Semi-Widely Linear Estimation Algorithms of Quaternion Signals with Missing Observations and Correlated Noises

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

Department of Statistics and Operations Research, University of Jaén, 23071 Jaén, Spain. E-mails: {rmfernan,jnavarro,jdomingo,jcruiz}@ujaen.es. Tel:+34 953212449, Fax:+34 953212034.

Abstract

The paper deals with the estimation problem of a discrete-time vectorial quaternion signal which is observed through a linear dynamic system with intermittent observations and autocorrelated and cross-correlated noises. Under $\mathbb{C}^\eta$-properness conditions, a semi-widely linear processing is considered to provide filtering, fixed-point and fixed-lag smoothing algorithms for estimating the quaternion state. The proposed solutions give a substantial reduction in computational burden in relation to the widely linear estimation techniques, this benefit being impossible to be fully captured by using a real number framework. The feasibility and performance of the aforementioned algorithms are illustrated by means of simulation examples.

Keywords: $\mathbb{C}^\eta$-properness, filtering, fixed-point and fixed-lag smoothing algorithms, intermittent observations, quaternion signals, semi-widely linear processing.

1. Introduction

Traditionally, Kalman filter and its different extensions have been analyzed for real-valued signals (e.g., [1]–[5]). Nevertheless, the progress in technology and vector sensors have brought to light the necessity of processing multidimensional and multivariate signals, where the enhancement of using the quaternion

algebra has been highlighted as an important tool in this task (see, e.g., [6]-[10]). For example, the quaternion domain has usually been employed to model three-dimensional rotations and orientation in a compact and computationally efficient fashion. Actually, 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]-[15]. Other applications where the quaternion domain has been shown to perform admirably include color image processing [16], computer graphics [17], aeronautics [18, 19], seismology [20], meteorology [21], quantum physics [22], etc.

In the quaternion domain, there are different types of processing whose use depends on the kind of properness of the signal to be processed [23]. The most general processing is the so-called widely linear (WL), in which we must simultaneously operate on the quaternion signal and its three involutions. Such a processing is the one suitable for improper signals, being equivalent to the linear processing in $\mathbb{R}^4$ [7]. In this framework, several WL Kalman algorithms have been proposed in the literature (e.g., [9], [19], [24]-[26]).

Other types of processing are appropriate when the signal presents 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. The main practical repercussion of both SL and SWL processing is the significative reduction in the computational burden of the associated algorithms in comparison with the WL ones [27]. Notably, this saving in computational complexity cannot be attained when the algorithms are formulated in the real domain.

It is well-known that proper signals need a particular treatment. In fact, algorithms adapted for improper signals can fail or suffer from slow convergence when they are used for proper signals [28]. In this way, since the seminal work of Vakhania [29], extensive research has been done on studying the properness

properties of quaternion valued signals and on devising specific algorithms. For example, the concept of properness has been explored in [7, 23, 30, 31] and several tests to check experimentally whether a quaternion random vector is proper have been given in [32]-[34]. The relevance of this research field is also confirmed by the number of its applications such as signal detection [35, 36], ICA [37, 38], signal estimation [6],[39]-[41], random monogenic signal [42], quaternion VAR modelling [43], Gaussian graphical models [44], directionality in random fields [45], classification of polarized signals [46], etc.

On the other hand, in the conventional formulation of the estimation problem it is assumed that all the observations are available during the state estimation time, since Kalman filters are sensitive to incomplete or missing measurements [47]. However, in many real-world applications, such as in the case of failures in the measuring sensors, high noise environments, and poor communication resources [48], there is a nonzero probability that the observed state contains only noise and the state is absent. Systems suffering from uncertainty in the observation process are commonly referred to as systems with intermittent observations. This class of problems was first solved by Nahi [49] by assuming that the uncertainty in the observation is characterized by an independent and identically distributed Bernoulli process, and it has been studied extensively in recent years (see, e.g., [50]-[61]).

The task of estimating the quaternion signal from intermittent observations presents similar drawbacks to those shown in the real domain, i.e., the classical estimation methods perform suboptimally and cannot be used directly. Several WL and SWL estimation algorithms for systems with intermittent observations have been investigated in [62]. The above WL algorithms were later extended to quaternion systems with mixed uncertainties of sensor delays, packet dropouts and uncertain observations in [63]. These two papers address the estimation problem with intermittent observations from a common setting: they are restricted to one-dimensional quaternion state vectors and the state and measurements noises are assumed to be white and non-cross-correlated. Obviously, these assumptions may not be realistic and can be a limitation in many real

world problems in which noise correlation may be present and/or a vectorial state is needed. Such problems arise, for example, when a target is taking an electronic countermeasure (e.g., noise jamming) or if the signal noise and the sensor measurement noises are dependent on the system state. Also, if all the sensors are observed in the same noisy environment, the measurement noises of different sensors are usually correlated (see [60, 64] and the references therein). Hence, an extension of these approaches dealing with multidimensional quaternion systems with colored and correlated noises is required.

Motivated by the above considerations, this paper addresses the optimal SWL estimation problem for multidimensional systems with intermittent observations when the state and observation noises are correlated and colored. Specifically, filtering, fixed-point and fixed-lag smoothing algorithms are devised under the hypothesis of $\mathbb{C}^\eta$-properness and by using quaternion Bernoulli variables to model the missing measurements phenomenon. For that, we propose a quaternion system in which the missing measurement model considers at each sensor the possibility of observation containing only partial information about the state, or even only noise. Moreover, the process and the measurement noises present contemporaneous autocorrelation and cross-correlation as well as one-step autocorrelation and cross-correlation. As remarked above, the properness hypotheses implies that the suggested SWL recursive procedures provide a substantial saving in computational burden in comparison with their WL counterparts (and thus, with their real-valued versions). Really, since the SWL processing is not equivalent to a real-valued processing, this saving cannot be attained by using the real number domain.

Then, the rest of the paper is organized as follows. In Section 2 we introduce the notation and operations of quaternion algebra utilized throughout. Section 3 defines a SWL state-space model with intermittent observations, suitable for modeling $\mathbb{C}^\eta$-proper vectorial signals affected by colored and correlated noises. Section 4 addresses the corresponding methods for the SWL filtering, fixed-point, and fixed-lag smoothing algorithms. Additionally, simulation examples are provided in Section 5 where the effectiveness of the proposed algorithms

is analyzed in terms of the uncertainty. Finally, our concluding remarks are included in Section 6. To improve readability, all technical proofs have been deferred to an Appendix.

2. Preliminaries

Throughout this paper, all the random variables are assumed to have zero-mean. Next, we introduce the notation and then some basic concepts and properties.

2.1. Notation

We use boldfaced upper case letters to denote matrices, boldfaced lower case letters for column vectors, and lightfaced letters for scalar quantities. $\mathbf{I}_m$ denotes identity matrix of dimension m, $\mathbf{0}_{l\times m}$ denotes the $l \times m$ zero matrix and $\mathbf{0}_m$ a zero vector of length m. Superscripts $(\cdot)^*$, $(\cdot)^{\mathtt{T}}$ and $(\cdot)^{\mathtt{H}}$ represent complex conjugate, transpose and conjugate transpose, respectively. Moreover, subscripts $(\cdot)_r$ and $(\cdot)_\nu$, for $\nu = \eta, \eta', \eta''$, represent the real and imaginary parts of a quaternion. The notation $\mathbf{A} \in \mathbb{R}^{l\times m}$ (respectively $\mathbf{A} \in \mathbb{H}^{l\times m}$) means that $\mathbf{A}$ is a real (respectively quaternion) $l \times m$ matrix. Similarly, $\mathbf{r} \in \mathbb{R}^m$ (respectively, $\mathbf{r} \in \mathbb{H}^m$) denotes the m-dimensional real (respectively quaternion) vector. $E[\cdot]$ and $\mathrm{Cov}(\cdot)$ are the expectation and covariance operators and $\mathrm{diag}(\cdot)$ is a diagonal matrix with the elements specified on the main diagonal. $\delta_{n,l}$ is the Kronecker delta function, which is equal to one if $l = n$, and zero otherwise. Finally, "$\circ$" and "$\otimes$" denote the Hadamard and the Kronecker product, respectively.

2.2. Basic Concepts and Properties

The next property of the Hadamard product will be useful: if $\mathbf{A} \in \mathbb{R}^{m\times m}$ and $\mathbf{b} \in \mathbb{R}^m$, then $\mathrm{diag}(\mathbf{b})\mathbf{A}\,\mathrm{diag}(\mathbf{b}) = (\mathbf{b}\mathbf{b}^{\mathtt{T}}) \circ \mathbf{A}$.

Let $\mathbf{x}_n \in \mathbb{H}^m$ be a m-dimensional quaternion random signal given by

$$\mathbf{x}_n = \mathbf{x}_{n,r} + \eta\mathbf{x}_{n,\eta} + \eta'\mathbf{x}_{n,\eta'} + \eta''\mathbf{x}_{n,\eta''}$$

where $\mathbf{x}_{n,\nu} \in \mathbb{R}^m$, $\nu = r, \eta, \eta', \eta''$, are zero-mean real random signals and $\{1, \eta, \eta', \eta''\}$ obeys the following rules:

$$\begin{aligned}
\eta^2 = \eta'^2 = \eta''^2 = \eta\eta'\eta'' = -1 \\
\eta\eta' = \eta'' = -\eta'\eta \\
\eta'\eta'' = \eta = -\eta''\eta' \\
\eta''\eta = \eta' = -\eta\eta''
\end{aligned}$$

In the quaternion domain, the complete description of the second-order statistical properties of $\mathbf{x}_n$ is given by any combination of four elements among the signal, the three perpendicular quaternion involutions $\mathbf{x}_n^\nu = -\nu\mathbf{x}_n\nu$, $\nu = \eta, \eta', \eta''$ or their conjugates [23]. For example, the augmented quaternion signal vector $\bar{\mathbf{x}}_n = [\mathbf{x}_n^{\mathtt{T}}, \mathbf{x}_n^{\eta\mathtt{T}}, \mathbf{x}_n^{\eta'\mathtt{T}}, \mathbf{x}_n^{\eta''\mathtt{T}}]^{\mathtt{T}}$ can be used for this purpose.

Note that the following relationship between the augmented vector and the real vector $\mathbf{x}_n^r = [\mathbf{x}_{n,r}^{\mathtt{T}}, \mathbf{x}_{n,\eta}^{\mathtt{T}}, \mathbf{x}_{n,\eta'}^{\mathtt{T}}, \mathbf{x}_{n,\eta''}^{\mathtt{T}}]^{\mathtt{T}}$ can be established

$$\bar{\mathbf{x}}_n = 2\boldsymbol{\mathcal{T}}_m \mathbf{x}_n^r \tag{1}$$

where $\boldsymbol{\mathcal{T}}_m = \frac{1}{2}\boldsymbol{\mathcal{A}} \otimes \mathbf{I}_m$ and

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

with $\boldsymbol{\mathcal{T}}_m^{\mathtt{H}}\boldsymbol{\mathcal{T}}_m = \mathbf{I}_{4m}$.

Definition 1. *Given two quaternion random signals $\mathbf{x}_n, \mathbf{y}_l \in \mathbb{H}^m$, the product $\odot$ between them is defined as*

$$\mathbf{x}_n \odot \mathbf{y}_l = \mathbf{x}_{n,r} \circ \mathbf{y}_{l,r} + \eta\mathbf{x}_{n,\eta} \circ \mathbf{y}_{l,\eta} + \eta'\mathbf{x}_{n,\eta'} \circ \mathbf{y}_{l,\eta'} + \eta''\mathbf{x}_{n,\eta''} \circ \mathbf{y}_{l,\eta''}, \quad \forall n, l \tag{2}$$

The following property of the product $\odot$ is easy to check.

Property 1. *The augmented vector of $\mathbf{x}_n \odot \mathbf{y}_l$ is $\overline{\mathbf{x}_n \odot \mathbf{y}_l} = \boldsymbol{\mathcal{T}}_m \operatorname{diag}(\mathbf{x}_n^r)\boldsymbol{\mathcal{T}}_m^H \bar{\mathbf{y}}_l$.*

On the other hand, for a quaternion signal there exist three different kinds of quaternion properness ($\mathbb{Q}$-properness, $\mathbb{C}^{\eta}$-properness and improperness), which are based on the vanishing of three different complementary functions [33]. This paper focuses on $\mathbb{C}^{\eta}$-properness which is defined below.

Definition 2. *A quaternion random signal* $\mathbf{x}_n \in \mathbb{H}^m$ *is said to be* $\mathbb{C}^{\eta}$*-proper if the complementary matrices* $E\left[\mathbf{x}_n \mathbf{x}_k^{\eta' H}\right]$ *and* $E\left[\mathbf{x}_n \mathbf{x}_k^{\eta'' H}\right]$ *vanish for all* n *and* k*. Similarly, two quaternion random signals* $\mathbf{x}_n \in \mathbb{H}^m$ *and* $\mathbf{z}_k \in \mathbb{H}^l$ *are cross* $\mathbb{C}^{\eta}$*-proper if the complementary matrices* $E\left[\mathbf{x}_n \mathbf{z}_k^{\eta' H}\right]$ *and* $E\left[\mathbf{x}_n \mathbf{z}_k^{\eta'' H}\right]$ *vanish for all* n *and* k*.*

The kind of properness determines the type of linear processing that one has to follow in the quaternion domain. The most general linear processing approach, called WL processing, requires the operation on the augmented quaternion signal vector $\bar{\mathbf{x}}_n$, which offers performance advantages in relation to the conventional processing (see, e.g., [7] and [9]). The WL processing is also 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 [7]. The condition of $\mathbb{C}^{\eta}$-properness gives rise to the so-called SWL processing which only makes use of the quaternion signal and its involution over the pure unit quaternion η, i.e., the $2m$-dimensional augmented vector $\tilde{\mathbf{x}}_n = [\mathbf{x}_n^{\mathrm{T}}, \mathbf{x}_n^{\eta \mathrm{T}}]^{\mathrm{T}}$. SWL processing presents several mathematical advantages with respect to WL processing since it alleviates the computational burden involved [27]. Moreover, such advantages cannot be attained when a real-valued signal processing is used [1].

[1] $\mathbb{C}^{\eta}$-properness implies that the second-order statistical properties of $\bar{\mathbf{x}}_n$ are completely determined by $E\left[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_k^{\mathrm{H}}\right]$ [23]. However, the orthogonality properties in the components of $\bar{\mathbf{x}}_n$ under this condition are not transferred to the components of $\mathbf{x}_n^r$ and hence $\mathbb{C}^{\eta}$-properness does not lead to a simplified structure of $E[\mathbf{x}_n^r \mathbf{x}_k^{r\mathrm{T}}]$ as compared to $E\left[\bar{\mathbf{x}}_n \bar{\mathbf{x}}_k^{\mathrm{H}}\right]$ [62]. As a consequence, it is not possible to express a SWL algorithm equivalently through a real formalism while maintaining the same computational complexity.

