filtering and fixed-point smoothing estimates when $\bar{\theta} = \bar{\lambda} = 0.5$.

Influence of the probability $\bar{\theta}$. Assuming, as in the above figures, that the attack probability is $\bar{\lambda} = 0.5$, now we compare the performance of the estimators considering different values of the probability, $\bar{\theta}$, that the signal is present in the actual measurements. Since, from $k = 50$ onwards the estimation error variances show a similar behaviour, only the variances at a specific iteration $k = 100$ are considered. To illustrate the influence of the probability $\bar{\theta}$, Figure 3 depicts the comparative results between the filtering and smoothing error variances for both linear and quadratic estimators, considering several values of the probability $\bar{\theta}$ (namely, $\bar{\theta} = 0.1$ to 0.9). This figure shows that $\bar{\theta}$ —the probability that the observations contain the signal— or, equivalently, the probability $1 - \bar{\theta}$ that the signal is missing in the actual measurements, indeed influence the performance of the estimators. Actually, as expected, both linear and quadratic estimation error variances decrease as $\bar{\theta}$ increases and, consequently, the filtering and smoothing estimators perform better when the probability that the signal is missing in the actual measurements, $1 - \bar{\theta}$, decreases. As in Figure 1, this figure also shows that, for all the values of $\bar{\theta}$, the quadratic estimation error variances are smaller than the linear ones; besides, it is observed that the smoothing estimation error variances, for both the linear and quadratic estimators, are lower than those of the filters, and that the smoother performs better as the number of available observations increases. It is also inferred that, as the values of the probability $\bar{\theta}$ increase, a higher reduction in the estimation error variances is yielded by the quadratic filtering and smoothing estimators over the linear ones.

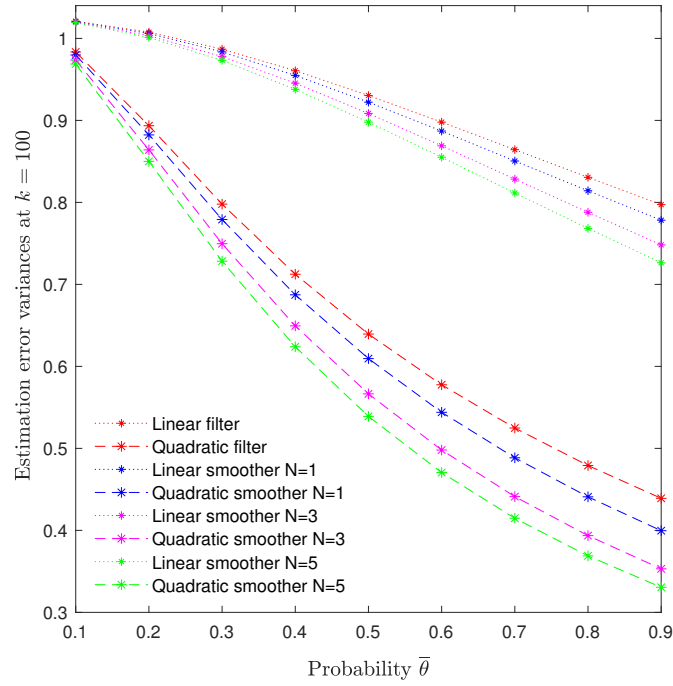


Figure 3. Linear and quadratic filtering and smoothing error variances versus $\bar{\theta}$, when $\bar{\lambda} = 0.5$.

Effect of the attack probability $\bar{\lambda}$. Assuming again, as in Figure 1, that $\bar{\theta} = 0.5$, we examine the impact of the deception attacks on the estimation accuracy. More precisely, we compare the performance of the estimators considering several values of the successful deception attack probability $\bar{\lambda} = 0.1$ to 0.9 . As $\bar{\lambda}$ increases, the number of successful attack is expected to be greater and, consequently, a higher number of available measurements used for estimation will be only noise; so, worse estimations will be obtained and, hence, the error variances are expected to be higher. Figure 4 confirms this fact, showing that the filtering and smoothing error variances at $k = 100$, of both linear and quadratic estimators, become smaller as the successful deception attack probability $\bar{\lambda}$ decreases. This figure also shows that, in the case of quadratic estimators, similar increments in the values of the probability $\bar{\lambda}$ produce essentially the same increase of the estimation error variances. However, in the linear estimation problem, such increase is more significant for small values of $\bar{\lambda}$.

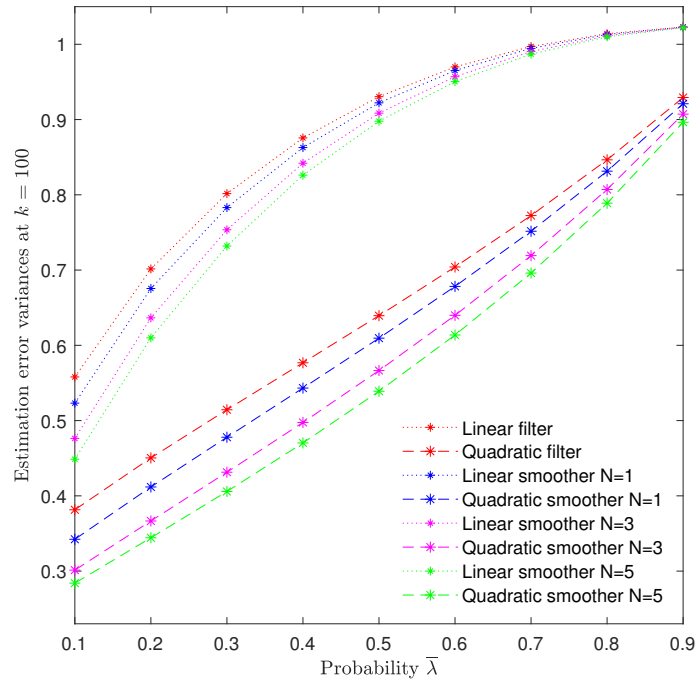


Figure 4. Linear and quadratic filtering and smoothing error variances versus $\bar{\lambda}$, when $\bar{\theta} = 0.5$.

