Wide-Sense Markov Signals on the Tessarine Domain. A Study under Properness Conditions

Jesús Navarro-Moreno[a], Juan Carlos Ruiz-Molina[a,*]

[a]*Department of Statistics and Operations Research, University of Jaén, 23071 Jaén, Spain.*

Abstract

The quaternion algebra is not always the best choice for processing 4D hypercomplex signals. This paper aims to explore tessarines as an alternative algebra to solve the estimation problem. More concretely, wide-sense Markov signals in the tessarine domain are introduced and their properties under properness properties are analyzed. Firstly, the $\mathbb{T}_2$-properness condition in the tessarine setting is defined and then, the linear estimation problem under tessarine processing is addressed. The equivalence between the optimal estimator based on tessarine widely linear processing and the one based on tessarine $\mathbb{T}_2$ processing is proved, thus attaining a notable reduction in computational burden. Next, the $\mathbb{T}_i$-proper wide-sense Markov signals, $i = 1, 2$, are defined and a forwards representation for modeling them is suggested. Finally, the estimation problem with intermittent observations for this class of signals is tackled. Specifically, based on the forwards representation, two algorithms for the problems of filtering, prediction and fixed-interval smoothing are devised. Numerical simulations are developed where the superiority of the $\mathbb{T}_i$ estimators, $i = 1, 2$, over their counterparts in the quaternion domain is shown.

Keywords: Estimation, Intermittent observations, Tessarine processing, $\mathbb{T}_i$-properness, Wide-sense Makov signals.

---

*Corresponding author

Email addresses: `jnavarro@ujaen.es` (Jesús Navarro-Moreno), `jcruiz@ujaen.es` (Juan Carlos Ruiz-Molina)

1. Introduction

Estimation is a fundamental problem in signal processing. In linear minimum-mean square error (MMSE) estimation theory, when the processes under consideration are not Gaussian, the classes of stochastic processes which are of practical importance are wide-sense Markov (WSM) processes [1]. The equivalence between the WSM condition and the state-space representation for the signal is really what makes WSM signals especially attractive in signal processing. In general, the papers devoted to studying WSM signals are restricted to the real or complex domains (see, e.g., [1]-[5]) while other mathematical domains offer sound representations [6]. For example, and to our knowledge, WSM signals of order n have not been studied in the four-dimensional (4D) hypercomplex domains yet.

In the classical formulation of the estimation problem it is assumed that all the observations are available during the state estimation time, since Kalman estimators are sensitive to missing measurements, that is, they perform suboptimally [7]. Nevertheless, in many real-world applications, such as in the case of failures in the measuring sensors, high noise environments, and poor communication resources [8], 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. They have been studied extensively for real-valued signals (see, e.g., [9]-[16]) and recently for quaternion-valued signals [17]-[19].

Quaternions and tessarines are 4D hypercomplex algebras that extend the system of complex numbers [20]. Both are originated from a similar mathematical construction method which endows them with interesting features for the development of signal processing tools [21]-[23]. Until now, however, quaternions has been the most commonly used hypercomplex algebra by the signal processing community. The reason for which quaternions is the predominant domain is twofold. On the one hand, this system is constructed to be a normed division algebra, that is, it is closed with an inverse operation and it is equipped with a norm. These properties make them naturally suited to address benchmark problems of signal processing (e.g., the signal estimation problem [24]). On the other hand, quaternions have shown its superiority over the real field in some applied problems [25]. For instance, quaternions provide mathematical robustness to represent rotations in space since they are immune to gimbal lock singularity [26]. In contrast to quaternions, the analysis of tessarine data has the limitation of not being a normed space. Nevertheless, this drawback has been recently overcome in the context of the signal estimation problem [27].

Adopting a 4D hypercomplex system as the mathematical framework has some important repercussions in the processing involved. Indeed, it implies renouncing to some usual properties of the real and complex fields. Specifically, the quaternions algebra is not commutative and the tessarines is not a division algebra (it contains some divisors of zero) [20]. Thus, the techniques developed on 4D hypercomplex domains are not direct extensions of the vectorial treatment of the problem. Recent literature has shown how such obstacles have been overcome and multiple applications have been suggested (see, e.g., [21]-[37]). Interestingly, comparative studies between the use of quaternion and tessarine algebras have evidenced that the choice of a specific algebra may condition the behavior of the proposed methods [27], [38, 39]. In fact, neither of the two 4D hypercomplex algebras always performs better than the other.

It is well-known that quaternions, tessarines and $\mathbb{R}^4$ are isomorphic spaces [20]. Therefore, the signal processing in $\mathbb{R}^4$ is equivalent to both the quaternion widely linear (QWL) processing [40] and the tessarine widely linear (TWL) processing [27], with the devised algorithms having the same computational burden regardless of the chosen processing. However, this characteristic is lost when the signals present properness properties [27], [41]-[43]. Quaternion properness has been widely studied in the past and the background theory is well developed [44]. Knowing that a quaternion is proper has two important effects: first, either a quaternion strictly linear (QSL) or semi-widely linear (QSWL) processing should be used; second, a notable reduction in the computational burden involved in comparison with the real processing is attained.

Unlike quaternion properness, tessarine properness has been much less studied. The $\mathbb{T}_1$-properness concept and the tessarine strictly linear (TSL) processing were introduced[1] in [27]. This paper also proves that, under $\mathbb{T}_1$-properness conditions, TSL processing can outperform QSL processing. In our opinion, the reported results are promising and set the way for a deeper study of new types of tessarine properness and of the associated processing.

The aim of this work is to study the estimation of WSM tessarine-valued signals with intermittent observations and to show their advantages when used under properness conditions. For that, firstly, a new type of properness for tessarine signals (called $\mathbb{T}_2$-properness), which is less restrictive than the $\mathbb{T}_1$-properness given in [27], is introduced. We prove that, under $\mathbb{T}_2$-properness conditions, $\mathbb{T}_2$ processing is the most convenient one since it provides estimators with equal performance as TWL processing while attaining a lower computational cost. Then, we define the class of $\mathbb{T}_i$-proper, $i = 1, 2$, WSM signals and characterize them by means of their forwards representation. This representation is the basis for devising two Kalman-like algorithms for the problems of filtering, prediction and fixed-interval-smoothing with intermittent observations. Finally, the effectiveness of the developed algorithms is demonstrated through simulations. Remarkably, the numerical examples show that the $\mathbb{T}_i$-proper, $i = 1, 2$, estimators outperform their counterparts in the quaternion domain.

The rest of this paper is organized as follows. Section 2 revisits the tessarine processing. Then, Section 3 introduces the concept of $\mathbb{T}_2$-properness and studies the associated $\mathbb{T}_2$ processing. The $\mathbb{T}_i$-proper, $i = 1, 2$, WSM signals are analyzed in Section 4. Two algorithms for filtering, prediction and fixed-interval smoothing with intermittent observations are also devised in this section and their performance is compared with that of the quaternion estimators by means of numerical examples. A Section of conclusions ends the paper. To preserve

[1] Currently, $\mathbb{T}_1$-properness is called $\mathbb{T}$-properness in [27]. We change the denomination to avoid confusion with the one used for the new properness concept suggested here.

continuity in our presentation, all proofs are deferred to an Appendix.

2. Preliminaries

2.1. Notation

Throughout this paper, all the random variables are assumed to have zero-mean unless otherwise stated. Next we introduce the basic notation. We use boldface uppercase letters to denote matrices, boldface lowercase letters for column vectors, and lightface lowercase letters for scalar quantities. Superscripts "$*$", "$\mathtt{T}$" and "$\mathtt{H}$" represent the tessarine conjugate, transpose and Hermitian transpose, respectively. The real part of a tessarine will be denoted by $\mathcal{R}\{\cdot\}$.

Moreover, the notation $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{T}$ is used to denote the set of integer, real and tessarine numbers, respectively. In this sense, $\mathbf{A} \in \mathbb{R}^{p\times q}$ (respectively, $\mathbf{A} \in \mathbb{T}^{p\times q}$) means that $\mathbf{A}$ is a real (respectively, tessarine) $p \times q$ matrix, and similarly $\mathbf{r} \in \mathbb{R}^p$ (respectively, $\mathbf{r} \in \mathbb{T}^p$) means that $\mathbf{r}$ is a p-dimensional real (respectively, tessarine) vector. $\mathbf{r}^{(m)}$ denotes the vector formed by the first m elements of $\mathbf{r}$; similarly, $\mathbf{A}^{(m)}$ denotes the matrix formed by the first m rows of $\mathbf{A}$ and $\mathbf{A}^{(m,n)}$ the matrix formed by the first m rows and the first n columns of $\mathbf{A}$. $\text{diag}\{\mathbf{r}\}$ represents the diagonal matrix created with the elements of $\mathbf{r}$ and $\text{diag}\{\mathbf{A}\}$ the main diagonal of $\mathbf{A}$. δ_{nl} is the Kronecker delta function, which is equal to one if $l = n$, and zero otherwise. Finally, $E[\cdot]$ is the expectation operator, $\mathbf{0}_{p\times q}$ denotes the $p \times q$ zero matrix, $\mathbf{I}_p$ represents the identity matrix of dimension p and "$\otimes$", "$\odot$" stand for the Kronecker and Hadamard products. The next property of the Hadamard product will be useful: if $\mathbf{A} \in \mathbb{R}^{p\times p}$ and $\mathbf{r} \in \mathbb{R}^p$ then,

$$\text{diag}\{\mathbf{r}\}\mathbf{A}\,\text{diag}\{\mathbf{r}\} = \mathbf{r}\mathbf{r}^{\mathtt{T}} \odot \mathbf{A} \tag{1}$$

2.2. Review of the Tessarine Processing

In this subsection, the basic concepts in the tessarine domain as well as some properties of interest are established.

Consider a random variable $x \in \mathbb{T}$ defined as [39]

$$x = a + \mathrm{i}b + \mathrm{j}c + \mathrm{k}d$$

where the random variables $a, b, c, d \in \mathbb{R}$ and the imaginary units $(\mathrm{i}, \mathrm{j}, \mathrm{k})$ satisfy the identities

$$\mathrm{ij} = \mathrm{k}, \quad \mathrm{jk} = \mathrm{i}, \quad \mathrm{ki} = -\mathrm{j}, \quad \mathrm{i}^2 = \mathrm{k}^2 = -1, \quad \mathrm{j}^2 = 1$$

The real vector of x is defined by $\mathbf{x}_r = [a, b, c, d]^{\mathtt{T}}$. The conjugate of x is

$$x^* = a - \mathrm{i}b + \mathrm{j}c - \mathrm{k}d$$

and we define the following auxiliary tessarines:

$$x^{\mathrm{i}} = a + \mathrm{i}b - \mathrm{j}c - \mathrm{k}d$$
$$x^{\mathrm{k}} = a - \mathrm{i}b - \mathrm{j}c + \mathrm{k}d$$

Consider a random signal $\{\mathbf{x}(t) \in \mathbb{T}^p, t \in \mathbb{Z}\}$, with components $x_i(t) = a_i(t) + \mathrm{i}b_i(t) + \mathrm{j}c_i(t) + \mathrm{k}d_i(t)$, $i = 1, \dots, p$, where $a_i(t)$, $b_i(t)$, $c_i(t)$, and $d_i(t)$ real random signals. The real vectors associated to $a_i(t)$, $b_i(t)$, $c_i(t)$, and $d_i(t)$ are denoted by

$$\mathbf{a}(t) = [a_1(t), \dots, a_p(t)]^{\mathtt{T}}, \qquad \mathbf{b}(t) = [b_1(t), \dots, b_p(t)]^{\mathtt{T}}$$
$$\mathbf{c}(t) = [c_1(t), \dots, c_p(t)]^{\mathtt{T}}, \qquad \mathbf{d}(t) = [d_1(t), \dots, d_p(t)]^{\mathtt{T}}$$

and the real vector formed with the components of $\mathbf{x}(t)$ is denoted by

$$\mathbf{x}_r(t) = [\mathbf{a}^{\mathtt{T}}(t), \mathbf{b}^{\mathtt{T}}(t), \mathbf{c}^{\mathtt{T}}(t), \mathbf{d}^{\mathtt{T}}(t)]^{\mathtt{T}}, \quad t \in \mathbb{Z} \tag{2}$$

Moreover, the following auxiliary tessarine vectors are defined:

$$\mathbf{x}^{\nu}(t) = [x_1^{\nu}(t), \dots, x_p^{\nu}(t)]^{\mathtt{T}}, \quad \nu = \mathrm{i}, *, \mathrm{k}, \quad t \in \mathbb{Z}$$

The pseudo autocorrelation and pseudo cross-correlation matrices of the random vectors $\mathbf{x} \in \mathbb{T}^{p_1}$ and $\mathbf{y} \in \mathbb{T}^{p_2}$ are denoted by $\mathbf{\Gamma}_{\mathbf{x}} = E[\mathbf{x}\mathbf{x}^{\mathtt{H}}]$ and $\mathbf{\Gamma}_{\mathbf{xy}} = E[\mathbf{x}\mathbf{y}^{\mathtt{H}}]$, respectively. Similarly, the pseudo autocorrelation function of the random signal $\mathbf{x}(t) \in \mathbb{T}^p$ is defined as $\mathbf{\Gamma}_{\mathbf{x}}(t,s) = E[\mathbf{x}(t)\mathbf{x}^{\mathtt{H}}(s)]$, $\forall t, s \in \mathbb{Z}$.