3. WL and SWL State-Space Models

Consider a quaternion state $\mathbf{x}_n \in \mathbb{H}^m$ defined by the following WL state model:

$$\mathbf{x}_n = \mathbf{F}_n\mathbf{x}_{n-1} + \mathbf{G}_n\mathbf{x}_{n-1}^{\eta} + \mathbf{H}_n\mathbf{x}_{n-1}^{\eta'} + \mathbf{E}_n\mathbf{x}_{n-1}^{\eta''} + \mathbf{u}_{n-1}, \quad n \geq 1 \tag{3}$$

where $\mathbf{F}_n, \mathbf{G}_n, \mathbf{H}_n, \mathbf{E}_n \in \mathbb{H}^{m\times m}$ are deterministic matrices and $\mathbf{u}_n$ is a quaternion noise. The state cannot be observed directly but it is obtained from the following quaternion observation equation:

$$\mathbf{z}_n = \boldsymbol{\gamma}_n \odot \mathbf{x}_n + \mathbf{v}_n, \quad n \geq 1 \tag{4}$$

where the product $\odot$ is defined by (2), being $\boldsymbol{\gamma}_n = [\gamma_{n,1}, \ldots, \gamma_{n,m}]^{\mathsf{T}} \in \mathbb{H}^m$ a vector of quaternion Bernoulli random variables and $\mathbf{v}_n$ is a quaternion noise. Each component $\gamma_{n,i}$ of $\boldsymbol{\gamma}_n$ is composed by independent Bernoulli random variables, $\gamma_{n,i,\nu}$, $\nu = r, \eta, \eta', \eta''$, with probability $p_{n,i,\nu}$, respectively. Moreover, the Bernoulli variables $\boldsymbol{\gamma}_n$ are independent of $\mathbf{x}_n$, $\mathbf{u}_n$ and $\mathbf{v}_n$. The initial state $\mathbf{x}_0$ is also independent of the additive noises $\{\mathbf{u}_n; n \geq 0\}$ and $\{\mathbf{v}_n; n \geq 1\}$.

Notice that the quaternion Bernoulli variables $\gamma_{n,i}$ indicate the presence or absence of each state quaternion component $x_{n,i,\nu}$ in the observation and this one is determined by the probabilities $p_{n,i,\nu}$ which are assumed to be known. In fact, as $p_{n,i,\nu}$ is closer to 1, there is a greater probability that the observation at time instant n, $z_{n,i,\nu}$, contains the quaternion component $x_{n,i,\nu}$ (and, conversely, as $p_{n,i,\nu}$ is closer to 0), the extreme cases being: $p_{n,i,\nu} = 1$ for $\nu = r, \eta, \eta', \eta''$, which indicates that $x_{n,i,\nu}$ is always present in $z_{n,i,\nu}$, and $p_{n,i,\nu} = 0$ for $\nu = r, \eta, \eta', \eta''$, it is always absent. This formulation allows us to assign different probabilities to the quaternion components of being affected by this uncertainty at every time instant n and depending on the sensor involved in the transmission [62, 63].

By considering the augmented state, $\bar{\mathbf{x}}_n$, and the augmented observation, $\bar{\mathbf{z}}_n$, and applying Property 1 on system (3)-(4), the following WL state-space

model is derived:

$$\bar{\mathbf{x}}_n = \boldsymbol{\Phi}_n \bar{\mathbf{x}}_{n-1} + \bar{\mathbf{u}}_{n-1} \tag{5}$$

$$\bar{\mathbf{z}}_n = \boldsymbol{\Gamma}_n \bar{\mathbf{x}}_n + \bar{\mathbf{v}}_n \tag{6}$$

where

$$\boldsymbol{\Phi}_n = \begin{bmatrix} \mathbf{F}_n & \mathbf{G}_n & \mathbf{H}_n & \mathbf{E}_n \\ \mathbf{G}_n^{\eta} & \mathbf{F}_n^{\eta} & \mathbf{E}_n^{\eta} & \mathbf{H}_n^{\eta} \\ \mathbf{H}_n^{\eta'} & \mathbf{E}_n^{\eta'} & \mathbf{F}_n^{\eta'} & \mathbf{G}_n^{\eta'} \\ \mathbf{E}_n^{\eta''} & \mathbf{H}_n^{\eta''} & \mathbf{G}_n^{\eta''} & \mathbf{F}_n^{\eta''} \end{bmatrix},$$

and $\boldsymbol{\Gamma}_n = \boldsymbol{\mathcal{T}}_m \operatorname{diag}(\boldsymbol{\gamma}_n^r) \boldsymbol{\mathcal{T}}_m^{\mathrm{H}}$, with $\boldsymbol{\mathcal{T}}_m$ defined in (1). Additionally, the augmented noises $\bar{\mathbf{u}}_n$ and $\bar{\mathbf{v}}_n$ are assumed to be cross-correlated and autocorrelated. Specifically, the following statistical properties are assumed:

i) $E[\bar{\mathbf{u}}_n \bar{\mathbf{u}}_l^{\mathrm{H}}] = \mathbf{Q}_n \delta_{n,l} + \mathbf{U}_n \delta_{n,l+1}$

ii) $E[\bar{\mathbf{v}}_n \bar{\mathbf{v}}_l^{\mathrm{H}}] = \mathbf{R}_n \delta_{n,l} + \mathbf{V}_n \delta_{n,l+1}$

iii) $E[\bar{\mathbf{u}}_n \bar{\mathbf{v}}_l^{\mathrm{H}}] = \mathbf{T}_n \delta_{n,l} + \mathbf{W}_n \delta_{n,l-1}$

Assumptions *i)* and *ii)* imply that both the state and observation noises are autocorrelated at the same time and one time step apart, while assumption *iii)* involves contemporaneous as well as one lagged correlation between $\bar{\mathbf{u}}_n$ and $\bar{\mathbf{v}}_n$.

An example of the above scenario can be found in the problem of dynamic targeting tracking (see, e.g., [60]) and, in particular, in the problem of tracking the rotations of an aircraft [19], where the quaternion domain has provided a robust mathematical framework for problems associated with gimbal lock.

The standard quaternion Kalman filter devised in [9] is based on a particular case of model (5)-(6) obtained by setting $p_{n,i,\nu} = 1$ for $\nu = r, \eta, \eta', \eta''$ and by removing the correlation hypotheses on the noises, i.e., they assume that the observations are always present and the state and observation noises are both white and uncorrelated.

With the purpose of a SWL processing, we state the following property whose proof is similar to that of Proposition 1 in [62].

Property 2. *Consider the quaternion state* $\mathbf{x}_n$ *defined in* (3).

i) *If* $\mathbf{x}_0$ *and* $\mathbf{u}_n$ *are* $\mathbb{C}^\eta$*-proper then,* $\mathbf{x}_n$ *is* $\mathbb{C}^\eta$*-proper.*

ii) *If* $\mathbf{x}_n$ *is* $\mathbb{C}^\eta$*-proper and cross* $\mathbb{C}^\eta$*-proper with* $\mathbf{u}_n$ *then,* $\mathbf{u}_n$ *is* $\mathbb{C}^\eta$*-proper and* $\boldsymbol{\Phi}_n$ *in* (5) *is block-diagonal, i.e,* $\boldsymbol{\Phi}_n = \begin{bmatrix} \tilde{\boldsymbol{\Phi}}_n & \mathbf{0}_{2m\times 2m} \\ \mathbf{0}_{2m\times 2m} & \tilde{\boldsymbol{\Phi}}_n^{\eta'} \end{bmatrix}$ *with*

$$\tilde{\boldsymbol{\Phi}}_n = \begin{bmatrix} \mathbf{F}_n & \mathbf{G}_n \\ \mathbf{G}_n^\eta & \mathbf{F}_n^\eta \end{bmatrix} \tag{7}$$

iii) *If* $\mathbf{x}_n$ *is* $\mathbb{C}^\eta$*-proper and cross* $\mathbb{C}^\eta$*-proper with* $\mathbf{z}_n$ *and* $\mathbf{v}_n$ *then,* $\boldsymbol{\Pi}_n = E[\boldsymbol{\Gamma}_n]$ *is block-diagonal, i.e.,*

$$\boldsymbol{\Pi}_n = \begin{bmatrix} \tilde{\boldsymbol{\Pi}}_n & \mathbf{0}_{2m\times 2m} \\ \mathbf{0}_{2m\times 2m} & \tilde{\boldsymbol{\Pi}}_n \end{bmatrix}$$

with

$$\tilde{\boldsymbol{\Pi}}_n = \frac{1}{2}\begin{bmatrix} \tilde{\boldsymbol{\Pi}}_n^{(1)} & \tilde{\boldsymbol{\Pi}}_n^{(2)} \\ \tilde{\boldsymbol{\Pi}}_n^{(2)} & \tilde{\boldsymbol{\Pi}}_n^{(1)} \end{bmatrix} \tag{8}$$

$\tilde{\boldsymbol{\Pi}}_n^{(1)} = \mathrm{diag}(p_{n,1,r} + p_{n,1,\eta'}, \ldots, p_{n,m,r} + p_{n,m,\eta'})$ *and* $\tilde{\boldsymbol{\Pi}}_n^{(2)} = \mathrm{diag}(p_{n,1,r} - p_{n,1,\eta'}, \ldots, p_{n,m,r} - p_{n,m,\eta'})$.

From now on, we assume that $\mathbf{x}_n$ is $\mathbb{C}^\eta$-proper and cross $\mathbb{C}^\eta$-proper with $\mathbf{z}_n$, $\mathbf{u}_n$ and $\mathbf{v}_n$. Define

$$\tilde{\boldsymbol{\Gamma}}_n = \tilde{\boldsymbol{\mathcal{T}}}_m \,\mathrm{diag}(\boldsymbol{\gamma}_n^r)\boldsymbol{\mathcal{T}}_m^{\mathtt{H}} \tag{9}$$

where $\tilde{\boldsymbol{\mathcal{T}}}_m = \frac{1}{2}\boldsymbol{\mathcal{B}} \otimes \mathbf{I}_m$ with

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

and consider the following observation equation

$$\tilde{\mathbf{z}}_n = \tilde{\boldsymbol{\Gamma}}_n \bar{\mathbf{x}}_n + \tilde{\mathbf{v}}_n \tag{10}$$

with $\tilde{\mathbf{z}}_n = [\mathbf{z}_n^{\mathtt{T}}, \mathbf{z}_n^{\eta\mathtt{T}}]^{\mathtt{T}}$, $\tilde{\mathbf{v}}_n = [\mathbf{v}_n^{\mathtt{T}}, \mathbf{v}_n^{\eta\mathtt{T}}]^{\mathtt{T}}$, $\tilde{\mathbf{R}}_n = E[\tilde{\mathbf{v}}_n\tilde{\mathbf{v}}_n^{\mathtt{H}}]$ and $\tilde{\mathbf{V}}_n = E[\tilde{\mathbf{v}}_n\tilde{\mathbf{v}}_{n-1}^{\mathtt{H}}]$. Also, the correlation hypotheses imposed above on $\bar{\mathbf{u}}_n$ and $\bar{\mathbf{v}}_n$ give rise to $E[\bar{\mathbf{u}}_n\tilde{\mathbf{v}}_l^{\mathtt{H}}] = \begin{bmatrix} \tilde{\mathbf{T}}_n \\ \tilde{\mathbf{S}}_n \end{bmatrix}\delta_{n,l} + \begin{bmatrix} \tilde{\mathbf{W}}_n \\ \tilde{\mathbf{X}}_n \end{bmatrix}\delta_{n,l-1}$, where $\tilde{\mathbf{T}}_n = E[\tilde{\mathbf{u}}_n\tilde{\mathbf{v}}_n^{\mathtt{H}}]$, $\tilde{\mathbf{S}}_n = E[\tilde{\mathbf{u}}_n^{\eta'}\tilde{\mathbf{v}}_n^{\mathtt{H}}]$, $\tilde{\mathbf{W}}_n = E[\tilde{\mathbf{u}}_n\tilde{\mathbf{v}}_{n+1}^{\mathtt{H}}]$ and $\tilde{\mathbf{X}}_n = E[\tilde{\mathbf{u}}_n^{\eta'}\tilde{\mathbf{v}}_{n+1}^{\mathtt{H}}]$.

The state-space system given by equations (5) and (10) is the starting point to devise our SWL estimation algorithms. This system generalizes the one defined in [62] in three ways: *i)* it is formulated for vectorial quaternion signals, *ii)* correlation between the state noise and the measurement noise is admitted, and now *iii)* both noises are colored.

Remark 1. *Notice that the state equation* (5) *is equivalent to the SWL state equation:*

$$\tilde{\mathbf{x}}_n = \tilde{\boldsymbol{\Phi}}_n \tilde{\mathbf{x}}_{n-1} + \tilde{\mathbf{u}}_{n-1}$$

where $\tilde{\boldsymbol{\Phi}}_n$ *is given in* (7) *and* $\tilde{\mathbf{u}}_n = [\mathbf{u}_n^T, \mathbf{u}_n^{\eta T}]^T$. *This equation cannot be used as a state equation together with* (10) *since the latter involves the augmented vector* $\bar{\mathbf{x}}_n$ *as state. However, it allows us to obtain the following recursion for the correlation matrix of* $\tilde{\mathbf{x}}_n$ *which will be valuable in our proposed algorithms:*

$$E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^H] = \tilde{\boldsymbol{\Phi}}_n E[\tilde{\mathbf{x}}_{n-1} \tilde{\mathbf{x}}_{n-1}^H] \tilde{\boldsymbol{\Phi}}_n^H + \tilde{\boldsymbol{\Phi}}_n \tilde{\mathbf{U}}_{n-1}^H + \tilde{\mathbf{U}}_{n-1} \tilde{\boldsymbol{\Phi}}_n^H + \tilde{\mathbf{Q}}_{n-1} \qquad (11)$$

with $\tilde{\mathbf{Q}}_n = E[\tilde{\mathbf{u}}_n \tilde{\mathbf{u}}_n^H]$, $\tilde{\mathbf{U}}_n = E[\tilde{\mathbf{u}}_n \tilde{\mathbf{u}}_{n-1}^H]$ *and* $E[\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_1^H] = \tilde{\boldsymbol{\Phi}}_1 E[\tilde{\mathbf{x}}_0 \tilde{\mathbf{x}}_0^H] \tilde{\boldsymbol{\Phi}}_1^H + \tilde{\mathbf{Q}}_0$.

4. SWL Estimation Algorithms

In this section, the optimal (in the least-squares sense) filtering, fixed-point and fixed-lag smoothing problems are addressed for a signal $\mathbf{x}_n$ satisfying the state equation (3) and subject to randomly missing observations obeying equation (4). The approach that we use is a SWL processing, i.e., we aim to devise recursive algorithms to obtain the optimal estimator of $\mathbf{x}_n$ on the basis of the observations $\{\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2, \ldots, \tilde{\mathbf{z}}_l\}$, denoted by $\hat{\mathbf{x}}_{n|l}$. For that, we follow the next steps. Firstly, we operate on the system (5) and (10) to derive the optimal estimators of the augmented state vector. Then we apply the block-diagonality properties of the state and correlation matrices involved, which are induced by the $\mathbb{C}^\eta$-properness hypotheses assumed above, and remove the redundant parts giving rise to the estimator $\hat{\tilde{\mathbf{x}}}_{n|l}$. Finally, $\hat{\mathbf{x}}_{n|l}$ is obtained as a subvector of $\hat{\tilde{\mathbf{x}}}_{n|l}$.

