

## Article

# $\mathbb{T}$ -proper Hypercomplex Centralized Fusion Estimation for Randomly Multiple Sensor Delays Systems with Correlated Noises

Rosa M. Fernández-Alcalá<sup>1,\*</sup>, Jesús Navarro-Moreno<sup>1,†</sup> and Juan C. Ruiz-Molina<sup>1,†</sup>

<sup>1</sup> Department of Statistics and Operations Research, University of Jaén, Paraje Las Lagunillas, 23071 Jaén; {rmfernan, jnavarro, jcruiz@ujaen.es}

\* Correspondence: rmfernan@ujaen.es; Tel.: +34-953-212-449

† These authors contributed equally to this work.

**Abstract:** The centralized fusion estimation problem for discrete-time vectorial tessarine signals in multiple sensor stochastic systems with random one-step delays and correlated noises is analyzed under different  $\mathbb{T}$ -properness conditions. Based on  $\mathbb{T}_k$ ,  $k = 1, 2$ , linear processing, new centralized fusion filtering, prediction, and fixed-point smoothing algorithms are devised. These algorithms have the advantage of providing optimal estimators with a significant reduction in computational cost compared to that obtained through a real or a widely linear processing approach. Simulation examples illustrate the effectiveness and applicability of the algorithms proposed, in which the superiority of the  $\mathbb{T}_k$  linear estimators over their counterparts in the quaternion domain is apparent.

**Keywords:** Centralized fusion estimation, Random delay systems, Tessarine processing,  $\mathbb{T}_k$  properness.

## 1. Introduction

Multi-sensor systems and related information fusion estimation theory have attracted much attention over the last few decades due to their wide range of applications in many fields, including target tracking, robotics, navigation, big data and signal processing [1–7].

In practice, failures during data transmission are unavoidable and lead to uncertain systems. In this regards, a significant problem is the estimation of the state from systems with random sensor delays (see, for example, [8–13]). Such delays may be mainly caused by computational load, heavy network traffic, and the limited bandwidth of the communication channel, as well as other limitations which mean that the measurements are not always up to date [8]. It is commonly assumed that measurement delays can be described by Bernoulli distributed random variables with known conditional probabilities, where the values 1 and 0 of these variables indicate the presence or absence of measurement delays in the corresponding sensor [10].

Traditionally, there have been two basic approaches to process the information from multiple sensors, centralized and distributed fusion. In the former approach, all the measurement data from each sensor are collected in a fusion center where they are fused and processed, whereas in the distributed fusion method, the measurements of each sensor are transmitted to a local processor where they are independently processed before being transmitted to the fusion center. It is well known that centralized fusion methods lead to the best (optimal) solution when all sensors work healthily [14,15]. The strength of this approach lies in the fact that it is easy to implement, and it makes possible the best use of the available information. Accordingly, with the purpose of optimal estimation, centralized fusion methodology has received increased attention in recent literature related to multi-sensor fusion estimation (see for example, [9,16–18]). Notwithstanding the foregoing, the main disadvantage of this approach is the high

computational load that may be required, especially when the number of sensors is too large. Alternatively, distributed fusion methodologies are developed with the purpose of designing solutions that are robust to failures and have a reduced computational load. The main handicap of these solutions is that they are suboptimal and, hence, it is desirable to explore other alternatives that can alleviate the computational burden. In this respect, the use of hypercomplex algebras may well offer an ideal framework in which to analyze the properness characteristics of the signals which lead to lower computational costs without losing optimality.

In general, the use of hypercomplex algebras has proliferated considerably in signal processing problems because of their natural ability to model multi-dimensional data giving rise to better geometrical interpretations. In this vein, quaternions and tessarines appear as 4D hypercomplex algebras composed of a real part and three imaginary parts which endows them with the ideal structure to describe three and four-dimensional signals. Nowadays, they play a fundamental role in a variety of applications such as robotics, avionics, 3D graphics, and virtual reality [19]. In principle, the use of quaternions or tessarines means renouncing some of the usual properties of the real or complex fields. Thus, while quaternion algebra is non-commutative, tessarines become a non-division algebra. These properties make each algebra more appropriate for every specific problem. With this in mind, in [20–23] the use of these two isodimensional algebras are compared with the objective of showing how the choice of a particular algebra may determine the proposed method performance.