The particular case $t = s$ is denoted by $\mathbf{\Gamma}_{\mathbf{x}}(t) = E[\mathbf{x}(t)\mathbf{x}^{\mathtt{H}}(t)]$. The pseudo cross-correlation function of the random signals $\mathbf{x}(t) \in \mathbb{T}^{p_1}$ and $\mathbf{y}(t) \in \mathbb{T}^{p_2}$ is defined as $\mathbf{\Gamma}_{\mathbf{xy}}(t,s) = E[\mathbf{x}(t)\mathbf{y}^{\mathtt{H}}(s)]$, $\forall t, s \in \mathbb{Z}$.

We define the augmented vector of $\mathbf{x}(t) \in \mathbb{T}^p$ as

$$\bar{\mathbf{x}}(t) = [\mathbf{x}^{\mathtt{T}}(t), \mathbf{x}^{\mathtt{H}}(t), \mathbf{x}^{\mathrm{i}^{\mathtt{T}}}(t), \mathbf{x}^{\mathrm{k}^{\mathtt{T}}}(t)]^{\mathtt{T}}, \quad t \in \mathbb{Z} \tag{3}$$

and thus,

$$\mathbf{\Gamma}_{\bar{\mathbf{x}}}(t,s) = \begin{pmatrix} \mathbf{\Gamma}_{\mathbf{x}}(t,s) & \mathbf{\Gamma}_{\mathbf{xx}^*}(t,s) & \mathbf{\Gamma}_{\mathbf{xx}^{\mathrm{i}}}(t,s) & \mathbf{\Gamma}_{\mathbf{xx}^{\mathrm{k}}}(t,s) \\ \mathbf{\Gamma}^*_{\mathbf{xx}^*}(t,s) & \mathbf{\Gamma}^*_{\mathbf{x}}(t,s) & \mathbf{\Gamma}^*_{\mathbf{xx}^{\mathrm{k}}}(t,s) & \mathbf{\Gamma}^*_{\mathbf{xx}^{\mathrm{i}}}(t,s) \\ \mathbf{\Gamma}^{\mathrm{i}}_{\mathbf{xx}^{\mathrm{i}}}(t,s) & \mathbf{\Gamma}^{\mathrm{i}}_{\mathbf{xx}^{\mathrm{k}}}(t,s) & \mathbf{\Gamma}^{\mathrm{i}}_{\mathbf{x}}(t,s) & \mathbf{\Gamma}^{\mathrm{i}}_{\mathbf{xx}^*}(t,s) \\ \mathbf{\Gamma}^{\mathrm{k}}_{\mathbf{xx}^{\mathrm{k}}}(t,s) & \mathbf{\Gamma}^{\mathrm{k}}_{\mathbf{xx}^{\mathrm{i}}}(t,s) & \mathbf{\Gamma}^{\mathrm{k}}_{\mathbf{xx}^*}(t,s) & \mathbf{\Gamma}^{\mathrm{k}}_{\mathbf{x}}(t,s) \end{pmatrix}, \quad t, s \in \mathbb{Z} \tag{4}$$

Finally, the following relationships between the augmented vector (3) and the real vector (2) can be established:

$$\bar{\mathbf{x}}(t) = 2\mathcal{T}_p \mathbf{x}_r(t) \tag{5}$$

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

$$\mathcal{B} = \begin{pmatrix} 1 & \mathrm{i} & \mathrm{j} & \mathrm{k} \\ 1 & -\mathrm{i} & \mathrm{j} & -\mathrm{k} \\ 1 & \mathrm{i} & -\mathrm{j} & -\mathrm{k} \\ 1 & -\mathrm{i} & -\mathrm{j} & \mathrm{k} \end{pmatrix}$$

and $\mathcal{T}_p^{\mathtt{H}}\mathcal{T}_p = \mathbf{I}_{4p}$.

$\mathbb{T}_1$-properness was introduced in [27]. Specifically, a random signal $\mathbf{x}(t) \in \mathbb{T}^p$ is said to be $\mathbb{T}_1$-proper if, and only if, the functions $\mathbf{\Gamma}_{\mathbf{xx}^\nu}(t,s)$, $\nu = *, \mathrm{i}, \mathrm{k}$, vanish $\forall t, s \in \mathbb{Z}$. Similarly, two random signal $\mathbf{x}(t) \in \mathbb{T}^{p_1}$ and $\mathbf{y}(t) \in \mathbb{T}^{p_2}$ are cross $\mathbb{T}_1$-proper if, and only if, the functions $\mathbf{\Gamma}_{\mathbf{xy}^\nu}(t,s)$, $\nu = *, \mathrm{i}, \mathrm{k}$, vanish $\forall t, s \in \mathbb{Z}$. Finally, $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly $\mathbb{T}_1$-proper if, and only if, are $\mathbb{T}_1$-proper and cross $\mathbb{T}_1$-proper. Also, a statistical test to check experimentally if a tessarine random vector is $\mathbb{T}_1$-proper or improper can be found in the above paper.

$\mathbb{T}_1$-properness has an important effect on estimation. A definition of metric in the tessarine domain is required for solving this problem. The distance between the random variables $x, y \in \mathbb{T}$ is defined as $d(x,y) = \|x-y\|$, where $\|x\| = (\mathcal{R}\{E[xx^*]\})^{\frac{1}{2}}$. In general, in a metric space, neither the existence nor the uniqueness of the projection of an element on a set is assured. However, the properties of the distance $d(\cdot,\cdot)$ guarantee the existence and uniqueness of the projection [27].

Let $\mathcal{G}_{\mathcal{C}_1}$ be the closed linear subspace associated to the set of random vectors $\mathcal{C}_1 = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$, with $\mathbf{x}_i \in \mathbb{T}^p$, $i = 1, \ldots, m$. The projection of a given random variable $y \in \mathbb{T}$ onto the set $\mathcal{G}_{\mathcal{C}_1}$ is called the $\mathbb{T}_1$ estimator of y respect to $\mathcal{C}_1$ and it is denoted by $\hat{y}^{\mathtt{T}_1}$. If the information set is formed by augmented vectors defined in (3), i.e., if we consider $\mathcal{C}_{\mathtt{WL}} = \{\bar{\mathbf{x}}_1, \ldots, \bar{\mathbf{x}}_m\}$ with $\bar{\mathbf{x}}_i \in \mathbb{T}^{4p}$, $i = 1, \ldots, m$, then, the projection of y onto the set $\mathcal{G}_{\mathcal{C}_{\mathtt{WL}}}$ is called TWL estimator of y respect to $\mathcal{C}_1$ and it is denoted by $\hat{y}^{\mathtt{TWL}}$. It should be noted that $\hat{y}^{\mathtt{T}_1}$ and $\hat{y}^{\mathtt{TWL}}$ are the linear MMSE estimators of y from the information supplied by $\mathcal{C}_1$ and $\mathcal{C}_{\mathtt{WL}}$, respectively. In [27], it is shown that[2]

$$\begin{aligned} \hat{y}^{\mathtt{T}_1} &= \mathbf{l}_2^{\mathtt{T}} \mathbf{z} \\ \hat{y}^{\mathtt{TWL}} &= \mathbf{l}_1^{\mathtt{T}} \bar{\mathbf{z}} \end{aligned} \tag{6}$$

with $\mathbf{z} = [\mathbf{x}_1^{\mathtt{T}}, \ldots, \mathbf{x}_m^{\mathtt{T}}]^{\mathtt{T}}$, $\bar{\mathbf{z}} = [\mathbf{x}_1^{\mathtt{T}}, \ldots, \mathbf{x}_m^{\mathtt{T}}, \mathbf{x}_1^{\mathtt{H}}, \ldots, \mathbf{x}_m^{\mathtt{H}}, \mathbf{x}_1^{\mathtt{iT}}, \ldots, \mathbf{x}_m^{\mathtt{iT}}, \mathbf{x}_1^{\mathtt{kT}}, \ldots, \mathbf{x}_m^{\mathtt{kT}}]^{\mathtt{T}}$ and the deterministic vectors $\mathbf{l}_i$, $i = 1, 2$, are computed through the equations $\mathbf{l}_1^{\mathtt{T}} = \mathbf{\Gamma}_{y\bar{\mathbf{z}}} \mathbf{\Gamma}_{\bar{\mathbf{z}}}^{-1}$ and $\mathbf{l}_2^{\mathtt{T}} = \mathbf{\Gamma}_{y\mathbf{z}} \mathbf{\Gamma}_{\mathbf{z}}^{-1}$. Moreover, TWL processing is the most suitable in terms of performance whenever the signal is improper. In fact, since $\mathcal{G}_{\mathcal{C}_1} \subseteq \mathcal{G}_{\mathcal{C}_{\mathtt{WL}}}$ then $\epsilon^{\mathtt{TWL}} \leq \epsilon^{\mathtt{T}_1}$, where

$$\begin{aligned} \epsilon^{\mathtt{TWL}} &= \|y - \hat{y}^{\mathtt{TWL}}\|^2 \\ \epsilon^{\mathtt{T}_1} &= \|y - \hat{y}^{\mathtt{T}_1}\|^2 \end{aligned} \tag{7}$$

and thus, in general, the TWL estimation error is lower than the $\mathbb{T}_1$ estimation error. However, the computation of the $\mathbb{T}_1$ estimator entails a lower computational cost than the one associated with the TWL estimator. Fortunately, under $\mathbb{T}_1$-properness conditions, both TWL and $\mathbb{T}_1$ estimators coincide and hence, the $\mathbb{T}_1$ estimator has minimum error with a lower computational cost in this case.

[2]Note that $\hat{y}^{\mathtt{T}_1}$ is denoted by $\hat{y}^{\mathtt{TSL}}$ in [27].

3. $\mathbb{T}_2$-Properness

This section is devoted to introduce a new concept of tessarine properness, called $\mathbb{T}_2$-properness, and the implications of the $\mathbb{T}_2$ processing are also studied. Similar to the $\mathbb{C}$-properness property given in [44] for quaternions, we consider the following notion of properness.