The SWL filtering, fixed-point smoothing and fixed-lag smoothing algorithms are presented in detail below.

Theorem 1 (SWL filter). *The optimal SWL filter of* $\mathbf{x}_n$ *is obtained as*

$$\hat{\mathbf{x}}_{n|n} = [\mathbf{I}_m, \mathbf{0}_{m\times m}]\hat{\tilde{\mathbf{x}}}_{n|n}$$

where $\hat{\tilde{\mathbf{x}}}_{n|n}$ *is recursively computed from the expression*

$$\begin{aligned} \hat{\tilde{\mathbf{x}}}_{n|n} &= \hat{\tilde{\mathbf{x}}}_{n|n-1} + \tilde{\mathbf{\Theta}}_n \tilde{\mathbf{\Omega}}_n^{-1} \tilde{\mathbf{y}}_{n|n-1}, \quad n \geq 1 \\ \hat{\tilde{\mathbf{x}}}_{0|0} &= \mathbf{0}_{2m}, \end{aligned} \tag{12}$$

the one-stage predictor is given by the equation

$$\hat{\tilde{\mathbf{x}}}_{n|n-1} = \tilde{\mathbf{\Phi}}_n \hat{\tilde{\mathbf{x}}}_{n-1|n-1} + \tilde{\mathbf{\Psi}}_{n-1} \tilde{\mathbf{y}}_{n-1|n-2}, \quad n \geq 1 \tag{13}$$

where $\tilde{\mathbf{\Psi}}_n = [\tilde{\mathbf{U}}_n \tilde{\mathbf{\Pi}}_n + \tilde{\mathbf{T}}_n]\tilde{\mathbf{\Omega}}_n^{-1}$ *and the innovations are calculated as*

$$\begin{aligned} \tilde{\mathbf{y}}_{n|n-1} &= \tilde{\mathbf{z}}_n - \tilde{\mathbf{\Pi}}_n \hat{\tilde{\mathbf{x}}}_{n|n-1} - \tilde{\mathbf{V}}_n \tilde{\mathbf{\Omega}}_{n-1}^{-1} \tilde{\mathbf{y}}_{n-1|n-2}, \quad n \geq 1 \\ \tilde{\mathbf{y}}_{0|-1} &= \mathbf{0}_{2m} \end{aligned} \tag{14}$$

with $\tilde{\mathbf{\Pi}}_n$ *given in* (8).

The matrix $\tilde{\mathbf{\Theta}}_n$ *is computed through the equation*

$$\tilde{\mathbf{\Theta}}_n = \tilde{\mathbf{P}}_{n|n-1}\tilde{\mathbf{\Pi}}_n + \tilde{\mathbf{W}}_{n-1} - [\tilde{\mathbf{\Phi}}_n \tilde{\mathbf{\Theta}}_{n-1} \tilde{\mathbf{\Omega}}_{n-1}^{-1} + \tilde{\mathbf{\Psi}}_{n-1}]\tilde{\mathbf{V}}_n^H, \quad n > 1 \tag{15}$$

$$\tilde{\mathbf{\Theta}}_1 = \tilde{\mathbf{P}}_{1|0}\tilde{\mathbf{\Pi}}_1 + \tilde{\mathbf{W}}_0 \tag{16}$$

and where the innovations covariance matrix is obtained as

$$\begin{aligned} \tilde{\mathbf{\Omega}}_n = {} & \tilde{\boldsymbol{\mathcal{T}}}_m \left\{ \mathrm{Cov}(\boldsymbol{\gamma}_n^r) \circ \left(\boldsymbol{\mathcal{T}}_m^H \tilde{\mathbf{\Sigma}}_n \boldsymbol{\mathcal{T}}_m \right) \right\} \tilde{\boldsymbol{\mathcal{T}}}_m^H \\ & + \tilde{\mathbf{R}}_n - \tilde{\mathbf{V}}_n \tilde{\mathbf{\Omega}}_{n-1}^{-1} \tilde{\mathbf{V}}_n^H - \tilde{\mathbf{\Pi}}_n \tilde{\mathbf{P}}_{n|n-1} \tilde{\mathbf{\Pi}}_n + \tilde{\mathbf{\Pi}}_n \tilde{\mathbf{\Theta}}_n + \tilde{\mathbf{\Theta}}_n^H \tilde{\mathbf{\Pi}}_n, \quad n > 1 \end{aligned} \tag{17}$$

with

$$\tilde{\mathbf{\Sigma}}_n = \begin{bmatrix} E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^H] & \mathbf{0}_{2m\times 2m} \\ \mathbf{0}_{2m\times 2m} & (E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^H])^{\eta'} \end{bmatrix} \tag{18}$$

where $E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^H]$ *is updated by means of equation* (11) *and*

$$\tilde{\mathbf{\Omega}}_1 = \tilde{\boldsymbol{\mathcal{T}}}_m \left\{ \mathrm{Cov}(\boldsymbol{\gamma}_1^r) \circ \left(\boldsymbol{\mathcal{T}}_m^H \tilde{\mathbf{\Sigma}}_1 \boldsymbol{\mathcal{T}}_m \right) \right\} \tilde{\boldsymbol{\mathcal{T}}}_m^H + \tilde{\mathbf{R}}_1 - \tilde{\mathbf{\Pi}}_1 \tilde{\mathbf{P}}_{1|0} \tilde{\mathbf{\Pi}}_1 + \tilde{\mathbf{\Pi}}_1 \tilde{\mathbf{\Theta}}_1 + \tilde{\mathbf{\Theta}}_1^H \tilde{\mathbf{\Pi}}_1 \tag{19}$$

Finally, the SWL filtering error is

$$\mathbf{P}_{n|n} = [\mathbf{I}_m, \mathbf{0}_{m\times m}]\tilde{\mathbf{P}}_{n|n}[\mathbf{I}_m, \mathbf{0}_{m\times m}]^T$$

where $\tilde{\mathbf{P}}_{n|n}$ *can be recursively obtained as a function of the one-stage prediction error covariance matrix as follows*

$$\tilde{\mathbf{P}}_{n|n} = \tilde{\mathbf{P}}_{n|n-1} - \tilde{\mathbf{\Theta}}_n \tilde{\mathbf{\Omega}}_n^{-1} \tilde{\mathbf{\Theta}}_n^H \tag{20}$$

with

$$\begin{aligned}\tilde{\mathbf{P}}_{n|n-1} = \tilde{\mathbf{\Phi}}_n \tilde{\mathbf{P}}_{n-1|n-1} \tilde{\mathbf{\Phi}}_n^H + \tilde{\mathbf{\Phi}}_n [\tilde{\mathbf{U}}_{n-1}^H - \tilde{\mathbf{\Theta}}_{n-1} \tilde{\mathbf{\Psi}}_{n-1}^H] \\ + [\tilde{\mathbf{U}}_{n-1}^H - \tilde{\mathbf{\Theta}}_{n-1} \tilde{\mathbf{\Psi}}_{n-1}^H]^H \tilde{\mathbf{\Phi}}_n^H - \tilde{\mathbf{\Psi}}_{n-1} \tilde{\mathbf{\Omega}}_{n-1} \tilde{\mathbf{\Psi}}_{n-1}^H + \tilde{\mathbf{Q}}_{n-1}\end{aligned} \tag{21}$$

and the initialization $\tilde{\mathbf{P}}_{1|0} = \tilde{\mathbf{\Phi}}_1 \tilde{\mathbf{P}}_{0|0} \tilde{\mathbf{\Phi}}_1^H + \tilde{\mathbf{Q}}_0$ *and* $\tilde{\mathbf{P}}_{0|0} = E[\tilde{\mathbf{x}}_0 \tilde{\mathbf{x}}_0^H]$.

Remark 2. *The SWL filter in Theorem 1 presents the classical structure of a Kalman filter, i.e., it is an iterative process that follows two steps: prediction and updating. Moreover, the error covariance matrix associated with the filter can be precomputed, as can the filter gain, before processing the data. The effects on the estimation due to the presence of measurement losses are picked up by the innovations* (14) *and their covariance matrices* (17)*. Finally, and for application purposes, the SWL filter is given in Algorithm 1.*

Theorem 2 (SWL fixed-point smoother). *The optimal SWL fixed-point smoother is obtained as*

$$\hat{\mathbf{x}}_{n|L} = [\mathbf{I}_m, \mathbf{0}_{m\times m}]\hat{\tilde{\mathbf{x}}}_{n|L}$$

where $\hat{\tilde{\mathbf{x}}}_{n|L}$*, for a fixed instant* $n < L$*, can be recursively calculated as follows*

$$\hat{\tilde{\mathbf{x}}}_{n|L} = \hat{\tilde{\mathbf{x}}}_{n|L-1} + \tilde{\mathbf{\Theta}}_{n,L} \tilde{\mathbf{\Omega}}_L^{-1} \tilde{\mathbf{y}}_{L|L-1}, \quad L > n \tag{22}$$

with initial condition $\hat{\tilde{\mathbf{x}}}_{n|n}$ *given by* (12)*. The matrix* $\tilde{\mathbf{\Theta}}_{n,L}$*,* $L > n+1$*, obeys the equation*

$$\tilde{\mathbf{\Theta}}_{n,L} = \tilde{\mathbf{\Xi}}_{n,L-1} \tilde{\mathbf{\Phi}}_L^H \tilde{\mathbf{\Pi}}_L - \tilde{\mathbf{\Theta}}_{n,L-1} \tilde{\mathbf{\Psi}}_{L-1}^H \tilde{\mathbf{\Pi}}_L - \tilde{\mathbf{\Theta}}_{n,L-1} \tilde{\mathbf{\Omega}}_{L-1}^{-1} \tilde{\mathbf{V}}_L \tag{23}$$

$$\tilde{\mathbf{\Theta}}_{n,n+1} = \tilde{\mathbf{P}}_{n|n} \tilde{\mathbf{\Phi}}_{n+1}^H \tilde{\mathbf{\Pi}}_{n+1} - \tilde{\mathbf{\Theta}}_n \tilde{\mathbf{\Psi}}_n^H \tilde{\mathbf{\Pi}}_{n+1} - \tilde{\mathbf{\Theta}}_n \tilde{\mathbf{\Omega}}_n^{-1} \tilde{\mathbf{V}}_{n+1} + \tilde{\mathbf{U}}_n^H \tilde{\mathbf{\Pi}}_{n+1} \tag{24}$$

Algorithm 1 SWL filter

Require: $\{\tilde{\mathbf{\Phi}}_i, \tilde{\mathbf{z}}_i, \tilde{\mathbf{\Pi}}_i, \tilde{\mathbf{R}}_i, \tilde{\mathbf{V}}_i, \mathrm{Cov}(\boldsymbol{\gamma}_i^r)\}_{i=1}^{n}$, $\{\tilde{\mathbf{Q}}_i, \tilde{\mathbf{W}}_i\}_{i=0}^{n-1}$, $\{\tilde{\mathbf{U}}_i, \tilde{\mathbf{T}}_i\}_{i=1}^{n-1}$, and $\tilde{\mathbf{P}}_{0|0}$

Ensure: $\hat{\mathbf{x}}_{n|n}$ and $\mathbf{P}_{n|n}$

$\tilde{\mathbf{P}}_{1|0} \leftarrow \tilde{\mathbf{\Phi}}_1 \tilde{\mathbf{P}}_{0|0} \tilde{\mathbf{\Phi}}_1^{\mathtt{H}} + \tilde{\mathbf{Q}}_0$

$E[\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_1^{\mathtt{H}}] \leftarrow \tilde{\mathbf{P}}_{1|0}$

$\tilde{\mathbf{\Sigma}}_1 \leftarrow \begin{bmatrix} E[\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_1^{\mathtt{H}}] & \mathbf{0}_{2m\times 2m} \\ \mathbf{0}_{2m\times 2m} & (E[\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_1^{\mathtt{H}}])^{\eta'} \end{bmatrix}$

$\tilde{\mathbf{\Theta}}_1 \leftarrow \tilde{\mathbf{P}}_{1|0} \tilde{\mathbf{\Pi}}_1 + \tilde{\mathbf{W}}_0$

$\tilde{\mathbf{\Omega}}_1 \leftarrow \tilde{\boldsymbol{\mathcal{T}}}_m \left\{ \mathrm{Cov}(\boldsymbol{\gamma}_1^r) \circ \left(\boldsymbol{\mathcal{T}}_m^{\mathtt{H}} \tilde{\mathbf{\Sigma}}_1 \boldsymbol{\mathcal{T}}_m \right) \right\} \tilde{\boldsymbol{\mathcal{T}}}_m^{\mathtt{H}} + \tilde{\mathbf{R}}_1 - \tilde{\mathbf{\Pi}}_1 \tilde{\mathbf{P}}_{1|0} \tilde{\mathbf{\Pi}}_1 + \tilde{\mathbf{\Pi}}_1 \tilde{\mathbf{\Theta}}_1 + \tilde{\mathbf{\Theta}}_1^{\mathtt{H}} \tilde{\mathbf{\Pi}}_1$

$\tilde{\mathbf{y}}_{1|0} \leftarrow \tilde{\mathbf{z}}_1$

$\hat{\tilde{\mathbf{x}}}_{1|1} \leftarrow \tilde{\mathbf{\Theta}}_1 \tilde{\mathbf{\Omega}}_1^{-1} \tilde{\mathbf{z}}_1$

$\tilde{\mathbf{P}}_{1|1} \leftarrow \tilde{\mathbf{P}}_{1|0} - \tilde{\mathbf{\Theta}}_1 \tilde{\mathbf{\Omega}}_1^{-1} \tilde{\mathbf{\Theta}}_1^{\mathtt{H}}$

$\hat{\mathbf{x}}_{1|1} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m\times m}] \hat{\tilde{\mathbf{x}}}_{1|1}$

$\mathbf{P}_{1|1} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m\times m}] \tilde{\mathbf{P}}_{1|1} [\mathbf{I}_m, \mathbf{0}_{m\times m}]^{\mathtt{T}}$

for $n \geq 2$ **do**

  $\tilde{\mathbf{\Psi}}_{n-1} \leftarrow [\tilde{\mathbf{U}}_{n-1} \tilde{\mathbf{\Pi}}_{n-1} + \tilde{\mathbf{T}}_{n-1}] \tilde{\mathbf{\Omega}}_{n-1}^{-1}$

  $\tilde{\mathbf{P}}_{n|n-1} \leftarrow \tilde{\mathbf{\Phi}}_n \tilde{\mathbf{P}}_{n-1|n-1} \tilde{\mathbf{\Phi}}_n^{\mathtt{H}} + \tilde{\mathbf{\Phi}}_n [\tilde{\mathbf{U}}_{n-1}^{\mathtt{H}} - \tilde{\mathbf{\Theta}}_{n-1} \tilde{\mathbf{\Psi}}_{n-1}^{\mathtt{H}}] + [\tilde{\mathbf{U}}_{n-1}^{\mathtt{H}} - \tilde{\mathbf{\Theta}}_{n-1} \tilde{\mathbf{\Psi}}_{n-1}^{\mathtt{H}}]^{\mathtt{H}} \tilde{\mathbf{\Phi}}_n^{\mathtt{H}} - \tilde{\mathbf{\Psi}}_{n-1} \tilde{\mathbf{\Omega}}_{n-1} \tilde{\mathbf{\Psi}}_{n-1}^{\mathtt{H}} + \tilde{\mathbf{Q}}_{n-1}$