6. Conclusion

Recursive algorithms for the LS quadratic filtering and fixed-point smoothing estimation problems are proposed from measurements perturbed by random parameter matrices, time-correlated additive noises and random deception attacks. Unlike most studies on quadratic estimation, in which the linear estimator of the augmented signal is calculated and, from it, the estimator of the original signal is extracted, we deal with the direct estimation of the original signal based on the augmented observations. Some numerical results are used to examine the accuracy of the quadratic estimators, which reveal that the proposed estimators outperform the linear ones and illustrate how the theoretical system model under consideration covers the missing measurements phenomenon as a specific example. In addition, the effect of missing measurement and deception attack success probabilities on the estimation accuracy are analyzed in the context of the numerical simulation study.

Future research topics would include extending the proposed framework to deal with more sophisticated attack models, such as the important-data-based attack model used in [35]. It would also be interesting to consider the distributed estimation problem in the scenario of networked systems whose sensor nodes are spatially distributed and connected according to a predetermined topology (see [36]).

Funding

This research was supported by the “Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación” of Spain and the European Regional Development Fund [grant number PID2021-

Appendix A. Proof of Theorem 1

(a) Taking into account that $E[H_k \mathbf{x}_k \otimes \mathbf{v}_k] = 0$, it is clear that $\overline{\mathbf{V}}_k^* = 0$; then, taking expectations in (5), we have that $\overline{\mathbf{Z}}_k = \overline{\mathcal{H}}_k \overline{\mathbf{X}}_k + \overline{\mathbf{V}}_k$, $k \geq 1$. Hence, again from (5), by adding and subtracting $\mathcal{H}_k \overline{\mathbf{X}}_k$, we obtain

$$\mathbf{z}_k = \mathcal{H}_k \mathbf{x}_k + (\mathcal{H}_k - \overline{\mathcal{H}}_k) \overline{\mathbf{X}}_k + \mathbf{v}_k + \mathbf{V}_k^* - \nu_k + \nu_k, \quad k \geq 1,$$

where $\nu_k = \begin{pmatrix} 0 \\ K_{n_z^2}(\overline{\mathcal{H}}_k \mathbf{x}_k) \otimes \nu_k \end{pmatrix}$.

Clearly, $\mathbf{v}_k^* = \mathbf{v}_k + \nu_k$ and, taking into account that

$$(\mathcal{H}_k - \overline{\mathcal{H}}_k) \overline{\mathbf{X}}_k = \begin{pmatrix} 0 \\ (\mathbf{H}_k^{[2]} - \overline{\mathbf{H}}_k^{[2]}) \text{vec}(\mathbf{A}_k \mathbf{B}_k^T) \end{pmatrix}$$

it is obvious that $\mathbf{v}_k^* = (\mathcal{H}_k - \overline{\mathcal{H}}_k) \overline{\mathbf{X}}_k + \mathbf{V}_k^* - \nu_k$; so, equation (8) is directly obtained.

Finally, taking expectations in (6), we have $\overline{\mathbf{V}}_k = \mathcal{D}_{k-1} \overline{\mathbf{V}}_{k-1} + \overline{\mathbf{U}}_{k-1}$, $k \geq 1$, and equation (9) is straightforward.

(b) Taking expectations in equation (7), it is clear that $\overline{\mathbf{Y}}_k = (1 - \bar{\lambda}_k) \overline{\mathbf{Z}}_k + \bar{\lambda}_k \overline{\mathbf{W}}_k$, $k \geq 1$, and hence

$$\mathbf{y}_k = (1 - \lambda_k) \mathbf{z}_k - (1 - \bar{\lambda}_k) \overline{\mathbf{Z}}_k + \lambda_k \mathbf{w}_k - \bar{\lambda}_k \overline{\mathbf{W}}_k, \quad k \geq 1.$$

Now, by adding and subtracting both $\lambda_k \overline{\mathbf{W}}_k$ and $(1 - \lambda_k) \overline{\mathbf{Z}}_k$, we obtain

$$\mathbf{y}_k = (1 - \lambda_k) \mathbf{z}_k + \lambda_k \mathbf{w}_k - (\lambda_k - \bar{\lambda}_k) (\overline{\mathbf{Z}}_k - \overline{\mathbf{W}}_k), \quad k \geq 1,$$

and denoting

$$\mathbf{g}_k = \overline{\mathbf{Z}}_k - \overline{\mathbf{W}}_k = \begin{pmatrix} 0 \\ \overline{\mathbf{H}}_k^{[2]} \text{vec}(\mathbf{A}_k \mathbf{B}_k^T) + \text{vec}(\mathbf{D}_k \mathbf{F}_k^T) \end{pmatrix} - \begin{pmatrix} 0 \\ \text{vec}(\Sigma_k^w) \end{pmatrix},$$

expression (10) holds.