Definition 1. *A random signal $\mathbf{x}(t) \in \mathbb{T}^p$ is said to be $\mathbb{T}_2$-proper if, and only if, the functions $\mathbf{\Gamma}_{\mathbf{x}\mathbf{x}^\nu}(t,s)$, $\nu = \mathrm{i}, \mathrm{k}$, vanish $\forall t, s \in \mathbb{Z}$. In like manner, two random signals $\mathbf{x}(t) \in \mathbb{T}^{p_1}$ and $\mathbf{y}(t) \in \mathbb{T}^{p_2}$ are cross $\mathbb{T}_2$-proper if, and only if, the functions $\mathbf{\Gamma}_{\mathbf{x}\mathbf{y}^\nu}(t,s)$, $\nu = \mathrm{i}, \mathrm{k}$, vanish $\forall t, s \in \mathbb{Z}$. Finally, $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly $\mathbb{T}_2$-proper if, and only if, are $\mathbb{T}_2$-proper and cross $\mathbb{T}_2$-proper.*

Notice that $\mathbb{T}_2$-properness is a less restrictive property than $\mathbb{T}_1$-properness. From Property 2 in [27] we have the following result.

Proposition 1. *A random signal $\mathbf{x}(t) \in \mathbb{T}^p$ is $\mathbb{T}_2$-proper if, and only if, the following relations hold:*

$$\begin{aligned} \mathbf{\Gamma}_{\mathbf{a}}(t,s) &= \mathbf{\Gamma}_{\mathbf{c}}(t,s), & \mathbf{\Gamma}_{\mathbf{b}}(t,s) &= \mathbf{\Gamma}_{\mathbf{d}}(t,s) \\ \mathbf{\Gamma}_{\mathbf{ab}}(t,s) &= \mathbf{\Gamma}_{\mathbf{cd}}(t,s), & \mathbf{\Gamma}_{\mathbf{ba}}(t,s) &= \mathbf{\Gamma}_{\mathbf{dc}}(t,s) \\ \mathbf{\Gamma}_{\mathbf{ac}}(t,s) &= \mathbf{\Gamma}_{\mathbf{ca}}(t,s), & \mathbf{\Gamma}_{\mathbf{bd}}(t,s) &= \mathbf{\Gamma}_{\mathbf{db}}(t,s) \\ \mathbf{\Gamma}_{\mathbf{ad}}(t,s) &= \mathbf{\Gamma}_{\mathbf{cb}}(t,s), & \mathbf{\Gamma}_{\mathbf{bc}}(t,s) &= \mathbf{\Gamma}_{\mathbf{da}}(t,s) \end{aligned} \tag{8}$$

As noted above, TWL processing is generally the most suitable since the estimator $\hat{y}^{\mathtt{TWL}}$ makes a full use of the second-order statistical information contained in the set $\mathcal{C}_1$. Also, we have seen that when the signals present $\mathbb{T}_1$-properness conditions then, $\hat{y}^{\mathtt{TWL}} = \hat{y}^{\mathtt{T}_1}$ and thus, the most efficient estimator is $\hat{y}^{\mathtt{T}_1}$. Now, we introduce a new estimator based on $\mathbb{T}_2$ processing which, under $\mathbb{T}_2$-properness conditions, has similar advantages to those obtained for $\hat{y}^{\mathtt{T}_1}$ under $\mathbb{T}_1$-properness conditions.

Definition 2. *Consider a random variable $y \in \mathbb{T}$ and the set of random vectors $\mathcal{C}_1 = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$, with $\mathbf{x}_i \in \mathbb{T}^p$, $i = 1, \dots, m$. The $\mathbb{T}_2$ estimator of*

y respect to $\mathcal{C}_1$ is defined as the projection of y onto the set $\mathcal{G}_{\mathcal{C}_2}$ with $\mathcal{C}_2 = \{\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{x}_1^*, \dots, \mathbf{x}_m^*\}$ and it is denoted by $\hat{y}^{T_2}$.

The concept of $\mathbb{T}_i$ estimator, $i = 1, 2$, is easily extended to the vectorial case. Specifically, the $\mathbb{T}_i$ estimator of the random vector $\mathbf{y} \in \mathbb{T}^p$ respect to $\mathcal{C}_1$ is $\hat{\mathbf{y}}^{T_i} = [\hat{y}_1^{T_i}, \dots, \hat{y}_p^{T_i}]^T$, where $\hat{y}_j^{T_i}$ is the projection of the component jth of $\mathbf{y}$ onto the set $\mathcal{G}_{\mathcal{C}_i}$, $i = 1, 2$.

Theorem 1. *The following statements hold:*

1. *The $\mathbb{T}_2$ estimator can be calculated as*

$$\hat{y}^{T_2} = \sum_{i=1}^{m} \left(\mathbf{k}_{1i}^T \mathbf{x}_i + \mathbf{k}_{2i}^T \mathbf{x}_i^* \right)$$

   where the deterministic vectors $\mathbf{k}_{ji} \in \mathbb{T}^p$ are computed through the equation

$$\left[\mathbf{k}_{11}^T, \dots, \mathbf{k}_{1m}^T, \mathbf{k}_{21}^T, \dots, \mathbf{k}_{2m}^T\right] = \mathbf{\Gamma}_{y\tilde{\mathbf{z}}} \mathbf{\Gamma}_{\tilde{\mathbf{z}}}^{-1}$$

   with $\tilde{\mathbf{z}} = [\mathbf{x}_1^T, \dots, \mathbf{x}_m^T, \mathbf{x}_1^H, \dots, \mathbf{x}_m^H]^T$.

2. *$\epsilon^{TWL} \leq \epsilon^{T_2} \leq \epsilon^{T_1}$, where ϵ^{TWL} and ϵ^{T_1} are given in (7), and $\epsilon^{T_2} = \|y - \hat{y}^{T_2}\|^2$.*
3. *If $\mathbf{x}_1, \dots, \mathbf{x}_m$ are jointly $\mathbb{T}_2$-proper and y is cross $\mathbb{T}_2$-proper with each element of $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ then, $\hat{y}^{TWL} = \hat{y}^{T_2}$, with $\hat{y}^{TWL}$ given in (6).*

As a consequence, under $\mathbb{T}_2$-properness conditions, the $\mathbb{T}_2$ estimator has the same performance as the TWL estimator but with a lower computational complexity.

Remark 1. *A statistical test to determine whether a random vector $\mathbf{x} \in \mathbb{T}^p$ presents properness properties from the information supplied by a random sample can be provided following a similar reasoning as in [27]. Specifically, consider the following statistical hypothesis test:*

$$\begin{aligned} &H_0 : \mathbf{x} \text{ is } \mathbb{T}_2\text{-proper (i.e., } \mathbf{\Gamma}_{\mathbf{x}\mathbf{x}^{\mathrm{i}}} = \mathbf{\Gamma}_{\mathbf{x}\mathbf{x}^{\mathrm{k}}} = \mathbf{0}_{p\times p}) \\ &H_1 : \mathbf{x} \text{ is not } \mathbb{T}_2\text{-proper} \end{aligned} \tag{9}$$

and given n independent and identically distributed random samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ of a random vector $\mathbf{x} \in \mathbb{T}^p$ such that $\mathbf{x}_r$ follows a Gaussian distribution, then *the GLRT statistic for* (9) *is given by*

$$\phi_1(\mathbf{x}_1,\ldots,\mathbf{x}_n) = -n[\ln|\hat{\boldsymbol{\Gamma}}_{\bar{\mathbf{x}}}| - \ln|\boldsymbol{\Gamma}_{\mathbb{T}_2}|]$$

where $\hat{\boldsymbol{\Gamma}}_{\bar{\mathbf{x}}} = 4\mathcal{T}_p\hat{\boldsymbol{\Gamma}}_{\mathbf{x}_r}\mathcal{T}_p^H$ *with* $\hat{\boldsymbol{\Gamma}}_{\mathbf{x}_r}$ *the sample autocorrelation matrix and*

$$\boldsymbol{\Gamma}_{\mathbb{T}_2} = \begin{pmatrix} \hat{\boldsymbol{\Gamma}}_{\mathbf{x}} & \hat{\boldsymbol{\Gamma}}_{\mathbf{x}\mathbf{x}^*} & \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} \\ \hat{\boldsymbol{\Gamma}}^*_{\mathbf{x}\mathbf{x}^*} & \hat{\boldsymbol{\Gamma}}^*_{\mathbf{x}} & \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} & \hat{\boldsymbol{\Gamma}}^{\mathrm{i}}_{\mathbf{x}} & \hat{\boldsymbol{\Gamma}}^{\mathrm{i}}_{\mathbf{x}\mathbf{x}^*} \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} & \hat{\boldsymbol{\Gamma}}^{\mathrm{k}}_{\mathbf{x}\mathbf{x}^*} & \hat{\boldsymbol{\Gamma}}^{\mathrm{k}}_{\mathbf{x}} \end{pmatrix}$$

Also, assuming that H_0 *is true, the distribution of the statistic* $\phi_1(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ *tends towards a chi-squared distribution with degrees of freedom equal to* $4p^2$ *as the sample size tends to infinity.*

4. $\mathbb{T}_i$-proper Wide-Sense Markov Signals

This section first introduces the $\mathbb{T}_i$WSM signals and studies their properties. Next, two estimation algorithms for this kind of signals which are valid when a phenomenon of intermittent observations is present are devised. These algorithms solve the prediction, filtering and fixed-interval smoothing problems.

Consider a random signal $\{\mathbf{x}(t) \in \mathbb{T}^p, t \in \mathbb{Z}\}$. We define the following forwards vectors:

$$\boldsymbol{\xi}_1(t) = [\mathbf{x}^{\mathtt{T}}(t), \mathbf{x}^{\mathtt{T}}(t-1),\ldots,\mathbf{x}^{\mathtt{T}}(t-n+1)]^{\mathtt{T}}$$
$$\boldsymbol{\xi}_2(t) = [\mathbf{x}^{\mathtt{T}}(t), \mathbf{x}^{\mathtt{H}}(t), \mathbf{x}^{\mathtt{T}}(t-1), \mathbf{x}^{\mathtt{H}}(t-1),\ldots,\mathbf{x}^{\mathtt{T}}(t-n+1), \mathbf{x}^{\mathtt{H}}(t-n+1)]^{\mathtt{T}}$$

and assume that $\det\{\boldsymbol{\Gamma}_{\boldsymbol{\xi}_i}(t)\} \neq 0$, $\forall t$, $i = 1,2$.

The $\mathbb{T}_i$ estimators of $\mathbf{x}(t)$ and $\boldsymbol{\xi}_i(t)$ obtained from the information supplied by the set $\{\mathbf{x}(t_1),\ldots,\mathbf{x}(t_n)\}$ will be denoted by $\hat{\mathbf{x}}^{\mathtt{T}_i}(t|t_1,\ldots,t_n)$ and $\hat{\boldsymbol{\xi}}_i^{\mathtt{T}_i}(t|t_1,\ldots,t_n)$, respectively, $i = 1,2$. Analogously, $\hat{\mathbf{x}}^{\mathtt{T}_i}(t|\tau \leq s)$ and $\hat{\boldsymbol{\xi}}_i^{\mathtt{T}_i}(t|\tau \leq s)$ denote the $\mathbb{T}_i$ estimators of $\mathbf{x}(t)$ and $\boldsymbol{\xi}_i(t)$ based on $\{\mathbf{x}(\tau), \tau \leq s\}$, respectively.

Definition 3. *A random signal* $\mathbf{x}(t) \in \mathbb{T}^p$ *is said to be* $\mathbb{T}_i$WSM *of order* $n \geq 1$, *briefly a* $\mathbb{T}_i$WSM(n) *signal, if it is* $\mathbb{T}_i$*-proper and the following condition holds:*

$$\hat{\mathbf{x}}^{T_i}(t|\tau \leq s) = \hat{\mathbf{x}}^{T_i}(t|s, s-1,\ldots,s-n+1), \quad \forall s \leq t, \quad i = 1,2 \qquad (10)$$

In the following result we give a characterization for $\mathbb{T}_i$WSM(n) signals.