  $E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^{\mathtt{H}}] \leftarrow \tilde{\mathbf{\Phi}}_n E[\tilde{\mathbf{x}}_{n-1} \tilde{\mathbf{x}}_{n-1}^{\mathtt{H}}] \tilde{\mathbf{\Phi}}_n^{\mathtt{H}} + \tilde{\mathbf{\Phi}}_n \tilde{\mathbf{U}}_{n-1}^{\mathtt{H}} + \tilde{\mathbf{U}}_{n-1} \tilde{\mathbf{\Phi}}_n^{\mathtt{H}} + \tilde{\mathbf{Q}}_{n-1}$

  $\tilde{\mathbf{\Sigma}}_n \leftarrow \begin{bmatrix} E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^{\mathtt{H}}] & \mathbf{0}_{2m\times 2m} \\ \mathbf{0}_{2m\times 2m} & (E[\tilde{\mathbf{x}}_n \tilde{\mathbf{x}}_n^{\mathtt{H}}])^{\eta'} \end{bmatrix}$

  $\hat{\tilde{\mathbf{x}}}_{n|n-1} \leftarrow \tilde{\mathbf{\Phi}}_n \hat{\tilde{\mathbf{x}}}_{n-1|n-1} + \tilde{\mathbf{\Psi}}_{n-1} \tilde{\mathbf{y}}_{n-1|n-2}$

  $\tilde{\mathbf{y}}_{n|n-1} \leftarrow \tilde{\mathbf{z}}_n - \tilde{\mathbf{\Pi}}_n \hat{\tilde{\mathbf{x}}}_{n|n-1} - \tilde{\mathbf{V}}_n \tilde{\mathbf{\Omega}}_{n-1}^{-1} \tilde{\mathbf{y}}_{n-1|n-2}$

  $\tilde{\mathbf{\Theta}}_n \leftarrow \tilde{\mathbf{P}}_{n|n-1} \tilde{\mathbf{\Pi}}_n + \tilde{\mathbf{W}}_{n-1} - [\tilde{\mathbf{\Phi}}_n \tilde{\mathbf{\Theta}}_{n-1} \tilde{\mathbf{\Omega}}_{n-1}^{-1} + \tilde{\mathbf{\Psi}}_{n-1}] \tilde{\mathbf{V}}_n^{\mathtt{H}}$

  $\tilde{\mathbf{\Omega}}_n \leftarrow \tilde{\boldsymbol{\mathcal{T}}}_m \left\{ \mathrm{Cov}(\boldsymbol{\gamma}_n^r) \circ \left(\boldsymbol{\mathcal{T}}_m^{\mathtt{H}} \tilde{\mathbf{\Sigma}}_n \boldsymbol{\mathcal{T}}_m \right) \right\} \tilde{\boldsymbol{\mathcal{T}}}_m^{\mathtt{H}} + \tilde{\mathbf{R}}_n - \tilde{\mathbf{V}}_n \tilde{\mathbf{\Omega}}_{n-1}^{-1} \tilde{\mathbf{V}}_n^{\mathtt{H}} - \tilde{\mathbf{\Pi}}_n \tilde{\mathbf{P}}_{n|n-1} \tilde{\mathbf{\Pi}}_n + \tilde{\mathbf{\Pi}}_n \tilde{\mathbf{\Theta}}_n + \tilde{\mathbf{\Theta}}_n^{\mathtt{H}} \tilde{\mathbf{\Pi}}_n$

  $\hat{\tilde{\mathbf{x}}}_{n|n} \leftarrow \hat{\tilde{\mathbf{x}}}_{n|n-1} + \tilde{\mathbf{\Theta}}_n \tilde{\mathbf{\Omega}}_n^{-1} \tilde{\mathbf{y}}_{n|n-1}$

  $\tilde{\mathbf{P}}_{n|n} \leftarrow \tilde{\mathbf{P}}_{n|n-1} - \tilde{\mathbf{\Theta}}_n \tilde{\mathbf{\Omega}}_n^{-1} \tilde{\mathbf{\Theta}}_n^{\mathtt{H}}$

  $\hat{\mathbf{x}}_{n|n} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m\times m}] \hat{\tilde{\mathbf{x}}}_{n|n}$

  $\mathbf{P}_{n|n} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m\times m}] \tilde{\mathbf{P}}_{n|n} [\mathbf{I}_m, \mathbf{0}_{m\times m}]^{\mathtt{T}}$

end for

where $\tilde{\mathbf{\Psi}}_n$ *is given in Theorem 1,* $\tilde{\mathbf{P}}_{n|n}$ *is computed from* (20) *and* $\tilde{\mathbf{\Theta}}_n$ *by* (15). *Also,* $\tilde{\mathbf{\Xi}}_{n,L}$ *satisfies the recursive expression*

$$\tilde{\mathbf{\Xi}}_{n,L} = \left[\tilde{\mathbf{\Xi}}_{n,L-1}\tilde{\mathbf{\Phi}}_L^H - \tilde{\mathbf{\Theta}}_{n,L-1}\tilde{\mathbf{\Psi}}_{L-1}^H\right]\left[\mathbf{I}_{2m} - \tilde{\mathbf{\Pi}}_L\tilde{\mathbf{\Omega}}_L^{-1}\tilde{\mathbf{\Theta}}_L^H\right] + \tilde{\mathbf{\Theta}}_{n,L-1}\tilde{\mathbf{\Omega}}_{L-1}^{-1}\tilde{\mathbf{V}}_L^H\tilde{\mathbf{\Omega}}_L^{-1}\tilde{\mathbf{\Theta}}_L^H \quad (25)$$

for $L > n+1$ *and*

$$\tilde{\mathbf{\Xi}}_{n,n+1} = \left[\tilde{\mathbf{P}}_{n|n}\tilde{\mathbf{\Phi}}_{n+1}^H - \tilde{\mathbf{\Theta}}_n\tilde{\mathbf{\Psi}}_n^H + \tilde{\mathbf{U}}_n^H\right]\left[\mathbf{I}_{2m} - \tilde{\mathbf{\Pi}}_{n+1}\tilde{\mathbf{\Omega}}_{n+1}^{-1}\tilde{\mathbf{\Theta}}_{n+1}^H\right] + \tilde{\mathbf{\Theta}}_n\tilde{\mathbf{\Omega}}_n^{-1}\tilde{\mathbf{V}}_{n+1}^H\tilde{\mathbf{\Omega}}_{n+1}^{-1}\tilde{\mathbf{\Theta}}_{n+1}^H \quad (26)$$

Finally, the SWL fixed-point smoothing error is computed by

$$\mathbf{P}_{n|L} = [\mathbf{I}_m, \mathbf{0}_{m\times m}]\tilde{\mathbf{P}}_{n|L}[\mathbf{I}_m, \mathbf{0}_{m\times m}]^T$$

where $\tilde{\mathbf{P}}_{n|L}$ *can be recursively obtained as follows*

$$\tilde{\mathbf{P}}_{n|L} = \tilde{\mathbf{P}}_{n|L-1} - \tilde{\mathbf{\Theta}}_{n,L}\tilde{\mathbf{\Omega}}_L^{-1}\tilde{\mathbf{\Theta}}_{n,L}^H, \quad L > n \quad (27)$$

with the initial condition $\tilde{\mathbf{P}}_{n|n}$ *given by* (20).

Remark 3. *The SWL fixed-point smoother* $\hat{\tilde{\mathbf{x}}}_{n|L}$ *in* (22) *is driven from the innovations process* $\tilde{\mathbf{y}}_{L|L-1}$ *of the SWL filter in Theorem 1. Moreover, the gain matrix depends on the innovations covariance matrix and the Bernoulli probabilities. Both these facts reflect the way in which the missing measurements phenomenon affect the smoothing procedure. As in the case of the SWL filter, the error covariance is independent of the measurements and therefore can be computed beforehand. Finally, the smoothing method in Theorem 2 is presented in an algorithmic fashion in Algorithm 2.*

Theorem 3 (SWL fixed-lag smoother). *The optimal SWL fixed-lag smoother is obtained as*

$$\hat{\mathbf{x}}_{n+k|n} = [\mathbf{I}_m, \mathbf{0}_{m\times m}]\hat{\tilde{\mathbf{x}}}_{n+k|n}$$

where $\hat{\tilde{\mathbf{x}}}_{n+k|n}$ *is recursively computed from the expression*

$$\hat{\tilde{\mathbf{x}}}_{n+k|n} = \hat{\tilde{\mathbf{x}}}_{n+k|n-1} + \tilde{\mathbf{\Theta}}_{n+k,n}\tilde{\mathbf{\Omega}}_n^{-1}\tilde{\mathbf{y}}_{n|n-1}, \quad k < 0 \quad (28)$$

Algorithm 2 SWL fixed-point smoother

Require: $\tilde{\mathbf{U}}_n$, $\{\tilde{\boldsymbol{\Theta}}_i, \tilde{\boldsymbol{\Omega}}_i\}_{i=n}^{L}$, $\{\tilde{\boldsymbol{\Phi}}_i, \tilde{\boldsymbol{\Pi}}_i, \tilde{\mathbf{V}}_i, \tilde{\boldsymbol{\Psi}}_{i-1}, \tilde{\mathbf{y}}_{i|i-1}\}_{i=n+1}^{L}$, $\hat{\tilde{\mathbf{x}}}_{n|n}$, and $\tilde{\mathbf{P}}_{n|n}$

Ensure: $\hat{\mathbf{x}}_{n|L}$ and $\mathbf{P}_{n|L}$ for a fixed instant $n < L$

$\tilde{\boldsymbol{\Theta}}_{n,n+1} \leftarrow \tilde{\mathbf{P}}_{n|n}\tilde{\boldsymbol{\Phi}}_{n+1}^{\mathtt{H}}\tilde{\boldsymbol{\Pi}}_{n+1} - \tilde{\boldsymbol{\Theta}}_n\tilde{\boldsymbol{\Psi}}_n^{\mathtt{H}}\tilde{\boldsymbol{\Pi}}_{n+1} - \tilde{\boldsymbol{\Theta}}_n\tilde{\boldsymbol{\Omega}}_n^{-1}\tilde{\mathbf{V}}_{n+1} + \tilde{\mathbf{U}}_n^{\mathtt{H}}\tilde{\boldsymbol{\Pi}}_{n+1}$

$\tilde{\boldsymbol{\Xi}}_{n,n+1} \leftarrow \left[\tilde{\mathbf{P}}_{n|n}\tilde{\boldsymbol{\Phi}}_{n+1}^{\mathtt{H}} - \tilde{\boldsymbol{\Theta}}_n\tilde{\boldsymbol{\Psi}}_n^{\mathtt{H}} + \tilde{\mathbf{U}}_n^{\mathtt{H}}\right]\left[\mathbf{I}_{2m} - \tilde{\boldsymbol{\Pi}}_{n+1}\tilde{\boldsymbol{\Omega}}_{n+1}^{-1}\tilde{\boldsymbol{\Theta}}_{n+1}^{\mathtt{H}}\right] + \tilde{\boldsymbol{\Theta}}_n\tilde{\boldsymbol{\Omega}}_n^{-1}\tilde{\mathbf{V}}_{n+1}^{\mathtt{H}}\tilde{\boldsymbol{\Omega}}_{n+1}^{-1}\tilde{\boldsymbol{\Theta}}_{n+1}^{\mathtt{H}}$

$\hat{\tilde{\mathbf{x}}}_{n|n+1} \leftarrow \hat{\tilde{\mathbf{x}}}_{n|n} + \tilde{\boldsymbol{\Theta}}_{n,n+1}\tilde{\boldsymbol{\Omega}}_{n+1}^{-1}\tilde{\mathbf{y}}_{n+1|n}$

$\tilde{\mathbf{P}}_{n|n+1} \leftarrow \tilde{\mathbf{P}}_{n|n} - \tilde{\boldsymbol{\Theta}}_{n,n+1}\tilde{\boldsymbol{\Omega}}_{n+1}^{-1}\tilde{\boldsymbol{\Theta}}_{n,n+1}^{\mathtt{H}}$

for $i = n+2$ to L **do**

  $\tilde{\boldsymbol{\Theta}}_{n,i} \leftarrow \tilde{\boldsymbol{\Xi}}_{n,i-1}\tilde{\boldsymbol{\Phi}}_i^{\mathtt{H}}\tilde{\boldsymbol{\Pi}}_i - \tilde{\boldsymbol{\Theta}}_{n,i-1}\tilde{\boldsymbol{\Psi}}_{i-1}^{\mathtt{H}}\tilde{\boldsymbol{\Pi}}_i - \tilde{\boldsymbol{\Theta}}_{n,i-1}\tilde{\boldsymbol{\Omega}}_{i-1}^{-1}\tilde{\mathbf{V}}_i$

  $\tilde{\boldsymbol{\Xi}}_{n,i} \leftarrow \left[\tilde{\boldsymbol{\Xi}}_{n,i-1}\tilde{\boldsymbol{\Phi}}_i^{\mathtt{H}} - \tilde{\boldsymbol{\Theta}}_{n,i-1}\tilde{\boldsymbol{\Psi}}_{i-1}^{\mathtt{H}}\right]\left[\mathbf{I}_{2m} - \tilde{\boldsymbol{\Pi}}_i\tilde{\boldsymbol{\Omega}}_i^{-1}\tilde{\boldsymbol{\Theta}}_i^{\mathtt{H}}\right] + \tilde{\boldsymbol{\Theta}}_{n,i-1}\tilde{\boldsymbol{\Omega}}_{i-1}^{-1}\tilde{\mathbf{V}}_i^{\mathtt{H}}\tilde{\boldsymbol{\Omega}}_i^{-1}\tilde{\boldsymbol{\Theta}}_i^{\mathtt{H}}$

  $\hat{\tilde{\mathbf{x}}}_{n|i} \leftarrow \hat{\tilde{\mathbf{x}}}_{n|i-1} + \tilde{\boldsymbol{\Theta}}_{n,i}\tilde{\boldsymbol{\Omega}}_i^{-1}\tilde{\mathbf{y}}_{i|i-1}$

  $\tilde{\mathbf{P}}_{n|i} \leftarrow \tilde{\mathbf{P}}_{n|i-1} - \tilde{\boldsymbol{\Theta}}_{n,i}\tilde{\boldsymbol{\Omega}}_i^{-1}\tilde{\boldsymbol{\Theta}}_{n,i}^{\mathtt{H}}$

end for

$\hat{\mathbf{x}}_{n|L} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m\times m}]\hat{\tilde{\mathbf{x}}}_{n|L}$