Appendix B. Proof of Proposition 1

(a) Expression (11) for $\Sigma_{k,s}^x$ is obtained taking into account that, from hypothesis (H1),

$$\Sigma_{k,s}^x = \begin{pmatrix} \mathbf{A}_k \mathbf{B}_s^T & \mathbf{A}_{1,k} \mathbf{B}_{1,s}^T \\ \mathbf{A}_{2,k} \mathbf{B}_{2,s}^T & \check{\mathbf{A}}_k \check{\mathbf{B}}_s^T \end{pmatrix}, \quad s \leq k.$$

The factorization (12) for the expectations $E[x_k \mathbf{x}_s^T]$ is immediately deduced using that, again from hypothesis (H1),

$$E[x_k \mathbf{x}_s^T] = \begin{cases} (A_k B_s^T | A_{1,k} B_{1,s}^T), & s \leq k, \\ (B_k A_s^T | B_{2,k} B_{2,s}^T), & k \leq s. \end{cases}$$

(b) We write $\mathbf{v}_k^* = \mathbf{v}_k + v_k$, where $v_k = \begin{pmatrix} 0 \\ K_{n_z}(\bar{H}_k x_k) \otimes v_k \end{pmatrix}$.

Since, from the independence hypotheses on the model, the signal x_k and the noise v_s are uncorrelated, it is clear that \mathbf{v}_k and v_s are also uncorrelated and, hence, $\Sigma_{k,s}^{\mathbf{v}^*} = \Sigma_{k,s}^{\mathbf{v}} + \Sigma_{k,s}^v$. Next, these covariance matrices are obtained:

- From (9) and assumptions (H3) and (H4), the covariance function $\Sigma_{k,s}^{\mathbf{v}}$ is factorized in a separable form, $\Sigma_{k,s}^{\mathbf{v}} = \mathbb{D}_k \mathbb{F}_s^T$, $s \leq k$, in which $\mathbb{D}_k = \mathcal{D}_{k,0}$ and $\mathbb{F}_s^T = \mathcal{D}_{s,0}^{-1} \Sigma_s^{\mathbf{v}}$, with $\mathcal{D}_{k,0} = \mathcal{D}_{k-1} \cdots \mathcal{D}_0$. Again, from (9) and the model hypotheses, starting from $\Sigma_0^{\mathbf{v}}$, the matrices $\Sigma_s^{\mathbf{v}}$ are recursively computed by $\Sigma_s^{\mathbf{v}} = D_{s-1} \Sigma_{s-1}^{\mathbf{v}} D_{s-1}^T + \Sigma_{s-1}^{\mathbf{u}}$, $s \geq 1$.
- Taking into account that, from (H1), $E[x_k x_s^T] = A_k B_s^T$, $s \leq k$, and, from Remark 5, $E[v_k v_s^T] = \mathbf{D}_k \mathbf{F}_s^T$, $s \leq k$, the Kronecker product properties lead us to

$$\begin{aligned} E[(\bar{H}_k x_k \otimes v_k)(\bar{H}_s x_s \otimes v_s)^T] &= (\bar{H}_k A_k B_s^T \bar{H}_s^T) \otimes \mathbf{D}_k \mathbf{F}_s^T \\ &= ((\bar{H}_k A_k) \otimes \mathbf{D}_k)((\bar{H}_s B_s) \otimes \mathbf{F}_s)^T, \quad s \leq k. \end{aligned}$$

So, it is clear that $\Sigma_{k,s}^{\mathbf{v}}$ is factorized as $\Sigma_{k,s}^{\mathbf{v}} = \mathbb{C}_k \mathbb{E}_s^T$, $s \leq k$.

The above two items guarantee that $\Sigma_{k,s}^{\mathbf{v}^*} = \mathbb{D}_k \mathbb{F}_s^T + \mathbb{C}_k \mathbb{E}_s^T$, $s \leq k$, so the factorization (13) for $\Sigma_{k,s}^{\mathbf{v}^*}$ is immediately obtained.

(c) Clearly, $E[\mathbf{v}_k^{**}] = 0$. Next, using again the independence hypotheses on the model and the Kronecker product properties, we have:

- $E[(\bar{H}_k^{[2]} \text{vec}(A_k B_k^T))(\bar{H}_s^{[2]} \text{vec}(A_s B_s^T))^T] = E[\bar{H}_k^{[2]} \text{vec}(A_k B_k^T) \text{vec}^T(A_k B_k^T) \bar{H}_k^{[2]T}] \delta_{k,s}$.
- $E[(\bar{H}_k^{[2]} \text{vec}(A_k B_k^T))(\bar{H}_s x_s \otimes v_s)^T] = 0$.
- $E[(\bar{H}_k x_k \otimes v_k)(\bar{H}_s x_s \otimes v_s)^T] = E[\bar{H}_k E[x_k x_k^T] \bar{H}_k] \delta_{k,s} \otimes E[v_k v_s^T] = (E[\bar{H}_k A_k B_k^T \bar{H}_k] \otimes \mathbf{D}_k \mathbf{F}_k^T) \delta_{k,s}$.

From the above items, we conclude that, for $k \neq s$, $E[\mathbf{v}_k^{**} \mathbf{v}_s^{**T}] = 0$; so, the uncorrelation of the vectors is proven. Also, expression (14) for $(\Sigma_k^{\mathbf{v}^{**}})_{22}$, the (2, 2)-block of the matrix $\Sigma_k^{\mathbf{v}^{**}}$, is straightforward.

Appendix C. Proof of Theorem 2

Since the quadratic estimator, $\widehat{x}_{k/L}$, of the signal x_k based on the observations y_1, \dots, y_L is equal to the linear estimator of x_k based on the augmented observations $\mathbf{y}_1, \dots, \mathbf{y}_L$, according to the innovation approach it can be expressed as a linear combination of the innovations μ_1, \dots, μ_L ; namely,

$$\widehat{x}_{k/L} = \sum_{h=1}^L \check{S}_{k,h}^x \Pi_h^{-1} \mu_h, \quad k, L \geq 1, \quad (\text{C1})$$

where $\check{S}_{k,h}^x = E[x_k \mu_h^T]$, $\mu_h = \mathbf{y}_h - \widehat{\mathbf{y}}_{h/h-1}$ and $\Pi_h = E[\mu_h \mu_h^T]$. To begin with, we are going to derive a proper expression for the one-stage observation predictors $\widehat{\mathbf{y}}_{h/h-1}$, that allows

us to calculate the innovations μ_h and, from them, the coefficients $\check{S}_{k,h}^x$ and the innovation covariance matrices Π_h involved in the general expression of the estimators (C1).