Proposition 2. *Let* $\mathbf{x}(t) \in \mathbb{T}^p$ *be a* $\mathbb{T}_i$*-proper signal and consider the function*

$$\mathbf{K}_i(t,s) = \mathbf{\Gamma}_{\boldsymbol{\xi}_i}(t,s)\mathbf{\Gamma}_{\boldsymbol{\xi}_i}^{-1}(s) \tag{11}$$

Then, $\mathbf{x}(t)$ *is* $\mathbb{T}_i$WSM(n) *if, and only if,*

$$\mathbf{K}_i(t,s) = \mathbf{K}_i(t,\upsilon)\mathbf{K}_i(\upsilon,s), \qquad t \geq \upsilon \geq s \tag{12}$$

for $i = 1, 2$.

From now on, we denote $\mathbf{K}_i(t) = \mathbf{K}_i(t+1,t)$. Next, we provide a way for modeling $\mathbb{T}_i$WSM(n) signals based on the forwards vector.

Theorem 2. *A random signal* $\{\mathbf{x}(t) \in \mathbb{T}^p, 0 \leq t \leq m\}$ *is* $\mathbb{T}_i$WSM(n) *if, and only if,* $\mathbf{x}(t)$ *has forwards representation*

$$\mathbf{x}(t+1) = \mathbf{K}_i^{(p)}(t)\boldsymbol{\xi}_i(t) + \boldsymbol{\omega}_i(t), \quad t \geq n-1 \tag{13}$$

where $\boldsymbol{\omega}_i(t)$ *is a white noise such that for* $t \geq n-1$,

$$\begin{aligned} \mathbf{\Gamma}_{\boldsymbol{\omega}_i\boldsymbol{\xi}_i}(t,n-1) &= \mathbf{0}_{p\times inp} \\ \mathbf{\Gamma}_{\boldsymbol{\omega}_i}(t) &= \mathbf{\Gamma}_{\mathbf{x}}(t+1) - \mathbf{K}_i^{(p)}(t)\mathbf{\Gamma}_{\boldsymbol{\xi}_i}(t)\mathbf{K}_i^{(p)}(t)^H \end{aligned} \tag{14}$$

and being $\boldsymbol{\omega}_i(t)$ *and* $\boldsymbol{\varepsilon} = [\mathbf{x}^T(0), \mathbf{x}^T(1), \ldots, \mathbf{x}^T(n-1)]^T$ *jointly* $\mathbb{T}_i$*-proper for* $i = 1, 2$.

4.1. Estimation of $\mathbb{T}_i$WSM(n) *Signals*

We focus our attention on the problem of estimating a $\mathbb{T}_i$WSM(n) signal. The $\mathbb{T}_i$-properness condition together with the forwards representation (13) make the design of computationally efficient estimation algorithms feasible. Specifically, we provide the solution of the prediction, filtering and fixed-interval smoothing problems with intermittent observations. For that, we need some previous definitions.

Definition 4. *Given two tessarines* $x = a_1 + \mathrm{i}b_1 + \mathrm{j}c_1 + \mathrm{k}d_1$ *and* $y = a_2 + \mathrm{i}b_2 + \mathrm{j}c_2 + \mathrm{k}d_2$*, the* $\star$ *product between them is defined as*

$$x \star y = a_1a_2 + \mathrm{i}b_1b_2 + \mathrm{j}c_1c_2 + \mathrm{k}d_1d_2$$

Likewise, for two tessarine vectors $\mathbf{x} = [x_1, \ldots, x_p]^T$ *and* $\mathbf{y} = [y_1, \ldots, y_p]^T$ *the* $\star$ *product is defined as*

$$\mathbf{x} \star \mathbf{y} = [x_1 \star y_1, \ldots, x_p \star y_p]^T$$

Consider a $\mathbb{T}_i$WSM(n) signal $\{\mathbf{x}(t) \in \mathbb{T}^p, 0 \leq t \leq m\}$ with components given by $x_i(t) = x_{i1}(t) + \mathrm{i}x_{i2}(t) + \mathrm{j}x_{i3}(t) + \mathrm{k}x_{i4}(t)$, $i = 1, \ldots, p$. Suppose that $\mathbf{x}(t)$ cannot be observed directly but it is obtained from the following observation equation

$$\mathbf{y}(t) = \boldsymbol{\lambda}(t) \star \mathbf{x}(t) + \mathbf{v}(t), \qquad 0 \leq t \leq m \tag{15}$$

with $\mathbf{v}(t) \in \mathbb{T}^p$ a white noise and $\boldsymbol{\lambda}(t) \in \mathbb{T}^p$ a vector of tessarine Bernoulli random variables. The components of $\boldsymbol{\lambda}(t)$ are of the form $\lambda_i(t) = \lambda_{i1}(t) + \mathrm{i}\lambda_{i2}(t) + \mathrm{j}\lambda_{i3}(t) + \mathrm{k}\lambda_{i4}(t)$, being $\lambda_{ij}(t)$ independent Bernoulli random variables with parameter $\rho_{ij}(t)$ indicating the presence or absence of each state component $x_{ij}(t)$ in the observation. Moreover, $\boldsymbol{\lambda}(t)$, $\mathbf{x}(t)$ and $\mathbf{v}(t)$ are independent.

Notice that the observation equation (15) generalizes the measurements equation in the classical Kalman filter since we are assuming that, at each instant of time t there is a probability $\rho_{ij}(t)$ that the component $x_{ij}(t)$ is present or not in the observation. More specifically, as $\rho_{ij}(t)$ is closer to 1, there is a greater probability that the observation at time instant t, $\mathbf{y}(t)$, contains the signal $x_{ij}(t)$ and conversely, as $\rho_{ij}(t)$ is closer to 0.

With the purpose of a $\mathbb{T}_i$ processing we state the following result.

Proposition 3. *Let* $\{\mathbf{x}(t) \in \mathbb{T}^p, 0 \leq t \leq m\}$ *be a* $\mathbb{T}_i$WSM(n) *signal which is observed through the observation process* $\mathbf{y}(t)$ *given by* (15).

1. $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *are jointly* $\mathbb{T}_1$*-proper if, and only if,* $\mathbf{v}(t)$ *is* $\mathbb{T}_1$*-proper and* $\rho_{ij}(t) = \rho_i(t)$, $\forall t, i$ *and* $j = 1, 2, 3, 4$.
2. $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *are jointly* $\mathbb{T}_2$*-proper if, and only if,* $\mathbf{v}(t)$ *is* $\mathbb{T}_2$*-proper and* $\rho_{i1}(t) = \rho_{i3}(t)$ *and* $\rho_{i2}(t) = \rho_{i4}(t)$, $\forall t, i$.

4.1.1. Prediction and Filtering Problem

Our first aim is to devise the one-stage predictor and the filter of $\mathbf{x}(t)$ on the basis of the set of observations $\mathcal{D} = \{\mathbf{y}(0), \ldots, \mathbf{y}(t)\}$. We assume that $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly $\mathbb{T}_i$-proper. Recall that, under $\mathbb{T}_i$-properness conditions, the $\mathbb{T}_i$ processing is the most appropriate approach due to the computational saving involved (see Sections 2.2 and 3).

For simplicity, the $\mathbb{T}_i$ filter of $\mathbf{x}(t)$ respect to $\mathcal{D}$ is denoted by $\hat{\mathbf{x}}^{\mathtt{T}_i}(t|t) = [\hat{x}_1^{\mathtt{T}_i}(t|t), \ldots, \hat{x}_p^{\mathtt{T}_i}(t|t)]^{\mathtt{T}}$, where $\hat{x}_j^{\mathtt{T}_i}(t|t)$ is the $\mathbb{T}_i$ estimator of $x_j(t)$ obtained from the information supplied by $\mathcal{D}$. Similarly, the $\mathbb{T}_i$ one-stage predictor is denoted by $\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|t)$.

Now, we introduce some matrices and auxiliary stochastic processes which are necessary for posterior developments. Specifically, consider the following processes:

$$\mathbf{r}_i(t) = \mathbf{H}_i(t)\boldsymbol{\xi}_i(t) + \mathbf{u}_i(t), \qquad i = 1, 2 \tag{16}$$

with $\mathbf{u}_i(t)$ a white noise independent of $\boldsymbol{\xi}_i(t)$ such that $\boldsymbol{\Gamma}_{\mathbf{u}_i}(t) = \mathbf{R}_i(t) + \boldsymbol{\Sigma}_i(t)$, $i = 1, 2$, and where

$$\mathbf{H}_1(t) = [\mathrm{diag}\{\boldsymbol{\rho}(t)\}, \mathbf{0}_{p\times p(n-1)}], \quad \mathbf{R}_1(t) = \boldsymbol{\Gamma}_{\mathbf{v}}(t), \quad \boldsymbol{\Sigma}_1(t) = \mathrm{diag}\{\boldsymbol{\theta}(t)\}$$

with $\mathbf{v}(t)$ defined in (15), $\boldsymbol{\rho}(t) = [\rho_1(t), \ldots, \rho_p(t)]^{\mathtt{T}}$ with $\rho_i(t)$ given in Proposition 3 and $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_p(t)]^{\mathtt{T}}$ with $\theta_i(t) = 4\rho_i(t)(1 - \rho_i(t))E[x_{i1}^2(t)]$, $i = 1, \ldots, p$. Also, under $\mathbb{T}_1$-properness, we denote $\mathbf{z}_1(t) = \mathbf{y}(t)$.

Alternatively, for the case of $\mathbb{T}_2$-properness, we have to consider the following matrices in (16):

$$\mathbf{H}_2(t) = [\mathbf{S}(t), \mathbf{0}_{2p\times 2p(n-1)}], \quad \mathbf{R}_2(t) = \boldsymbol{\Gamma}_{\boldsymbol{v}}(t), \quad \boldsymbol{\Sigma}_2(t) = [\sigma_{ij}(t)]$$

where the elements of $\mathbf{S}(t)$ are

$$\begin{aligned} s_{ii}(t) = s_{i+p,i+p}(t) &= \frac{1}{2}(\rho_{i1}(t) + \rho_{i2}(t)), \quad i = 1, \ldots, p \\ s_{i,i+p}(t) = s_{i+p,i}(t) &= \frac{1}{2}(\rho_{i1}(t) - \rho_{i2}(t)), \quad i = 1, \ldots, p \\ s_{ij} &= 0 \quad \text{for the rest} \end{aligned}$$

$\boldsymbol{\upsilon}(t) = [\mathbf{v}^{\mathtt{T}}(t), \mathbf{v}^{\mathtt{H}}(t)]^{\mathtt{T}}$, and the elements of $\boldsymbol{\Sigma}_2(t)$ given by

$$\sigma_{ii}(t) = \sigma_{i+p,i+p}(t) = \phi_{i1} + \phi_{i2}, \quad i = 1, \ldots, p$$

$$\sigma_{i,i+p}(t) = \sigma_{i+p,i}(t) = \phi_{i1} - \phi_{i2}, \quad i = 1, \ldots, p$$

$$\sigma_{ij} = 0 \quad \text{for the rest}$$

with $\phi_{ij}(t) = 2\rho_{ij}(t)(1 - \rho_{ij}(t))E[x_{ij}^2(t)]$, $i = 1, \ldots, p$, $j = 1, 2$. Finally, in this case we consider $\mathbf{z}_2(t) = [\mathbf{y}^{\mathtt{T}}(t), \mathbf{y}^{\mathtt{H}}(t)]^{\mathtt{T}}$.

It should be noted that for the purposes of estimation working with $\{\mathbf{x}(t), \mathbf{z}_i(t)\}$ is equivalent to working with $\{\boldsymbol{\xi}_i(t), \mathbf{z}_i(t)\}$. Thus, the next result becomes a key tool to solve our estimation problem.