$\mathbf{P}_{n|L} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m\times m}]\tilde{\mathbf{P}}_{n|L}[\mathbf{I}_m, \mathbf{0}_{m\times m}]^{\mathtt{T}}$

with the initial condition the filter $\hat{\tilde{\mathbf{x}}}_{n+k|n+k}$. *Moreover, the matrix* $\tilde{\mathbf{\Theta}}_{n+k,n}$ *obeys the equation*

$$\tilde{\mathbf{\Theta}}_{n+k,n} = \tilde{\mathbf{\Lambda}}_{n+k,n}\tilde{\mathbf{\Pi}}_n - \tilde{\mathbf{\Theta}}_{n+k,n-1}\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{V}}_n^H \tag{29}$$

with

$$\begin{aligned}\tilde{\mathbf{\Lambda}}_{n-1,n} = \tilde{\mathbf{P}}_{n-1|n-2}\left[\tilde{\mathbf{\Phi}}_n - \tilde{\mathbf{\Theta}}_{n,n-1}\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{\Pi}}_{n-1}\right]^H \\ + \tilde{\mathbf{U}}_{n-1}^H - \tilde{\mathbf{W}}_{n-2}\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{\Theta}}_{n,n-1}^H + \tilde{\mathbf{\Theta}}_{n-1,n-2}\tilde{\mathbf{\Omega}}_{n-2}^{-1}\tilde{\mathbf{V}}_{n-1}^H\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{\Theta}}_{n,n-1}^H\end{aligned} \tag{30}$$

$$\begin{aligned}\tilde{\mathbf{\Theta}}_{n,n-1} = \tilde{\mathbf{\Phi}}_n\tilde{\mathbf{P}}_{n-1|n-2}\tilde{\mathbf{\Pi}}_{n-1} + \tilde{\mathbf{\Phi}}_n\tilde{\mathbf{W}}_{n-2} - \tilde{\mathbf{\Phi}}_n\tilde{\mathbf{\Theta}}_{n-1,n-2}\tilde{\mathbf{\Omega}}_{n-2}^{-1}\tilde{\mathbf{V}}_{n-1}^H \\ + \tilde{\mathbf{\Psi}}_{n-1}\tilde{\mathbf{\Omega}}_{n-1}\end{aligned} \tag{31}$$

where $\tilde{\mathbf{\Psi}}_n$ *is given in Theorem 1, and for* $k < -1$

$$\begin{aligned}\tilde{\mathbf{\Lambda}}_{n+k,n} = \tilde{\mathbf{\Lambda}}_{n+k,n-1}\left[\tilde{\mathbf{\Phi}}_n - \tilde{\mathbf{\Theta}}_{n,n-1}\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{\Pi}}_{n-1}\right]^H \\ + \tilde{\mathbf{\Theta}}_{n+k,n-2}\tilde{\mathbf{\Omega}}_{n-2}^{-1}\tilde{\mathbf{V}}_{n-1}^H\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{\Theta}}_{n,n-1}^H\end{aligned} \tag{32}$$

Additionally, the SWL fixed-lag smoothing error is

$$\mathbf{P}_{n+k|n} = [\mathbf{I}_m, \mathbf{0}_{m\times m}]\tilde{\mathbf{P}}_{n+k|n}[\mathbf{I}_m, \mathbf{0}_{m\times m}]^T$$

where $\tilde{\mathbf{P}}_{n+k|n}$ *can be recursively obtained as follows*

$$\tilde{\mathbf{P}}_{n+k|n} = \tilde{\mathbf{P}}_{n+k|n-1} - \tilde{\mathbf{\Theta}}_{n+k,n}\tilde{\mathbf{\Omega}}_n^{-1}\tilde{\mathbf{\Theta}}_{n+k,n}^H \tag{33}$$

with the initial condition $\tilde{\mathbf{P}}_{k|0}$ *computed from* (27).

Remark 4. *The SWL fixed-lag smoother is a very useful device since it allows "on-line" production of smoothed estimates (see Algorithm 3). The equations in Theorem 3 for the SWL fixed-lag smoother bear some resemblance to the fixed-point smoothing equations in Theorem 2 and both share similar characteristics in terms of the treatment of missing observations. More specifically, the fixed-lag smoother described by* (28) *may be viewed as being driven from the innovations process and the gain matrices depend on both the innovations covariance matrix and the Bernoulli probabilities. As in the case of the SWL estimators given in*

Algorithm 3 SWL fixed-lag smoother

Require: $\{\tilde{\mathbf{\Phi}}_i, \tilde{\mathbf{W}}_{i-2}, \tilde{\mathbf{V}}_i, \tilde{\mathbf{U}}_{i-1}, \tilde{\mathbf{\Psi}}_{i-1}, \tilde{\mathbf{y}}_{i|i-1}, \tilde{\mathbf{\Theta}}_{i-1}, \tilde{\mathbf{P}}_{i-1|i-2}\}_{i=2}^{n}$, $\{\tilde{\mathbf{\Omega}}_i, \tilde{\mathbf{\Pi}}_i\}_{i=1}^{n}$, and $\{\hat{\tilde{\mathbf{x}}}_{i|i}, \tilde{\mathbf{P}}_{i|i}\}_{i=1}^{n+k}$

Ensure: $\hat{\mathbf{x}}_{n+k|n}$ and $\mathbf{P}_{n+k|n}$ for a fixed-lag $k < 0$

$\tilde{\mathbf{\Theta}}_{2,1} \leftarrow \tilde{\mathbf{\Phi}}_2 \tilde{\mathbf{P}}_{1|0} \tilde{\mathbf{\Pi}}_1 + \tilde{\mathbf{\Phi}}_2 \tilde{\mathbf{W}}_0 + \tilde{\mathbf{\Psi}}_1 \tilde{\mathbf{\Omega}}_1$

for $i = 3$ to n **do**

  $\tilde{\mathbf{\Theta}}_{i,i-1} \leftarrow \tilde{\mathbf{\Phi}}_i \tilde{\mathbf{P}}_{i-1|i-2} \tilde{\mathbf{\Pi}}_{i-1} + \tilde{\mathbf{\Phi}}_i \tilde{\mathbf{W}}_{i-2} - \tilde{\mathbf{\Phi}}_i \tilde{\mathbf{\Theta}}_{i-1,i-2} \tilde{\mathbf{\Omega}}_{i-2}^{-1} \tilde{\mathbf{V}}_{i-1}^{\mathtt{H}} + \tilde{\mathbf{\Psi}}_{i-1} \tilde{\mathbf{\Omega}}_{i-1}$

end for

for $l = 2$ to $|k|+1$ **do**

  if $l = 2$ **then**

    $\tilde{\mathbf{\Lambda}}_{1,l} \leftarrow \tilde{\mathbf{P}}_{1|l-2} \left[\tilde{\mathbf{\Phi}}_l - \tilde{\mathbf{\Theta}}_{l,l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Pi}}_{l-1}\right]^{\mathtt{H}} + \tilde{\mathbf{U}}_{l-1}^{\mathtt{H}} - \tilde{\mathbf{W}}_{l-2} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Theta}}_{l,l-1}^{\mathtt{H}}$

    $\tilde{\mathbf{\Theta}}_{1,l} \leftarrow \tilde{\mathbf{\Lambda}}_{1,l} \tilde{\mathbf{\Pi}}_l - \tilde{\mathbf{\Theta}}_{l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{V}}_l^{\mathtt{H}}$

  else

    $\tilde{\mathbf{\Lambda}}_{1,l} \leftarrow \tilde{\mathbf{\Lambda}}_{1,l-1} \left[\tilde{\mathbf{\Phi}}_l - \tilde{\mathbf{\Theta}}_{l,l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Pi}}_{l-1}\right]^{\mathtt{H}} + \tilde{\mathbf{\Theta}}_{1,l-2} \tilde{\mathbf{\Omega}}_{l-2}^{-1} \tilde{\mathbf{V}}_{l-1}^{\mathtt{H}} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Theta}}_{l,l-1}^{\mathtt{H}}$

    $\tilde{\mathbf{\Theta}}_{1,l} \leftarrow \tilde{\mathbf{\Lambda}}_{1,l} \tilde{\mathbf{\Pi}}_l - \tilde{\mathbf{\Theta}}_{1,l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{V}}_l^{\mathtt{H}}$

  end if

  $\hat{\tilde{\mathbf{x}}}_{1|l} \leftarrow \hat{\tilde{\mathbf{x}}}_{1|l-1} + \tilde{\mathbf{\Theta}}_{1,l} \tilde{\mathbf{\Omega}}_l^{-1} \tilde{\mathbf{y}}_{l|l-1}$

  $\tilde{\mathbf{P}}_{1|l} \leftarrow \tilde{\mathbf{P}}_{1|l-1} - \tilde{\mathbf{\Theta}}_{1,l} \tilde{\mathbf{\Omega}}_l^{-1} \tilde{\mathbf{\Theta}}_{1,l}^{\mathtt{H}}$

end for

for $i = 2$ to $n + k$ **do**

  for $l = 1 + i$ to $|k|$+i **do**

    if $l - i = 1$ **then**

      $\tilde{\mathbf{\Lambda}}_{i,l} \leftarrow \tilde{\mathbf{P}}_{i|l-2} \left[\tilde{\mathbf{\Phi}}_l - \tilde{\mathbf{\Theta}}_{l,l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Pi}}_{l-1}\right]^{\mathtt{H}} + \tilde{\mathbf{U}}_{l-1}^{\mathtt{H}} - \tilde{\mathbf{W}}_{l-2} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Theta}}_{l,l-1}^{\mathtt{H}} + \tilde{\mathbf{\Theta}}_{l-1,l-2} \tilde{\mathbf{\Omega}}_{l-2}^{-1} \tilde{\mathbf{V}}_{l-1}^{\mathtt{H}} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Theta}}_{l,l-1}^{\mathtt{H}}$

      $\tilde{\mathbf{\Theta}}_{i,l} \leftarrow \tilde{\mathbf{\Lambda}}_{i,l} \tilde{\mathbf{\Pi}}_l - \tilde{\mathbf{\Theta}}_{l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{V}}_l^{\mathtt{H}}$

    else

      $\tilde{\mathbf{\Lambda}}_{i,l} \leftarrow \tilde{\mathbf{\Lambda}}_{i,l-1} \left[\tilde{\mathbf{\Phi}}_l - \tilde{\mathbf{\Theta}}_{l,l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Pi}}_{l-1}\right]^{\mathtt{H}} + \tilde{\mathbf{\Theta}}_{i,l-2} \tilde{\mathbf{\Omega}}_{l-2}^{-1} \tilde{\mathbf{V}}_{l-1}^{\mathtt{H}} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{\Theta}}_{l,l-1}^{\mathtt{H}}$

      $\tilde{\mathbf{\Theta}}_{i,l} \leftarrow \tilde{\mathbf{\Lambda}}_{i,l} \tilde{\mathbf{\Pi}}_l - \tilde{\mathbf{\Theta}}_{i,l-1} \tilde{\mathbf{\Omega}}_{l-1}^{-1} \tilde{\mathbf{V}}_l^{\mathtt{H}}$

    end if

    $\hat{\tilde{\mathbf{x}}}_{i|l} \leftarrow \hat{\tilde{\mathbf{x}}}_{i|l-1} + \tilde{\mathbf{\Theta}}_{i,l} \tilde{\mathbf{\Omega}}_l^{-1} \tilde{\mathbf{y}}_{l|l-1}$

    $\tilde{\mathbf{P}}_{i|l} \leftarrow \tilde{\mathbf{P}}_{i|l-1} - \tilde{\mathbf{\Theta}}_{i,l} \tilde{\mathbf{\Omega}}_l^{-1} \tilde{\mathbf{\Theta}}_{i,l}^{\mathtt{H}}$

  end for

end for

$\hat{\mathbf{x}}_{n+k|n} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m \times m}] \hat{\tilde{\mathbf{x}}}_{n+k|n}$

$\mathbf{P}_{n+k|n} \leftarrow [\mathbf{I}_m, \mathbf{0}_{m \times m}] \tilde{\mathbf{P}}_{n+k|n} [\mathbf{I}_m, \mathbf{0}_{m \times m}]^{\mathtt{T}}$

Theorems 1 and 2, the error covariances in Theorem 3 are independent of the measurements and therefore they can be computed beforehand.

Remark 5. *As noted above, the WL processing is isomorph to the real processing and thus, any WL algorithm could be equivalently expressed through a real formalism. However, when the quaternions involved present $\mathbb{C}^{\eta}$-properness properties and a SWL processing has to be used then, the real processing becomes inefficient due to a higher computational burden. Hence, whenever missing measurements are present, the algorithms in Theorems 1, 2 and 3 provide the optimal solution to the estimation problem with a notable reduction in the computational burden in relation to their WL counterparts that cannot be attained by using a real framework. Actually, the dimension of the optimal WL estimator would be $4m$ whereas the optimal SWL estimator in the algorithms given here is $2m$.*

5. Numerical Simulations

The aim of this section is to numerically demonstrate that the algorithms in Theorems 1-3 outperform their standard counterparts (i.e., the SWL filtering, fixed-point, and fixed-lag algorithms which ignore the presence of uncertainties in the observations) when a phenomenon of intermittent observations is present. Such standard estimation algorithms used for comparison purposes are obtained as particular cases of those given in Theorems 1-3 by setting the Bernoulli probabilities to be the unity. It is noteworthy that both techniques have the same computational complexity.

