Widely Linear Estimation for Multisensor Quaternion Systems with Mixed Uncertainties in the Observations

Jesús Navarro-Moreno, Rosa María Fernández-Alcalá, José Domingo Jiménez-López and Juan Carlos Ruiz-Molina

Department of Statistics and Operations Research, University of Jaén, 23071 Jaén, Spain.
E-mails: {jnavarro,rmfernan,jdomingo,jcruiz}@ujaen.es. Tel:+34 953212729.

Abstract

The optimal widely linear state estimation problem for quaternion systems with multiple sensors and mixed uncertainties in the observations is solved in a unified framework. For that, we devise a unified model to describe the mixed uncertainties of sensor delays, packet dropouts and uncertain observations by using three Bernoulli distributed quaternion random processes. The proposed model is valid for linear discrete-time quaternion stochastic systems measured by multiple sensors and it allows us to provide filtering, prediction and smoothing algorithms for estimating the quaternion state through a widely linear processing. Simulation results are employed to show the superior performance of such algorithms in comparison to standard widely linear methods when mixed uncertainties are present in the observations.

Keywords: Multisensor system, packet dropouts, quaternion state-space modeling, sensor delays, uncertain observations, widely linear state estimation.

1. Introduction

Sensor fusion is the process of integrating information obtained from different sensors into a consistent, accurate, and useful representation [1]. In recent years, fusion estimation problems have gained much attention due to the wide applications in multiple sensor systems such as sensor networks (see e.g. [2, 3, 4, 5]). Collaborative target tracking, distributed fault detection, control of unmanned aerial vehicles, and automated vehicle guidance technology are application fields where sensor networks have been used (see [6] and the references therein). In these applications, algorithms based on Kalman filtering have proven to be of great practical value due to their underlying state space model that accounts for observational noise.

The use of a communication network in a sensor network can lead to problems such as intermittent packet losses, delays of the communicated information, or missing observations (i.e., the data packets in the measurement contain only noise) [7]. These problems exist in data transmission by unreliable communication channels, which give rise to complicated stochastic uncertainties. Such uncertainties happen randomly and thus, they can degrade system performance and increase the difficulty of estimation. Thus, classical estimation methods cannot be used directly. Moreover, all of them can appear simultaneously in some practical systems such as wireless sensor networks [8]. The state estimation problem with mixed uncertainties, where all of the three uncertainties are admissible in the data received from the network, and based on a single sensor has been considered in [7, 9, 10] (see the references therein for the case of considering only one or two of the aforementioned uncertainties). The extension to systems with multiple sensors and mixed uncertainties of sensor delays, packet dropouts and uncertain observations has been tackled in [8].

Standard sensor fusion algorithms typically operate on real valued vector state representations while other mathematical domains offers sound representations [1]. For example, the quaternion domain has been traditionally employed to model three-dimensional rotations and orientation in a compact and computationally efficient fashion. In fact, quaternions provide mathematical robustness to represent rotations since they are immune to the gimbal lock singularity (a limitation of the Euler angle representations [11]), making them suitable for attitude control systems [12, 13, 14, 15]. Other applications where the quaternion domain has been shown to perform admirably include color image processing [16], computer graphics [17], aeronautics [18, 6], seismology [19], meteorology [20], quantum physics [21], etc.

Most estimation algorithms treat the quaternion as $\mathbb{R}^4$ and operate on that with operations for the real numbers ignoring its structure as a quaternion. It is theoretically interesting to formulate estimation algorithms entirely in quaternion arithmetic. A notable characteristic of quaternion signals is that the adequate linear processing that must be used depends on the kind of properness that the signal presents [22]. In general, the optimal linear processing is widely linear (WL), which means that we must simultaneously operate on the quaternion signal and its three involutions. In other words, the WL processing is equivalent to the linear processing in $\mathbb{R}^4$ and constitutes an adequate framework for exploiting the full second-order statistical information of quaternion-valued random signals [23]. Similar to the case of complex numbers, quaternions can also present properness properties. $\mathbb{H}$-properness or $\mathbb{C}^\eta$-properness constitute the two principal types of quaternion properness and the optimal processing is then reduced to the strictly linear (SL) or semi-widely linear (SWL) processing, respectively. SL processing ignores the involutions while SWL processing makes use of the quaternion and its involution over the pure unit quaternion η. As a consequence, such processings lead to a significative reduction in the computational burden of the estimation algorithms in comparison with the WL ones (see, e.g., [24]). This saving in computational complexity constitutes another advantage of the quaternion formulation that cannot be attained when the algorithms are formulated in the real domain.

Several fusion estimation algorithms have been proposed in the quaternion domain, such as the quaternion Kalman filter [25], the quaternion extended Kalman filter (EKF) [26], the quaternion unscented Kalman filter [1, 27] or the distributed quaternion Kalman filter [6]. Such quaternion-valued signal processing algorithms offer accurate and mathematically tractable solutions with fewer constraints than those obtained in the real domain through vector algebras. However, and to the best of our knowledge, fusion estimation problems with the three uncertainties of sensor delays, packet dropouts and uncertain observations have not been studied in the quaternion domain yet. The case of a single sensor and with only one of the aforementioned uncertainties (missing observations) for quaternion systems has been treated in [24]. Nevertheless, this last solution is suboptimal in those cases where mixed uncertainties are present in the data received from the network. For instance, in energy-constrained wireless sensor network for real-time body motion tracking it is usual that the transmitted packet might be lost and/or can suffer delay [28].

This paper aims to investigate algorithms for optimal WL estimation for quaternion systems with multiple sensors and mixed uncertainties of sensor delays, packet dropouts and uncertain observations. Firstly, we develop a unified model to describe the mixed uncertainties by using three Bernoulli distributed quaternion random processes and then, an augmented state-space model is considered. This new quaternion model assumes that each component of the quaternion is measured by different sensors, where each sensor may have different mixed uncertainties of sensor delays and/or packet dropouts and/or uncertain observations. Then, and on the basis of this unified model, we provide filtering, prediction and fixed-lag smoothing algorithms for estimating the quaternion state by using a WL process. Estimation problems with only one or two of the above uncertainties can be obtained as particular cases by assigning specific values to the Bernoulli probabilities. Additionally, the suggested algorithms can be adapted to operate on nonlinear models by means of the linearization strategy used by the EKF and to process proper signals, reducing in this last case the computational complexity involved. Finally, a numerical simulation example proves experimentally the superiority of the suggested algorithms in relation to the classical WL methods given in [25] when mixed uncertainties are present in the observations. To preserve continuity in our presentation, all proofs are deferred to an Appendix.

2. Preliminaries

The purpose of this section is to introduce the notation and basic concepts in this paper. In keeping with the notation established by [22], we use boldfaced upper case letters to denote matrices, boldfaced lower case letters for column vectors, and lightfaced letters for scalar quantities. For some special matrices we will also use boldfaced upper case italicized letters (e.g., $\boldsymbol{\mathcal{A}}$). Superscripts "$*$", "T" and "H" represent the quaternion conjugate, transpose and Hermitian, respectively. $\mathbf{I}_m$ denotes the identity matrix of dimension m and $\mathbf{0}$ a matrix zero with suitable dimensions. The notation $\mathbf{A} \in \mathbb{R}^{n \times m}$ (respectively $\mathbf{A} \in \mathbb{H}^{n \times m}$) means that $\mathbf{A}$ is a real (respectively quaternion) $n \times m$ matrix. We will remove a dimension in the case of vectors (e.g., $\mathbf{r} \in \mathbb{H}^m$ means that $\mathbf{r}$ is a m-dimensional quaternion vector). $\mathrm{E}[\cdot]$ is the expectation operator and $\mathrm{diag}(\cdot)$ is a diagonal matrix with the elements specified on the main diagonal. The $\mathrm{vec}(\cdot)$ operator for a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ stacks the columns of $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m]$ sequentially, one upon another, to form a $nm \times 1$ vector: $\mathrm{vec}(\mathbf{A}) = [\mathbf{a}_1^T, \mathbf{a}_2^T, \ldots, \mathbf{a}_m^T]^T$. $\mathbf{1}_i$ denotes a vector in $\mathbb{R}^3$ with a 1 in the ith position and 0's elsewhere and $\boldsymbol{\mathcal{I}}_{ij}$ represents an indicator matrix in $\mathbb{R}^{3 \times 3}$ that has zeros for all elements except for element $\mathcal{I}_{ij} = 1$. Finally, "$\odot$" and "$\otimes$" denote the Hadamard and the Kronecker product, respectively.

Definition 1. *A quaternion random signal is a stochastic process of the form [22]*

$$x(t) = x_r(t) + \eta x_\eta(t) + \eta' x_{\eta'}(t) + \eta'' x_{\eta''}(t), \qquad t \in T \tag{1}$$

where $x_r(t)$, $x_\eta(t)$, $x_{\eta'}(t)$ and $x_{\eta''}(t)$ are real random signals, $\{1, \eta, \eta', \eta''\}$ is a quaternion orthogonal basis and T a set of real discrete time values.

The suitable type of linear processing in the quaternion domain depends on the kind of quaternion properness [22]. The most general quaternion linear processing requires the operation on the quaternion signal and its involutions[1] over the three pure unit quaternions in an orthogonal basis $\{\eta, \eta', \eta''\}$ [23]. In other words, the complete description of the second-order statistical properties of $x(t)$ is determined by the autocorrelation of the augmented quaternion signal vector $\bar{\mathbf{x}}(t) = [x(t), x^\eta(t), x^{\eta'}(t), x^{\eta''}(t)]^T$. Such processing, based on augmented statis-

[1]The involution of $y \in \mathbb{H}$ over a pure unit quaternion ν is $y^\nu = -\nu y \nu$.

tics, is called WL processing and offers performance advantages when compared to SL processing (see, e.g., [23]).

The following relationship between the augmented vector of $x(t)$ in Eq. (1) and the real vector $\mathbf{x}^r(t) = [x_r(t), x_\eta(t), x_{\eta'}(t), x_{\eta''}(t)]^T$ can be established

$$\bar{\mathbf{x}}(t) = \boldsymbol{\mathcal{A}}\mathbf{x}^r(t) \tag{2}$$

where

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} 1 & \eta & \eta' & \eta'' \\ 1 & \eta & -\eta' & -\eta'' \\ 1 & -\eta & \eta' & -\eta'' \\ 1 & -\eta & -\eta' & \eta'' \end{bmatrix}$$

with $\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}} = 4\mathbf{I}_4$. Likewise, denote $\boldsymbol{\mathcal{T}} = \mathbf{I}_3 \otimes \boldsymbol{\mathcal{A}}$, which satisfies $\boldsymbol{\mathcal{T}}^H\boldsymbol{\mathcal{T}} = 4\mathbf{I}_{12}$.

Definition 2. *Given two random signals $x(t), y(s) \in \mathbb{H}$, the product $\star$ between them is defined as*

$$x(t) \star y(s) = x_r(t)y_r(s) + \eta x_\eta(t)y_\eta(s) + \eta' x_{\eta'}(t)y_{\eta'}(s) + \eta'' x_{\eta''}(t)y_{\eta''}(s)$$

Remark 1. *It is easy to check that the augmented vector of $x(t) \star y(s)$ is given by $\overline{\mathbf{x}(t) \star \mathbf{y}(s)} = \mathcal{D}(\mathbf{x}^r(t))\bar{\mathbf{y}}(s)$, where $\mathcal{D}(\mathbf{x}^r(t)) = \frac{1}{4}\boldsymbol{\mathcal{A}}\,\mathrm{diag}(\mathbf{x}^r(t))\boldsymbol{\mathcal{A}}^H$.*

3. Problem Formulation

Let us consider a networked system with a state $x(t)$ defined as in Eq. (1) and in which the components $z_j(t)$, $j = r, \eta, \eta', \eta''$, of the packets or measured outputs $z(t)$, are transmitted from multiple sensors to the processing center subject to possible communication failures: sensor delays, packet dropouts and uncertain observations. Assume that this system is represented through the following WL state-space model

$$\bar{\mathbf{x}}(t+1) = \mathbf{\Phi}(t)\bar{\mathbf{x}}(t) + \mathbf{\Gamma}(t)\bar{\mathbf{w}}(t) \tag{3}$$

$$\bar{\mathbf{z}}(t) = \bar{\mathbf{x}}(t) + \bar{\mathbf{v}}(t) \tag{4}$$

where $\mathbf{\Phi}(t), \mathbf{\Gamma}(t) \in \mathbb{H}^{4\times 4}$ are deterministic and $\bar{\mathbf{w}}(t), \bar{\mathbf{v}}(t) \in \mathbb{H}^4$ are mutually uncorrelated white noises with zero mean and variances $\mathbf{Q}_{\bar{\mathbf{w}}}$ and $\mathbf{Q}_{\bar{\mathbf{v}}}$. The initial state $\bar{\mathbf{x}}(0)$ is uncorrelated with $\bar{\mathbf{w}}(t)$ and $\bar{\mathbf{v}}(t)$ and satisfies that $E[\bar{\mathbf{x}}(0)] = \boldsymbol{\mu}_0$ and $E[(\bar{\mathbf{x}}(0) - \boldsymbol{\mu}_0)(\bar{\mathbf{x}}(0) - \boldsymbol{\mu}_0)^H] = \mathbf{P}_0$.