Proposition 4. *The following relations hold:*

$$\begin{aligned} \boldsymbol{\Gamma}_{\mathbf{z}_i\boldsymbol{\xi}_i}(t,s) &= \boldsymbol{\Gamma}_{\mathbf{r}_i\boldsymbol{\xi}_i}(t,s) \\ \boldsymbol{\Gamma}_{\mathbf{z}_i}(t,s) &= \boldsymbol{\Gamma}_{\mathbf{r}_i}(t,s) \end{aligned} \tag{17}$$

By combining the forwards representation (13), the previous Proposition and the classical Kalman filter we suggest the following procedure to provide $\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|t)$ and $\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|t+1)$ and their associated errors in an efficient way.

i) Derive the Kalman filter recursions for the state-space model given by (27) and (16).

ii) Since, from Proposition 4, the second-order properties of $\mathbf{r}_i(t)$ and $\mathbf{z}_i(t)$ are equal then, replace $\mathbf{r}_i(t)$ by $\mathbf{z}_i(t)$ in such equations. The algorithm should be initialized by using the $\mathbb{T}_i$ estimate of $\boldsymbol{\xi}_i(n-1)$ respect to the set $\{\mathbf{y}(0), \ldots, \mathbf{y}(n-1)\}$, denoted by $\mathbf{n}_{2i}(0)$, and its associated error, $\mathbf{P}_{2i}(0)$. This step provides the one-stage predictor $\mathbf{n}_{1i}(t)$ and the filter $\mathbf{n}_{2i}(t)$ and their error covariances $\mathbf{P}_{1i}(t)$ and $\mathbf{P}_{2i}(t)$, respectively.

iii) The desired estimators, $\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|t)$ and $\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|t+1)$, are obtained by extracting the p first components of the optimal estimators, $\mathbf{n}_{1i}(t)$ and $\mathbf{n}_{2i}(t)$, derived in the previous step. Similarly, their estimation errors, $\boldsymbol{\epsilon}^{\mathtt{T}_i}(t+1|t)$ and $\boldsymbol{\epsilon}^{\mathtt{T}_i}(t+1)$, whose components are denoted by

$$\begin{aligned} \epsilon_j^{\mathtt{T}_i}(t+1|t) &= \|x_j(t+1) - \hat{x}_j^{\mathtt{T}_i}(t+1|t)\|^2, \qquad j = 1, \ldots, p \\ \epsilon_j^{\mathtt{T}_i}(t+1) &= \|x_j(t+1) - \hat{x}_j^{\mathtt{T}_i}(t+1|t+1)\|^2, \qquad j = 1, \ldots, p, \end{aligned} \tag{18}$$

are derived from the error matrices, $\mathbf{P}_{1i}(t)$ and $\mathbf{P}_{2i}(t)$, calculated above.

Also, this procedure is described in algorithmic form in Algorithm 1.

Algorithm 1 $\mathbb{T}_i$ one-stage predictor and filter

Require: $\{\mathbf{K}_i(t), \mathbf{\Gamma}_{\mathbf{w}_i}(t)\}_{t=n-1}^{k}$, $\{\mathbf{z}_i(t), \mathbf{H}_i(t), \mathbf{R}_i(t), \mathbf{\Sigma}_i(t)\}_{t=n}^{k+1}$, $\mathbf{n}_{2i}(0)$, $\mathbf{P}_{2i}(0)$

Ensure: $\hat{\mathbf{x}}^{\mathbb{T}_i}(k+1|k)$, $\hat{\mathbf{x}}^{\mathbb{T}_i}(k+1|k+1)$, $\boldsymbol{\epsilon}^{\mathbb{T}_i}(k+1|k)$, $\boldsymbol{\epsilon}^{\mathbb{T}_i}(k+1)$

for $t = 0$ **to** $k-n+1$ **do**

  $\mathbf{n}_{1i}(t+1) \leftarrow \mathbf{K}_i(t+n-1)\mathbf{n}_{2i}(t)$

  $\mathbf{P}_{1i}(t+1) \leftarrow \mathbf{K}_i(t+n-1)\mathbf{P}_{2i}(t)\mathbf{K}_i^{\mathrm{H}}(t+n-1) + \mathbf{\Gamma}_{\mathbf{w}_i}(t+n-1)$

  $\mathbf{F}_i(t+1) \leftarrow \mathbf{P}_{1i}(t+1)\mathbf{H}_i^{\mathrm{H}}(t+n)[\mathbf{H}_i(t+n)\mathbf{P}_{1i}(t+1)\mathbf{H}_i^{\mathrm{H}}(t+n) + \mathbf{R}_i(t+n) + \mathbf{\Sigma}_i(t+n)]^{-1}$

  $\mathbf{n}_{2i}(t+1) \leftarrow \mathbf{n}_{1i}(t+1) + \mathbf{F}_i(t+1)[\mathbf{z}_i(t+n) - \mathbf{H}_i(t+n)\mathbf{n}_{1i}(t+1)]$

  $\mathbf{P}_{2i}(t+1) \leftarrow \mathbf{P}_{1i}(t+1) - \mathbf{F}_i(t+1)\mathbf{H}_i(t+n)\mathbf{P}_{1i}(t+1)$

end for

$\hat{\mathbf{x}}^{\mathbb{T}_i}(k+1|k) \leftarrow \mathbf{n}_{1i}^{(p)}(k-n+2)$

$\hat{\mathbf{x}}^{\mathbb{T}_i}(k+1|k+1) \leftarrow \mathbf{n}_{2i}^{(p)}(k-n+2)$

$\boldsymbol{\epsilon}^{\mathbb{T}_i}(k+1|k) \leftarrow \mathcal{R}\{\mathrm{diag}\{\mathbf{P}_{1i}^{(p,p)}(k-n+2)\}\}$

$\boldsymbol{\epsilon}^{\mathbb{T}_i}(k+1) \leftarrow \mathcal{R}\{\mathrm{diag}\{\mathbf{P}_{2i}^{(p,p)}(k-n+2)\}\}$

Remark 2. *Theorem 5 in [27] is a particular case of this algorithm for a* $\mathbb{T}_1$WSM(1) *signal with the observation process of the form* $\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{v}(t)$.

4.1.2. Fixed-Interval Smoothing Problem

Experimental data is often noisy and available only over a fixed time interval. In this section, we consider the smoothing of such data with the additional characteristic that the observations can be purely noise. Specifically, we wish to compute the $\mathbb{T}_i$ smoothed estimator of $\mathbf{x}(t)$ based on uncertain future data, i.e., we want to estimate $\mathbf{x}(t)$ based on the information supplied by the set $\mathcal{D}_1 = \{\mathbf{y}(0), \ldots, \mathbf{y}(m_1)\}$ with $t < m_1 < m$ and where the observation process is given by (15). Similarly to the previous section, we assume that $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly $\mathbb{T}_i$-proper. For simplicity, the $\mathbb{T}_i$ smoothed estimator of $\mathbf{x}(t)$ respect to $\mathcal{D}_1$ is denoted by $\hat{\mathbf{x}}^{\mathrm{T}_i}(t|m_1) = [\hat{x}_1^{\mathrm{T}_i}(t|m_1), \ldots, \hat{x}_p^{\mathrm{T}_i}(t|m_1)]^{\mathrm{T}}$, where $\hat{x}_j^{\mathrm{T}_i}(t|m_1)$ is the $\mathbb{T}_i$ smoothed estimator of $x_j(t)$ based on $\mathcal{D}_1$.

The following procedure achieves $\hat{\mathbf{x}}^{\mathrm{T}_i}(t|m_1)$ in a computationally efficient way. Such a procedure makes use of forwards representation (13), Proposition 4 and the Rauch-Tung-Striebel (RTS) formulas [1].

i) The first step coincides with the typical forward pass in the RTS algorithm, i.e., the filter and predictor, $\mathbf{n}_{1i}(t)$ and $\mathbf{n}_{2i}(t)$, and their error matrices, $\mathbf{P}_{1i}(t)$ and $\mathbf{P}_{2i}(t)$, calculated in Algorithm 1 are saved for use in the backwards pass.

ii) In the backwards pass, we compute the smoothed state estimator $\mathbf{n}_{3i}(t)$ and the associated error covariance $\mathbf{P}_{3i}(t)$.

iii) The smoothed estimator $\hat{\mathbf{x}}^{\mathrm{T}_i}(t|m_1)$ is available as a subvector of the smoothed state estimator provided in the previous step. Also, the error associated with $\hat{\mathbf{x}}^{\mathrm{T}_i}(t|m_1)$, denoted by $\boldsymbol{\epsilon}^{\mathrm{T}_i}(t|m_1)$ and with components defined as

$$\epsilon_j^{\mathrm{T}_i}(t|m_1) = \|x_j(t) - \hat{x}_j^{\mathrm{T}_i}(t|m_1)\|^2, \qquad j = 1, \ldots, p \tag{19}$$

can be extracted from the error covariance matrices computed in step *ii)*.

The procedure is also presented in algorithmic form in Algorithm 2.

4.1.3. Example 1

With the aim of assessing the effectiveness of the Algorithms 1 and 2, we consider a simulation example where the behavior of the proposed $\mathbb{T}_i$ estimators is compared with their counterparts in the quaternion domain. We show that the $\mathbb{T}_i$ processing is the most adequate approach to solve the estimation problem when $\mathbb{T}_i$-properness conditions are present.

Consider a tessarine random signal $x(t)$ with forwards representation

$$x(t+1) = 0.5x(t) + fx(t-1) + \omega(t), \quad 1 \leq t \leq 100 \tag{20}$$

where f is a real parameter, $\omega(t)$ is a tessarine white noise whose real vector

Algorithm 2 $\mathbb{T}_i$ fixed-interval smoother

Require: $\{\mathbf{K}_i(t)\}_{t=k+n-1}^{m_1-1}$, $\{\mathbf{n}_{1i}(t), \mathbf{P}_{1i}(t)\}_{t=k+1}^{m_1-n+1}$, $\{\mathbf{n}_{2i}(t), \mathbf{P}_{2i}(t)\}_{t=k}^{m_1-n+1}$

Ensure: $\hat{\mathbf{x}}^{\mathtt{T}_i}(k|m_1)$, $\boldsymbol{\epsilon}^{\mathtt{T}_i}(k|m_1)$

$\mathbf{n}_{3i}(m_1-n+1) \leftarrow \mathbf{n}_{2i}(m_1-n+1)$

$\mathbf{P}_{3i}(m_1-n+1) \leftarrow \mathbf{P}_{2i}(m_1-n+1)$

if $k > n-1$ **then**

  for $t = m_1 - n$ **to** $k-n+1$ **do**

    $\mathbf{J}_i(t) \leftarrow \mathbf{P}_{2i}(t)\mathbf{K}_i^{\mathtt{H}}(t+n-1)\mathbf{P}_{1i}^{-1}(t+1)$

    $\mathbf{n}_{3i}(t) \leftarrow \mathbf{n}_{2i}(t) + \mathbf{J}_i(t)\left[\mathbf{n}_{3i}(t+1) - \mathbf{n}_{1i}(t+1)\right]$

    $\mathbf{P}_{3i}(t) \leftarrow \mathbf{P}_{2i}(t) + \mathbf{J}_i(t)\left[\mathbf{P}_{3i}(t+1) - \mathbf{P}_{1i}(t+1)\right]\mathbf{J}_i^{\mathtt{H}}(t)$

  end for

  $\hat{\mathbf{x}}^{\mathtt{T}_i}(k|m_1) \leftarrow \mathbf{n}_{3i}^{(p)}(k-n+1)$

  $\boldsymbol{\epsilon}^{\mathtt{T}_i}(k|m_1) \leftarrow \mathcal{R}\{\mathrm{diag}\{\mathbf{P}_{3i}^{(p,p)}(k-n+1)\}\}$

else

  for $t = m_1 - n$ **to** 0 **do**

    $\mathbf{J}_i(t) \leftarrow \mathbf{P}_{2i}(t)\mathbf{K}_i^{\mathtt{H}}(t+n-1)\mathbf{P}_{1i}^{-1}(t+1)$