We consider a SWL state-space model as in (5) where $\mathbf{x}_n = [x_{n,1}, x_{n,2}]^{\mathrm{T}}$, $\mathbf{H}_n = \mathbf{E}_n = \mathbf{0}_{2\times 2}$ and

$$\mathbf{F}_n = \begin{bmatrix} 0.1 - 0.3\eta + 0.2\eta' + 0.1\eta'' & 0.2 - 0.1\eta + 0.4\eta' + 0.5\eta'' \\ -0.1 - 0.1\eta - 0.1\eta' + 0.1\eta'' & 0.1 - 0.1\eta + 0.1\eta' - 0.1\eta'' \end{bmatrix}$$

$$\mathbf{G}_n = \begin{bmatrix} 0.2 + 0.3\eta - 0.2\eta' + 0.1\eta'' & 0.2 - 0.1\eta + 0.3\eta' + 0.09\eta'' \\ 0.1 + 0.3\eta - 0.2\eta' - 0.2\eta'' & 0.1 + 0.1\eta - 0.2\eta' - 0.1\eta'' \end{bmatrix}$$

The state noise is defined as $\mathbf{u}_{n-1} = \mathbf{w}_{n-1} + \lambda \mathbf{w}_n$, $\lambda \in \mathbb{R}$, being $\mathbf{w}_n$ a quaternion white Gaussian noise with real covariance matrix given by

$$E[\mathbf{w}_n^r \mathbf{w}_n^{r\mathrm{T}}] = \begin{bmatrix} 10 & 1 & 0 & 0 & 1 & -2 & 1 & 3 \\ 1 & 5 & 0 & 0 & 1 & -2 & 2 & 1 \\ 0 & 0 & 10 & 1 & 1 & 3 & -1 & 2 \\ 0 & 0 & 1 & 5 & 2 & 1 & -1 & 2 \\ 1 & 1 & 1 & 2 & 3 & 2 & 0 & 0 \\ -2 & -2 & 3 & 1 & 2 & 10 & 0 & 0 \\ 1 & 2 & -1 & -1 & 0 & 0 & 3 & 2 \\ 3 & 1 & 2 & 2 & 0 & 0 & 2 & 10 \end{bmatrix}$$

and the measurement noise in (10) is defined by $\mathbf{v}_n = \alpha \mathbf{w}_n$, $\alpha \in \mathbb{R}$. Finally, the initial state is a quaternion Gaussian vector determined by the following real covariance matrix

$$E[\mathbf{x}_0^r \mathbf{x}_0^{r\mathrm{T}}] = \begin{bmatrix} 9 & -1 & 0 & 0 & 1 & -2 & 2 & 0.3 \\ -1 & 6 & 0 & 0 & 2 & 1 & 0 & 3 \\ 0 & 0 & 9 & -1 & 2 & 0.3 & -1 & 2 \\ 0 & 0 & -1 & 6 & 0 & 3 & -2 & -1 \\ 1 & 2 & 2 & 0 & 12 & 1 & 0 & 0 \\ -2 & 1 & 0.3 & 3 & 1 & 8 & 0 & 0 \\ 2 & 0 & -1 & -2 & 0 & 0 & 12 & 1 \\ 0.3 & 3 & 2 & -1 & 0 & 0 & 1 & 8 \end{bmatrix}$$

In order to compare the performance of the optimal estimators in Theorems 1-3 with their standard versions for the state-space model defined above, the error variances associated with the estimations of each component, $x_{n,1}$ and $x_{n,2}$, of $\mathbf{x}_n$ are compared under different scenarios. For that, by using Monte Carlo simulation we have generated 1000 values of $\bar{\mathbf{x}}_n$ and $\tilde{\mathbf{z}}_n$ for each $n = 1, \ldots, 100$ and for several values of the Bernoulli probabilities. Denote by $\bar{\mathbf{x}}_n^{(j,p)}$ and $\tilde{\mathbf{z}}_n^{(j,p)}$ the value of $\bar{\mathbf{x}}_n$ and $\tilde{\mathbf{z}}_n$, respectively, generated in the jth simulation and for a specific combination of probabilities p. The standard SWL estimators based on the observations $\tilde{\mathbf{z}}_n^{(j,p)}$ in the jth simulation and for probabilities p

are denoted by $\check{\mathbf{x}}_{n|n}^{(j,p)}$ (filter), $\check{\mathbf{x}}_{9|L}^{(j,p)}$ (fixed-point smoother at the instant $n = 9$), and $\check{\mathbf{x}}_{n-3|n}^{(j,p)}$ (fixed-lag 3-step smoother). Also, $x_{n,l}^{(j,p)}$, $l = 1, 2$, stand for the two first components of $\bar{\mathbf{x}}_n^{(j,p)}$ and a similar notation is used for the components of the estimators (e.g., $\check{x}_{n|n,2}^{(j,p)}$ is the second component of the filter $\check{\mathbf{x}}_{n|n}^{(j,p)}$). The accuracy of such estimators has been assessed by computing their associated mean square errors (MSEs). For example, the MSEs for every component in the filtering problem are calculated in the following way[2]:

$$MSE_{n|n,l}^{(p)} = \frac{1}{1000} \sum_{j=1}^{1000} \|x_{n,l}^{(j,p)} - \check{x}_{n|n,l}^{(j,p)}\|^2, \qquad l = 1, 2; \quad n = 1, \ldots, 100$$

The MSEs for the other estimators are defined in a similar way and are denoted by $MSE_{9|L,l}^{(p)}$ for the fixed-point smoother and $MSE_{n-3|n,l}^{(p)}$ for the fixed-lag 3-step smoother. Such quantities are compared with the error variances associated with the optimal estimators in Theorems 1-3. Denote the filtering error variance of the lth component and computed for probabilities p by $P_{n|n,l}^{(p)}$, $l = 1, 2$. Similar notations are used for the error variances in the other two estimation problems, i.e., $P_{9|L,l}^{(p)}$ and $P_{n-3|n,l}^{(p)}$. Moreover, denote by $\mathbf{p}_{n,l} = [p_{n,l,r}, p_{n,l,\eta}, p_{n,l,\eta'}, p_{n,l,\eta''}]^{\mathsf{T}}$, with $l = 1, 2$, the Bernoulli probabilities affecting each component $x_{n,l}$ of $\mathbf{x}_n$.

We have considered five different situations:

1. Case 1: $\mathbf{p}_{n,l} = [0.1, 0.1, 0.1, 0.1]^{\mathsf{T}} = \mathbf{0.1}$, $l = 1, 2$.
2. Case 2: $\mathbf{p}_{n,l} = \mathbf{0.5}$, $l = 1, 2$.
3. Case 3: $\mathbf{p}_{n,l} = \mathbf{0.9}$, $l = 1, 2$.
4. Case 4: $\mathbf{p}_{n,1} = \mathbf{0.1}$ and $\mathbf{p}_{n,2} = \mathbf{0.9}$.
5. Case 5: $\mathbf{p}_{n,1} = \mathbf{0.9}$ and $\mathbf{p}_{n,2} = \mathbf{0.1}$.

These five cases represent different levels of uncertainties which, together with the choice of different degrees of autocorrelation and cross-correlation of the system noises, allow us to carry out a representative performance comparison. Specifically, Case 1 represents the most unfavorable situation in which the

[2]The norm of x_n is $\|x_n\| = \sqrt{x_{n,r}^2 + x_{n,\eta}^2 + x_{n,\eta'}^2 + x_{n,\eta''}^2}$.

probabilities that the components of the state are present are the smallest. Case 3 is the opposite situation and here there is a high probability that the information received is successful, i.e., a low probability of missed data. Case 2 is an intermediate situation where the information transmitted is likely to be missing (every component has a probability of 0.5 to be successful or to be missing). On the other hand, Cases 4 and 5 aim to evaluate the impact on the estimators performance of high probabilities of missing data in one component and low probabilities in the other component. The question is to determine whether the useful information available about a state can compensate to some extent for the lack of data in the other, enabling us to achieve a higher estimation accuracy than in Case 1. Also, we have considered two correlation scenarios: low correlations ($\lambda = 0.5$ and $\alpha = 0.75$) and high correlations ($\lambda = 5$ and $\alpha = 7.5$).

Fig. 1 depicts the differences of the MSEs ($MSE_{n,l}^{(p)} - P_{n|n,l}^{(p)}$, $l = 1, 2$) for the filtering problem in the above five cases and under the two correlation scenarios. Likewise, Fig. 2 and Fig. 3 show the same quantities for the fixed-point smoothing in the instant $n = 9$ ($MSE_{9|L,l}^{(p)} - P_{9|L,l}^{(p)}$, $l = 1, 2$) and fixed-lag 3-step smoothing ($MSE_{n-3|n,l}^{(p)} - P_{n-3|n,l}^{(p)}$, $l = 1, 2$) problems, respectively.

As we can see on the above numerical results, the accuracy of the suggested estimators is always better than that of the standard estimators which ignore the uncertainty in the measurements. This superiority in performance increases when the uncertainty in the measurements becomes higher, i.e, as the probabilities that the state is present becomes smaller. In fact, Case 3, which closely resembles the conditions for the ideal application of the standard methods, registers the least difference in performance between both estimation techniques and under both correlation scenarios. The results for Case 2 and especially for Case 1 show that the performance of the standard estimators is substantially degraded when the uncertainty becomes higher and when correlations increase. In particular, the notable improvement in performance of the suggested estimators while correlations increase is remarkable (compare figures on the right and left).

Finally, by comparing in all figures Cases 4 and 1 for the first component

(in both $\mathbf{p}_{n,1} = \mathbf{0.1}$), we can check that the reduction of the uncertainty in the second component implied by Case 4 (where $\mathbf{p}_{n,2} = \mathbf{0.9}$) decreases the difference in performance between both estimation techniques in the first component (similar comments are valid comparing Cases 5 and 1 for the second component). This is due to the fact that, unlike Case 1, Cases 4 and 5 represent schemes where a component has a high probability of being present and its relation with the other component established by the state equation compensates the missing information up to a point. Likewise, by comparing the performance differences for the second component in Cases 4 and 3 (in both $\mathbf{p}_{n,2} = \mathbf{0.9}$), we observe that the performance of the standard methods get worse in this component in relation to the suggested solutions when the uncertainty increases in the first component ($\mathbf{p}_{n,1} = \mathbf{0.1}$ in Case 4 against $\mathbf{p}_{n,1} = \mathbf{0.9}$ in Case 3). Such behavior is also observed for the first component in Cases 5 and 3.

In summary, the suggested solutions outperform the standard techniques whenever some degree of uncertainty is present in one component or simultaneously in both components. Actually, the performance of the standard techniques is seriously affected by the intermittent observation phenomenon, a problem aggravated if the correlations of the system noises increase.

6. Conclusions

The optimal SWL estimation problem from missing measurements has been investigated for discrete-time vectorial quaternion systems with autocorrelated and cross-correlated noises. Specifically, filtering, fixed-point and fixed-lag smoothing algorithms are devised. The approach used to model the possibility of lack of data (i.e., observations containing only partial information about the state or even only noise) consists in introducing a sequence of independent quaternion Bernoulli variables with known probabilities. The proposed formulation implied by the properness properties significantly reduces the computational burden in comparison with the analogous algorithms based on either a real or a WL processing. Furthermore, the superiority in performance of the proposed estimators

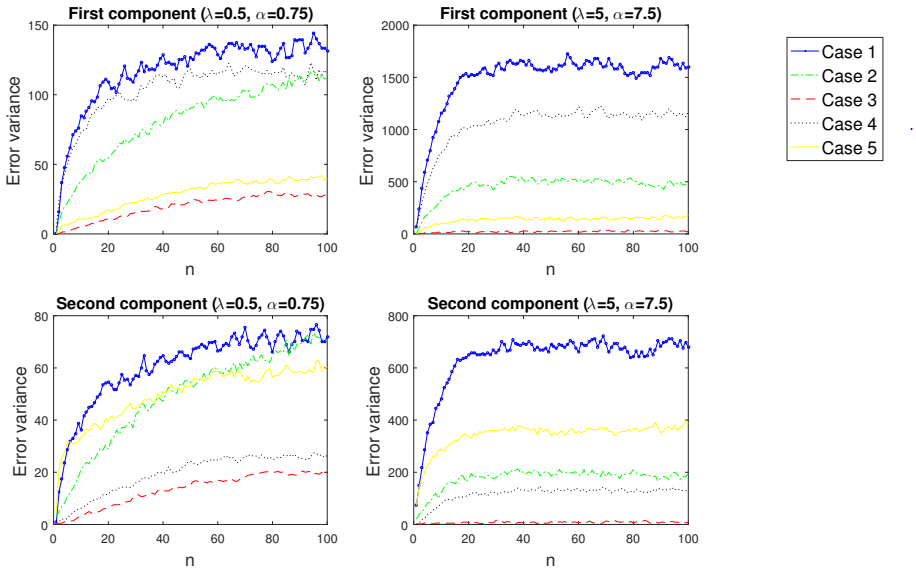

Figure 1: Differences of MSEs associated with each component in the filtering problem for Cases 1-5 and the two correlation scenarios.

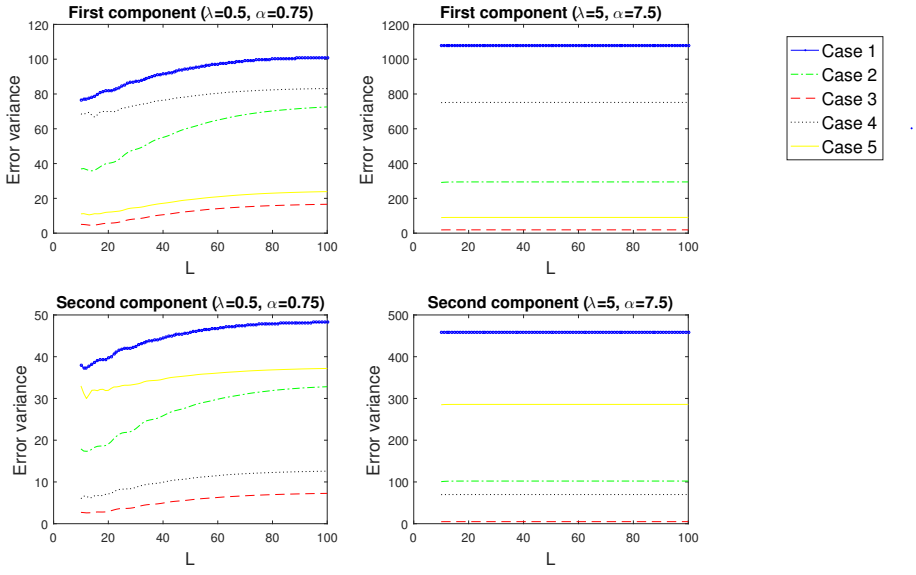

Figure 2: Differences of MSEs associated with each component in the fixed-point smoothing problem for Cases 1-5 and the two correlation scenarios.

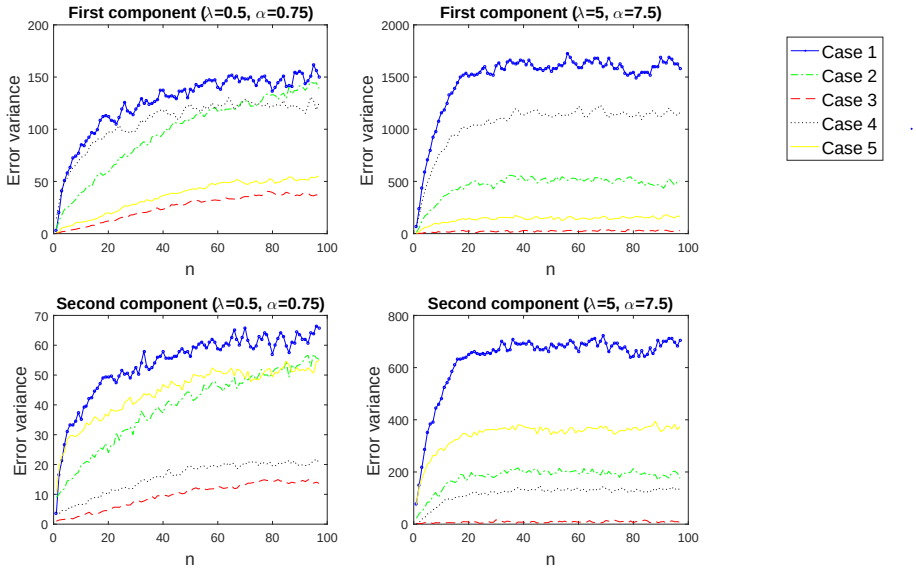

Figure 3: Differences of MSEs associated with each component in the fixed-lag smoothing problem for Cases 1-5 and the two correlation scenarios.

over more traditional estimators has been shown by means of a numerical application.