The processed measurement $y(t)$ is modeled by the following equation which describes the mixed uncertainties:

$$y(t) = \gamma_1(t)\star z(t) + \gamma_2(t)\star z(t-1) + \gamma_3(t)\star y(t-1) + [\mathit{1} - \gamma_1(t) - \gamma_2(t) - \gamma_3(t)]\star v(t) \tag{5}$$

where $\mathit{1} = 1 + \eta + \eta' + \eta''$ and $\gamma_i(t) \in \mathbb{H}$, $i = 1, 2, 3$, with components $\gamma_{j,i}(t)$, $j = r, \eta, \eta', \eta''$, are, for every time instant t, Bernoulli random variables with known parameters $p_{j,i}(t)$. As can be seen from Eq. (5), each component of $z(t)$, $z(t-1)$ and $y(t-1)$ is multiplied by a different Bernoulli variable and thus, different sensors can be affected by different mixed situations of uncertain observations and/or delays and/or packet dropouts. We assume the following hypotheses on such random variables:

1. $\gamma_i(t)$ is independent of $\gamma_k(s)$ for $t \neq s$ and $i, k = 1, 2, 3$.
2. For a given i, the components $\gamma_{j,i}(t)$, $j = r, \eta, \eta', \eta''$, are independent.
3. They satisfy either $\sum_{i=1}^{3} \gamma_{j,i}(t) = 1$ or $\sum_{i=1}^{3} \gamma_{j,i}(t) = 0$ at every time instant, i.e., if $\gamma_{j,i}(t) = 1$, then $\gamma_{j,k}(t) = 0$ for all $k \neq i$.
4. $\sum_{i=1}^{3} p_{j,i}(t) \leq 1$, $j = r, \eta, \eta', \eta''$.

Moreover, the Bernoulli random variables are independent of $\bar{\mathbf{x}}(t)$, $\bar{\mathbf{w}}(t)$ and $\bar{\mathbf{v}}(t)$. The model defined by Eq. (5) shows that:

1. The current measurement $z(t)$ is received by the estimator when $\gamma_1(t) = \mathit{1}$.
2. The one-step delayed measurement $z(t-1)$ is received when $\gamma_2(t) = \mathit{1}$.
3. The packet at time t is lost and the latest packet received previously $y(t-1)$ is used when $\gamma_3(t) = \mathit{1}$.
4. Only measurement noise is received (i.e., $y(t) = v(t)$) when $\gamma_1(t) = \gamma_2(t) = \gamma_3(t) = 0$.

This formulation allows us to assign different probabilities to the quaternion components of being affected by the same uncertainties (i.e., sensor delays, packet dropouts or uncertain observations) at every time instant t, and depending on the sensor involved in the transmission (in Section 4.4 some interesting particular cases are discussed).

From Eq. (5) and Remark 1, we get the WL observation equation

$$
\begin{aligned}
\bar{\mathbf{y}}(t) = \mathcal{D}(\boldsymbol{\gamma}_1^r(t))\bar{\mathbf{z}}(t) + \mathcal{D}(\boldsymbol{\gamma}_2^r(t))\bar{\mathbf{z}}(t-1) + \mathcal{D}(\boldsymbol{\gamma}_3^r(t))\bar{\mathbf{y}}(t-1) \\
+ \mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_1^r(t) - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))\bar{\mathbf{v}}(t) \quad (6)
\end{aligned}
$$

where $\mathbf{1} = [1,1,1,1]^T$.

By denoting $\underline{\mathbf{x}}(t) = [\bar{\mathbf{x}}^T(t), \bar{\mathbf{x}}^T(t-1), \bar{\mathbf{y}}^T(t-1)]^T$, the system defined by Eqs. (3), (4) and (6) can be represented through an augmented model in the following way

$$
\underline{\mathbf{x}}(t+1) = \tilde{\boldsymbol{\Phi}}(t)\underline{\mathbf{x}}(t) + \tilde{\boldsymbol{\Gamma}}(t)\underline{\mathbf{w}}(t) \quad (7)
$$

$$
\bar{\mathbf{y}}(t) = \tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t) + \left[\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t)), \mathcal{D}(\boldsymbol{\gamma}_2^r(t))\right] \begin{bmatrix} \bar{\mathbf{v}}(t) \\ \bar{\mathbf{v}}(t-1) \end{bmatrix} \quad (8)
$$

with
$$
\tilde{\boldsymbol{\Phi}}(t) = \begin{bmatrix} \boldsymbol{\Phi}(t) & \mathbf{0} & \mathbf{0} \\ \mathbf{I}_4 & \mathbf{0} & \mathbf{0} \\ \mathcal{D}(\boldsymbol{\gamma}_1^r(t)) & \mathcal{D}(\boldsymbol{\gamma}_2^r(t)) & \mathcal{D}(\boldsymbol{\gamma}_3^r(t)) \end{bmatrix}, \underline{\mathbf{w}}(t) = \begin{bmatrix} \bar{\mathbf{w}}(t) \\ \bar{\mathbf{v}}(t) \\ \bar{\mathbf{v}}(t-1) \end{bmatrix}, \tilde{\mathbf{H}}(t)^T = \begin{bmatrix} \mathcal{D}(\boldsymbol{\gamma}_1^r(t)) \\ \mathcal{D}(\boldsymbol{\gamma}_2^r(t)) \\ \mathcal{D}(\boldsymbol{\gamma}_3^r(t)) \end{bmatrix},
$$
and $\tilde{\boldsymbol{\Gamma}}(t) = \begin{bmatrix} \boldsymbol{\Gamma}(t) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t)) & \mathcal{D}(\boldsymbol{\gamma}_2^r(t)) \end{bmatrix}$.

Our aim is to obtain recursive algorithms for calculating the optimal (in the least-squares sense) WL estimator of $x(t)$ on the basis of the information supplied by the received measurements $\{\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(1), \dots, \bar{\mathbf{y}}(s)\}$, denoted by $\hat{x}(t|s)$, in the following estimation problems: the filtering ($t = s$), the prediction ($t > s$), and the smoothing ($t < s$) problems. To this end, we use the state-space model given by Eqs. (7)-(8) and devise optimal linear estimators for $\underline{\mathbf{x}}(t)$, denoted as $\hat{\underline{\mathbf{x}}}(t|s)$. Finally, we obtain $\hat{x}(t|s) = [1, 0, 0, \dots, 0]\hat{\underline{\mathbf{x}}}(t|s)$.

Remark 2. *The process noise in Eq.* (7) *and the measurement noise in Eq.* (8) *are correlated, and* $\underline{\mathbf{w}}(t)$ *is a colored noise. Specifically,*

$$\mathrm{E}\left[\underline{\mathbf{w}}(t)\underline{\mathbf{w}}^H(s)\right] = \mathbf{Q}_{\underline{\mathbf{w}}}\delta_{ts} + \mathbf{Q}_{\underline{\mathbf{w}}_1}\delta_{t,s+1} + \mathbf{Q}^H_{\underline{\mathbf{w}}_1}\delta_{t,s-1} \tag{9}$$

$$\mathrm{E}\left[\underline{\mathbf{w}}(t)\bar{\mathbf{v}}^H(s)\right] = \mathbf{S}\delta_{ts} + \mathbf{T}\delta_{t,s+1} \tag{10}$$

where $\mathbf{Q}_{\underline{\mathbf{w}}} = \begin{bmatrix} \mathbf{Q}_{\bar{\mathbf{w}}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\bar{\mathbf{v}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{\bar{\mathbf{v}}} \end{bmatrix}$, $\mathbf{Q}_{\underline{\mathbf{w}}_1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\bar{\mathbf{v}}} & \mathbf{0} \end{bmatrix}$, $\mathbf{S} = \begin{bmatrix} \mathbf{0} \\ \mathbf{Q}_{\bar{\mathbf{v}}} \\ \mathbf{0} \end{bmatrix}$, *and* $\mathbf{T} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{Q}_{\bar{\mathbf{v}}} \end{bmatrix}$.

4. Optimal Widely Linear Estimation

The optimal WL estimator $\hat{x}(t|s)$ of $x(t)$, which obeys the state-space model given by Eqs. (3), (4) and (6), is obtained as $\hat{x}(t|s) = [1, 0, 0, \ldots, 0]\hat{\underline{\mathbf{x}}}(t|s)$ and its error variance as $P(t|s) = [1, 0, 0, \ldots, 0]\mathbf{P}(t|s)[1, 0, 0, \ldots, 0]^T$, where $\hat{\underline{\mathbf{x}}}(t|s)$ and $\mathbf{P}(t|s)$ are the optimal linear estimators for $\underline{\mathbf{x}}(t)$ and its associated error obtained from the state-space model given by Eqs. (7)-(8). Theorems 1, 2, and 3 below provide algorithms for computing both $\hat{\underline{\mathbf{x}}}(t|s)$ and $\mathbf{P}(t|s)$ in the filtering, prediction, and smoothing problems, respectively.

The following results will be necessary for finding such estimators.

Property 1. *Denote* $\triangle\tilde{\mathbf{H}}(t) = \tilde{\mathbf{H}}(t) - \mathrm{E}[\tilde{\mathbf{H}}(t)]$ *and* $\triangle\tilde{\mathbf{\Phi}}(t) = \tilde{\mathbf{\Phi}}(t) - \mathrm{E}[\tilde{\mathbf{\Phi}}(t)]$, *then*

1. $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t) \triangle \tilde{\mathbf{H}}^H(t)]$

   $$= \frac{\boldsymbol{\mathcal{A}}}{16}\Big\{\big([1,1,1]\otimes\mathbf{I}_4\big)\Big(\mathbf{\Sigma}^r(t)\odot\Big[\boldsymbol{\mathcal{T}}^H\,\mathrm{E}[\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)]\boldsymbol{\mathcal{T}}\Big]\Big)\big([1,1,1]^T\otimes\mathbf{I}_4\big)\Big\}\boldsymbol{\mathcal{A}}^H$$

   with $\mathbf{\Sigma}^r(t) = [\mathbf{\Sigma}^r_{ij}(t)]$, $i, j = 1, 2, 3$, *and the block matrices are*

   $$\mathbf{\Sigma}^r_{ii}(t) = \mathrm{diag}\left(\mathbf{p}_i(t)\odot(\mathbf{1}-\mathbf{p}_i(t))\right)$$

   $$\mathbf{\Sigma}^r_{ij}(t) = -\,\mathrm{diag}\left(\mathbf{p}_i(t)\odot\mathbf{p}_j(t)\right), \quad i\neq j$$

   where $\mathbf{p}_i(t) = [p_{r,i}(t), p_{\eta,i}(t), p_{\eta',i}(t), p_{\eta'',i}(t)]^T$, $i = 1, 2, 3$.

2. $\mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)] = \mathbf{1}_3 \otimes \mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)]$.
3. $\mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{\Phi}}^H(t)] = \boldsymbol{\mathcal{I}}_{33} \otimes \mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)]$.

Property 2.

$$\begin{aligned}\mathrm{E}[\underline{\mathbf{x}}(t+1)\underline{\mathbf{x}}^H(t+1)] = \bar{\mathbf{\Phi}}(t)\,\mathrm{E}[\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)]\bar{\mathbf{\Phi}}^H(t) + \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{\Phi}}^H(t)] \\ + \mathbf{Q}(t) + \mathbf{\Upsilon}(t) + \mathbf{\Upsilon}^H(t)\end{aligned}$$

where $\bar{\mathbf{\Phi}}(t) = \mathrm{E}[\tilde{\mathbf{\Phi}}(t)]$*, the second summand is given in Property 1.3,*

$$\begin{aligned}\mathbf{Q}(t) = \boldsymbol{\mathcal{I}}_{11} \otimes (\mathbf{\Gamma}(t)\mathbf{Q}_{\bar{\mathbf{w}}}\mathbf{\Gamma}^H(t)) + \boldsymbol{\mathcal{I}}_{33} \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\big(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)] \\ + \mathrm{E}[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T]\big) \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big)\end{aligned} \quad (11)$$

$$\begin{aligned}\mathbf{\Upsilon}(t) = \boldsymbol{\mathcal{I}}_{33} \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[\boldsymbol{\gamma}_3^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \\ \odot \Big[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\Big]\Big\}\boldsymbol{\mathcal{A}}^H\Big)\end{aligned} \quad (12)$$

with $\mathbf{Q}_{\mathbf{v}^r} = \mathrm{E}[\mathbf{v}^r(t)\mathbf{v}^{rT}(t)]$ *and initial conditions* $\mathbf{\Upsilon}(0) = \mathbf{0}$, $\mathrm{E}[\underline{\mathbf{x}}(0)\underline{\mathbf{x}}^H(0)] = \begin{bmatrix}\mathbf{P}_0 + \boldsymbol{\mu}_0\boldsymbol{\mu}_0^H & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}$*, and* $\mathbf{Q}(0) = \boldsymbol{\mathcal{I}}_{11} \otimes \big(\mathbf{\Gamma}(0)\mathbf{Q}_{\bar{\mathbf{w}}}\mathbf{\Gamma}^H(0)\big) + \boldsymbol{\mathcal{I}}_{33} \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\mathrm{E}[(\mathbf{1} - \boldsymbol{\gamma}_2^r(0) - \boldsymbol{\gamma}_3^r(0))(\mathbf{1} - \boldsymbol{\gamma}_2^r(0) - \boldsymbol{\gamma}_3^r(0))^T] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big)$.