    $\mathbf{n}_{3i}(t) \leftarrow \mathbf{n}_{2i}(t) + \mathbf{J}_i(t)\left[\mathbf{n}_{3i}(t+1) - \mathbf{n}_{1i}(t+1)\right]$

    $\mathbf{P}_{3i}(t) \leftarrow \mathbf{P}_{2i}(t) + \mathbf{J}_i(t)\left[\mathbf{P}_{3i}(t+1) - \mathbf{P}_{1i}(t+1)\right]\mathbf{J}_i^{\mathtt{H}}(t)$

  end for

  $\mathbf{B}_i(k) \leftarrow [\mathbf{0}_{p\times pi(n-1-k)}, \mathbf{I}_p, \mathbf{0}_{p\times pi(k+\frac{i-1}{2})}]$

  $\hat{\mathbf{x}}^{\mathtt{T}_i}(k|m_1) \leftarrow \mathbf{B}_i(k)\mathbf{n}_{3i}(0)$

  $\boldsymbol{\epsilon}^{\mathtt{T}_i}(k|m_1) \leftarrow \mathcal{R}\{\mathrm{diag}\{\mathbf{B}_i(k)\mathbf{P}_{3i}(0)\mathbf{B}_i(k)^{\mathtt{T}}\}\}$

end if

has covariance matrix

$$\mathbf{\Gamma}_{\boldsymbol{\omega}_r}(t,s)=\begin{pmatrix}3 & 0 & 2.5 & 0\\ 0 & 3 & 0 & 2.5\\ 2.5 & 0 & 3 & 0\\ 0 & 2.5 & 0 & 3\end{pmatrix}\delta_{ts}$$

and the autocorrelation matrix of the real vector of $[x(1), x(0)]^{\mathsf{T}}$ is given by

$$\begin{pmatrix}4 & 0.5+c & a & -0.8 & 1.5 & 0.3 & 0 & -0.1\\ 0.5+c & 4 & 0.8 & a & 0.3 & 1.5 & 0.1 & 0\\ a & 0.8 & 4 & 0.5 & 0 & 0.1 & 0.5+b & 0.3\\ -0.8 & a & 0.5 & 4 & -0.1 & 0 & 0.3 & 0.5+b\\ 1.5 & 0.3 & 0 & -0.1 & 4 & 0.5 & a & -0.8\\ 0.3 & 1.5 & 0.1 & 0 & 0.5 & 4 & 0.8 & a\\ 0 & 0.1 & 0.5+b & 0.3 & a & 0.8 & 4 & 0.5\\ -0.1 & 0 & 0.3 & 0.5+b & -0.8 & a & 0.5 & 4\end{pmatrix} \tag{21}$$

with a, b and c real parameters.

Suppose that $x(t)$ is observed through the equation

$$y(t)=\lambda(t)\star x(t)+v(t)$$

with $\lambda(t)=\lambda_1(t)+\mathrm{i}\lambda_2(t)+\mathrm{j}\lambda_3(t)+\mathrm{k}\lambda_4(t)$, being $\lambda_i(t)$ independent Bernoulli random variables with parameter ρ, and $v(t)$ a tessarine white noise whose real vector has the following covariance matrix:

$$\mathbf{\Gamma}_{\mathbf{v}_r}(t,s)=\begin{pmatrix}e & 0 & 0.55 & 0\\ 0 & 0.75 & 0 & 0.55\\ 0.55 & 0 & e & 0\\ 0 & 0.55 & 0 & 0.75\end{pmatrix}\delta_{ts} \tag{22}$$

with e a real parameter.

Firstly, we treat the $\mathbb{T}_1$ case. For that, we assume $a=c=0$, $b=1$ in (21), $f=0.25$ in (20) and $e=0.75$ in (22). From Theorem 2, it is straightforward to demonstrate that $x(t)$ is a $\mathbb{T}_1$WSM(2) signal under these conditions. Next, we compare the proposed $\mathbb{T}_1$ estimators with the QSL estimators. We denote the QSL filter by $\hat{x}^{\mathtt{QSL}}(t|t)$ and its error by $\epsilon^{\mathtt{QSL}}(t)$. Similarly, we denote by $\epsilon^{\mathtt{QSL}}(t|m_1)$ the error for the fixed-interval smoothing case. To assess the performance of $\hat{\mathbf{x}}^{\mathtt{T}_1}(t|t)$ in relation to $\hat{x}^{\mathtt{QSL}}(t|t)$, we consider the following behavioral measure:

$$DF_1(t) = \epsilon^{\mathtt{QSL}}(t) - \epsilon^{\mathtt{T}_1}(t) \tag{23}$$

with $\epsilon^{\mathtt{T}_1}(t)$ given in (18). For the smoothing case, we consider $m_1 = 99$ and a similar performance measure:

$$DS_1(t) = \epsilon^{\mathtt{QSL}}(t|99) - \epsilon^{\mathtt{T}_1}(t|99)$$

with $\epsilon^{\mathtt{T}_1}(t|99)$ given in (19). Such measures are displayed for $\rho = 0.7$ and $\rho = 0.9$ in Figure 1. Both graphics confirm that, under $\mathbb{T}_1$-properness conditions, $\mathbb{T}_1$ processing is a better approach than QSL processing in terms of performance. Likewise, in this particular case, as ρ increases then the error differences decreases, i.e., the performance of the $\mathbb{T}_1$ estimators is improved in relation to that of QSL estimators when the probability of missing observations becomes greater (that is, when the Bernoulli probabilities become smaller).

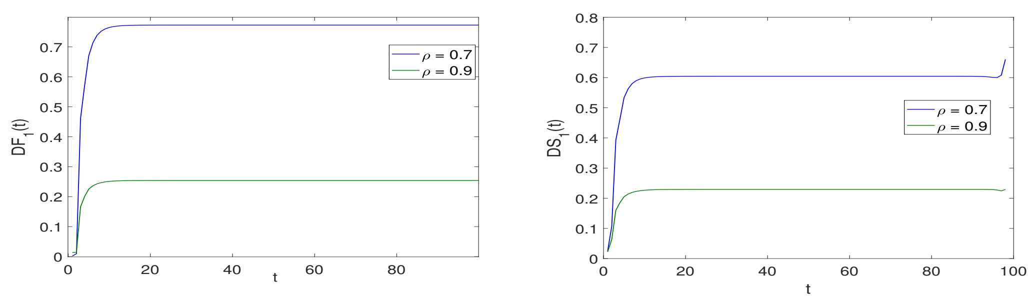

Figure 1: $DF_1(t)$ (left) and $DS_1(t)$ (right) for $\rho = 0.7, 0.9$.

Finally, we study the $\mathbb{T}_2$ case. We assume $a = 0.6$, $b = c = 0$ in (21), $f = 0$ in (20) and $e = 3.25$ in (22). It is easy to check that $x(t)$ is a $\mathbb{T}_2$WSM(1) signal under these conditions. We aim to assess the behavior of the proposed $\mathbb{T}_2$ estimators in relation to their counterparts QSWL estimators. To this end, we can choose between three possibilities in the quaternion domain. Specifically, the QSWL processing can be carried out by using any of the following three pairs: $\{x(t), x^{\mathrm{i}}(t)\}$, $\{x(t), x^{*}(t)\}$ or $\{x(t), x^{\mathrm{k}}(t)\}$. Thus, we denote the errors associated to the QSWL filters and the smoothed estimators for the three options by $\epsilon^{\mathtt{QSWL}_{\nu}}(t)$ and $\epsilon^{\mathtt{QSWL}_{\nu}}(t|99)$, $\nu = \mathrm{i}, *, \mathrm{k}$, respectively. Moreover, we define the following performance measures:

$$DF_2(t,\nu) = \epsilon^{\mathtt{QSWL}_{\nu}}(t) - \epsilon^{\mathtt{T}_2}(t)$$
$$DS_2(t,\nu) = \epsilon^{\mathtt{QSWL}_{\nu}}(t|99) - \epsilon^{\mathtt{T}_2}(t|99)$$

for $\nu = \mathrm{i}, *, \mathrm{k}$. Such measures are depicted for $\rho = 0.7$ and $\rho = 0.9$ in Figure 2. All the graphics show a superior performance of the $\mathbb{T}_2$ processing in comparison with QSWL processing. In other words, under $\mathbb{T}_2$-properness conditions, the suggested $\mathbb{T}_2$ estimators outperform their counterpart in the quaternion domain. In fact, the performance differences found depend on both, the kind of QSWL processing chosen and the level of uncertainty in the observations. As in the $\mathbb{T}_1$ case, a higher uncertainty in the observations (lower Bernoulli probabilities) leads to a better performance of the $\mathbb{T}_2$ estimators. This being so even if the pair $\{x(t), x^{\mathrm{k}}(t)\}$ is used, with which the QSWL processing achieves its best results.

4.1.4. Example 2

In this second example we consider the general model of motion, applicable in a large number of scenarios, such as bearings-only and rotation tracking [45]. The equations are given by

$$\frac{\partial \varphi}{\partial t} = \phi, \qquad \frac{\partial \phi}{\partial t} = \nu \tag{24}$$

where ν is the input of the system and φ represents the variable of interest with ϕ indicating its rate of change. The equivalent discrete-time model of (24) is

$$\mathbf{x}(t+1) = \begin{pmatrix} 1 & \Delta T \\ 0 & 1 \end{pmatrix} \mathbf{x}(t) + \begin{bmatrix} \Delta T^2/2 \\ \Delta T \end{bmatrix} \nu(t), \qquad t = 1 \dots 100$$

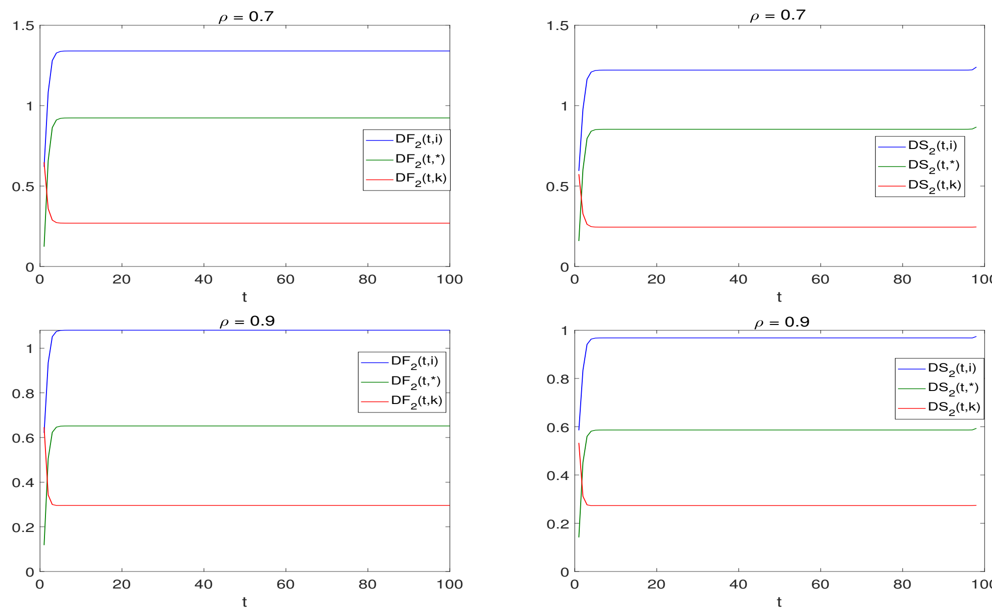

Figure 2: $DF_2(t,\nu)$ and $DS_2(t,\nu)$, $\nu = \mathrm{i}, *, \mathrm{k}$, for $\rho = 0.7$ (top) and $\rho = 0.9$ (bottom).

where $\mathbf{x}(t) = [\varphi(t), \phi(t)]^{\mathtt{T}}$ and $\Delta T = 0.04$ denotes the sampling interval. Consider the initial condition $\mathbf{x}(0) = \mathbf{0}_{2\times 1}$ and that the input $\nu(t)$ is a tessarine white noise whose real vector verifies

$$\mathbf{\Gamma}_{\nu_r}(t,s) = \begin{pmatrix} 3 & 0 & 2 & 0 \\ 0 & 3 & 0 & 2 \\ 2 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix} \delta_{ts}$$

In our study, we assume that the observation process is given by (15) being $\lambda_{ij}(t)$ independent Bernoulli random variables with parameter $\rho_{ij}(t) = \rho_i$, $i = 1, 2$. Moreover, $\mathbf{v}(t) = [v_1(t), v_2(t)]^{\mathtt{T}}$ is a tessarine white noise such that $v_1(t)$

and $v_2(t)$ are independent and

$$\mathbf{\Gamma}_{\mathbf{v}_{ir}}(t,s) = \begin{pmatrix} 6.5 & 0 & 0.1 & 0 \\ 0 & 6.5 & 0 & 0.1 \\ 0.1 & 0 & 6.5 & 0 \\ 0 & 0.1 & 0 & 6.5 \end{pmatrix} \delta_{ts}, \qquad i = 1, 2$$

Under these conditions, $\mathbf{x}(t)$ is a $\mathbb{T}_1 WSM(1)$ signal. Next, a performance comparison between the proposed $\mathbb{T}_1$ filter and the QSL filter of $\varphi(t)$ is detailed. Figure 3 depicts $DF_1(t)$ given in (23) for several values of ρ_1 and ρ_2. As in Example 1, the performance superiority of the tessarine processing over the quaternion processing is shown. Moreover, a greater performance difference is observed for smaller values of ρ_2 and when the filters execution time grows.