Taking into account equations (8) and (10), together with the incorrelation properties established in Proposition 3.3 and the Orthogonal Projection Lemma (OPL), it is clear that

$$\widehat{\mathbf{y}}_{k/k-1} = (1 - \bar{\lambda}_k)(\overline{\mathcal{H}}_k \widehat{\mathbf{x}}_{k/k-1} + \widehat{\mathbf{v}}_{k/k-1}^*), \quad k \geq 1. \quad (\text{C2})$$

So, the one-stage predictor of both the augmented signal, $\widehat{\mathbf{x}}_{k/k-1}$, and the augmented noise, $\widehat{\mathbf{v}}_{k/k-1}^*$, must be calculated. Similarly to (C1), denoting $\mathcal{S}_{k,h}^x = E[\mathbf{x}_k \mu_h^T]$ and $\mathcal{S}_{k,h}^{y^*} = E[\mathbf{v}_k^* \mu_h^T]$, these estimators can be obtained by the following general expressions:

$$\widehat{\mathbf{x}}_{k/L} = \sum_{h=1}^L \mathcal{S}_{k,h}^x \Pi_h^{-1} \mu_h, \quad \widehat{\mathbf{v}}_{k/L}^* = \sum_{h=1}^L \mathcal{S}_{k,h}^{y^*} \Pi_h^{-1} \mu_h; \quad k, L \geq 1. \quad (\text{C3})$$

- *One-stage predictor and filter of the signal:* $\widehat{x}_{k/s}$, $s \leq k$.

From (C2), it is straightforward to see that

$$\check{S}_{k,h}^x = E[x_k \mathbf{y}_h^T] - (1 - \bar{\lambda}_h) \left(E[x_k \widehat{\mathbf{x}}_{h/h-1}^T] \overline{\mathcal{H}}_h^T + E[x_k \widehat{\mathbf{v}}_{h/h-1}^{*T}] \right), \quad 1 \leq h \leq k.$$

Now, taking into account Proposition 3.2(a) and Proposition 3.3, it is clear that

$$E[x_k \mathbf{y}_h^T] = (1 - \bar{\lambda}_h) \check{\mathbb{A}}_k \mathbb{B}_h^T \overline{\mathcal{H}}_h^T, \quad 1 \leq h \leq k,$$

and, using (C3), we can write

$$E[x_k \widehat{\mathbf{x}}_{h/h-1}^T] = \sum_{j=1}^{h-1} \check{S}_{k,j}^x \Pi_j^{-1} \mathcal{S}_{h,j}^{xT}, \quad E[x_k \widehat{\mathbf{v}}_{h/h-1}^{*T}] = \sum_{j=1}^{h-1} \check{S}_{k,j}^x \Pi_j^{-1} \mathcal{S}_{h,j}^{y^*T}, \quad h \geq 2.$$

Consequently,

$$\check{S}_{k,h}^x = (1 - \bar{\lambda}_h) \left[\check{\mathbb{A}}_k \mathbb{B}_h^T \overline{\mathcal{H}}_h^T - (1 - \delta_{h,1}) \sum_{j=1}^{h-1} \check{S}_{k,j}^x \Pi_j^{-1} (\overline{\mathcal{H}}_h \mathcal{S}_{h,j}^x + \mathcal{S}_{h,j}^{y^*})^T \right], \quad h \geq 1.$$

Hence, if we define

$$\Psi_h^x = (1 - \bar{\lambda}_h) \left[\mathbb{B}_h^T \overline{\mathcal{H}}_h^T - (1 - \delta_{h,1}) \sum_{j=1}^{h-1} \Psi_j^x \Pi_j^{-1} (\overline{\mathcal{H}}_h \mathcal{S}_{h,j}^x + \mathcal{S}_{h,j}^{y^*})^T \right], \quad h \geq 1, \quad (\text{C4})$$

we can write $\check{S}_{k,h}^x = \check{\mathbb{A}}_k \Psi_h^x$, $h \leq k$. So, denoting

$$e_k^x = \sum_{h=1}^k \Psi_h^x \Pi_h^{-1} \mu_h, \quad k \geq 1; \quad e_0^x = 0, \quad (\text{C5})$$

and using (C1), we conclude that

$$\widehat{\mathbf{x}}_{k/s} = \check{\mathbb{A}}_k \mathbf{e}_s^{\mathbf{x}}, \quad 1 \leq s \leq k. \quad (\text{C6})$$

• *One-stage predictor and filter of the augmented signal and the augmented noise: $\widehat{\mathbf{x}}_{k/s}$ and $\widehat{\mathbf{v}}_{k/s}^*$, $s \leq k$.*

Reasoning as above, it is proven that $\mathbf{S}_{k,h}^{\mathbf{x}} = \mathbb{A}_k \Psi_h^{\mathbf{x}}$, $h \leq k$, and the following expression for the augmented signal estimators is obtained:

$$\widehat{\mathbf{x}}_{k/s} = \mathbb{A}_k \mathbf{e}_s^{\mathbf{x}}, \quad 1 \leq s \leq k. \quad (\text{C7})$$