7. Acknowledgments

This work has been supported in part by I+D+i Project with reference number 1256911, under "Programa Operativo FEDER Andalucía 20014-2020", Junta de Andalucía, and the Project EI-FQM2-2019 of "Plan de Apoyo a la Investigación" of the University of Jaén.

8. Appendix

8.1. Proof of Theorem 1

We use the innovations technique to show the result. Define the innovations by $\tilde{\mathbf{y}}_{j|j-1} = \tilde{\mathbf{z}}_j - \hat{\tilde{\mathbf{z}}}_{j|j-1}$. By applying the orthogonal projection theorem we

obtain

$$\hat{\bar{\mathbf{x}}}_{n|n} = \hat{\bar{\mathbf{x}}}_{n|n-1} + \boldsymbol{\Theta}_n \tilde{\boldsymbol{\Omega}}_n^{-1} \tilde{\mathbf{y}}_{n|n-1} \tag{34}$$

with $\boldsymbol{\Theta}_n = E[\bar{\mathbf{x}}_n \tilde{\mathbf{y}}_{n|n-1}^{\mathtt{H}}]$ and $\tilde{\boldsymbol{\Omega}}_n = E[\tilde{\mathbf{y}}_{n|n-1} \tilde{\mathbf{y}}_{n|n-1}^{\mathtt{H}}]$ and then, Eq. (12) is immediately devised.

Taking projections on both sides of (5) onto the linear space spanned by $\{\tilde{\mathbf{y}}_{1|0}, \ldots, \tilde{\mathbf{y}}_{n-1|n-2}\}$ and taking into account that

$$E[\bar{\mathbf{u}}_{n-1} \tilde{\mathbf{y}}_{n-1|n-2}^{\mathtt{H}}] = \mathbf{U}_{n-1} [\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m \times 2m}]^{\mathtt{T}} + [\tilde{\mathbf{T}}_{n-1}^{\mathtt{H}}, \tilde{\mathbf{S}}_{n-1}^{\mathtt{H}}]^{\mathtt{H}}$$

we have

$$\hat{\bar{\mathbf{x}}}_{n|n-1} = \boldsymbol{\Phi}_n \hat{\bar{\mathbf{x}}}_{n-1|n-1} + \bar{\boldsymbol{\Psi}}_{n-1} \tilde{\mathbf{y}}_{n-1|n-2} \tag{35}$$

where $\bar{\boldsymbol{\Psi}}_{n-1} = \left[\mathbf{U}_{n-1} [\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m \times 2m}]^{\mathtt{T}} + [\tilde{\mathbf{T}}_{n-1}^{\mathtt{H}}, \tilde{\mathbf{S}}_{n-1}^{\mathtt{H}}]^{\mathtt{H}}\right] \tilde{\boldsymbol{\Omega}}_{n-1}^{-1}$. Hence, from the block-diagonality of $\boldsymbol{\Phi}_n$, (13) follows.

On the other hand, taking similar projections on both sides of (10), it follows that $\hat{\tilde{\mathbf{z}}}_{n|n-1} = [\tilde{\boldsymbol{\Pi}}_n, \mathbf{0}_{2m \times 2m}] \hat{\bar{\mathbf{x}}}_{n|n-1} + \tilde{\mathbf{V}}_n \tilde{\boldsymbol{\Omega}}_{n-1}^{-1} \tilde{\mathbf{y}}_{n-1|n-2}$, and then

$$\tilde{\mathbf{y}}_{n|n-1} = \tilde{\mathbf{z}}_n - [\tilde{\boldsymbol{\Pi}}_n, \mathbf{0}_{2m \times 2m}] \hat{\bar{\mathbf{x}}}_{n|n-1} - \tilde{\mathbf{V}}_n \tilde{\boldsymbol{\Omega}}_{n-1}^{-1} \tilde{\mathbf{y}}_{n-1|n-2} \tag{36}$$

from which (14) holds. Denote the prediction error by $\bar{\boldsymbol{\epsilon}}_{n|n-1} = \bar{\mathbf{x}}_n - \hat{\bar{\mathbf{x}}}_{n|n-1}$ and the prediction error covariance matrix by $\bar{\mathbf{P}}_{n|n-1} = E[\bar{\boldsymbol{\epsilon}}_{n|n-1} \bar{\boldsymbol{\epsilon}}_{n|n-1}^{\mathtt{H}}]$. From (36), (10), (5), we have

$$\boldsymbol{\Theta}_n = [\tilde{\mathbf{W}}_{n-1}^{\mathtt{H}}, \tilde{\mathbf{X}}_{n-1}^{\mathtt{H}}]^{\mathtt{H}} + \bar{\mathbf{P}}_{n|n-1} [\tilde{\boldsymbol{\Pi}}_n, \mathbf{0}_{2m \times 2m}]^{\mathtt{T}} - \boldsymbol{\Phi}_n \boldsymbol{\Theta}_{n-1} \tilde{\boldsymbol{\Omega}}_{n-1}^{-1} \tilde{\mathbf{V}}_n^{\mathtt{H}} - \bar{\boldsymbol{\Psi}}_{n-1} \tilde{\mathbf{V}}_n^{\mathtt{H}}$$

and thus, from the block-diagonality of $\boldsymbol{\Phi}_n$ and denoting by $\tilde{\mathbf{P}}_{n|n-1}$ the first $2m \times 2m$ submatrix of $\bar{\mathbf{P}}_{n|n-1}$, we derive (15). Also, taking into account that $\tilde{\mathbf{y}}_{0|-1} = \mathbf{0}_{2m}$, Eq. (16) is similarly obtained.

Next, we express the innovations in the following way

$$\tilde{\mathbf{y}}_{n|n-1} = (\tilde{\boldsymbol{\Gamma}}_n - [\tilde{\boldsymbol{\Pi}}_n, \mathbf{0}_{2m \times 2m}]) \bar{\mathbf{x}}_n + \tilde{\mathbf{v}}_n + [\tilde{\boldsymbol{\Pi}}_n, \mathbf{0}_{2m \times 2m}] \bar{\boldsymbol{\epsilon}}_{n|n-1} - \tilde{\mathbf{V}}_n \tilde{\boldsymbol{\Omega}}_{n-1}^{-1} \tilde{\mathbf{y}}_{n-1|n-2} \tag{37}$$

Thus, since $E[\tilde{\mathbf{\Gamma}}_n] = [\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]$, we have

$$E\left[(\tilde{\mathbf{\Gamma}}_n - [\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}])\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^{\mathtt{H}}(\tilde{\mathbf{\Gamma}}_n - [\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}])^{\mathtt{H}}\right] =$$
$$E\left[\tilde{\mathbf{\Gamma}}_n\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^{\mathtt{H}}\tilde{\mathbf{\Gamma}}_n^{\mathtt{H}}\right] - [\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]E\left[\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^{\mathtt{H}}\right][\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} =$$
$$\tilde{\boldsymbol{\mathcal{T}}}_m\left\{\mathrm{Cov}(\boldsymbol{\gamma}_n^r)\circ\left(\boldsymbol{\mathcal{T}}_m^{\mathtt{H}}\tilde{\mathbf{\Sigma}}_n\boldsymbol{\mathcal{T}}_m\right)\right\}\tilde{\boldsymbol{\mathcal{T}}}_m^{\mathtt{H}}$$

where $\tilde{\mathbf{\Sigma}}_n$ is given in (18), and we have applied the property of the Hadamard product stated in the Preliminaries section, besides Eqs. (1) and (9). Moreover, we have that $E[\bar{\boldsymbol{\epsilon}}_{n|n-1}\tilde{\mathbf{v}}_n^{\mathtt{H}}] = \mathbf{\Theta}_n - \bar{\mathbf{P}}_{n|n-1}[\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}}$, $E[\tilde{\mathbf{v}}_n\tilde{\mathbf{y}}_{n-1|n-2}^{\mathtt{H}}] = \tilde{\mathbf{V}}_n$, and $\bar{\boldsymbol{\epsilon}}_{n|n-1}\perp\tilde{\mathbf{y}}_{n-1|n-2}$. Thus, the innovations covariance matrix is obtained from (37) as

$$\tilde{\mathbf{\Omega}}_n = \tilde{\boldsymbol{\mathcal{T}}}_m\left\{\mathrm{Cov}(\boldsymbol{\gamma}_n^r)\circ\left(\boldsymbol{\mathcal{T}}_m^{\mathtt{H}}\tilde{\mathbf{\Sigma}}_n\boldsymbol{\mathcal{T}}_m\right)\right\}\tilde{\boldsymbol{\mathcal{T}}}_m^{\mathtt{H}} + \tilde{\mathbf{R}}_n - \tilde{\mathbf{V}}_n\tilde{\mathbf{\Omega}}_{n-1}^{-1}\tilde{\mathbf{V}}_n^{\mathtt{H}}$$
$$-[\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]\bar{\mathbf{P}}_{n|n-1}[\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} + [\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]\mathbf{\Theta}_n + \mathbf{\Theta}_n^{\mathtt{H}}[\tilde{\mathbf{\Pi}}_n, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}}$$

from which we get (17). Eq. (19) follows similarly from (37) and taking into account that $\tilde{\mathbf{y}}_{0|-1} = \mathbf{0}_{2m}$.

Next, consider the filtering error covariance matrix $\bar{\mathbf{P}}_{n|n} = E[\bar{\boldsymbol{\epsilon}}_{n|n}\bar{\boldsymbol{\epsilon}}_{n|n}^{\mathtt{H}}]$ with $\bar{\boldsymbol{\epsilon}}_{n|n} = \bar{\mathbf{x}}_n - \hat{\bar{\mathbf{x}}}_{n|n}$. From (34), it is easy to check that $\bar{\mathbf{P}}_{n|n} = \bar{\mathbf{P}}_{n|n-1} - \mathbf{\Theta}_n\tilde{\mathbf{\Omega}}_n^{-1}\mathbf{\Theta}_n^{\mathtt{H}}$ and thus (20) holds.

Finally, from (5) and (35) we get

$$\bar{\boldsymbol{\epsilon}}_{n|n-1} = \mathbf{\Phi}_n\bar{\boldsymbol{\epsilon}}_{n-1|n-1} - \bar{\mathbf{\Psi}}_{n-1}\tilde{\mathbf{y}}_{n-1|n-2} + \bar{\mathbf{u}}_{n-1} \tag{38}$$

Due to $\bar{\boldsymbol{\epsilon}}_{n-1|n-1}\perp\tilde{\mathbf{y}}_{n-1|n-2}$, $E[\bar{\mathbf{u}}_{n-1}\bar{\boldsymbol{\epsilon}}_{n-1|n-1}^{\mathtt{H}}] = \mathbf{U}_{n-1} - \bar{\mathbf{\Psi}}_{n-1}\mathbf{\Theta}_{n-1}^{\mathtt{H}}$ and $E[\bar{\mathbf{u}}_{n-1}\tilde{\mathbf{y}}_{n-1|n-2}^{\mathtt{H}}] = \bar{\mathbf{\Psi}}_{n-1}\tilde{\mathbf{\Omega}}_{n-1}$ then, we have

$$\bar{\mathbf{P}}_{n|n-1} = \mathbf{\Phi}_n\bar{\mathbf{P}}_{n-1|n-1}\mathbf{\Phi}_n^{\mathtt{H}} + \mathbf{\Phi}_n\left[\mathbf{U}_{n-1}^{\mathtt{H}} - \mathbf{\Theta}_{n-1}\bar{\mathbf{\Psi}}_{n-1}^{\mathtt{H}}\right]$$
$$+\left[\mathbf{U}_{n-1}^{\mathtt{H}} - E[\bar{\mathbf{x}}_{n-1}\tilde{\mathbf{y}}_{n-1|n-2}^{\mathtt{H}}]\bar{\mathbf{\Psi}}_{n-1}^{\mathtt{H}}\right]^{\mathtt{H}}\mathbf{\Phi}_n^{\mathtt{H}} - \bar{\mathbf{\Psi}}_{n-1}\tilde{\mathbf{\Omega}}_{n-1}\bar{\mathbf{\Psi}}_{n-1}^{\mathtt{H}} + \mathbf{Q}_{n-1}$$

and from the block-diagonality of $\mathbf{\Phi}_n$, (21) follows.

8.2. Proof of Theorem 2

Defining $\mathbf{\Theta}_{n,L} = E[\bar{\mathbf{x}}_n\tilde{\mathbf{y}}_{L|L-1}^{\mathtt{H}}]$ then, Eq. (22) is derived from a similar reasoning to that employed in obtaining (12). Moreover, from (36), (10), (35), and

(5) we have

$$\boldsymbol{\Theta}_{n,L} = E[\bar{\mathbf{x}}_n \bar{\boldsymbol{\epsilon}}_{L-1|L-1}^{\mathtt{H}}] \boldsymbol{\Phi}_L^{\mathtt{H}} [\tilde{\boldsymbol{\Pi}}_L, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} - \boldsymbol{\Theta}_{n,L-1} \bar{\boldsymbol{\Psi}}_{L-1}^{\mathtt{H}} [\tilde{\boldsymbol{\Pi}}_L, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} - \boldsymbol{\Theta}_{n,L-1} \tilde{\boldsymbol{\Omega}}_{L-1}^{-1} \mathbf{V}_L$$

and from the block-diagonality of $\boldsymbol{\Phi}_L$, (23) is obtained being $\tilde{\boldsymbol{\Xi}}_{n,L} = E[\tilde{\mathbf{x}}_n \tilde{\boldsymbol{\epsilon}}_{L|L}^{\mathtt{H}}]$. Eq. (24) for $\tilde{\boldsymbol{\Theta}}_{n,n+1}$ is similarly devised.

Now, from (34), (35), (36), and (5) we express the filtering error as

$$\bar{\boldsymbol{\epsilon}}_{L|L} = \boldsymbol{\Phi}_L \bar{\boldsymbol{\epsilon}}_{L-1|L-1} + \bar{\mathbf{u}}_{L-1} - \left[\bar{\boldsymbol{\Psi}}_{L-1} - \boldsymbol{\Theta}_L \tilde{\boldsymbol{\Omega}}_L^{-1} \tilde{\mathbf{V}}_L \tilde{\boldsymbol{\Omega}}_{L-1}^{-1}\right] \tilde{\mathbf{y}}_{L-1|L-2} - \boldsymbol{\Theta}_L \tilde{\boldsymbol{\Omega}}_L^{-1} \left[\tilde{\boldsymbol{\Gamma}}_L \bar{\mathbf{x}}_L + \tilde{\mathbf{v}}_L - [\tilde{\boldsymbol{\Pi}}_L, \mathbf{0}_{2m\times 2m}] \hat{\bar{\mathbf{x}}}_{L|L-1}\right] \quad (39)$$

and from (38), it follows that

$$E[\bar{\mathbf{x}}_n \bar{\boldsymbol{\epsilon}}_{L|L-1}^{\mathtt{H}}] = \bar{\boldsymbol{\Xi}}_{n,L-1} \boldsymbol{\Phi}_L^{\mathtt{H}} - \boldsymbol{\Theta}_{n,L-1} \bar{\boldsymbol{\Psi}}_{L-1}^{\mathtt{H}} \quad (40)$$

Thus, taking (39) and (40) into account, we have

$$\bar{\boldsymbol{\Xi}}_{n,L} = \left[\bar{\boldsymbol{\Xi}}_{n,L-1} \boldsymbol{\Phi}_L^{\mathtt{H}} - \boldsymbol{\Theta}_{n,L-1} \bar{\boldsymbol{\Psi}}_{L-1}^{\mathtt{H}}\right] \left[\mathbf{I}_{4m} - [\tilde{\boldsymbol{\Pi}}_L, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} \tilde{\boldsymbol{\Omega}}_L^{-1} \boldsymbol{\Theta}_L^{\mathtt{H}}\right] + \boldsymbol{\Theta}_{n,L-1} \tilde{\boldsymbol{\Omega}}_{L-1}^{-1} \tilde{\mathbf{V}}_L^{\mathtt{H}} \tilde{\boldsymbol{\Omega}}_L^{-1} \boldsymbol{\Theta}_L^{\mathtt{H}}$$

and from the block-diagonality of $\boldsymbol{\Phi}_L$, (25) holds. Also, (26) is similarly obtained.