In the related literature, quaternion algebra has been widely exploited as a signal processing tool and it is still a trending topic in different areas. In particular, in the area of multi-sensor fusion estimation, [24] and [25] proposed sensor fusion estimation algorithms based on a quaternion extended Kalman filter, [26] and [27] have provided robust distributed quaternion Kalman filtering algorithm for data fusion over sensor networks dealing with three-dimensional data, and [28] designed a linear quaternion fusion filter from multi-sensor observations. A common characteristic of all the estimation algorithms above is that their methodologies are based on a strictly linear (SL) processing. However, in the quaternion domain, optimal linear processing is widely linear (WL) which requires the consideration of the quaternion signal and its three involutions. In this framework, [29] devised WL filtering, prediction and smoothing algorithms for multi-sensor systems with mixed uncertainties of sensor delays, packet dropout and missing observations. Interestingly, when the signal presents properness properties (cancellation of one or more of the three complementary covariance matrices), the optimal processing is SL (if the signal is Q-proper) or semi-widely linear (if the signal is C-proper) which amounts to operate on a vector with reduced dimension, which means a significative reduction in the computational burden of the associated algorithms (please review [30–33] for further details).

On the other hand, the use of tessarines is less common in the signal processing literature and, to the best of the author's knowledge, they have never been considered in multi-sensor fusion estimation problems. In general, the use of tessarines in estimation problems has been limited by the fact that it is not being a normed division algebra. This drawback was successfully overcome in [22] by introducing a metric which guarantees the existence and unicity of the optimal estimator. Moreover, although the optimal processing in the tessarine field is the WL processing, under properness conditions it is possible to get the optimal solution from estimation algorithms with lower computational costs. In this sense, [22] and [23] introduced the concept of  $\mathbb{T}_1$  and  $\mathbb{T}_2$ -properness and provided a statistical test to determine whether a signal presents one of these properness properties. According to the type of properness, the most suitable form of processing is  $\mathbb{T}_1$  linear processing, which supposes to operate on the signal itself, or  $\mathbb{T}_2$  linear processing, based on the augmented vector given by the signal and its conjugate. The application of both  $\mathbb{T}_1$  and  $\mathbb{T}_2$  linear processing to the estimation problem has provided optimal estimation algorithms of reduced dimension.

Motivated by the above discussions, in this paper we consider a tessarine multiple sensor system where each sensor may be delayed at any time independently from the others. The probability of the occurrence of each delay is dealt by a Bernoulli distribution. Moreover, unlike most sensor fusion estimation algorithms, the observation noises of different sensors can be correlated. In this context, new centralized fusion filtering, prediction and fixed-point smoothing algorithms are designed under both  $\mathbb{T}_1$  and  $\mathbb{T}_2$ -properness conditions. The algorithms proposed provide the optimal estimations of the state; meanwhile, the computational burden has been reduced with respect to the counterpart tessarine WL (TWL) estimation algorithms. It is important to note that such savings in computational burden cannot be achieved in the real field. The superiority of these algorithms obtained from a  $\mathbb{T}_k$  linear approach over those derived in the quaternion domain is numerically demonstrated under different conditions of properness.

The remainder of the paper is organized as follows. Section 2 introduces the notation used throughout the paper and briefly reviews the main concepts related to the processing of tessarine signals and their implications under  $\mathbb{T}_k$  properness. Then, in Section 3, the problem of estimating a tessarine signal in linear discrete stochastic systems with random state delays and multiple sensors is formulated. Concretely, under  $\mathbb{T}_k$ -properness conditions, a compact state-space model of reduced dimension is proposed. From this model, and based on  $\mathbb{T}_k$ -properness properties,  $\mathbb{T}_k$  centralized fusion filtering, step ahead prediction, and fixed-point smoothing algorithms are devised in Section 4. Furthermore, the goodness of these algorithms in performance is numerically analyzed in Section 5 by means of a simulation example, where the superiority of the  $\mathbb{T}_k$  estimation algorithms above over their counterparts in the quaternion domain is evidenced. The paper ends with a section of conclusions. In order to maintain continuity, all technical proofs have been deferred to the appendix section.