Now, by using (C2), the coefficients $\mathbf{S}_{k,h}^{\mathbf{v}^*} = E[\mathbf{v}_k^* \mu_h^T]$ can be expressed as

$$\mathbf{S}_{k,h}^{\mathbf{v}^*} = E[\mathbf{v}_k^* \mathbf{y}_h^T] - (1 - \bar{\lambda}_h) \left(E[\mathbf{v}_k^* \widehat{\mathbf{x}}_{h/h-1}^T] \bar{\mathcal{H}}_h^T + E[\mathbf{v}_k^* \widehat{\mathbf{v}}_{h/h-1}^T] \right), \quad 1 \leq h \leq k.$$

From Proposition 3.2(b) and Proposition 3.3, it is clear that

$$E[\mathbf{v}_k^* \mathbf{y}_h^T] = (1 - \bar{\lambda}_h) \Delta_k \Upsilon_h^T, \quad 1 \leq h \leq k,$$

and, in view of (C3), we can write

$$E[\mathbf{v}_k^* \widehat{\mathbf{x}}_{h/h-1}^T] = \sum_{j=1}^{h-1} \mathbf{S}_{k,j}^{\mathbf{v}^*} \Pi_j^{-1} \mathbf{S}_{h,j}^{\mathbf{x}T}, \quad E[\mathbf{v}_k^* \widehat{\mathbf{v}}_{h/h-1}^T] = \sum_{j=1}^{h-1} \mathbf{S}_{k,j}^{\mathbf{v}^*} \Pi_j^{-1} \mathbf{S}_{h,j}^{\mathbf{v}^*T}, \quad h \geq 2.$$

As a consequence,

$$\mathbf{S}_{k,h}^{\mathbf{v}^*} = (1 - \bar{\lambda}_h) \left[\Delta_k \Upsilon_h^T - (1 - \delta_{h,1}) \sum_{j=1}^{h-1} \mathbf{S}_{k,j}^{\mathbf{v}^*} \Pi_j^{-1} (\bar{\mathcal{H}}_h \mathbf{S}_{h,j}^{\mathbf{x}} + \mathbf{S}_{h,j}^{\mathbf{v}^*})^T \right], \quad h \geq 1,$$

so, we can write $\mathbf{S}_{k,h}^{\mathbf{v}^*} = \Delta_k \Psi_h^{\mathbf{v}^*}$, $h \leq k$, in which $\Psi_h^{\mathbf{v}^*}$ is a function satisfying

$$\Psi_h^{\mathbf{v}^*} = (1 - \bar{\lambda}_h) \left[\Upsilon_h^T - (1 - \delta_{h,1}) \sum_{j=1}^{h-1} \Psi_j^{\mathbf{v}^*} \Pi_j^{-1} (\bar{\mathcal{H}}_h \mathbf{S}_{h,j}^{\mathbf{x}} + \mathbf{S}_{h,j}^{\mathbf{v}^*})^T \right], \quad h \geq 1. \quad (\text{C8})$$

Hence, if we define

$$\mathbf{e}_k^{\mathbf{v}^*} = \sum_{h=1}^k \Psi_h^{\mathbf{v}^*} \Pi_h^{-1} \mu_h, \quad k \geq 1; \quad \mathbf{e}_0^{\mathbf{v}^*} = 0, \quad (\text{C9})$$

and we use (C3), it is concluded that

$$\widehat{\mathbf{v}}_{k/s}^* = \Delta_k \mathbf{e}_s^{\mathbf{v}^*}, \quad 1 \leq s \leq k. \quad (\text{C10})$$

• *Derivation of the filtering formulas: expressions (16)-(22).*

In what follows, for the sake of simplicity, we will denote:

$$\mathbf{e}_k = \begin{pmatrix} e_k^x \\ e_k^v \end{pmatrix}, \quad k \geq 0; \quad \Psi_k = \begin{pmatrix} \Psi_k^x \\ \Psi_k^v \end{pmatrix}, \quad k \geq 1,$$

from which expression (16) for the filter is immediate, just using (C6). Using the OPL, the filtering error covariances are expressed as $\widehat{\Sigma}_{k/k} = E[x_k x_k^T] - E[\widehat{x}_{k/k} \widehat{x}_{k/k}^T]$; so, using (H1) and defining $\mathbf{T}_k = E[\mathbf{e}_k \mathbf{e}_k^T]$, expression (17) is directly obtained.

The recursive relation (18) for \mathbf{e}_k is easily derived from (C5) and (C9). Then, using that \mathbf{e}_{k-1} is orthogonal to μ_k , the recursion (19) for \mathbf{T}_k is straightforward.

In view of (C2), (C7) and (C10), it is clear that

$$\widehat{\mathbf{y}}_{k/k-1} = (1 - \bar{\lambda}_k) (\overline{\mathcal{H}}_k \mathbb{A}_k | \Delta_k) \mathbf{e}_{k-1}, \quad k \geq 1, \quad (\text{C11})$$

which yields (20) for the innovation $\mu_k = \mathbf{y}_k - \widehat{\mathbf{y}}_{k/k-1}$. To obtain its covariance matrix, $\Pi_k = E[\mu_k \mu_k^T]$, we just observe that the OPL guarantees that $\Pi_k = \Sigma_k^y - E[\widehat{\mathbf{y}}_{k/k-1} \widehat{\mathbf{y}}_{k/k-1}^T]$, and using (C11), expression (22) for Π_k is easily proven.