4.1. Widely Linear Filter

Theorem 1. *The optimal linear filter obtained from the system defined by Eqs.* (7)-(8) *is calculated through the following equations*

$$\hat{\underline{\mathbf{x}}}(t|t) = \hat{\underline{\mathbf{x}}}(t|t-1) + \mathbf{K}(t)\boldsymbol{\varepsilon}(t) \quad (13)$$

$$\hat{\underline{\mathbf{x}}}(t+1|t) = \bar{\mathbf{\Phi}}(t)\hat{\underline{\mathbf{x}}}(t|t-1) + \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\boldsymbol{\varepsilon}(t-1) + \mathbf{L}(t)\boldsymbol{\varepsilon}(t) \quad (14)$$

$$\boldsymbol{\varepsilon}(t) = \bar{\mathbf{y}}(t) - \bar{\mathbf{H}}(t)\hat{\underline{\mathbf{x}}}(t|t-1) - \mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\mathbf{F}(t-1)\boldsymbol{\varepsilon}(t-1) \quad (15)$$

$$\mathbf{G}(t-1) = \mathbf{T}\,\mathcal{D}\left(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)]\right)\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t-1) \quad (16)$$

$$\mathbf{F}(t-1) = \mathbf{Q}_{\bar{\mathbf{v}}}\,\mathcal{D}\left(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)]\right)\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t-1) \quad (17)$$

with initial conditions $\hat{\underline{\mathbf{x}}}(0|-1) = \left[\boldsymbol{\mu}_0^H, \mathbf{0}^T\right]^H$, $\mathbf{G}(-1) = \mathbf{0}$ *y* $\mathbf{F}(-1) = \mathbf{0}$. $\boldsymbol{\varepsilon}(t)$ *are the innovations with covariance matrix* $\mathbf{Q}_{\boldsymbol{\varepsilon}}(t)$ *and initial condition* $\boldsymbol{\varepsilon}(-1) =$

$\mathbf{0}$. *Moreover,* $\bar{\mathbf{H}}(t)$, $\bar{\mathbf{\Phi}}(t)$, *and* $\bar{\mathbf{\Gamma}}(t)$ *are given in Eqs.* (32), (35), *and* (39), *respectively.*

The variance in the error associated with $\hat{\underline{\mathbf{x}}}(t|t)$ *is computed by*

$$\mathbf{P}(t|t) = \mathbf{P}(t|t-1) - \mathbf{K}(t)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t)\mathbf{K}^H(t) \tag{18}$$

$$\begin{aligned}\mathbf{P}(t+1|t) &= \mathbf{M}(t)\,\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t) \triangle \tilde{\mathbf{H}}^H(t)]\mathbf{M}^H(t) \\ &+ [\bar{\mathbf{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]\mathbf{P}(t|t-1)[\bar{\mathbf{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]^H + \mathbf{Q}(t) + \mathbf{\Pi}(t) + \mathbf{\Pi}^H(t) \\ &+ \mathbf{L}(t)\boldsymbol{\mathcal{A}}\Big\{\Big(\mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] \\ &+\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right]\Big)\odot\mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\mathbf{L}^H(t) - \Big[\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) - \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\Big] \\ &\times \mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\Big[\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) - \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\Big]^H\end{aligned} \tag{19}$$

where $\mathbf{M}(t) = \left[\mathbf{0}_{4\times 8}, \mathbf{I}_4\right]^T - \mathbf{L}(t)$, $\mathbf{Q}(t)$ *is given in Eq.* (11) *and*

$$\begin{aligned}\mathbf{\Pi}(t) &= \mathbf{M}(t)\Big[\mathbf{1}_3^T \otimes \boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\ &\odot\left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\Big] \\ &- \mathbf{M}(t)\boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\ &\odot\left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\mathbf{L}^H(t) \\ &+ [\bar{\mathbf{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]\Big[\bar{\mathbf{\Gamma}}(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H \\ &- \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{T}^H\Big]\bar{\mathbf{\Gamma}}^H(t) \\ &- [\bar{\mathbf{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]\Big[\bar{\mathbf{\Gamma}}(t-1)\mathbf{S} \\ &- \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\bar{\mathbf{v}}}\Big]\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{L}^H(t) \\ &- \Big(\mathbf{1}_3 \otimes \Big[\boldsymbol{\mathcal{A}}\Big\{\Big(\mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] \\ &+ \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right]\Big)\odot\mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big]\Big)\mathbf{L}^H(t)\end{aligned}$$

The gain matrices $\mathbf{K}(t)$ *and* $\mathbf{L}(t)$ *are computed by*

$$\begin{aligned}\mathbf{K}(t) &= \Big[\mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t) + \Big(\mathbf{1}_3 \otimes \left[\mathcal{D}\left(\mathbf{1} - \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t-1)\right] - \mathrm{E}\left[\boldsymbol{\gamma}_3^r(t-1)\right]\right)\mathbf{Q}_{\bar{\mathbf{v}}}\right] \\ &- \mathbf{L}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\Big)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\Big]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t)\end{aligned} \tag{20}$$

$$
\begin{aligned}
\mathbf{L}(t) = \Big\{ & \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t) \triangle \tilde{\mathbf{H}}^H(t)] + \bar{\mathbf{\Phi}}(t)\mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t) \\
+\mathbf{1}_3\otimes\Big[&\boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[\boldsymbol{\gamma}_3^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \Big[\mathrm{diag}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\Big]\Big\}\boldsymbol{\mathcal{A}}^H\Big] \\
& - \bar{\mathbf{\Phi}}(t)\mathbf{L}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right) \\
+\mathbf{1}_3\otimes\boldsymbol{\mathcal{A}}\Big\{ & \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)(\boldsymbol{\gamma}_3^r(t)-\mathrm{E}[\boldsymbol{\gamma}_3^r(t)])^T\right]\odot\left[\mathbf{Q}_{\mathbf{v}^r}\,\mathrm{diag}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\right]\Big\}\boldsymbol{\mathcal{A}}^H \\
+\bar{\mathbf{\Gamma}}(t)\Big[&\mathbf{Q}_{\underline{\mathbf{w}}_1}\bar{\mathbf{\Gamma}}^H(t-1) - \mathbf{T}\,\mathcal{D}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{L}^H(t-1)\Big]\bar{\mathbf{H}}^H(t) \\
+\mathbf{1}_3\otimes\boldsymbol{\mathcal{A}}\Big\{&\Big(\mathrm{E}\left[(\mathbf{1}-\boldsymbol{\gamma}_2^r(t)-\boldsymbol{\gamma}_3^r(t))(\mathbf{1}-\boldsymbol{\gamma}_2^r(t)-\boldsymbol{\gamma}_3^r(t))^T\right]+\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right]\Big) \\
&\odot\mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H - \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\Big\}\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t) \qquad (21)
\end{aligned}
$$

and where the innovation covariance matrix is obtained as

$$
\begin{aligned}
\mathbf{Q}_{\boldsymbol{\varepsilon}}(t) = {} & \mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)] + \bar{\mathbf{H}}(t)\mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t) \\
& + \boldsymbol{\mathcal{A}}\Big\{\Big(\mathrm{E}\left[(\mathbf{1}-\boldsymbol{\gamma}_2^r(t)-\boldsymbol{\gamma}_3^r(t))(\mathbf{1}-\boldsymbol{\gamma}_2^r(t)-\boldsymbol{\gamma}_3^r(t))^T\right] \\
& \qquad + \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right]\Big)\odot\mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H \\
& - \mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\mathbf{F}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]) + \mathbf{\Xi}(t) + \mathbf{\Xi}^H(t) \qquad (22)
\end{aligned}
$$

with

$$
\begin{aligned}
\mathbf{\Xi}(t) = {} & \boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t)-\mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right]\odot\left[\mathrm{diag}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H \\
& + \bar{\mathbf{H}}(t)\Big[\bar{\mathbf{\Gamma}}(t-1)\mathbf{S} - \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\bar{\mathbf{v}}}\Big]\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])
\end{aligned}
$$

and with initial conditions $\mathbf{P}(0|-1) = \boldsymbol{\mathcal{I}}_{11}\otimes\mathbf{P}_0$, $\mathbf{L}(-1) = \mathbf{0}$ *and* $\mathbf{Q}_{\boldsymbol{\varepsilon}}(-1) = \mathbf{0}$.

4.2. Widely Linear Predictor

Theorem 2. *The optimal linear predictor obtained from the system defined by Eqs.* (7)-(8) *is computed as*

$$
\hat{\underline{\mathbf{x}}}(t+2|t) = \bar{\mathbf{\Phi}}(t+1)\hat{\underline{\mathbf{x}}}(t+1|t) + \bar{\mathbf{\Gamma}}(t+1)\mathbf{G}(t)\boldsymbol{\varepsilon}(t) \qquad (23)
$$

$$
\hat{\underline{\mathbf{x}}}(t+\tau|t) = \bar{\mathbf{\Phi}}(t+\tau-1)\hat{\underline{\mathbf{x}}}(t+\tau-1|t), \qquad \tau > 2 \qquad (24)
$$

The values of its prediction error covariance matrix are updated through

$$
\begin{aligned}
\mathbf{P}(t+2|t) = {} & \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t+1)\underline{\mathbf{x}}(t+1)\underline{\mathbf{x}}^H(t+1)\triangle\tilde{\mathbf{\Phi}}^H(t+1)] \\
& + \bar{\mathbf{\Phi}}(t+1)\mathbf{P}(t+1|t)\bar{\mathbf{\Phi}}^H(t) + \mathbf{Q}(t+1) + \mathbf{\Upsilon}(t+1) + \mathbf{\Upsilon}^H(t+1) \\
& - \bar{\mathbf{\Gamma}}(t+1)\mathbf{G}(t)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t)\mathbf{G}^H(t)\bar{\mathbf{\Gamma}}^H(t+1) \\
& - \bar{\mathbf{\Phi}}(t+1)\mathbf{L}(t)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\mathbf{T}^H\bar{\mathbf{\Gamma}}^H(t+1) \\
& - \bar{\mathbf{\Gamma}}(t+1)\mathbf{T}\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\mathbf{L}^H(t)\bar{\mathbf{\Phi}}^H(t+1) \qquad (25)
\end{aligned}
$$

with $\mathbf{Q}(t+1)$ *given in Eq.* (11) *and* $\mathbf{\Upsilon}(t+1)$ *in Eq.* (12)*. Likewise, for* $\tau > 2$*, the error covariance matrix satisfies*

$$
\begin{aligned}
\mathbf{P}(t+\tau|t) = {} & \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t+\tau-1)\underline{\mathbf{x}}(t+\tau-1)\underline{\mathbf{x}}^H(t+\tau-1)\triangle\tilde{\mathbf{\Phi}}^H(t+\tau-1)] + \mathbf{\Upsilon}(t+\tau-1) \\
& + \bar{\mathbf{\Phi}}(t+\tau-1)\mathbf{P}(t+\tau-1|t)\bar{\mathbf{\Phi}}^H(t+\tau-1) + \mathbf{Q}(t+\tau-1) + \mathbf{\Upsilon}^H(t+\tau-1) \qquad (26)
\end{aligned}
$$

The initial conditions $\hat{\underline{\mathbf{x}}}(t+1|t)$ *and* $\mathbf{P}(t+1|t)$ *are computed by Theorem 1.*

4.3. Widely Linear Smoother

Theorem 3. *The optimal linear fixed-lag* τ*-step smoother obtained from the system defined by Eqs.* (7)*-*(8) *is computed through*

$$
\hat{\underline{\mathbf{x}}}(t+\tau|t) = \hat{\underline{\mathbf{x}}}(t+\tau|t-1) + \mathbf{N}(t+\tau|t)\boldsymbol{\varepsilon}(t) \qquad (27)
$$

The smoothing gain matrix is computed by

$$
\mathbf{N}(t+\tau|t) = \Big[\mathbf{\Lambda}_\tau(t)\bar{\mathbf{H}}^H(t) - \mathbf{N}(t+\tau|t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\Big]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t) \qquad (28)
$$

where $\mathbf{\Lambda}_{-1}(1) = \mathbf{P}(0|-1)[\bar{\mathbf{\Phi}}(0) - \mathbf{L}(0)\bar{\mathbf{H}}(0)]^H$ *and*

$$
\begin{aligned}
\mathbf{\Lambda}_{-1}(t) = {} & \mathbf{P}(t-1|t-2)\big[\bar{\mathbf{\Phi}}(t-1) - \mathbf{L}(t-1)\bar{\mathbf{H}}(t-1)\big]^H \\
& - \Big[\mathbf{1}_3 \otimes \big[\mathcal{D}\left(\mathbf{1} - \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t-2)\right] - \mathrm{E}\left[\boldsymbol{\gamma}_3^r(t-2)\right]\right)\mathbf{Q}_{\bar{\mathbf{v}}}\big]\Big]\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)])\mathbf{L}^H(t-1) \\
& + \mathbf{L}(t-2)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-2)\Big[\mathbf{L}(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]\right)\mathbf{F}(t-2) - \bar{\mathbf{\Gamma}}(t-1)\mathbf{G}(t-2)\Big]^H \\
& + \bar{\mathbf{\Gamma}}_3(t-2)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H\bar{\mathbf{\Gamma}}_2^H(t-1), \qquad t \geq 2 \qquad (29)
\end{aligned}
$$

$$\boldsymbol{\Lambda}_{\tau}(t) = \boldsymbol{\Lambda}_{\tau+1}(t-1)\left[\bar{\boldsymbol{\Phi}}(t-1) - \mathbf{L}(t-1)\bar{\mathbf{H}}(t-1)\right]^{H} + \mathbf{N}(t+\tau|t-2)\mathbf{Q}_{\boldsymbol{\epsilon}}(t-2)$$
$$\times\left[\mathbf{L}(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]\right)\mathbf{F}(t-2) - \bar{\boldsymbol{\Gamma}}(t-1)\mathbf{G}(t-2)\right]^{H}, \qquad \tau < -1 \quad (30)$$

The smoothing error covariance matrix is updated through

$$\mathbf{P}(t+\tau|t) = \mathbf{P}(t+\tau|t-1) - \mathbf{N}(t+\tau|t)\mathbf{Q}_{\boldsymbol{\epsilon}}(t)\mathbf{N}^{H}(t+\tau|t) \qquad (31)$$

The initial conditions $\hat{\underline{\mathbf{x}}}(t+\tau|t+\tau)$, $\mathbf{P}(t+\tau|t+\tau)$, $\mathbf{N}(t+\tau|t+\tau) = \mathbf{K}(t+\tau)$, *and* $\mathbf{P}(t+\tau|t+\tau-1)$ *are computed in Theorem 1.*

4.4. Particular Cases

When $\mathbf{p}_1(t) = \mathbf{1}$, Theorems 1, 2 and 3 provide the WL Kalman estimators given in [25]. When $\mathbf{p}_2(t) = \mathbf{p}_3(t) = \mathbf{0}$ the WL estimators are reduced to those for the case of uncertain observations [24]. When $\mathbf{p}_1(t) + \mathbf{p}_2(t) = \mathbf{1}$ and $\mathbf{p}_1(t) + \mathbf{p}_3(t) = \mathbf{1}$, the WL estimators are reduced to those for the case of sensor delays and multiple packet dropouts, respectively. When $\mathbf{p}_2(t) = \mathbf{0}$, then we obtain the case of multiple packet dropouts and uncertain observations. When $\mathbf{p}_3(t) = \mathbf{0}$, they are reduced to those for the case of sensor delays and uncertain observations. Finally, when $\mathbf{p}_1(t) + \mathbf{p}_2(t) + \mathbf{p}_3(t) = \mathbf{1}$, they are reduced to those for the case of sensor delays and multiple packet dropouts.