## 2. Preliminaries

Throughout this paper, and unless otherwise stated, all the random variables are assumed to have zero-mean. Moreover, the notation and terminology is fairly standard. They are summarized in the following two subsections.

### 2.1. Notation

Boldfaced upper case letters are used to denote matrices, boldfaced lower case letters for column vectors, and lightfaced lower case letters for scalar quantities.  $\mathbf{I}_n$  denotes identity matrix of dimension  $n$ ,  $\mathbf{0}_{n \times m}$  denotes the  $n \times m$  zero matrix, and  $\mathbf{1}_n$  and  $\mathbf{0}_n$  represents the vectors of length  $n$  whose elements are all ones and zeros, respectively.

Superscripts  $*$ ,  $T$  and  $H$  represent the tessarine conjugate, transpose, and Hermitian transpose, respectively. Subscripts  $r$  and  $v$ , for  $v = \eta, \eta', \eta''$ , represent the real and imaginary parts of a tessarine. Moreover,  $\mathbb{Z}$ ,  $\mathbb{R}$ , and  $\mathbb{T}$  are used to denote the set of integer, real, and tessarine field, respectively. According to this notation,  $\mathbf{A} \in \mathbb{R}^{n \times m}$  (respectively,  $\mathbf{A} \in \mathbb{T}^{n \times m}$ ) means that  $\mathbf{A}$  is a real (respectively, tessarine)  $n \times m$  matrix, and similarly  $\mathbf{r} \in \mathbb{R}^n$  (respectively,  $\mathbf{r} \in \mathbb{T}^n$ ) means that  $\mathbf{r}$  is a  $m$ -dimensional real (respectively, tessarine) vector.

Furthermore,  $E[\cdot]$  and  $\text{Cov}(\cdot)$  are the expectation and covariance operators and  $\text{diag}(\cdot)$  is a diagonal (or block diagonal) matrix with elements specified on the main diagonal.  $\mathbf{0}_{n \times m}$  denotes the  $n \times m$  zero matrix,  $\mathbf{I}_n$  represents the identity matrix of dimension  $n$  and  $\mathbf{1}_n$  is the  $n$ -vector whose elements are all ones.  $\delta_{n,l}$  is the Kronecker delta function, which is equal to one if  $l = n$ , and zero otherwise. Finally, the Hadamard and Kronecker products are denoted by  $\circ$  and  $\otimes$ , respectively.

### 2.2. Basic concepts and properties

The following property of the Hadamard product will be useful.

**Property 1.** If  $\mathbf{A} \in \mathbb{R}^{n \times n}$  and  $\mathbf{b} \in \mathbb{R}^n$ , then

$$\text{diag}(\mathbf{b})\mathbf{A}\text{diag}(\mathbf{b}) = (\mathbf{b}\mathbf{b}^T) \circ \mathbf{A}. \quad (1)$$

**Definition 1.** A tessarine random signal  $\mathbf{x}(t) \in \mathbb{T}^n$  is a stochastic process of the form [22]

$$\mathbf{x}(t) = \mathbf{x}_r(t) + \eta\mathbf{x}_\eta(t) + \eta'\mathbf{x}_{\eta'}(t) + \eta''\mathbf{x}_{\eta''}(t), \quad t \in \mathbb{Z}$$

where  $\mathbf{x}_v(t) \in \mathbb{R}^n$ , for  $v = r, \eta, \eta', \eta''$ , are real random signals and  $\{1, \eta, \eta', \eta''\}$  obeys the following rules:

$$\eta\eta' = \eta'', \quad \eta'\eta'' = \eta, \quad \eta''\eta = -\eta', \quad \eta^2 = \eta'^2 = -1, \quad \eta'^2 = 1.$$

The conjugate of a given tessarine random signal  $\mathbf{x}(t) \in \mathbb{T}^n$ , is:

$$\mathbf{x}^*(t) = \mathbf{x}_r(t) - \eta\mathbf{x}_\eta(t) + \eta'\mathbf{x}_{\eta'}(t) - \eta''\mathbf{x}_{\eta''}(t).$$