Expression (21) is derived just by combining (C4) and (C8), using in them that $\mathbf{S}_{h,j}^x = \mathbb{A}_h \Psi_j^x$ and $\mathbf{S}_{h,j}^v = \Delta_h \Psi_j^v$ to obtain

$$\Psi_k = (1 - \bar{\lambda}_k) \left((\overline{\mathcal{H}}_k \mathbb{B}_k | \Upsilon_k) - (\overline{\mathcal{H}}_k \mathbb{A}_k | \Delta_k) \sum_{h=1}^{k-1} \Psi_h \Pi_h^{-1} \Psi_h^T \right)^T, \quad k \geq 1,$$

and taking into account that $\mathbf{T}_k = E[\mathbf{e}_k \mathbf{e}_k^T] = \sum_{h=1}^k \Psi_h \Pi_h^{-1} \Psi_h^T$, $k \geq 1$.

• *Derivation of the fixed-point smoothing formulas: (23)-(26).*

The recursive relation (23) for the fixed-point smoothers yields directly from (C1) and, from it, the recursion for the error covariance matrices (24) is immediately obtained using the OPL.

In order to calculate the coefficients

$$\check{\Sigma}_{k,k+N}^x = E[x_k \mu_{k+N}^T] = E[x_k \mathbf{y}_{k+N}^T] - E[x_k \widehat{\mathbf{y}}_{k+N/k+N-1}^T],$$

let us note that, from Proposition 3.2(b) and Proposition 3.3,

$$E[x_k \mathbf{y}_{k+N}^T] = (1 - \bar{\lambda}_{k+N}) \check{\mathbb{B}}_k \mathbb{A}_{k+N}^T \overline{\mathcal{H}}_{k+N}^T$$

Using (C11) and denoting $\mathbf{M}_{k,k+N} = E[x_k \mathbf{e}_{k+N}^T]$, $N \geq 0$, it is clear that

$$E[x_k \widehat{\mathbf{y}}_{k+N/k+N-1}^T] = (1 - \bar{\lambda}_{k+N}) \mathbf{M}_{k,k+N-1} \left(\overline{\mathcal{H}}_{k+N} \mathbb{A}_{k+N} | \Delta_{k+N} \right)^T$$

and, combining both expressions, (25) is straightforward.

Finally, using the recursive relation (18) for \mathbf{e}_{k+N} , expression (26) for the matrices $\mathbf{M}_{k,k+N}$ is directly obtained; its initial condition is also easily derived by observing that, from the OPL, $\mathbf{M}_k = E[\widehat{x}_{k/k} \mathbf{e}_k^T]$ and using (16).

Appendix D. Linear estimation algorithms

In this appendix, we present a recursive algorithm to obtain the linear filter and fixed-point smoother; its derivation, via an innovation approach, is analogous to that of Theorem 2.

The LS linear filter, $\widehat{x}_{k/k}^L$, and the error covariance matrices, $\widehat{\Sigma}_{k/k}^L$, are obtained by

$$\begin{aligned}\widehat{x}_{k/k}^L &= (A_k | 0) \mathbf{e}_k^L, \quad k \geq 1, \\ \widehat{\Sigma}_{k/k}^L &= A_k \mathbf{B}_k^T - (A_k | 0) \mathbf{T}_k^L (A_k | 0)^T, \quad k \geq 1,\end{aligned}$$

where the vectors \mathbf{e}_k^L and the matrices \mathbf{T}_k^L satisfy

$$\begin{aligned}\mathbf{e}_k^L &= \mathbf{e}_{k-1}^L + \mathbf{\Psi}_k^L \mathbf{\Xi}_k^{-1} \eta_k, \quad k \geq 1; \quad \mathbf{e}_0^L = 0, \\ \mathbf{T}_k^L &= \mathbf{T}_{k-1}^L + \mathbf{\Psi}_k^L \mathbf{\Xi}_k^{-1} (\mathbf{\Psi}_k^L)^T, \quad k \geq 1; \quad \mathbf{T}_0^L = 0,\end{aligned}$$

in which the matrices $\mathbf{\Psi}_k^L$ are calculated from

$$\mathbf{\Psi}_k^L = (1 - \bar{\lambda}_k) \left((\bar{H}_k \mathbf{B}_k | \mathbf{F}_k) - (\bar{H}_k A_k | \mathbf{D}_k) \mathbf{T}_{k-1}^L \right)^T, \quad k \geq 1.$$

The innovations η_k and their covariance matrices, $\mathbf{\Xi}_k$, are calculated by

$$\begin{aligned}\eta_k &= y_k - (1 - \bar{\lambda}_k) (\bar{H}_k A_k | \mathbf{D}_k) \mathbf{e}_{k-1}^L, \quad k \geq 1. \\ \mathbf{\Xi}_k &= \Sigma_k^y - (1 - \bar{\lambda}_k)^2 (\bar{H}_k A_k | \mathbf{D}_k) \mathbf{T}_{k-1}^L (\bar{H}_k A_k | \mathbf{D}_k)^T, \quad k \geq 1,\end{aligned}$$

where $\Sigma_k^y = (1 - \bar{\lambda}_k) (E[H_k A_k \mathbf{B}_k^T H_k^T] + \mathbf{D}_k \mathbf{F}_k^T) + \bar{\lambda}_k \Sigma_k^w$, $k \geq 1$.