4.5. Multiple-step Delays

Multiple-step random delays can also exist in networked systems. The one-step sensor delay assumption is based on the fact that the induced data latency (i.e. the delay that occurs when a packet crosses a network connection from sensor to estimator) is sometimes restricted not to exceed the sampling period [29]. Actually, the current model constitutes the basis of a more general one in which two or more sample random delays can occur at each time. The extension of the proposed model to the case of two or more sampling delays is obtained by introducing additional Bernoulli random variables in Eq. (5) and by augmenting the state vector $\underline{\mathbf{x}}(t)$ in Eq. (7) appropriately. The optimal WL estimators for the case of multiple sampling delays can be derived by following a similar framework to that used in this paper.

4.6. Properness

The algorithms in Theorems 1, 2, and 3 can be adapted to $\mathbb{H}$-proper or $\mathbb{C}^\eta$-proper quaternion signals by using a SL or a SWL processing, respectively. The resultant algorithms entail a notable saving in computational burden with respect to those derived in the general or improper case. An example that illustrates this approach in the particular case of only uncertain observations ($\mathbf{p}_2(t) = \mathbf{p}_3(t) = \mathbf{0}$) can be seen in [24].

4.7. Nonlinear Systems

The suggested WL system and algorithms can also serve as a starting point to devise estimation algorithms for nonlinear systems with mixed uncertainties in the observations via the linearization approach used by the EKF [26]. This strategy allows us to apply our algorithms, for example, to solve the orientation estimation problem. For that, and in order to enforce unit-normalization of the estimator, we can normalize both the quaternion estimation and the error variance at the end of each update step. In this way, the complexity that the normalization introduces into the Kalman estimators derivation is avoided. As a result, although the Kalman filter is an optimal algorithm, this normalization procedure leads to a suboptimal algorithm. Nevertheless, such an approach has been shown to estimate efficiently (see, e.g., [30]).

5. Numerical Simulations

The aim of this example is to experimentally demonstrate that the algorithms in Theorems 1-3 outperform the standard WL Kalman estimators provided in [25] when mixed uncertainties in the observations are present[2]. For that, we consider the following state-space system [24]

$$\begin{aligned} x(t+1) =& f x(t) + g x^\eta(t) + w(t) \\ z(t) =& x(t) + v(t) \end{aligned}$$

[2]Notice that both techniques have the same computational complexity.

where $f = 0.8 - 0.3\eta + 0.2\eta' + 0.1\eta''$ and $g = -0.1\eta + 0.04\eta' - 0.05\eta''$. Assume that $x(0)$ is a $\mathbb{H}$-proper quaternion Gaussian variable with $E[x(0)] = 0$ and $E[x(0)x^*(0)] = 1$, the noise $w(t)$ has the following components: $w_r(t) = \epsilon_r(t)$, $w_\eta(t) = -0.6w_r(t) + \epsilon_\eta(t)$, $w_{\eta'}(t) = 0.8w_\eta(t) + \epsilon_{\eta'}(t)$, and $w_{\eta''}(t) = 0.8w_r(t) - 0.4w_\eta(t) + \epsilon_{\eta''}(\mathrm{t})$, where $\epsilon_r(t)$, $\epsilon_\eta(t)$, $\epsilon_{\eta'}(t)$, and $\epsilon_{\eta''}(t)$ are real independent white Gaussian noises (WGN) with variance parameter 0.25, and the measurement noise $v(t)$ is a $\mathbb{H}$-proper quaternion WGN with variance parameter $E[v(t)v^*(t)] = 0.4$.

By using Monte Carlo simulation we have generated 10000 values of $x(t)$ and $\bar{\mathbf{y}}(t)$ for each $t = 0, 1, \ldots, 99$ and for different values of the Bernoulli probabilities. Denote by $x_j^p(t)$ and $\bar{\mathbf{y}}_j^p(t)$ the value of $x(t)$ and $\bar{\mathbf{y}}(t)$, respectively, generated in the jth simulation and for a specific combination of probabilities p. Similarly, the standard WL Kalman estimators obtained from $\bar{\mathbf{y}}_j^p(t)$ in the jth simulation and for a specific combination of probabilities p are denoted by $\check{x}_j^p(t|t)$ (filter), $\check{x}_j^p(t+3|t)$ (3-step predictor), and $\check{x}_j^p(t-2|t)$ (fixed-lag 2-step smoother). The performance of such estimators has been assessed by computing the following mean square errors (MSEs):

$$MSE^p(s|t) = \frac{1}{10000}\sum_{j=1}^{10000} SSE_j^p(s|t), \qquad s = t, t+3, t-2$$

where the sum square errors (SSEs) are calculated as[3]

$$SSE_j^p(t|t) = \|x_j^p(t) - \check{x}_j^p(t|t)\|^2, \qquad t = 0, \ldots, 99$$

$$SSE_j^p(t+3|t) = \|x_j^p(t+3) - \check{x}_j^p(t+3|t)\|^2, \qquad t = 0, \ldots, 96$$

$$SSE_j^p(t-2|t) = \|x_j^p(t-2) - \check{x}_j^p(t-2|t)\|^2, \qquad t = 2, \ldots, 99$$

and with $j = 1, \ldots, 10000$.

$MSE^p(s|t)$, for $s = t, t+3, t-2$, are estimators of the population error variances, denoted by $V^p(s|t)$, associated to the standard WL Kalman estimators. These MSEs have been compared with the filtering, prediction and smoothing

[3]The norm of $x(t)$ is $\|x(t)\| = \sqrt{x_r^2(t) + x_\eta^2(t) + x_{\eta'}^2(t) + x_{\eta''}^2(t)}$.

error variances given in Theorems 1, 2 and 3 and denoted by $P^p(t|t)$, $P^p(t+3|t)$, and $P^p(t-2|t)$, respectively, as functions of both t and p. Specifically, 95% confidence intervals for the population error variances $V^p(s|t)$, $s = t, t+3, t-2$, have been obtained from the corresponding SSEs, $SSE_j^p(s|t)$, simulated by 10000 Monte Carlo runs at every instant t. Denote the lower and upper confidence bounds of such intervals by $L^p(s|t)$ and $U^p(s|t)$, respectively. In order to get clearer interpretations of these results, the following differences have been derived at each instant of time: $L^p(s|t) - P^p(s|t)$, $MSE^p(s|t) - P^p(s|t)$, and $U^p(s|t) - P^p(s|t)$. For instance, when $s = t+3$, $MSE^p(t+3|t) - P^p(t+3|t)$ represents the point estimation of $V^p(t+3|t) - P^p(t+3|t)$ in a given instant t. Additionally, $L^p(t+3|t) - P^p(t+3|t)$ and $U^p(t+3|t) - P^p(t+3|t)$ stand for the lower and upper 95%-confidence bounds, respectively, for $V^p(t+3|t) - P^p(t+3|t)$.

We have considered four different cases:

1. Case 1: $\mathbf{p}_1(t) = \mathbf{0.8} = [0.8, 0.8, 0.8, 0.8]^T$ and $\mathbf{p}_2(t) = \mathbf{p}_3(t) = \mathbf{0.1} = [0.1, 0.1, 0.1, 0.1]^T$.
2. Case 2: $\mathbf{p}_1(t) = \mathbf{p}_3(t) = \mathbf{0.1}$ and $\mathbf{p}_2(t) = \mathbf{0.8}$.
3. Case 3: $\mathbf{p}_1(t) = \mathbf{0.05}$, $\mathbf{p}_2(t) = \mathbf{0.1}$ and $\mathbf{p}_3(t) = \mathbf{0.8}$.
4. Case 4: $\mathbf{p}_1(t) = \mathbf{0.1}$ and $\mathbf{p}_2(t) = \mathbf{p}_3(t) = \mathbf{0.3}$.

These four cases represent various levels of mixed uncertainties which allow us to carry out a representative performance comparison. Indeed, they involve the three uncertainties in different degrees of intensity. The first two satisfy $\mathbf{p}_1(t) + \mathbf{p}_2(t) + \mathbf{p}_3(t) = \mathbf{1}$ and thus, they are cases of sensor delays and multiple packet dropouts which can appear, e.g., in an energy-constrained wireless sensor network [28]. Case 1 reproduces a scenario where the above mixed uncertainties are unlikely, being the most similar scenario to that where the standard WL Kalman estimators operate. Unlike Case 1, Case 2 gives high probabilities to delayed observations. Cases 3 and 4 are examples where all the three uncertainties appear. Case 3 assigns a high probability to a packet dropouts occurrence as well as to the fact that the data packets in the measurements contain only noise. Finally, Case 4 gives moderate probabilities to both delayed observations and

	Case 1		Case 2		Case 3		Case 4	
s	$\bar{P}^p(s\|t)$	$\bar{M}^p(s\|t)$	$\bar{P}^p(s\|t)$	$\bar{M}^p(s\|t)$	$\bar{P}^p(s\|t)$	$\bar{M}^p(s\|t)$	$\bar{P}^p(s\|t)$	$\bar{M}^p(s\|t)$
t	0.798	0.908	1.965	3.584	6.065	11.500	3.977	5.934
$t+3$	4.169	4.244	4.733	5.788	7.136	10.649	5.874	7.192
$t-2$	0.606	0.755	0.627	3.235	5.410	11.068	2.877	5.455

Table 1: Time means of the filtering, prediction and smoothing MSEs.

packet dropouts and a lower probability to receive the current measurement.

Fig. 1 depicts the 95% pointwise confidence bands for $V^p(t|t) - P^p(t|t)$, as well as the point estimations, $MSE^p(t|t) - P^p(t|t)$, for the filtering problem and the above four cases. Likewise, Fig. 2 and Fig. 3 show the same quantities for the 3-step prediction and the fixed-lag 2-step smoothing problems, respectively. On the other hand, Table 1 lists the time averages of the above MSEs. For example, $\bar{P}^p(t|t) = \frac{1}{100}\sum_{t=0}^{99} P^p(t|t)$ or $\bar{M}^p(t+3|t) = \frac{1}{97}\sum_{t=0}^{96} MSE^p(t+3|t)$.

Because the standard WL Kalman estimators ignore the presence of uncertainties in the observations, from Figs. 1-3 and Table 1, we can see that the proposed algorithms have in all cases better accuracy than the ones proposed in [25]. More specifically, the difference in performance between both kinds of estimators is increased as the probabilities of sensor delays and/or packet dropouts and/or uncertain observations become larger. Actually, this difference is more pronounced as the probabilities of packet dropouts increase (Case 3). However, from Table 1, the biggest relative performance difference is attained for the smoothing problem in Case 2 since the simulated MSEs are approximately five times higher than the error variances $P^p(t-2|t)$. Obviously, the minor performance difference for all the three estimation problems is observed in Case 1 since this one resembles the conditions under which the standard WL Kalman estimators operate.

6. Conclusions

The optimal WL estimation problem has been analyzed for quaternion systems with multiple sensors and mixed uncertainties of sensor delays, packet

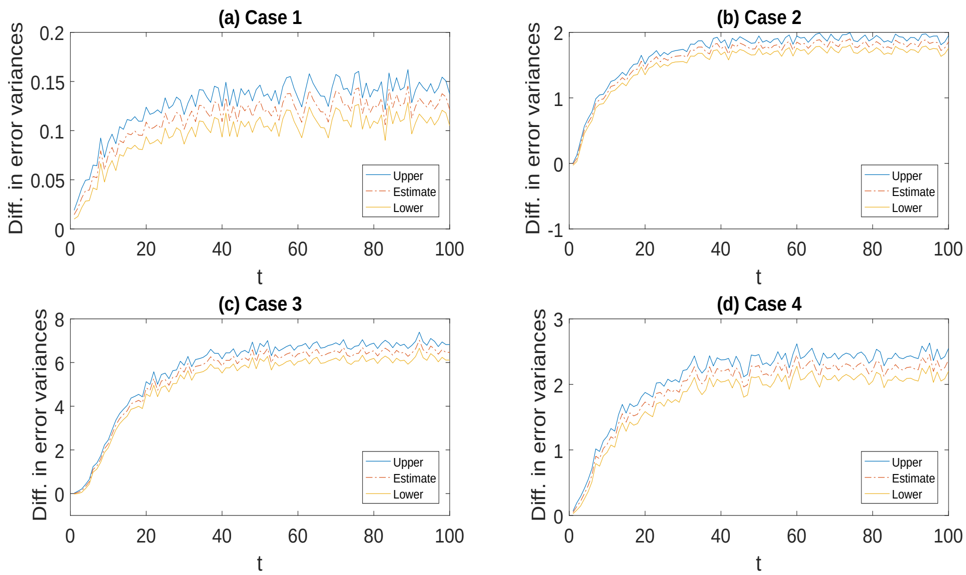

Figure 1: 95% pointwise confidence bands for $V^p(t|t) - P^p(t|t)$ in cases 1 (a), 2 (b), 3 (c), and 4 (d). The point estimations, $MSE^p(t|t) - P^p(t|t)$, are designated by "Estimate".

dropouts and uncertain observations. Specifically, filtering, prediction and fixed-lag smoothing algorithms have been given. The discrete-time system under study represents the three above uncertainties in a unified way and assumes that each component of the quaternion is measured by different sensors, where each sensor may be subject to different mixed uncertainties. A numerical example has been provided to show the effectiveness of the proposed algorithms.