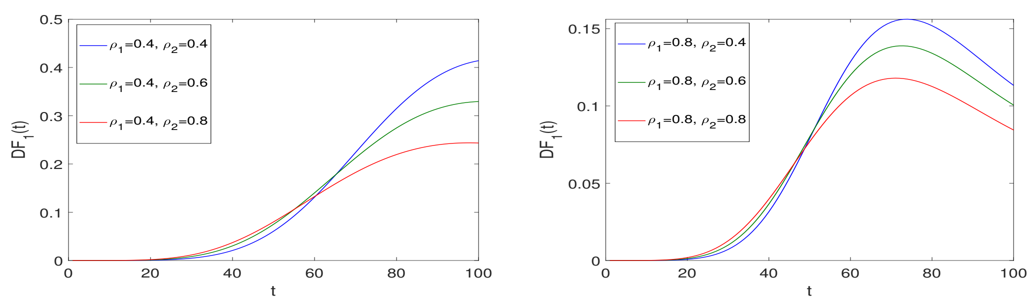

Figure 3: $DF_1(t)$ for $\rho_1 = 0.4, 0.8$ and $\rho_2 = 0.4, 0.6, 0.8$.

5. Conclusions

Tessarines represent a promising framework for solving signal processing problems, as confirmed by the growing number of papers published over recent years. Although quaternions has been the preferred hypercomplex approach in the literature, the finding that the behavior of the estimators is conditioned by the choice of a specific algebra can be a new impulse for the development of the tessarine signal processing. This paper has aimed to contribute to this field by paying attention to the effects of the properness properties over such a processing. In particular, a new concept of tessarine properness has been introduced and its application in the estimation problem with intermittent observations has been illustrated. For that, the class of $\mathbb{T}_i$WSM(n) signals has been considered and the superior performance of the $\mathbb{T}_i$ estimators in comparison with their counterparts in the quaternion domain has been shown experimentally. In light of these advances, it is expected that the tessarine signal processing will find increasing importance in coming years.

6. Acknowledgments

This work has been supported in part by the I+D+i Project with reference number 1256911, under "Programa Operativo FEDER Andalucía 2014–2020", Junta de Andalucía, and the Project EI_FQM2_2019 of "Plan de Apoyo a la Investigación", University of Jaén, Spain.

7. Appendix

7.1. Proof of Theorem 1

Point 1 is obtained from Corollary 1 in [27]. The proof of point 2 is immediate due to the fact that $\mathcal{G}_{\mathcal{C}_1} \subseteq \mathcal{G}_{\mathcal{C}_2} \subseteq \mathcal{G}_{\mathcal{C}_{\mathtt{WL}}}$. Now, we demonstrate point 3. From (4), (6) and Definition 1 we have that

$$\mathbf{l}_1^{\mathtt{T}} = [\mathbf{\Gamma}_{y\tilde{\mathbf{z}}}, \mathbf{0}_{1\times 2pm}] \begin{pmatrix} \mathbf{\Gamma}_{\tilde{\mathbf{z}}}^{-1} & \mathbf{0}_{2pm\times 2pm} \\ \mathbf{0}_{2pm\times 2pm} & \mathbf{\Gamma}_{\tilde{\mathbf{z}}}^{\mathrm{i}^{-1}} \end{pmatrix}$$

and thus, $\hat{y}^{\mathtt{TWL}} = \hat{y}^{\mathtt{T}_2}$.

7.2. Proof of Proposition 2

Suppose that $\mathbf{x}(t)$ is $\mathbb{T}_i$WSM(n), $i = 1, 2$. Then, from (10) and Corollary 1 in [27], we get

$$\hat{\boldsymbol{\xi}}_i^{\mathtt{T}_i}(t|\tau \leq \upsilon) = \mathbf{K}_i(t, \upsilon)\boldsymbol{\xi}_i(\upsilon) \tag{25}$$

Also, by using Theorem 3 of [27], it follows that $\boldsymbol{\xi}_i(t)-\hat{\boldsymbol{\xi}}_i^{\mathtt{T}_i}(t|\tau \leq \upsilon)$ is orthogonal to $\boldsymbol{\xi}_i(s)$, i.e.,

$$\boldsymbol{\xi}_i(t) - \hat{\boldsymbol{\xi}}_i^{\mathtt{T}_i}(t|\tau \leq \upsilon) \perp \boldsymbol{\xi}_i(s), \quad \forall s \leq \upsilon \tag{26}$$

Thus, from (25) and (26), we have

$$\boldsymbol{\Gamma}_{\boldsymbol{\xi}_i}(t,s) = \mathbf{K}_i(t,\upsilon)\boldsymbol{\Gamma}_{\boldsymbol{\xi}_i}(\upsilon,s)$$

and (12) holds. The proof of the sufficient condition is similar.

7.3. Proof of Theorem 2

From (10), (25) and Theorem 3 of [27] we obtain

$$\boldsymbol{\xi}_i(t+1) = \mathbf{K}_i(t)\boldsymbol{\xi}_i(t) + \mathbf{w}_i(t), \quad t \geq n-1 \tag{27}$$

where $\mathbf{w}_i(t) = \boldsymbol{\xi}_i(t+1) - \hat{\boldsymbol{\xi}}_i^{\mathtt{T}_i}(t+1|\tau \leq t)$ is the innovations process with

$$\boldsymbol{\Gamma}_{\mathbf{w}_i}(t) = \boldsymbol{\Gamma}_{\boldsymbol{\xi}_i}(t+1) - \mathbf{K}_i(t)\boldsymbol{\Gamma}_{\boldsymbol{\xi}_i}(t)\mathbf{K}_i^{\mathtt{H}}(t)$$

By construction, $\mathbf{w}_i(t)$ is uncorrelated with $\boldsymbol{\xi}_i(n-1)$, for $t \geq n-1$. Thus, by considering $\boldsymbol{\omega}_i(t) = \mathbf{w}_i^{(p)}(t)$ we get (13) and (14) for $i = 1,2$.

Now, from (27), Property 1 in [27] and the $\mathbb{T}_i$-properness condition of $\mathbf{x}(t)$ then, the jointly $\mathbb{T}_i$-properness of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\omega}_i(t)$ is easily checked for $i = 1,2$.

Conversely, suppose that $\mathbf{x}(t)$ has the representation (13). Then, proving (10) is equivalent to proving

$$\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|\tau \leq t) = \hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|t,t-1,\ldots,t-n+1), \quad \forall t \tag{28}$$

For that, from (13), we have

$$\hat{\mathbf{x}}^{\mathtt{T}_i}(t+1|\tau \leq t) = \mathbf{K}_i^{(p)}(t)\boldsymbol{\xi}_i(t) + \hat{\boldsymbol{\omega}}_i^{\mathtt{T}_i}(t|\tau \leq t)$$

and taking (14) into account we obtain that $\hat{\boldsymbol{\omega}}_i^{\mathtt{T}_i}(t|\tau \leq t) = \mathbf{0}_{ip\times 1}$, $i = 1,2$, and hence, (28) holds.

Finally, by applying the principle of recursion in (27), we get for $t > n-1$

$$\boldsymbol{\xi}_i(t) = \left(\prod_{s=n-1}^{t-1} \mathbf{K}_i(s)\right)\boldsymbol{\xi}_i(n-1) + \sum_{s=n-1}^{t-2}\left(\prod_{j=s+1}^{t-1}\mathbf{K}_i(j)\right)\mathbf{w}_i(s) + \mathbf{w}_i(t-1)$$

and, from the jointly $\mathbb{T}_i$-properness condition of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\omega}_i(t)$, we have that $\mathbf{x}(t)$ is $\mathbb{T}_i$-proper for $i = 1,2$.

7.4. Proof of Proposition 3

Firstly, from (5), we have that the augmented vector of $\boldsymbol{\lambda}(t) \star \mathbf{x}(t)$ is

$$\overline{\boldsymbol{\lambda}(t) \star \mathbf{x}(t)} = \mathcal{T}_p \operatorname{diag}\{\boldsymbol{\lambda}_r(t)\}\mathcal{T}_p^{\mathtt{H}}\bar{\mathbf{x}}(t)$$

and, taking (1) into account, it follows that

$$\boldsymbol{\Gamma}_{\overline{\boldsymbol{\lambda}\star\mathbf{x}}}(t,s) = 4\mathcal{T}_p \left(E[\boldsymbol{\lambda}_r(t)\boldsymbol{\lambda}_r^{\mathtt{T}}(s)] \odot E[\mathbf{x}_r(t)\mathbf{x}_r^{\mathtt{T}}(s)]\right) \mathcal{T}_p^{\mathtt{H}}$$

Furthermore, from (15), we get

$$\boldsymbol{\Gamma}_{\bar{\mathbf{y}}}(t,s) = \boldsymbol{\Gamma}_{\overline{\boldsymbol{\lambda}\star\mathbf{x}}}(t,s) + \boldsymbol{\Gamma}_{\bar{\mathbf{v}}}(t,s) \tag{29}$$

$$\boldsymbol{\Gamma}_{\bar{\mathbf{y}}\bar{\mathbf{x}}}(t,s) = \boldsymbol{\Pi}(t)\boldsymbol{\Gamma}_{\bar{\mathbf{x}}}(t,s) \tag{30}$$

with $\boldsymbol{\Pi}(t) = \mathcal{T}_p \operatorname{diag}\{E[\boldsymbol{\lambda}_r(t)]\}\mathcal{T}_p^{\mathtt{H}}$.

The proof is similar for the cases of $\mathbb{T}_1$-properness and $\mathbb{T}_2$-properness. For example, if $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly $\mathbb{T}_2$-proper then, from (4) and (30), we have that

$$\boldsymbol{\Pi}(t) = \begin{pmatrix} \boldsymbol{\Pi}_{11}(t) & \boldsymbol{\Pi}_{12}(t) \\ \boldsymbol{\Pi}_{21}(t) & \boldsymbol{\Pi}_{22}(t) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Pi}_{11}(t) & \mathbf{0}_{2p\times 2p} \\ \mathbf{0}_{2p\times 2p} & \boldsymbol{\Pi}_{22}(t) \end{pmatrix}$$

and thus, $\boldsymbol{\Pi}_{12}(t) = \boldsymbol{\Pi}_{21}(t) = \mathbf{0}_{2p\times 2p}$ implies that $\rho_{i1}(t) = \rho_{i3}(t)$ and $\rho_{i2}(t) = \rho_{i4}(t)$, $\forall t, i$. These equalities also imply that $E[\boldsymbol{\lambda}_r(t)\boldsymbol{\lambda}_r^{\mathtt{T}}(s)]$ satisfies the conditions given in (8) and hence $\boldsymbol{\lambda}(t) \star \mathbf{x}(t)$ is $\mathbb{T}_2$-proper. Finally, from (29), we have that $\mathbf{v}(t)$ is $\mathbb{T}_2$-proper.