The LS linear fixed-point smoothers, $\widehat{x}_{k/k+N}^L$, and their error covariances, $\widehat{\Sigma}_{k/k+N}^L$, are recursively obtained by

$$\begin{aligned}\widehat{x}_{k/k+N}^L &= \widehat{x}_{k/k+N-1}^L + \mathbf{S}_{k,k+N}^L \mathbf{\Xi}_{k+N}^{-1} \eta_{k+N}, \quad N \geq 1, \\ \widehat{\Sigma}_{k/k+N}^L &= \widehat{\Sigma}_{k/k+N-1}^L - \mathbf{S}_{k,k+N}^L \mathbf{\Xi}_{k+N}^{-1} (\mathbf{S}_{k,k+N}^L)^T, \quad N \geq 1,\end{aligned}$$

where

$$\mathbf{S}_{k,k+N}^L = (1 - \bar{\lambda}_{k+N}) \left((\mathbf{B}_k | 0) - \mathbf{M}_{k,k+N-1}^L \right) (\bar{H}_{k+N} A_{k+N} | \mathbf{D}_{k+N})^T, \quad N \geq 1,$$

with

$$\begin{aligned}\mathbf{M}_{k,k+N}^L &= \mathbf{M}_{k,k+N-1}^L + \mathbf{S}_{k,k+N}^L \mathbf{\Xi}_{k+N}^{-1} (\mathbf{\Psi}_{k+N}^L)^T, \quad N \geq 1; \\ \mathbf{M}_k^L &= (A_k | 0) \mathbf{T}_k^L, \quad k \geq 1.\end{aligned}$$

References

- [1] U. Singh, A. Abraham, A. Kaklauskas, T. Hong, *Smart Sensor Networks. Analytics, Sharing and Control*, Springer, Switzerland, 2022. <https://doi.org/10.1007/978-3-030-77214-7>.

- [2] Q. Liu, Z. Wang, X. He, *Stochastic Control and Filtering over Constrained Communication Networks*, Springer, Switzerland, 2019. <https://doi.org/10.1007/978-3-030-00157-5>.
- [3] Y. Yang, Y. Liang, Q. Pan, Y. Qin, F. Yang, Distributed fusion estimation with square-root array implementation for Markovian jump linear systems with random parameter matrices and cross-correlated noises, *Inf. Sci.* 370–371 (2016) 446–462. <https://doi.org/10.1016/j.ins.2016.08.020>.
- [4] W. Wang, J. Zhou, Optimal linear filtering design for discrete time systems with cross-correlated stochastic parameter matrices and noises, *IET Control Theory Appl.* 11 (2017), 3353–3362. <https://doi.org/10.1049/iet-cta.2017.0425>.
- [5] F. Han, H. Dong, Z. Wang, G. Li, F.E. Alsaadi, Improved Tobit Kalman filtering for systems with random parameters via conditional expectation, *Signal Process.* 147 (2018) 35–45. <http://dx.doi.org/10.1016/j.sigpro.2018.01.015>.
- [6] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez Centralized filtering and smoothing algorithms from outputs with random parameter matrices transmitted through uncertain communication channels, *Digit. Signal Process.* 85 (2019) 77–85. <https://doi.org/10.1016/j.dsp.2018.11.010>
- [7] W. Liu, X. Xie, W. Qian, X. Xu, Y. Shi, Optimal linear filtering for networked control systems with random matrices, correlated noises, and packet dropouts, *IEEE Access.* 8 (2020) 59987–59997. <http://dx.doi.org/10.1109/ACCESS.2020.2983122>.
- [8] S. Sun, Distributed optimal linear fusion predictors and filters for systems with random parameter matrices and correlated noises, *IEEE Trans. Signal Process.* 68 (2020) 1064–1074. <https://doi.org/10.1109/TSP.2020.2967180>.
- [9] R. Caballero-Águila, J. Linares-Pérez, Distributed fusion filtering for uncertain systems with coupled noises, random delays and packet loss prediction compensation, *Int. J. Syst. Sci.* 54 (2023) 371–390. <https://doi.org/10.1080/00207721.2022.2122905>.
- [10] W. Li, Y. Jia, J. Du, Distributed filtering for discrete-time linear systems with fading measurements and time-correlated noise, *Digit. Signal Process.* 60 (2017) 211–219. <https://doi.org/10.1016/j.dsp.2016.10.003>.
- [11] W. Liu, P. Shi, J.S. Pan, State estimation for discrete-time Markov jump linear systems with time-correlated and mode-dependent measurement noise, *Automatica.* 85 (2017) 9–21. <https://doi.org/10.1016/j.automatica.2017.07.025>.
- [12] H. Geng, Z. Wang, Y. Cheng, F. Alsaadi, A.M. Dobaie, State estimation under non-Gaussian Lévy and time-correlated additive sensor noises: A modified Tobit Kalman filtering approach, *Signal Process.* 154 (2019) 120–128. <https://doi.org/10.1016/j.sigpro.2018.08.005>.
- [13] W. Liu, P. Shi, Convergence of optimal linear estimator with multiplicative and time-correlated additive measurement noises, *IEEE Trans. Autom. Control* 64 (2019) 2190–2197. <https://doi.org/10.1109/TAC.2018.2869467>.
- [14] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, Networked fusion estimation with multiple uncertainties and time-correlated channel noise, *Inf. Fusion* 54 (2020) 161–171. <https://doi.org/10.1016/j.inffus.2019.07.008>.
- [15] J. Ma, S. Sun, 2020. Optimal linear recursive estimators for stochastic uncertain systems with time-correlated additive noises and packet dropout compensations, *Signal Process.* 176, 107704. <https://doi.org/10.1016/j.sigpro.2020.107704>.
- [16] R. Caballero-Águila, J. Hu, J. Linares-Pérez, 2022. Two Compensation Strategies for Optimal Estimation in Sensor Networks with Random Matrices, Time-Correlated Noises, Deception Attacks and Packet Losses, *Sensors* 22, 8505. <https://doi.org/10.3390/s22218505>.
- [17] M.S. Mahmoud, M.M. Hamdan, U.A. Baroudi, Modeling and control of Cyber-Physical