7. Appendix

7.1. Preliminary Result

The following property, stated without proof, will be necessary to prove the main results in this paper.

Property 3. *Let* $\mathbf{a}_i, \mathbf{b}_i, \mathbf{x}_i, \mathbf{y}_i \in \mathbb{R}^4$ *and* $\mathbf{A}_i = \mathrm{diag}(\mathbf{a}_i)$, $\mathbf{B}_i = \mathrm{diag}(\mathbf{b}_i)$, $i = 1, 2, 3$. *Define* $\mathbf{a} = \mathrm{vec}([\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3])$, $\mathbf{b} = \mathrm{vec}([\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3])$, $\mathbf{x} = \mathrm{vec}([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3])$, *and* $\mathbf{y} = \mathrm{vec}([\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3])$, *then*

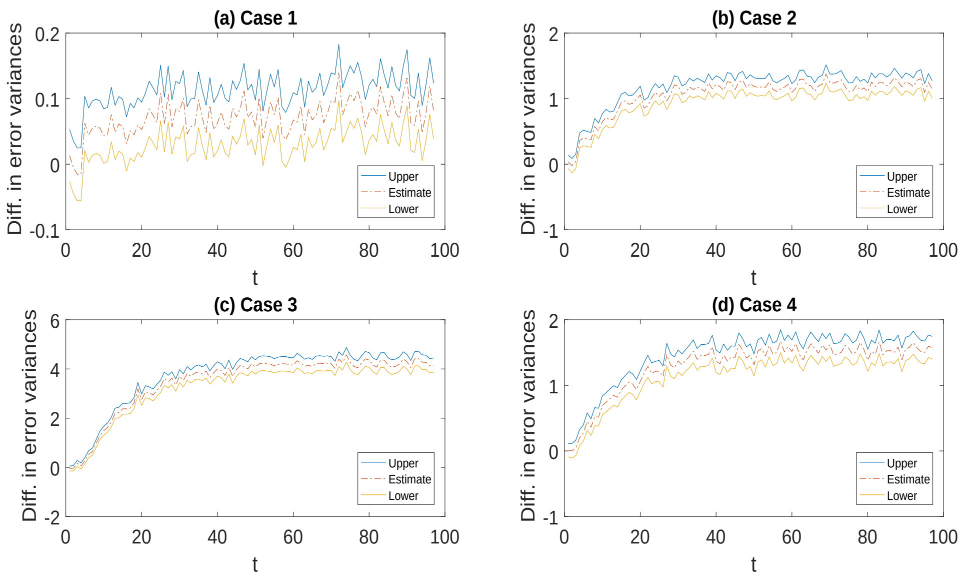

Figure 2: 95% pointwise confidence bands for $V^p(t+3|t) - P^p(t+3|t)$ in cases 1 (a), 2 (b), 3 (c), and 4 (d). The point estimations, $MSE^p(t+3|t) - P^p(t+3|t)$, are designated by "Estimate".

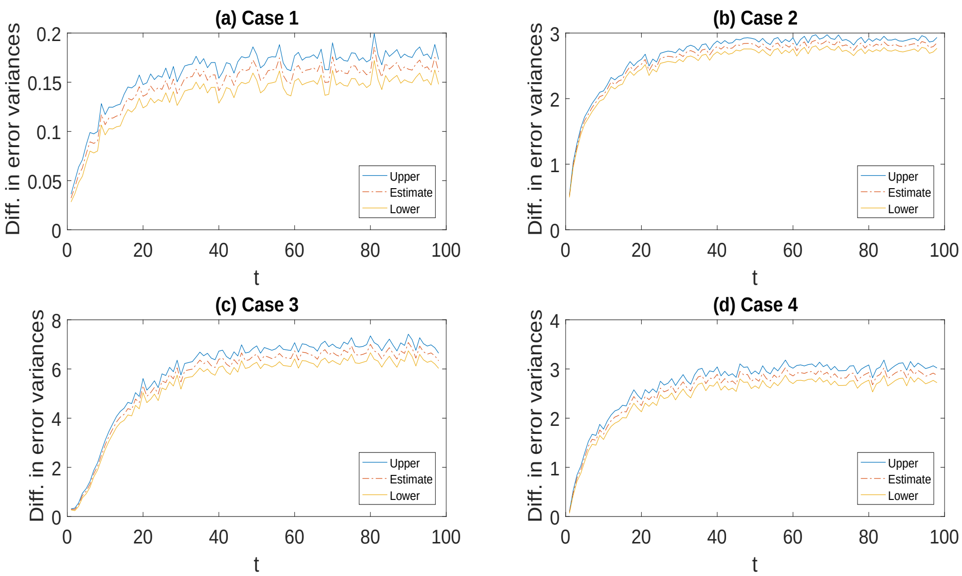

Figure 3: 95% pointwise confidence bands for $V^p(t-2|t) - P^p(t-2|t)$ in cases 1 (a), 2 (b), 3 (c), and 4 (d). The point estimations, $MSE^p(t-2|t) - P^p(t-2|t)$, are designated by "Estimate".

1. $\mathbf{A}_i\mathbf{x}_l\mathbf{y}_j^T\mathbf{B}_k = (\mathbf{a}_i\mathbf{b}_k^T) \odot (\mathbf{x}_l\mathbf{y}_j^T)$, *for all* i, j, k, l.
2. $[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3]\, \mathbf{x}\mathbf{y}_j^T\mathbf{B}_k = \sum_{i=1}^3 (\mathbf{a}_i\mathbf{b}_k^T) \odot (\mathbf{x}_i\mathbf{y}_j^T)$, *for all* j, k.
3. $[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3]\, \mathbf{x}\mathbf{y}^T\, [\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3]^T = \sum_{i,j=1}^3 (\mathbf{a}_i\mathbf{b}_j^T) \odot (\mathbf{x}_i\mathbf{y}_j^T)$
4. $\mathrm{diag}(\mathbf{a})\mathbf{x}\mathbf{y}_j^T\mathbf{B}_k = \sum_{i=1}^3 \mathbf{1}_i \otimes \{(\mathbf{a}_i\mathbf{b}_k^T) \odot (\mathbf{x}_i\mathbf{y}_j^T)\}$, *for all* j, k.
5. $[\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3]\, \mathbf{x}\mathbf{y}^T\, \mathrm{diag}(\mathbf{b}) = \sum_{i=1}^3 [(\mathbf{a}_i\mathbf{b}_1^T) \odot (\mathbf{x}_i\mathbf{y}_1^T), (\mathbf{a}_i\mathbf{b}_2^T) \odot (\mathbf{x}_i\mathbf{y}_2^T), (\mathbf{a}_i\mathbf{b}_3^T) \odot (\mathbf{x}_i\mathbf{y}_3^T)]$.

7.2. Proof of Property 1

It is easy to check that $\tilde{\mathbf{H}}(t) = \sum_{i=1}^3 \mathbf{1}_i^T \otimes \mathcal{D}(\boldsymbol{\gamma}_i^r(t))$ and

$$\bar{\mathbf{H}}(t) = \mathrm{E}[\tilde{\mathbf{H}}(t)] = \sum_{i=1}^3 \mathbf{1}_i^T \otimes \mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_i^r(t)]) \tag{32}$$

$$\triangle\tilde{\mathbf{H}}(t) = \tilde{\mathbf{H}}(t) - \bar{\mathbf{H}}(t) = \frac{1}{4}\boldsymbol{\mathcal{A}}\left[\mathbf{D}_1(t), \mathbf{D}_2(t), \mathbf{D}_3(t)\right]\boldsymbol{\mathcal{T}}^H \tag{33}$$

with $\mathbf{D}_i(t) = \mathrm{diag}(\boldsymbol{\gamma}_i^r(t) - \mathrm{E}[\boldsymbol{\gamma}_i^r(t)])$, $i = 1, 2, 3$. Similarly, we get that

$$\tilde{\boldsymbol{\Phi}}(t) = \begin{bmatrix} \boldsymbol{\Psi}(t) \\ \tilde{\mathbf{H}}(t) \end{bmatrix} = \sum_{i=1}^3 \boldsymbol{\mathcal{I}}_{3i} \otimes \mathcal{D}(\boldsymbol{\gamma}_i^r(t)) + \begin{bmatrix} \boldsymbol{\Psi}(t) \\ \mathbf{0} \end{bmatrix} = \sum_{i=1}^4 \tilde{\boldsymbol{\Phi}}_i(t) \tag{34}$$

with $\boldsymbol{\Psi}(t) = \begin{bmatrix} \boldsymbol{\Phi}(t) & \mathbf{0} & \mathbf{0} \\ \mathbf{I}_4 & \mathbf{0} & \mathbf{0} \end{bmatrix}$, $\tilde{\boldsymbol{\Phi}}_i(t) = \boldsymbol{\mathcal{I}}_{3i} \otimes \mathcal{D}(\boldsymbol{\gamma}_i^r(t))$, $i = 1, 2, 3$, and $\tilde{\boldsymbol{\Phi}}_4'(t) = [\boldsymbol{\Psi}'(t), \mathbf{0}']$. Also

$$\bar{\boldsymbol{\Phi}}(t) = \mathrm{E}[\tilde{\boldsymbol{\Phi}}(t)] = \begin{bmatrix} \boldsymbol{\Psi}(t) \\ \bar{\mathbf{H}}(t) \end{bmatrix} \tag{35}$$

$$\triangle\tilde{\boldsymbol{\Phi}}(t) = \tilde{\boldsymbol{\Phi}}(t) - \bar{\boldsymbol{\Phi}}(t) = \begin{bmatrix} \mathbf{0} \\ \triangle\tilde{\mathbf{H}}(t) \end{bmatrix} = \frac{1}{4}\boldsymbol{\mathcal{T}}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{D}_1(t) & \mathbf{D}_2(t) & \mathbf{D}_3(t) \end{bmatrix}\boldsymbol{\mathcal{T}}^H \tag{36}$$

Note that the real random vector $\boldsymbol{\gamma}^r(t) = \mathrm{vec}([\boldsymbol{\gamma}_1^r(t), \boldsymbol{\gamma}_2^r(t), \boldsymbol{\gamma}_3^r(t)])$ has covariance matrix $\boldsymbol{\Sigma}^r(t) = [\boldsymbol{\Sigma}_{ij}^r(t)]$, $i, j = 1, 2, 3$, with block matrices given by

$$\boldsymbol{\Sigma}_{ii}^r(t) = \mathrm{E}\left[(\boldsymbol{\gamma}_i^r(t) - \mathrm{E}[\boldsymbol{\gamma}_i^r(t)])(\boldsymbol{\gamma}_i^r(t) - \mathrm{E}[\boldsymbol{\gamma}_i^r(t)])^T\right] = \mathrm{diag}\left(\mathbf{p}_i(t) \odot (\mathbf{1} - \mathbf{p}_i(t))\right),$$

$$\boldsymbol{\Sigma}_{ij}^r(t) = \mathrm{E}\left[(\boldsymbol{\gamma}_i^r(t) - \mathrm{E}[\boldsymbol{\gamma}_i^r(t)])(\boldsymbol{\gamma}_j^r(t) - \mathrm{E}[\boldsymbol{\gamma}_j^r(t)])^T\right] = -\,\mathrm{diag}\left(\mathbf{p}_i(t) \odot \mathbf{p}_j(t)\right)$$

Denote

$$\underline{\mathbf{x}}^r(t) = \begin{bmatrix} \mathbf{x}^r(t) \\ \mathbf{x}^r(t-1) \\ \mathbf{y}^r(t-1) \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{x}}_1^r(t) \\ \underline{\mathbf{x}}_2^r(t) \\ \underline{\mathbf{x}}_3^r(t) \end{bmatrix} \tag{37}$$

Taking Eq. (33) into account and that $\underline{\mathbf{x}}(t) = \boldsymbol{\mathcal{T}}\underline{\mathbf{x}}^r(t)$, it follows that

$$\begin{aligned} &\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)] \\ &= \boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[\left[\mathbf{D}_1(t),\mathbf{D}_2(t),\mathbf{D}_3(t)\right]\underline{\mathbf{x}}^r(t)\underline{\mathbf{x}}^{rT}(t)\left[\mathbf{D}_1(t),\mathbf{D}_2(t),\mathbf{D}_3(t)\right]^T\right]\Big\}\boldsymbol{\mathcal{A}}^H \\ &= \boldsymbol{\mathcal{A}}\Big\{\sum_{i,j=1}^{3}\left(\boldsymbol{\Sigma}_{ij}^r(t)\odot\mathrm{E}[\underline{\mathbf{x}}_i^r(t)\underline{\mathbf{x}}_j^{rT}(t)]\right)\Big\}\boldsymbol{\mathcal{A}}^H \end{aligned}$$

where the last equality is a consequence of Property 3.3. Now, the result is immediate. Properties 1.2 and 1.3 follow easily from the above result and Eq. (36).