Conversely, $\rho_{i1}(t) = \rho_{i3}(t)$ and $\rho_{i2}(t) = \rho_{i4}(t)$, $\forall t, i$, imply that $\boldsymbol{\Pi}_{12}(t) = \boldsymbol{\Pi}_{21}(t) = \mathbf{0}_{2p\times 2p}$ and since $\mathbf{x}(t)$ is a $\mathbb{T}_i$WSM(n) signal then, from (30), $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are cross $\mathbb{T}_2$-proper and, from (29), $\mathbf{y}(t)$ is also $\mathbb{T}_2$-proper.

7.5. Proof of Proposition 4

Define the vector $\boldsymbol{\zeta}(t) = [\boldsymbol{\zeta}_1^{\mathtt{T}}(t), \boldsymbol{\zeta}_1^{\mathtt{T}}(t-1), \ldots, \boldsymbol{\zeta}_1^{\mathtt{T}}(t-n+1)]^{\mathtt{T}}$, with $\boldsymbol{\zeta}_1(t) = [\mathbf{x}_{1r}^{\mathtt{T}}(t), \mathbf{x}_{2r}^{\mathtt{T}}(t), \ldots, \mathbf{x}_{pr}^{\mathtt{T}}(t)]^{\mathtt{T}}$, and the matrix $\mathbf{T}(t) = \operatorname{diag}\{\boldsymbol{\beta}(t)\}$ where $\boldsymbol{\beta}(t) = [\boldsymbol{\beta}_1^{\mathtt{T}}(t), \boldsymbol{\beta}_1^{\mathtt{T}}(t-1), \ldots, \boldsymbol{\beta}_1^{\mathtt{T}}(t-n+1)]^{\mathtt{T}}$ being $\boldsymbol{\beta}_1(t) = [\boldsymbol{\lambda}_{1r}^{\mathtt{T}}(t), \boldsymbol{\lambda}_{2r}^{\mathtt{T}}(t), \ldots, \boldsymbol{\lambda}_{pr}^{\mathtt{T}}(t)]^{\mathtt{T}}$. Also, define the following matrices $\mathbf{A} = \mathbf{I}_p \otimes (1, \mathrm{i}, \mathrm{j}, \mathrm{k})$, $\mathbf{C}_1 = \mathbf{I}_n^{(1)} \otimes \mathbf{A}$, and

$\mathbf{C}_2 = \mathbf{I}_n^{(1)} \otimes [\mathbf{A}^{\mathtt{H}}, \mathbf{A}^{\mathtt{T}}]^{\mathtt{H}}$. Then, the processes $\boldsymbol{\xi}_i(t)$ and $\mathbf{z}_i(t)$ can be rewritten for $i = 1, 2$ as

$$\boldsymbol{\xi}_i(t) = \mathbf{L}_i \boldsymbol{\zeta}(t)$$
$$\mathbf{z}_i(t) = \mathbf{C}_i \mathbf{T}(t) \boldsymbol{\zeta}(t) + \boldsymbol{\varpi}_i(t)$$

with $\boldsymbol{\varpi}_1(t) = \mathbf{v}(t)$ and $\boldsymbol{\varpi}_2(t) = [\mathbf{v}^{\mathtt{T}}(t), \mathbf{v}^{\mathtt{H}}(t)]^{\mathtt{T}}$. Hence

$$\begin{aligned} \boldsymbol{\Gamma}_{\mathbf{z}_i \boldsymbol{\xi}_i}(t,s) =& \mathbf{C}_i E[\mathbf{T}(t)] \boldsymbol{\Gamma}_{\boldsymbol{\zeta}}(t,s) \mathbf{L}_i^{\mathtt{H}} \\ \boldsymbol{\Gamma}_{\mathbf{z}_i}(t,s) =& E[\mathbf{C}_i \mathbf{T}(t) \boldsymbol{\zeta}(t) \boldsymbol{\zeta}^{\mathtt{H}}(s) \mathbf{T}(s) \mathbf{C}_i^{\mathtt{H}}] + \mathbf{R}_i(t) \delta_{ts} \\ =& E[\mathbf{C}_i \mathbf{T}(t) \boldsymbol{\zeta}(t) \boldsymbol{\zeta}^{\mathtt{H}}(s) \mathbf{T}(s) \mathbf{C}_i^{\mathtt{H}}] - \mathbf{H}_i(t) \mathbf{L}_i \boldsymbol{\Gamma}_{\boldsymbol{\zeta}}(t,s) \mathbf{L}_i^{\mathtt{H}} \mathbf{H}_i^{\mathtt{H}}(s) \\ & + \mathbf{H}_i(t) \mathbf{L}_i \boldsymbol{\Gamma}_{\boldsymbol{\zeta}}(t,s) \mathbf{L}_i^{\mathtt{H}} \mathbf{H}_i^{\mathtt{H}}(s) + \mathbf{R}_i(t) \delta_{ts} \end{aligned} \tag{31}$$

where $\mathbf{L}_1 = \mathbf{I}_n \otimes \mathbf{A}$ and $\mathbf{L}_2 = \mathbf{I}_n \otimes [\mathbf{A}^{\mathtt{H}}, \mathbf{A}^{\mathtt{T}}]^{\mathtt{H}}$.

Now, taking Proposition 3 into account, we get that $\mathbf{C}_i E[\mathbf{T}(t)] = \mathbf{H}_i(t) \mathbf{L}_i$, $i = 1, 2$. From this relation, (31) and (1) we obtain after some algebra (17).

References

[1] T. Kailath, A.H. Sayed, and B. Hassibi, Linear Estimation, Pretince Hall, New Yersey, 2000.

[2] F. Beutler, Multivariate wide-sense Markov processes and prediction theory, Ann. Math. Statist. 34 (1963) 424-438.

[3] V. Mandrekar, On multivariate wide-sense Markov processes, Nagoya Math. J. 3 (1968) 7-19.

[4] A. Kasprzyk and W. Szczotka, Covariance structure of wide-sense Markov processes of order $k \geq 1$, Applicationes Mathematicae 33 (2) (2006) 129-143.

[5] J.A. Espinosa-Pulido, J. Navarro-Moreno, R.M. Fernández-Alcalá and J.C. Ruiz-Molina, Widely linear Markov signals, EURASIP J. Advan. Signal Process. (2012) http://asp.eurasipjournals.com/content/2012/1/256.

[6] 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.

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

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

[15] 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.

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

[17] 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.

[18] 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.

[19] R.M. Fernández-Alcalá, J. Navarro-Moreno, J.D. Jiménez López and J.C. Ruiz-Molina, Semi-Widely linear estimation algorithms of quaternion signals with missing observations and correlated noises, J. Frankl. Inst. 357 (2020) 3075–3096.

[20] F. Catoni, D. Boccaletti, R. Cannata, V. Catoni, E. Nichelatti and P. Zampetti, The Mathematics of Minkowski Space-Time. With an Introduction to Commutative Hypercomplex Numbers, Birkhäuser Verlag, Berlin, 2008.

[21] D. Alfsmann, On families of 2N-dimensional hypercomplex algebras suitable for digital signal processing, in: 14th European Signal Processing Conference, 2006, pp. 1-4.

[22] D. Alfsmann, H.G. Göckler, S.J. Sangwine and T.A. Ell, Hypercomplex algebras in digital signal processing: benefits and drawbacks, in: 15th European Signal Processing Conference, 2007, pp. 1322-1326.

[23] S.L. Hahn and K.M. Snopek, Complex and Hypercomplex Analytic Signals: Theory and Applications, Artech House, Norwood, 2016.

[24] F. Tobar and D.P. Mandic, Quaternion reproducing kernel Hilbert spaces: existence and uniqueness conditions, IEEE Trans. Inform. Theory 60 (9) (2014) 5736-5749.

[25] F. Zanetti de Castro and M. Eduardo Valle, A broad class of discrete-time hypercomplex-valued Hopfield neural networks, Neural Netw. 122 (2020) 54-67.

[26] J.B. Kuipers, Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality, Princeton University Press, New Jersey, 2002.

[27] J. Navarro-Moreno, R.M. Fernández-Alcalá J.D. López-Jiménez and J.C. Ruiz-Molina, Tessarine signal processing under the $\mathbb{T}$-properness condition, J. Frankl. Inst. 357 (14) (2020) 10099-10125.

[28] A.J. Hanson, Visualizing Quaternions, Morgan Kaufmann, San Francisco, 2005.

[29] E. Chen, Y. Xu, X. Yang and W. Zhang, Quaternion based optical flow estimation for robust object tracking, Dig. Signal Process. 23 (1) (2013) 118-125.

[30] H.D. Schütte and J. Wenzel, Hypercomplex numbers in digital signal processing, in: International Symposium on Circuits and Systems, 1990, pp. 1557-1560.

[31] S.H. Pei, J.H. Chang and J.J. Ding, Commutative reduced biquaternions and their Fourier transform for signal and image processing applications, IEEE Trans. Signal Process. 52 (7) (2004) 2012-2031.

[32] S. Miron, N. Le Bihan, and J. Mars, High resolution vector sensor array processing based on biquaternions, in: IEEE International Conference on Acoustics, Speech and Signal Processing, 2006, pp. 1077-1080.

[33] S.H. Pei, J.H. Chang, J.J. Ding and M.Y. Chen, Eigenvalues and singular value decompositions of reduced biquaternion matrices, IEEE Trans. Circuits Syst. I 525 (2008) 2673-2685.

[34] L. Guo, M. Zhu and X. Ge, Reduced biquaternion canonical transform, convolution and correlation, Signal Process. 91 (2011) 2147-2153.

[35] A.T. Kamal and M.T. El-Melegy, Color image processing using reduced biquaternions with applications to face recognition in a PCA framework, in: IEEE International Conference on Computer Vision Workshops, 2017, pp. 3039-3046.

[36] A. Esposito, M. Faundez, F.C. Morabito and E. Pasero, Neural Advances in Processing Nonlinear Dynamic Signals, Springer, Switzerland, 2018.

[37] M. Kobayashi, Twin-multistate commutative quaternion Hopfield neural networks, Neurocomputing 320 (2018) 150-156.

[38] F. Ortolani, M. Scarpiniti, D. Comminiello and A. Uncini, On the influence of microphone array geometry on the behavior of hypercomplex adaptive filters, in: 5th IEEE Microwaves, Radar and Remote Sensing Symposium, 2017, pp. 37-42.

[39] F. Ortolani, D. Comminiello, M. Scarpiniti and A. Uncini, On 4-Dimensional Hypercomplex Algebras in Adaptive Signal Processing, Neural Advances in Processing Nonlinear Dynamic Signals, 2017, pp. 131-140.

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

[41] J. Navarro-Moreno, R.M. Fernández-Alcalá and J.C. Ruiz-Molina, Semi-widely simulation and estimation of continuous-time $\mathbb{C}^{\eta}$-proper quaternion random signals, IEEE Trans. Signal Process. 63 (18) (2015) 4999-5012.

[42] 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.

[43] 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.

[44] 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.

[45] S.P. Talebi, S. Werner and D.P. Mandic, Quaternion-valued distributed filtering and control, IEEE Trans. Autom. Control. 65 (10) (2020) 4246-4256.