- Systems subject to cyber attacks: A survey of recent advances and challenges, *Neurocomputing*. 338 (2019) 101–115. <https://doi.org/10.1016/j.neucom.2019.01.099>.
- [18] H.S. Sánchez, D. Rotondo, T. Escobet, V. Puig, J. Quevedo, Bibliographical review on cyber attacks from a control oriented perspective, *Annu. Rev. Control*. 48 (2019) 103–128. <https://doi.org/10.1016/j.arcontrol.2019.08.002>.
- [19] Z. Wang, D. Wang, B. Shen, F.E. Alsaadi, Centralized security-guaranteed filtering in multirate-sensor fusion under deception attacks, *J. Frankl. Inst.* 355 (2018) 406–420. <https://doi.org/10.1016/j.jfranklin.2017.11.010>.
- [20] F. Han, H. Dong, Z. Wang, G. Li, Local design of distributed H_∞ -consensus filtering over sensor networks under multiplicative noises and deception attacks, *Int. J. Robust Nonlinear Control*. 29 (2019) 2296–2314. <https://doi.org/10.1002/rnc.4493>.
- [21] F. Qu, E. Tian, X. Zhao, 2022. Chance-Constrained H_∞ State Estimation for Recursive Neural Networks Under Deception Attacks and Energy Constraints: The Finite-Horizon Case, *IEEE Trans. Neural Netw. Learn. Syst.* <https://doi.org/10.1109/TNNLS.2021.3137426>.
- [22] W. Yang, Y. Zhang, G. Chen, C. Yang, L. Shi, Distributed filtering under false data injection attacks, *Automatica*. 102 (2019) 34–44. <https://doi.org/10.1016/j.automatica.2018.12.027>.
- [23] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, 2020. A two-phase distributed filtering algorithm for networked uncertain systems with fading measurements under deception attacks, *Sensors*. 20, 6445. <https://doi.org/10.3390/s20226445>.
- [24] S. Xiao, Q. Han, X. Ge, Y. Zhang, Secure distributed finite-time filtering for positive systems over sensor networks under deception attacks, *IEEE Trans. Cybern.* 50 (2020) 1200–1228. <https://doi.org/10.1109/tcyb.2019.2900478>.
- [25] Y. Ma, S. Sun, 2023. Distributed Optimal and Self-Tuning Filters Based on Compressed Data for Networked Stochastic Uncertain Systems with Deception Attacks, *Sensors*. 23, 335. <https://doi.org/10.3390/s23010335>.
- [26] L. Ma, Z. Wang, Y. Chen, X. Yi, Probability-guaranteed distributed secure estimation for nonlinear systems over sensor networks under deception attacks on innovations, *IEEE Trans. Signal Inf. Proc. Netw.* 7 (2021) 465–477. <https://doi.org/10.1109/TSIPN.2021.3097217>.
- [27] H. Zhao, C. Zhang, Non-Gaussian noise quadratic estimation for linear discrete-time time-varying systems, *Neurocomputing*. 174 (2016) 921–927. <http://dx.doi.org/10.1016/j.neucom.2015.10.015>.
- [28] R. Caballero-Águila, I. García-Garrido, J. Linares-Pérez, Quadratic estimation problem in discrete-time stochastic systems with random parameter matrices, *Appl. Math. Comp.* 273 (2016) 308–320. <http://dx.doi.org/10.1016/j.amc.2015.10.005>
- [29] F. Cacace, F. Conte, A. Germani, G. Palombo, Feedback quadratic filtering, *Automatica*. 82 (2017) 158–164. <https://doi.org/10.1016/j.automatica.2017.04.046>.
- [30] L. Li, L. Tan, X. Song, X. Yan, 2020. Quadratic Filtering for Discrete-Time Systems with Measurement Delay and Packet Dropping, *Complexity*. 2020, 1725121. <https://doi.org/10.1155/2020/1725121>.
- [31] Q. Liu, Z. Wang, Q. Han, C. Jiang, 2020. Quadratic estimation for discrete time-varying non-Gaussian systems with multiplicative noises and quantization effects, *Automatica*. 113, 108714. <https://doi.org/10.1016/j.automatica.2019.108714>.
- [32] S. Wang, Z. Wang, H. Dong, Y. Chen, F.E. Alsaadi, Recursive Quadratic Filtering for Linear Discrete Non-Gaussian Systems Over Time-Correlated Fading Channels, *IEEE Trans. Signal Process.* 70 (2022) 3343–3356. <https://doi.org/10.1109/TSP.2022.3182511>.
- [33] S. Wang, Z. Wang, H. Dong, Y. Chen, 2023. Recursive state estima-

- tion for stochastic nonlinear non-Gaussian systems using energy-harvesting sensors: A quadratic estimation approach, *Automatica*, 147, 110671. <https://doi.org/10.1016/j.automatica.2022.110671>.
- [34] Y.S. Shmaliy, S. Zhao, C.K. Ahn, Kalman and UFIR state estimation with coloured measurement noise using backward Euler method, *IET Signal Process.* 14 (2020) 64–71. <https://doi.org/10.1049/iet-spr.2019.0166>.
- [35] X. Wang, E. Tian, B. Wei, J. Liu, Novel attack-defense framework for nonlinear complex networks: An important-data-based method, *Int J Robust Nonlinear Control*, 33 (2023) 2861–2878. <https://doi.org/10.1002/rnc.6551>
- [36] Q. Liu, S. Zheng, Z. Wang, N. D. Alotaibi, F. E. Alsaadi, Distributed estimation for multi-agent systems with relative measurements and quantized communication: A feedback quadratic framework, *Int J Robust Nonlinear Control*, 33 (2023) 3164–3184. <https://doi.org/10.1002/rnc.6564>