7.3. Proof of Property 2

We first prove that the state vector and the noise are also correlated. For that, we note that

$$\tilde{\mathbf{\Gamma}}(t) = \boldsymbol{\mathcal{I}}_{11}\otimes\mathbf{\Gamma}(t) + \boldsymbol{\mathcal{I}}_{33}\otimes\mathcal{D}(\boldsymbol{\gamma}_2^r(t)) + \boldsymbol{\mathcal{I}}_{32}\otimes\mathcal{D}(\mathbf{1}-\boldsymbol{\gamma}_2^r(t)-\boldsymbol{\gamma}_3^r(t)) \tag{38}$$

$$\bar{\mathbf{\Gamma}}(t) = \boldsymbol{\mathcal{I}}_{11}\otimes\mathbf{\Gamma}(t) + \boldsymbol{\mathcal{I}}_{33}\otimes\mathcal{D}\left(\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\right]\right) + \boldsymbol{\mathcal{I}}_{32}\otimes\mathcal{D}\left(\mathrm{E}\left[\mathbf{1}-\boldsymbol{\gamma}_2^r(t)-\boldsymbol{\gamma}_3^r(t)\right]\right) \tag{39}$$

From Eq. (9), it follows that

$$\mathrm{E}\left[\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\right] = \mathrm{E}\left[\tilde{\mathbf{\Gamma}}(t-1)\underline{\mathbf{w}}(t-1)\underline{\mathbf{w}}^H(t)\right] = \bar{\mathbf{\Gamma}}(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H = \bar{\mathbf{\Gamma}}_3(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H \tag{40}$$

where we have taken into account that $\bar{\mathbf{\Gamma}}_1(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H = \bar{\mathbf{\Gamma}}_2(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H = \mathbf{0}$. From Eq. (40) we get $\mathrm{E}\left[\underline{\mathbf{x}}(t)\bar{\mathbf{w}}^H(t)\right] = \mathrm{E}\left[\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t)\right] = \mathbf{0}$ and thus

$$\mathrm{E}\left[\mathbf{x}^r(t)\mathbf{v}^{rT}(t)\right] = \mathrm{E}\left[\mathbf{x}^r(t-1)\mathbf{v}^{rT}(t)\right] = \mathrm{E}\left[\mathbf{y}^r(t-1)\mathbf{v}^{rT}(t)\right] = \mathbf{0} \tag{41}$$

and $\mathrm{E}\left[\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t-1)\right] = \mathbf{1}_3\otimes\left[\mathcal{D}\left(\mathbf{1}-\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t-1)\right]-\mathrm{E}\left[\boldsymbol{\gamma}_3^r(t-1)\right]\right)\mathbf{Q}_{\bar{\mathbf{v}}}\right]$, and then

$$\mathrm{E}\left[\mathbf{x}^r(t)\mathbf{v}^{rT}(t-1)\right] = \mathrm{E}\left[\mathbf{x}^r(t-1)\mathbf{v}^{rT}(t-1)\right] = \mathbf{0} \tag{42}$$

$$\mathrm{E}\left[\mathbf{y}^r(t-1)\mathbf{v}^{rT}(t-1)\right] = \mathrm{diag}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r} \tag{43}$$

with $\mathbf{Q}_{\mathbf{v}^r} = \mathrm{E}[\mathbf{v}^r(t)\mathbf{v}^{rT}(t)]$. Now, we have that

$$\mathrm{E}[\underline{\mathbf{x}}(t+1)\underline{\mathbf{x}}^H(t+1)] = \bar{\mathbf{\Phi}}(t)\,\mathrm{E}[\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)]\bar{\mathbf{\Phi}}^H(t) + \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{\Phi}}^H(t)]$$
$$+\mathrm{E}[\tilde{\mathbf{\Gamma}}(t)\underline{\mathbf{w}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] + \mathrm{E}[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] + \mathrm{E}[\tilde{\mathbf{\Gamma}}(t)\underline{\mathbf{w}}(t)\underline{\mathbf{x}}^H(t)\tilde{\mathbf{\Phi}}^H(t)] \tag{44}$$

Denoting $\mathbf{Q}(t) = \mathrm{E}[\tilde{\mathbf{\Gamma}}(t)\underline{\mathbf{w}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)]$ and taking Eq. (38) into account we get $\mathbf{Q}(t) = \sum_{i,j=1}^3 \mathrm{E}\left[\tilde{\mathbf{\Gamma}}_i(t)\underline{\mathbf{w}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_j^H(t)\right] = \sum_{i,j=1}^3 \mathbf{Q}_{ij}(t)$. It is easy to check that $\mathbf{Q}_{ij}(t) = \mathbf{Q}_{ji}^H(t) = \mathbf{0}$, for $i \neq j$. Likewise, it follows that

$$\begin{aligned}
&\mathbf{Q}_{11}(t) = \boldsymbol{\mathcal{I}}_{11} \otimes (\mathbf{\Gamma}(t)\mathbf{Q}_{\bar{\mathbf{w}}}\mathbf{\Gamma}^H(t))\\
&\mathbf{Q}_{22}(t) = \boldsymbol{\mathcal{I}}_{33} \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big)\\
&\mathbf{Q}_{33}(t) = \boldsymbol{\mathcal{I}}_{33} \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big)
\end{aligned}$$

where the last two equalities follow from Property 3.1. Hence, Eq. (11) holds.

Denote $\mathbf{\Upsilon}(t) = \mathrm{E}[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)]$. Then, from Eqs. (34) and (38), we get

$$\mathbf{\Upsilon}(t) = \sum_{i,j=1}^{3} \mathrm{E}[\tilde{\mathbf{\Phi}}_i(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_j^H(t)] + \mathrm{E}[\tilde{\mathbf{\Phi}}_4(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] \tag{45}$$

Taking Eq. (40) into account, the second summand of Eq. (45) satisfies

$$\mathrm{E}[\tilde{\mathbf{\Phi}}_4(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] = \begin{bmatrix}\mathbf{\Psi}(t)\\ \mathbf{0}\end{bmatrix}\bar{\mathbf{\Gamma}}_3(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H\bar{\mathbf{\Gamma}}^H(t) = \mathbf{0}$$

where we have applied that $[\mathbf{\Psi}^T(t), \mathbf{0}^T]^T\bar{\mathbf{\Gamma}}_3(t-1) = \mathbf{0}$. Consider the first summand in Eq. (45) and denote by $\mathbf{\Upsilon}_{ij}(t) = \mathrm{E}[\tilde{\mathbf{\Phi}}_i(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_j^H(t)]$, then by using the uncorrelation properties of the noises in the model given by Eqs. (3)-(4), Property 3, Eqs. (41), (42), and (43) we have $\mathbf{\Upsilon}_{ij}(t) = \mathbf{0}$, except for

$$\mathbf{\Upsilon}_{32}(t) = \boldsymbol{\mathcal{I}}_{33}\otimes\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[\boldsymbol{\gamma}_3^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right]\odot\Big[\,\mathrm{diag}(\mathbf{1}-\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\Big]\Big\}\boldsymbol{\mathcal{A}}^H$$

from which the result follows.

7.4. Proof of Theorem 1

Define the innovations by $\boldsymbol{\varepsilon}(t) = \bar{\mathbf{y}}(t) - \hat{\bar{\mathbf{y}}}(t|t-1)$. By applying the projection theorem we obtain Eq. (13) with $\mathbf{K}(t) = \mathrm{E}[\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t)]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t)$. Similarly, it follows that

$$\hat{\underline{\mathbf{x}}}(t+1|t) = \hat{\underline{\mathbf{x}}}(t+1|t-1) + \mathbf{L}(t)\boldsymbol{\varepsilon}(t) \tag{46}$$

with

$$\mathbf{L}(t) = \mathrm{E}[\underline{\mathbf{x}}(t+1)\boldsymbol{\varepsilon}^H(t)]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t) \tag{47}$$

Taking projections on both sides of Eqs. (7) and (8) onto the linear space spanned by $\{\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(1), \ldots, \bar{\mathbf{y}}(t-1)\}$, we have

$$\hat{\underline{\mathbf{x}}}(t+1|t-1) = \bar{\boldsymbol{\Phi}}(t)\hat{\underline{\mathbf{x}}}(t|t-1) + \bar{\boldsymbol{\Gamma}}(t)\hat{\underline{\mathbf{w}}}(t|t-1) \tag{48}$$

$$\hat{\bar{\mathbf{y}}}(t|t-1) = \bar{\mathbf{H}}(t)\hat{\underline{\mathbf{x}}}(t|t-1) + \mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\hat{\bar{\mathbf{v}}}(t-1|t-1) \tag{49}$$

where $\hat{\underline{\mathbf{w}}}(t|t-1)$ and $\hat{\bar{\mathbf{v}}}(t-1|t-1)$ are computed by

$$\hat{\underline{\mathbf{w}}}(t|t-1) = \hat{\underline{\mathbf{w}}}(t|t-2) + \mathbf{G}(t-1)\boldsymbol{\varepsilon}(t-1) = \mathbf{G}(t-1)\boldsymbol{\varepsilon}(t-1) \tag{50}$$

$$\hat{\bar{\mathbf{v}}}(t-1|t-1) = \hat{\bar{\mathbf{v}}}(t-1|t-2) + \mathbf{F}(t-1)\boldsymbol{\varepsilon}(t-1) = \mathbf{F}(t-1)\boldsymbol{\varepsilon}(t-1) \tag{51}$$

with

$$\mathbf{G}(t-1) = \mathrm{E}[\underline{\mathbf{w}}(t)\boldsymbol{\varepsilon}^H(t-1)]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t-1) \tag{52}$$

$$\mathbf{F}(t-1) = \mathrm{E}[\bar{\mathbf{v}}(t-1)\boldsymbol{\varepsilon}^H(t-1)]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t-1) \tag{53}$$

Thus, by substituting Eqs. (48) and (50) into Eq. (46) we get Eq. (14). Likewise, from Eqs. (49) and (51) we derive Eq. (15). Denoting the error $\tilde{\underline{\mathbf{x}}}(t+1|t) = \underline{\mathbf{x}}(t+1) - \hat{\underline{\mathbf{x}}}(t+1|t)$ and from Eqs. (8) and (15) it follows that

$$\begin{aligned}\boldsymbol{\varepsilon}(t) = {} & \triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t) + \bar{\mathbf{H}}(t)\tilde{\underline{\mathbf{x}}}(t|t-1) + \mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))\bar{\mathbf{v}}(t) \\ & + \mathcal{D}\left(\boldsymbol{\gamma}_2^r(t)\right)\bar{\mathbf{v}}(t-1) - \mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\mathbf{F}(t-1)\boldsymbol{\varepsilon}(t-1)\end{aligned} \tag{54}$$

and from Eq. (54) we have

$$\mathrm{E}[\underline{\mathbf{w}}(t)\boldsymbol{\varepsilon}^H(t-1)] = \mathbf{T}\,\mathcal{D}\left(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)]\right) \tag{55}$$

where the last equality is a consequence of $\underline{\mathbf{w}}(t) \perp \underline{\mathbf{x}}(t-1)$, $\underline{\mathbf{w}}(t) \perp \underline{\tilde{\mathbf{x}}}(t-1|t-2)$, $\underline{\mathbf{w}}(t) \perp \bar{\mathbf{v}}(t-2)$, $\underline{\mathbf{w}}(t) \perp \boldsymbol{\varepsilon}(t-2)$ and Eq. (10). Hence, by substituting Eq. (55) into Eq. (52) we get Eq. (16). By a similar reasoning, it follows that

$$\mathrm{E}[\bar{\mathbf{v}}(t-1)\boldsymbol{\varepsilon}^H(t-1)] = \mathbf{Q}_{\bar{\mathbf{v}}}\,\mathcal{D}\left(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)]\right) \tag{56}$$

and by substituting Eq. (56) into Eq. (53), we obtain Eq. (17).

Now, consider the error covariance matrix $\mathbf{P}(t|t) = \mathrm{E}[\underline{\tilde{\mathbf{x}}}(t|t)\underline{\tilde{\mathbf{x}}}^H(t|t)]$. From Eq. (13) we have that $\underline{\tilde{\mathbf{x}}}(t|t-1) = \underline{\tilde{\mathbf{x}}}(t|t) + \mathbf{K}(t)\boldsymbol{\varepsilon}(t)$ and noting that $\underline{\tilde{\mathbf{x}}}(t|t) \perp \boldsymbol{\varepsilon}(t)$ then Eq. (18) holds.

Let $\mathbf{M}(t) = \left[\mathbf{0}_{4\times 8}, \mathbf{I}_4\right]^T - \mathbf{L}(t)$, by subtracting Eq. (14) from Eq. (7) and taking Eq. (54) into account we obtain the following expression for the error

$$\begin{aligned}
\underline{\tilde{\mathbf{x}}}(t+1|t) &= \mathbf{M}(t) \triangle \tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t) + [\bar{\boldsymbol{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]\underline{\tilde{\mathbf{x}}}(t|t-1) + \tilde{\boldsymbol{\Gamma}}(t)\underline{\mathbf{w}}(t) \\
&\quad - \mathbf{L}(t)\,\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))\bar{\mathbf{v}}(t) - \mathbf{L}(t)\,\mathcal{D}(\boldsymbol{\gamma}_2^r(t))\bar{\mathbf{v}}(t-1) \\
&\quad + [\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) - \bar{\boldsymbol{\Gamma}}(t)\mathbf{G}(t-1)]\boldsymbol{\varepsilon}(t-1)
\end{aligned} \tag{57}$$

that we use for devising a recursion for $\mathbf{P}(t+1|t)$. To this end, noting that $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)] = \mathbf{0}$, then

$$\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\tilde{\mathbf{x}}}^H(t|t-1)] = \mathbf{0} \tag{58}$$

$$\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t-1)] = \mathbf{0} \tag{59}$$

Moreover, since $\underline{\tilde{\mathbf{x}}}(t|t-1) \perp \bar{\mathbf{v}}(t)$, $\bar{\mathbf{v}}(t) \perp \boldsymbol{\varepsilon}(t-1)$ and $\underline{\tilde{\mathbf{x}}}(t|t-1) \perp \boldsymbol{\varepsilon}(t-1)$, then

$$\mathrm{E}[\underline{\tilde{\mathbf{x}}}(t|t-1)\bar{\mathbf{v}}^H(t)\,\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))] = \mathbf{0} \tag{60}$$

$$\mathrm{E}[\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))\bar{\mathbf{v}}(t)\boldsymbol{\varepsilon}^H(t-1)] = \mathbf{0} \tag{61}$$

$$\mathrm{E}[\underline{\tilde{\mathbf{x}}}(t|t-1)\boldsymbol{\varepsilon}^H(t-1)] = \mathbf{0} \tag{62}$$

On the other hand, applying Property 3.2 and Eqs. (41) and (37), it follows that

$$\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t)\,\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))] = \mathbf{0} \tag{63}$$

By a similar reasoning, we have

$$\mathrm{E}[\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))\bar{\mathbf{v}}(t)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}(\boldsymbol{\gamma}_2^r(t))] = \mathbf{0} \tag{64}$$