Finally, (27) can be easily derived from (22).

8.3. Proof of Theorem 3

Eq. (28) is obtained similarly to (12). Moreover, by defining $\boldsymbol{\Lambda}_{n+k,n} = E[\bar{\mathbf{x}}_{n+k} \bar{\boldsymbol{\epsilon}}_{n|n-1}^{\mathtt{H}}]$ and from (37) it is easy to check that

$$\boldsymbol{\Theta}_{n+k,n} = \boldsymbol{\Lambda}_{n+k,n} [\tilde{\boldsymbol{\Pi}}_n, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} - \boldsymbol{\Theta}_{n+k,n-1} \tilde{\boldsymbol{\Omega}}_{n-1}^{-1} \tilde{\mathbf{V}}_n^{\mathtt{H}}$$

from which (29) holds with $\tilde{\boldsymbol{\Lambda}}_{n+k,n} = E[\tilde{\mathbf{x}}_{n+k} \tilde{\boldsymbol{\epsilon}}_{n|n-1}^{\mathtt{H}}]$. Moreover, since

$$\hat{\bar{\mathbf{x}}}_{n|n-1} = \boldsymbol{\Phi}_n \hat{\bar{\mathbf{x}}}_{n-1|n-2} + \boldsymbol{\Theta}_{n,n-1} \tilde{\boldsymbol{\Omega}}_{n-1}^{-1} \tilde{\mathbf{y}}_{n-1|n-2}$$

then, from (5) and (37), the prediction error can be expressed as

$$\bar{\boldsymbol{\epsilon}}_{n|n-1} = \boldsymbol{\Phi}_n \bar{\boldsymbol{\epsilon}}_{n-1|n-2} + \bar{\mathbf{u}}_{n-1} - \boldsymbol{\Theta}_{n,n-1}\tilde{\boldsymbol{\Omega}}_{n-1}^{-1}\Big[(\tilde{\boldsymbol{\Gamma}}_{n-1} - [\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m\times 2m}])\bar{\mathbf{x}}_{n-1} + \tilde{\mathbf{v}}_{n-1} + [\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m\times 2m}]\bar{\boldsymbol{\epsilon}}_{n-1|n-2} - \tilde{\mathbf{V}}_{n-1}\tilde{\boldsymbol{\Omega}}_{n-2}^{-1}\tilde{\mathbf{y}}_{n-2|n-3}\Big]$$

Hence, we have

$$\boldsymbol{\Lambda}_{n-1,n} = \bar{\mathbf{P}}_{n-1|n-2}\Big[\boldsymbol{\Phi}_n - \boldsymbol{\Theta}_{n,n-1}\tilde{\boldsymbol{\Omega}}_{n-1}^{-1}[\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m\times 2m}]\Big]^{\mathtt{H}} + \mathbf{U}_{n-1}^{\mathtt{H}} - [\tilde{\mathbf{W}}_{n-2}^{\mathtt{H}}, \tilde{\mathbf{X}}_{n-2}^{\mathtt{H}}]^{\mathtt{H}}\tilde{\boldsymbol{\Omega}}_{n-1}^{-1}\boldsymbol{\Theta}_{n,n-1}^{\mathtt{H}} + \boldsymbol{\Theta}_{n-1,n-2}\tilde{\boldsymbol{\Omega}}_{n-2}^{-1}\tilde{\mathbf{V}}_{n-1}^{\mathtt{H}}\tilde{\boldsymbol{\Omega}}_{n-1}^{-1}\boldsymbol{\Theta}_{n,n-1}^{\mathtt{H}}$$

and from the block-diagonality of $\boldsymbol{\Phi}_n$, (30) follows. Moreover, Eq. (32) is similarly devised.

On the other hand, from (5), we have

$$\begin{aligned}\boldsymbol{\Theta}_{n,n-1} &= \boldsymbol{\Phi}_n E[\bar{\mathbf{x}}_{n-1}\tilde{\mathbf{y}}_{n-1|n-2}^{\mathtt{H}}] + \mathbf{U}_{n-1}[\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} + [\tilde{\mathbf{T}}_{n-1}^{\mathtt{H}}, \tilde{\mathbf{S}}_{n-1}^{\mathtt{H}}]^{\mathtt{H}} \\ &= \boldsymbol{\Phi}_n \bar{\mathbf{P}}_{n-1|n-2}[\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} + \boldsymbol{\Phi}_n[\tilde{\mathbf{W}}_{n-2}^{\mathtt{H}}, \tilde{\mathbf{X}}_{n-2}^{\mathtt{H}}]^{\mathtt{H}} \\ &\quad - \boldsymbol{\Phi}_n\boldsymbol{\Theta}_{n-1,n-2}\tilde{\boldsymbol{\Omega}}_{n-2}^{-1}\tilde{\mathbf{V}}_{n-1}^{\mathtt{H}} + \mathbf{U}_{n-1}[\tilde{\boldsymbol{\Pi}}_{n-1}, \mathbf{0}_{2m\times 2m}]^{\mathtt{T}} + [\tilde{\mathbf{T}}_{n-1}^{\mathtt{H}}, \tilde{\mathbf{S}}_{n-1}^{\mathtt{H}}]^{\mathtt{H}}\end{aligned}$$

where (37) has been applied in the last equality. Thus, Eq. (31) follows.

Finally, (33) is easily derived from (28).

References

[1] A. Jazwinski, Stochastic Processes and Filtering Theory, San Diego, Academic Press, 1970.

[2] S. Julier and J. Uhlmann, Unscented filtering and nonlinear estimation, Proc. IEEE 92 (3) (2004) 401–422.

[3] G. Chang, On Kalman filter for linear system with colored measurement noise, J. Geod. 88 (2014) 1163–1170.

[4] L. Ma, H. Wang, and J. Chen, Analysis of Kalman filter with correlated noises under different dependence, J. Inform. Comput. Sci. 7 (5) (2010) 1147–1154.

[5] H. Zhao and C. Zhang, Non-Gaussian noise quadratic estimation for linear discrete-time time-varying systems, Neurocomputing 174 (2016) 921–927.

[6] C.C. Took, D.P. Mandic, A quaternion widely linear adaptive filter, IEEE Trans. Signal Process. 58 (8) (2010) 4427–4431.

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

[8] J. Navarro-Moreno, R.M. Fernández-Alcalá, and J.C. Ruiz-Molina, Quaternion widely linear series expansion and its applications, IEEE Signal Process. Letters 19 (12) (2012) 868–871.

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

[10] J. Navarro-Moreno, R.M. Fernández-Alcalá, and J.C. Ruiz-Molina, A quaternion widely linear model for nonlinear Gaussian estimation, IEEE Trans. Signal Process. 62 (24) (2014) 6414–6424.

[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. Frankl. 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] 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. Netw. 2 (4) (2016) 477–488.

[20] 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: Proc. 17th European Signal Process. Conf. (2009) 2037–2041.

[21] 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.

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

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

[24] 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.

[25] 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.

[26] 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.

[27] T. Nitta, M. Kobayashi and D.P. Mandic, Hypercomplex widely linear estimation through the lens of underpinning geometry, IEEE Trans. Signal Process. 67 (15) (2019) 3985–3994.

[28] J.P. Delmas, A. Oukaci, and P. Chevalier, On the asymptotic distribution of GLR for impropriety of complex signals, Signal Process. 91 (10) (2011) 2259–2267.

[29] N. Vakhania, Random vectors with values in quaternions Hilbert spaces, Th. Prob. App. 43 (1) (1998) 99–115.

[30] P.O. Amblard, N. Le Bihan, On properness of quaternion valued random variables, IMA Conf. Math. Signal Process. (2004).

[31] N. Le Bihan, The geometry of proper quaternion random variables, Signal Process. 138 (2017) 106–116.

[32] P. Ginzberg and A.T. Walden, Testing for quaternion propriety, IEEE Trans. Signal Process. 59 (7) (2011) 3025–3034.

[33] J. Vía, D.P. Palomar, and L. Vielva, Generalized likelihood ratios for testing the properness of quaternion Gaussian vectors, IEEE Trans. Signal Process. 59 (4) (2011) 1356–1370.

[34] J. Vía, and L. Vielva, Locally most powerful invariant tests for the properness of quaternion Gaussian vectors, IEEE Trans. Signal Process. 60 (3) (2012) 997–1009.

[35] N. Le Bihan and P.O. Amblard, Detection and estimation of Gaussian proper quaternion valued random processes, in: Proc. IMA Conf. Math. Signal Process. (2006) 23–26.

[36] Y. Wang, W. Xia, Z. He, Polarimetric detection for vector-sensor array in quaternion Gaussian proper noise, Int. Congress Image Signal Process. 2 (2013) 1164–1168.

[37] J. Via, D. Palomar, L. Vielva, and I. Santamaria, Quaternion ICA from second-order statistics, IEEE Trans. Signal Process. 59 (4) (2011) 1586–1600.

[38] S. Javidi, C. Took, D. Mandic, Fast independent component analysis algorithm for quaternion valued signals, IEEE Trans. Neural Netw. 22 (12) (2011) 1967–1978.

[39] J. Navarro-Moreno and J.C. Ruiz-Molina, Semi-widely linear estimation of $\mathbb{C}^{\eta}$-proper quaternion random signal vectors under Gaussian and stationary conditions, Signal Process. 119 (2016) 56–66.

[40] Y. Xia, S. Tao, Z. Li, M. Xiang, W. Pei and D. P. Mandic, Full mean square performance bounds on quaternion estimators for improper data, IEEE Trans. Signal Process. 67 (15) (2019) 4093–4106.

[41] M. Xiang, B. Scalzo-Dees, and D. P. Mandic, Multiple model adaptive estimation for 3D and 4D signals: A widely linear quaternion approach", IEEE Trans. Neural Netw. Learn. Syst 30 (1) (2019) 72–84.

[42] S. Olhede, D. Ramirez, and P. Schreier, The random monogenic signal, in: Proc. IEEE Int. Conf. Image Process. (2012) 2493–2496.

[43] P. Ginzberg and A.T. Walden, Quaternion VAR modelling and estimation, IEEE Trans. Signal Process. 61 (1) (2013) 154–158.

[44] A. Sloin, A. Wiesel, Proper quaternion Gaussian graphical models, IEEE Trans. Signal Process. 62 (20) (2014) 5487–5496.

[45] S. Olhede, D. Ramirez, and P. Schreier, Detecting directionality in random fields using the monogenic signal, IEEE Trans. Inform. Theory 60 (10) (2014) 6491–6510.

[46] S. Buchholz and N. Le Bihan, Polarized signal classification by complex and quatenionic multi-layer perceptrons, Int. J. Neural Systems, 18 (2) (2008) 75–85.

[47] A.V. Savkin and I.R. Petersen, Robust filtering with missing data and a deterministic description of noise and uncertainty, Int. J. of Systems Sci. 28 (4) (1997) 373–390.

[48] S.M.K. Mohamed and S. Nahavandi, Robust filtering for uncertain discrete-time systems with uncertain noise covariance and uncertain observations, in: Proc. 6th IEEE Int. Conf. Ind. Informatics 59 (11) (2008) 667–672.

[49] N. Nahi, Optimal recursive estimation with uncertain observation, IEEE Trans. Inf. Theory 15 (6) (1969) 457–462.

[50] 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.

[51] 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.

[52] J.D. Jiménez-López, J. Linares-Pérez, S. Nakamori, R. Caballero-Águila, and A. Hermoso-Carazo, Signal estimation based on covariance information from observations featuring correlated uncertainty and coming from multiple sensors, Signal Process. 88 (2008) 2998–3006.

[53] F. Cacace, A. Fasano, and A. Germani, Quadratic filtering of non-Gaussian systems with intermittent observations, in: Proc. IEEE 52nd Annual Conf. Decision Control (2013) 4024–4029.

[54] S. Ko, S. Kang, and J. Chang, Performance comparison of covariance-assignment state estimators with intermittent observations, Int. J. Control, Automation, and Systems 13 (86) (2015) 1–11.

[55] H. Dong, Z. Wang, D.W.C. Ho, and H. Gao, Variance-constrained $\mathcal{H}_\infty$ filtering for a class of nonlinear time-varying systems with multiple missing measurements: the finite-horizon case, IEEE Trans. Signal Process. 58 (5) (2010) 2534-2543.

[56] J. Hu, Z. Wang, H. Gao, and L.K. Stergioulas, Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements, Automatica 48 (2012) 2007–2015.

[57] S. Zhang, Y. Zhao, F. Wu, and J. Zhao, Robust recursive filtering for uncertain systems with finite-step correlated noises, stochastic nonlinearities and autocorrelated missing measurements, Aerospace Sci. Tech. 39 (2014) 272–280.

[58] G. Deng and H. Zhao, H^∞ fault detection for linear discrete time-varying descriptor systems with missing measurements, Discrete Dyn. Nat. Soc. (2015).

[59] L. Yan, L. Jiang, J. Liu, Y. Xia and M. Fu, Optimal distributed Kalman filtering fusion for multirate multisensor dynamic systems with correlated noise and unreliable measurements, IET Signal Process. 12 (4) (2018) 522-531.

[60] Z. Xing, Y. Xia, L. Yan, K. Lu and Q. Gong, Multisensor distributed weighted Kalman filter fusion with network delays, stochastic uncertainties, autocorrelated, and cross-correlated noises, IEEE Trans. Syst. Man Cyber.: Systems 48 (5) (2018) 716-726.

[61] H.S. Karimi and B. Natarajan, Kalman filtered compressive sensing with intermittent observations, Signal Process. 163 (2019) 49–58.

[62] 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.

[63] J. Navarro-Moreno, R.M. Fernández-Alcalá, J.D. Jiménez López and J.C. Ruiz-Molina, Widely linear estimation for multisensor quaternion systems with mixed uncertainties in the observations, J. Frankl. Inst. 356 (2019) 3115–3138.

[64] R. Caballero-Águila, I. García-Garrido, and J. Linares-Pérez, Optimal fusion filtering in multisensor stochastic systems with missing measurements and correlated noises, Math. Problems Eng. (2013) 1–14.