We consider now the non-null terms. Since $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)] = \mathbf{0}$, then

$$\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] = \mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_2^H(t)]+\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_3^H(t)] \tag{65}$$

From Eq. (37) and Property 3.5, the first summand in Eq. (65) can be rewritten as

$$\begin{aligned}\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_2^H(t)] = \mathbf{1}_3^T \otimes \boldsymbol{\mathcal{A}}\Big\{ &\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\ &\odot \left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\end{aligned} \tag{66}$$

where the last equality is a consequence of Eqs. (42) and (43). As for the second summand in Eq. (65), by Property 3.1 and Eq. (41), we get

$$\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}_3^H(t)] = \mathbf{0} \tag{67}$$

Hence, by substituting Eqs. (66) and (67) into Eq. (65), it follows that

$$\begin{aligned}\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] = \mathbf{1}_3^T \otimes \boldsymbol{\mathcal{A}}\Big\{ &\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\ &\odot \left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\end{aligned} \tag{68}$$

By a similar reasoning, we have

$$\begin{aligned}\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}(\boldsymbol{\gamma}_2^r(t))] = \boldsymbol{\mathcal{A}}\Big\{ &\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\ &\odot \left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\end{aligned} \tag{69}$$

From Eq. (57) and taking into account that $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t-1)] = \mathbf{0}$, $\tilde{\underline{\mathbf{x}}}(t-1|t-2)\perp\underline{\mathbf{w}}(t)$, $\underline{\mathbf{w}}(t)\perp\boldsymbol{\varepsilon}(t-2)$ and $\underline{\mathbf{w}}(t)\perp\bar{\mathbf{v}}(t-2)$, it follows that

$$\begin{aligned}\mathrm{E}[\tilde{\underline{\mathbf{x}}}(t|t-1)\underline{\mathbf{w}}^H(t)\tilde{\mathbf{\Gamma}}^H(t)] = {}&\bar{\mathbf{\Gamma}}(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H\bar{\mathbf{\Gamma}}^H(t) \\ &- \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{T}^H\bar{\mathbf{\Gamma}}^H(t)\end{aligned} \tag{70}$$

Analogously, since $\tilde{\underline{\mathbf{x}}}(t-1|t-2)\perp\bar{\mathbf{v}}(t-1)$ and $\bar{\mathbf{v}}(t-1)\perp\boldsymbol{\varepsilon}(t-2)$, we get

$$\begin{aligned}\mathrm{E}[\tilde{\underline{\mathbf{x}}}(t|t-1)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}(\boldsymbol{\gamma}_2^r(t))] = {}&\bar{\mathbf{\Gamma}}(t-1)\mathbf{S}\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]) \\ &- \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\bar{\mathbf{v}}}\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\end{aligned} \tag{71}$$

Taking into account that $\tilde{\mathbf{\Gamma}}_1(t)\mathbf{S} = \mathbf{0}$, Property 3.4 and that $\bar{\mathbf{v}}(t)$ is a white noise, then

$$\begin{aligned}
&\mathrm{E}[\tilde{\mathbf{\Gamma}}(t)\underline{\mathbf{w}}(t)\bar{\mathbf{v}}^H(t)\,\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))] \\
&= \mathbf{1}_3 \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big)
\end{aligned} \tag{72}$$

By similar arguments and because of $\tilde{\mathbf{\Gamma}}_1(t)\mathbf{T} = \mathbf{0}$, we deduce that

$$\mathrm{E}[\tilde{\mathbf{\Gamma}}(t)\underline{\mathbf{w}}(t)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}(\boldsymbol{\gamma}_2^r(t))] = \mathbf{1}_3 \otimes \Big(\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big) \tag{73}$$

From Eq. (52), we have

$$\mathrm{E}[\tilde{\mathbf{\Gamma}}(t)\underline{\mathbf{w}}(t)\boldsymbol{\varepsilon}^H(t-1)] = \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1) \tag{74}$$

$$\begin{aligned}
&\mathrm{E}[\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))\bar{\mathbf{v}}(t)\bar{\mathbf{v}}^H(t)\,\mathcal{D}(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))] \\
&\qquad = \boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H
\end{aligned} \tag{75}$$

$$\mathrm{E}[\mathcal{D}(\boldsymbol{\gamma}_2^r(t))\bar{\mathbf{v}}(t-1)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}(\boldsymbol{\gamma}_2^r(t))] = \boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H \tag{76}$$

Finally, taking Eq. (53) into account, it follows that

$$\mathrm{E}[\mathcal{D}(\boldsymbol{\gamma}_2^r(t))\bar{\mathbf{v}}(t-1)\boldsymbol{\varepsilon}^H(t-1)] = \mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1) \tag{77}$$

In conclusion, from Eqs. (58)-(77), we have

$$\begin{aligned}
\mathbf{P}(t+1|t) &= \mathbf{M}(t)\,\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)]\mathbf{M}^H(t) + \mathbf{Q}(t) + \mathbf{\Omega}(t) \\
&\quad + \mathbf{\Omega}^H(t) + [\bar{\mathbf{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]P(t|t-1)[\bar{\mathbf{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]^H \\
&\quad + \mathbf{L}(t)\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\mathbf{L}^H(t) \\
&\quad + \mathbf{L}(t)\boldsymbol{\mathcal{A}}\Big\{\,\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\mathbf{L}^H(t) + \Big[\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) \\
&\quad - \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\Big]\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\Big[\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) - \bar{\mathbf{\Gamma}}(t)\mathbf{G}(t-1)\Big]^H
\end{aligned} \tag{78}$$

where $\mathbf{Q}(t)$ is given in Eq. (11) and

$$
\begin{aligned}
&\boldsymbol{\Omega}(t) = \mathbf{M}(t)\Big[\mathbf{1}_3^T \otimes \boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\
&\odot \left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\Big] - \mathbf{M}(t)\boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t)\right. \\
&\left. - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H\mathbf{L}^H(t) \\
&+ [\bar{\boldsymbol{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]\Big[\bar{\boldsymbol{\Gamma}}(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1}^H\bar{\boldsymbol{\Gamma}}^H(t) - \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] \\
&- \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{T}^H\bar{\boldsymbol{\Gamma}}^H(t)\Big] - [\bar{\boldsymbol{\Phi}}(t) - \mathbf{L}(t)\bar{\mathbf{H}}(t)]\Big[\bar{\boldsymbol{\Gamma}}(t-1)\mathbf{S}\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]) \\
&- \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\bar{\mathbf{v}}}\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\Big]\mathbf{L}^H(t) \\
&- \mathbf{1}_3 \otimes \Big[\boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big]\mathbf{L}^H(t) \\
&- \mathbf{1}_3 \otimes \Big[\boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H\Big]\mathbf{L}^H(t) \\
&+ \bar{\boldsymbol{\Gamma}}(t)\mathbf{G}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\Big[\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) - \bar{\boldsymbol{\Gamma}}(t)\mathbf{G}(t-1)\Big]^H \\
&- \mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\Big[\mathbf{L}(t)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\mathbf{F}(t-1) - \bar{\boldsymbol{\Gamma}}(t)\mathbf{G}(t-1)\Big]^H
\end{aligned}
\tag{79}
$$

By algebraically manipulating Eqs. (78) and (79) we derive Eq. (19).

Consider now the gain matrix $\mathbf{K}(t) = \mathrm{E}[\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t)]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t)$ in Eq. (13). Applying Eq. (54) and taking into account that $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)] = \mathbf{0}$, $\tilde{\underline{\mathbf{x}}}(t|t-1)\perp\hat{\underline{\mathbf{x}}}(t|t-1)$, Eqs. (40) and (47), we have

$$
\begin{aligned}
&\mathrm{E}[\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t)] = \mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t) + \mathbf{1}_3 \otimes \left[\mathcal{D}\left(\mathbf{1} - \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t-1)\right]\right.\right. \\
&\left.\left. - \mathrm{E}\left[\boldsymbol{\gamma}_3^r(t-1)\right]\right)\mathbf{Q}_{\bar{\mathbf{v}}}\right]\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right) - \mathbf{L}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)
\end{aligned}
$$

from which Eq. (20) holds. Taking Eq. (47) into account, we consider

$$
\mathrm{E}[\underline{\mathbf{x}}(t+1)\boldsymbol{\varepsilon}^H(t)] = \mathrm{E}[\tilde{\boldsymbol{\Phi}}(t)\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t)] + \mathrm{E}[\tilde{\boldsymbol{\Gamma}}(t)\underline{\mathbf{w}}(t)\boldsymbol{\varepsilon}^H(t)] \tag{80}
$$

From Eq. (54) and since $\mathrm{E}[\bar{\boldsymbol{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t) \triangle \tilde{\mathbf{H}}^H(t)] = \mathbf{0}$, the first summand in Eq. (80) can be rewritten as

$$\begin{aligned}\mathrm{E}[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t)] =& \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t) \triangle \tilde{\mathbf{H}}^H(t)] + \bar{\mathbf{\Phi}}(t)\mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t)\\ &+ \mathrm{E}\left[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t)\,\mathcal{D}\left(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)]\right)\right]\\ &+ \mathrm{E}\left[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\right]\\ &- \mathrm{E}\left[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\right] \qquad (81)\end{aligned}$$

Following a similar reasoning to that used for deriving the expression of $\mathbf{\Upsilon}(t)$ in Eq. (45), we get that the third and fourth summands in Eq. (81) are

$$\mathrm{E}\left[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t)\,\mathcal{D}\left(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)]\right)\right] = \mathbf{0}$$

$$\begin{aligned}&E\Big[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\Big] = \sum_{i=1}^{4} E\Big[\tilde{\mathbf{\Phi}}_i(t)\underline{\mathbf{x}}(t)\bar{\mathbf{v}}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\Big] =\\ &= \mathbf{1}_3 \otimes \Big[\boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[\boldsymbol{\gamma}_3^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \Big[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\Big]\Big\}\boldsymbol{\mathcal{A}}^H\Big]\end{aligned}$$

Moreover, taking Eq. (47) into account, the last summand in Eq. (81) is obtained from

$$\mathrm{E}\left[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t-1)\mathbf{F}^H(t-1)\right] = \bar{\mathbf{\Phi}}(t)\mathbf{L}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)$$

Thus, Eq. (81) is of the form

$$\begin{aligned}&\mathrm{E}[\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\boldsymbol{\varepsilon}^H(t)] = \mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t) \triangle \tilde{\mathbf{H}}^H(t)] + \bar{\mathbf{\Phi}}(t)\mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t)\\ &+\mathbf{1}_3 \otimes \Big[\boldsymbol{\mathcal{A}}\Big\{\mathrm{E}\left[\boldsymbol{\gamma}_3^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \Big[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\Big]\Big\}\boldsymbol{\mathcal{A}}^H\Big]\\ &- \bar{\mathbf{\Phi}}(t)\mathbf{L}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)\right) \qquad (82)\end{aligned}$$

As for the second summand in Eq. (80), it follows that

$$
\begin{aligned}
\mathrm{E}[\tilde{\boldsymbol{\Gamma}}(t)\underline{\mathbf{w}}(t)\boldsymbol{\varepsilon}^H(t)] = \mathbf{1}_3 \otimes \boldsymbol{\mathcal{A}}\Big\{ &\mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])^T\right] \\
&\odot \left[\mathbf{Q}_{\mathbf{v}^r}\, \mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\right]\Big\}\boldsymbol{\mathcal{A}}^H \\
+ \Big[\bar{\boldsymbol{\Gamma}}(t)\mathbf{Q}_{\underline{\mathbf{w}}_1}\bar{\boldsymbol{\Gamma}}^H(t-1) - \bar{\boldsymbol{\Gamma}}(t)\mathbf{T}\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] &- \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{L}^H(t-1)\Big]\bar{\mathbf{H}}^H(t) \\
+ \mathbf{1}_3 \otimes \boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right] &\odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H \\
+ \mathbf{1}_3 \otimes \boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \odot \mathbf{Q}_{\mathbf{v}^r}\Big\}&\boldsymbol{\mathcal{A}}^H \\
- \bar{\boldsymbol{\Gamma}}(t)\mathbf{G}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,&\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]) \quad (83)
\end{aligned}
$$

where the last equality is a consequence of Eqs. (54), (68), (70), (72), (73) and (52). By substituting Eqs. (82) and (83) into Eq. (80) we obtain Eq. (21).

Finally, we devise the updating equation for the innovation covariance matrix. For that, we use Eq. (54) and take into account that $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)] = \mathbf{0}$, Eqs. (63), (69), (59), (60), (71), $\tilde{\underline{\mathbf{x}}}(t|t-1)\perp\boldsymbol{\varepsilon}(t-1)$, Eqs. (75), (64), (61), (76) and (77):

$$
\begin{aligned}
\mathbf{Q}_{\boldsymbol{\varepsilon}}(t) = \mathrm{E}[\triangle\tilde{\mathbf{H}}(t)\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^H(t)\triangle\tilde{\mathbf{H}}^H(t)] + \bar{\mathbf{H}}(t)\mathbf{P}(t|t-1)\bar{\mathbf{H}}^H(t) + \boldsymbol{\Theta}(t) + \boldsymbol{\Theta}^H(t) \\
+\boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))(\mathbf{1} - \boldsymbol{\gamma}_2^r(t) - \boldsymbol{\gamma}_3^r(t))^T\right]\odot\mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H + \boldsymbol{\mathcal{A}}\Big\{ \mathrm{E}\left[\boldsymbol{\gamma}_2^r(t)\boldsymbol{\gamma}_2^{rT}(t)\right] \\
\odot \mathbf{Q}_{\mathbf{v}^r}\Big\}\boldsymbol{\mathcal{A}}^H + \mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\mathbf{F}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]) \quad (84)
\end{aligned}
$$

with

$$
\begin{aligned}
\boldsymbol{\Theta}(t) = \boldsymbol{\mathcal{A}}\Big\{ &\mathrm{E}\left[(\boldsymbol{\gamma}_3^r(t) - \mathrm{E}[\boldsymbol{\gamma}_3^r(t)])\boldsymbol{\gamma}_2^{rT}(t)\right] \\
&\odot \left[\mathrm{diag}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] - \mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\mathbf{v}^r}\right]\Big\}\boldsymbol{\mathcal{A}}^H \\
&+ \bar{\mathbf{H}}(t)\Big[\bar{\boldsymbol{\Gamma}}(t-1)\mathbf{S}\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]) - \mathbf{L}(t-1)\,\mathcal{D}(\mathbf{1} - \mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)] \\
-\mathrm{E}[\boldsymbol{\gamma}_3^r(t-1)])\mathbf{Q}_{\bar{\mathbf{v}}}\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])\Big] &- \mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\mathbf{F}(t-1)\mathbf{Q}_{\boldsymbol{\varepsilon}}(t-1)\mathbf{F}^H(t-1)\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)])
\end{aligned}
$$

From Eq. (84), Eq. (22) holds.

7.5. Proof of Theorem 2

By projecting $\underline{\mathbf{x}}(t+\tau)$, with $\tau \geq 2$, onto the linear space spanned by $\{\bar{\mathbf{y}}(0), \bar{\mathbf{y}}(1), \ldots, \bar{\mathbf{y}}(t)\}$ and taking Eq. (7) into account, we have

$$\hat{\underline{\mathbf{x}}}(t+\tau|t) = \bar{\mathbf{\Phi}}(t+\tau-1)\hat{\underline{\mathbf{x}}}(t+\tau-1|t) + \bar{\mathbf{\Gamma}}(t+\tau-1)\hat{\underline{\mathbf{w}}}(t+\tau-1|t) \tag{85}$$

with $\hat{\underline{\mathbf{w}}}(t+\tau-1|t) = \mathbf{G}(t)\boldsymbol{\varepsilon}(t)$ for $\tau = 2$ (see Eq. (50)) and $\hat{\underline{\mathbf{w}}}(t+\tau-1|t) = \mathbf{0}$ for $\tau > 2$. From which Eqs. (23) and (24) hold. Moreover, from Eq. (23) it follows that

$$\begin{aligned}\tilde{\underline{\mathbf{x}}}(t+2|t) = \underline{\mathbf{x}}(t+2) - \hat{\underline{\mathbf{x}}}(t+2|t) = \triangle\tilde{\mathbf{\Phi}}(t+1)\underline{\mathbf{x}}(t+1) + \bar{\mathbf{\Phi}}(t+1)\tilde{\underline{\mathbf{x}}}(t+1|t)\\ + \tilde{\mathbf{\Gamma}}(t+1)\underline{\mathbf{w}}(t+1) - \bar{\mathbf{\Gamma}}(t+1)\mathbf{G}(t)\boldsymbol{\varepsilon}(t)\end{aligned} \tag{86}$$

From Eq. (86) and taking into account that $\mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t+1)] = \mathbf{0}$, $\tilde{\underline{\mathbf{x}}}(t+1|t) \perp \boldsymbol{\varepsilon}(t)$, Eqs. (40), (12), (70) and (74) we get Eq. (25). On the other hand, from Eq. (24) we have, for $\tau > 2$, that

$$\begin{aligned}\tilde{\underline{\mathbf{x}}}(t+\tau|t) = \underline{\mathbf{x}}(t+\tau) - \hat{\underline{\mathbf{x}}}(t+\tau|t) = \triangle\tilde{\mathbf{\Phi}}(t+\tau-1)\underline{\mathbf{x}}(t+\tau-1)\\ + \bar{\mathbf{\Phi}}(t+\tau-1)\tilde{\underline{\mathbf{x}}}(t+\tau-1|t) + \tilde{\mathbf{\Gamma}}(t+\tau-1)\underline{\mathbf{w}}(t+\tau-1)\end{aligned} \tag{87}$$

From Eq. (87), and since $\mathrm{E}[\triangle\tilde{\mathbf{\Phi}}(t+\tau-1)] = \mathbf{0}$, $\hat{\underline{\mathbf{x}}}(t+\tau-1|t) \perp \underline{\mathbf{w}}(t+\tau-1)$ for $\tau > 2$, Eqs. (40) and (12), then Eq. (26) holds.

7.6. Proof of Theorem 3

By a similar procedure to that applied to derive Eq. (13) we get Eq. (27) with

$$\mathbf{N}(t+\tau|t) = \mathrm{E}[\underline{\mathbf{x}}(t+\tau)\boldsymbol{\varepsilon}^H(t)]\mathbf{Q}_{\boldsymbol{\varepsilon}}^{-1}(t) \tag{88}$$

Take the expectation in Eq. (88). From Eq. (54) and taking into account that $\mathrm{E}[\triangle\tilde{\mathbf{H}}(t)] = \mathbf{0}$ and Eq. (40), it follows

$$\begin{aligned}\mathrm{E}[\underline{\mathbf{x}}(t+\tau)\boldsymbol{\varepsilon}^H(t)] = \mathrm{E}[\underline{\mathbf{x}}(t+\tau)\tilde{\underline{\mathbf{x}}}^H(t|t-1)]\bar{\mathbf{H}}^H(t)\\ - \mathrm{E}[\underline{\mathbf{x}}(t+\tau)\boldsymbol{\varepsilon}^H(t-1)]\mathbf{F}^H(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t)]\right)\end{aligned}$$

from which, by denoting $\mathbf{\Lambda}_\tau(t) = \mathrm{E}[\underline{\mathbf{x}}(t+\tau)\tilde{\underline{\mathbf{x}}}^H(t|t-1)]$ and taking Eq. (88) into account, we obtain Eq. (28). Now, consider $\mathbf{\Lambda}_\tau(t)$. For $\tau = -1$ and from Eq. (57), we have

$$\begin{aligned}\mathbf{\Lambda}_{-1}(t) &= \mathrm{E}[\underline{\mathbf{x}}(t-1)\tilde{\underline{\mathbf{x}}}^H(t-1|t-2)]\big[\bar{\mathbf{\Phi}}(t-1) - \mathbf{L}(t-1)\bar{\mathbf{H}}(t-1)\big]^H \\ &+ \mathrm{E}[\underline{\mathbf{x}}(t-1)\underline{\mathbf{w}}^H(t-1)]\bar{\mathbf{\Gamma}}^H(t-1) - \mathrm{E}[\underline{\mathbf{x}}(t-1)\bar{\mathbf{v}}^H(t-2)]\,\mathcal{D}(\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)])\mathbf{L}^H(t-1) \\ &+ \mathrm{E}[\underline{\mathbf{x}}(t-1)\boldsymbol{\varepsilon}^H(t-2)]\Big[\mathbf{L}(t-1)\,\mathcal{D}\left(\mathrm{E}[\boldsymbol{\gamma}_2^r(t-1)]\right)\mathbf{F}(t-2) - \bar{\mathbf{\Gamma}}(t-1)\mathbf{G}(t-2)\Big]^H\end{aligned}$$

and taking into account Eq. (40) and that $\bar{\mathbf{\Gamma}}_1(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1} = \bar{\mathbf{\Gamma}}_3(t-1)\mathbf{Q}_{\underline{\mathbf{w}}_1} = \mathbf{0}$ we derive Eq. (29). Analogously, since $\underline{\mathbf{x}}(t+\tau)\perp\underline{\mathbf{w}}(t-1)$ and $\underline{\mathbf{x}}(t+\tau)\perp\bar{\mathbf{v}}(t-2)$ for $\tau < -1$, then Eq. (30) holds. Finally, from Eq. (27) we have that the smoothing error satisfies $\tilde{\underline{\mathbf{x}}}(t+\tau|t) = \tilde{\underline{\mathbf{x}}}(t+\tau|t-1) - \mathbf{N}(t+\tau|t)\boldsymbol{\varepsilon}(t)$, where $\tilde{\underline{\mathbf{x}}}(t+\tau|t) = \underline{\mathbf{x}}(t+\tau) - \hat{\underline{\mathbf{x}}}(t+\tau|t)$. Since $\mathbf{P}(t+\tau|t) = \mathrm{E}[\tilde{\underline{\mathbf{x}}}(t+\tau|t)\tilde{\underline{\mathbf{x}}}^H(t+\tau|t)]$, then Eq. (31) follows.

8. Acknowledgements

The work was supported by the University of Jaén (grant number EI-FQM2-2017).

References

[1] C. Hertzberg, R. Wagner, U. Frese, L. Schröder, Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds, Inf. Fusion 4 (1) (2013) 57–77.

[2] X. R. Li, Y. M. Zhu, J. Wang, C. Z. Han, Optimal linear estimation fusion. i. unified fusion rules, IEEE Trans. Inf. Theory 49 (9) (2003) 2192—-2208.

[3] S. L. Sun, Multi-sensor optimal information fusion kalman filter with application, Aerosp. Sci. Technol. 8 (1) (2004) 57—-62.

[4] E. Song, Y. Zhu, J. Zhou, Z. You, Optimal kalman filtering fusion with cross-correlated sensor noises, Automatica 43 (8) (2007) 1450—-1456.

[5] S. S. Gao, Y. M. Zhong, W. Li, Random weighting method for multisensor data fusion, IEEE Sensors J. 11 (9) (2011) 1955—-1961.

[6] S. P. Talebi, S. Kanna, D. P. Mandic, A distributed quaternion kalman filter with applications to smart grid and target tracking, IEEE Trans. Signal Inform. Process. Networks 2 (4) (2016) 477–488.

[7] M. Moayedi, Y. K. Foo, Y. C. Soh, Adaptive kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements, IEEE Trans. Signal Process. 58 (3) (2010) 1577–1588.

[8] J. Ma, S. Sun, Centralized fusion estimators for multisensor systems with random sensor delays, multiple packet dropouts and uncertain observations, IEEE Sensors J. 13 (4) (2013) 1228–1235.

[9] J. Ma, S. Sun, Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations, IEEE Trans. Signal Process. 59 (11) (2011) 5181–5192.

[10] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, Covariance-based estimation algorithms in networked systems with mixed uncertainties in the observations, IEEE Trans. Signal Process. 94 (2014) 163–173.

[11] J. B. Kuipers, Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality, Princenton Univ. Press, 2002.

[12] F. Lizarralde, J. T. Wen, Attitude control without angular velocity measurement: a passivity approach, IEEE Trans. Autom. Control 41 (3) (1996) 468–472.

[13] A. Mehrabian, K. Khorasani, Distributed and cooperative quaternion-based attitude synchronization and tracking control for a network of heterogeneous spacecraft formation flying mission, J. Franklin Inst. 352 (9) (2015) 3885–3913.

[14] L. Chang, F. Qin, F. Zha, Pseudo open-loop unscented quaternion estimator for attitude estimation, IEEE Sensors J. 16 (11) (2016) 4460–4469.

[15] M. B. Del Rosario, N. H. Lovell, S. J. Redmond, Quaternion-based complementary filter for attitude determination of a smartphone, IEEE Sensors J. 16 (15) (2016) 6008–6017.

[16] C. E. Moxey, S. J. Sangwine, T. Ell, Hypercomplex correlation techniques for vector images, IEEE Trans. Signal Process. 51 (7) (2003) 1941–1953.

[17] A. J. Hanson, Visualizing Quaternions, Morgan Kaufmann, 2006.

[18] D. Choukroun, I. Bar-Itzhack, Y. Oshman, Novel quaternion kalman filter, IEEE Trans. Aerosp. Electron. Syst. 42 (1) (2006) 174—-190.

[19] M. Hobiger, N. Le Bihan, C. Cornou, P. Y. Bard, Rayleigh wave ellipticity estimation from ambient seismic noise using single and multiple vector-sensor techniques, in: 17th European Signal Process. Conf., 2009, pp. 2037–2041.

[20] C. C. Took, K. Strbac, K. Aihara, D. P. Mandic, Quaternion-valued short term forecasting of three-dimensional wind and atmospheric parameters, Renewable Energy 36 (2) (2011) 1754–1760.

[21] F. Tobar, D. P. Mandic, Design of positive-definite quatenion kernels, IEEE Signal Process. Letters 22 (11) (2015) 2117–2121.

[22] J. Vía, D. Ramírez, I. Santamaría, Properness and widely linear processing of quaternion random vectors, IEEE Trans. Inform. Theory 56 (7) (2010) 3502–3515.

[23] C. C. Took, D. P. Mandic, Augmented second-order statistics of quaternion random signals, Signal Process. 91 (2011) 214–224.

[24] J. D. Jiménez-López, R. M. Fernández-Alcalá, J. Navarro-Moreno, J. C. Ruiz-Molina, Widely linear estimation of quaternion signals with intermittent observations, Signal Process. 136 (2017) 92–101.

[25] C. Jahanchahi, D. P. Mandic, A class of quaternion kalman filters, IEEE Trans. Neural Netw. Learn. Syst. 25 (3) (2014) 533–544.

[26] D. H. Dini, C. Jahanchahi, D. P. Mandic, Kalman filtering for widely linear complex and quaternion valued bearings only tracking, IET Signal Process. 6 (5) (2012) 435—-445.

[27] X. Yuan, S. Yu, S. Zhang, G. Wang, S. Liu, Quaternion-based unscented kalman filter for accurate indoor heading estimation using wearable multi-sensor system, Sensors 15 (5) (2015) 10872–10890.

[28] J. Chen, L. Zhou, Y. Zhang, D. F. Ferreiro, Human motion tracking with wireless wearable sensor network: experience and lessons, KSII Trans. Internet Inform. Systems 7 (5) (2013) 998–1013.

[29] R. Caballero-Águila, A. Hermoso-Carazo, J. Linares-Pérez, Covariance-based estimation algorithms in networked systems with mixed uncertainties in the observations, Signal Process. 94 (2014) 163–173.

[30] X. Yun, E. R. Bachmann, Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking, IEEE Trans. Robotics 22 (6) (2006) 1216–1227.