Moreover, the following two auxiliary tessarine vectors are defined as:

$$\mathbf{x}^\eta(t) = \mathbf{x}_r(t) + \eta\mathbf{x}_\eta(t) - \eta'\mathbf{x}_{\eta'}(t) - \eta''\mathbf{x}_{\eta''}(t),$$

$$\mathbf{x}^{\eta''}(t) = \mathbf{x}_r(t) - \eta\mathbf{x}_\eta(t) - \eta'\mathbf{x}_{\eta'}(t) + \eta''\mathbf{x}_{\eta''}(t).$$

For a complete description of the second-order statistical properties of  $\mathbf{x}(t)$ , we need to consider the augmented tessarine signal vector  $\bar{\mathbf{x}}(t) = [\mathbf{x}^T(t), \mathbf{x}^H(t), \mathbf{x}^{\eta T}(t), \mathbf{x}^{\eta'' T}(t)]^T$ . The following relationship between the augmented vector and the real vector  $\mathbf{x}^r(t) = [\mathbf{x}_r^T(t), \mathbf{x}_\eta^T(t), \mathbf{x}_{\eta'}^T(t), \mathbf{x}_{\eta''}^T(t)]^T$  can be established:

$$\bar{\mathbf{x}}(t) = 2\mathcal{T}_n\mathbf{x}^r(t),$$

where  $\mathcal{T}_n = \frac{1}{2}\mathcal{A} \otimes \mathbf{I}_n$

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

with  $\mathcal{T}_n^H \mathcal{T}_n = \mathbf{I}_{4n}$ .

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

$$\mathbf{x}(t) \star \mathbf{y}(s) = \mathbf{x}_r(t) \circ \mathbf{y}_r(s) + \eta\mathbf{x}_\eta(t) \circ \mathbf{y}_\eta(s) + \eta'\mathbf{x}_{\eta'}(t) \circ \mathbf{y}_{\eta'}(s) + \eta''\mathbf{x}_{\eta''}(t) \circ \mathbf{y}_{\eta''}(s). \quad (2)$$

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

**Property 2.** The augmented vector of  $\mathbf{x}(t) \star \mathbf{y}(s)$  is  $\overline{\mathbf{x}(t) \star \mathbf{y}(s)} = \bar{\mathbf{D}}^x(t)\bar{\mathbf{y}}(s)$ , where  $\bar{\mathbf{D}}^x(t) = \mathcal{T}_n \text{diag}(\mathbf{x}^r(t))\mathcal{T}_n^H$ .

**Definition 3.** The pseudo autocorrelation function of  $\mathbf{x}(t) \in \mathbb{T}^n$  is defined as  $\mathbf{R}_x(t, s) = E[\mathbf{x}(t)\mathbf{x}^H(s)]$ ,  $\forall t, s \in \mathbb{Z}$ , and the pseudo cross-correlation function of  $\mathbf{x}(t), \mathbf{y}(t) \in \mathbb{T}^n$  is defined as  $\mathbf{R}_{xy}(t, s) = E[\mathbf{x}(t)\mathbf{y}^H(s)]$ ,  $\forall t, s \in \mathbb{Z}$ .

Note that, depending on the vanishing of the different pseudo correlation functions  $\mathbf{R}_{xx^v}(t, s)$ ,  $v = *, \eta, \eta''$ , various kinds of tessarine properness can be defined. In particular, the following properness conditions in the tessarine domain has been recently introduced in [22] and [23].

**Definition 4.** A random signal  $\mathbf{x}(t) \in \mathbb{T}^n$  is said to be  $\mathbb{T}_1$ -proper (respectively,  $\mathbb{T}_2$ -proper) if, and only if, the functions  $\mathbf{R}_{\mathbf{x}\mathbf{x}^v}(t, s)$ , with  $v = *, \eta, \eta''$  (respectively,  $v = \eta, \eta''$ ), vanish  $\forall t, s \in \mathbb{Z}$ .

In a like manner, two random signals  $\mathbf{x}(t) \in \mathbb{T}^{n_1}$  and  $\mathbf{y}(t) \in \mathbb{T}^{n_2}$  are cross  $\mathbb{T}_1$ -proper, (respectively, cross  $\mathbb{T}_2$ -proper) if, and only if, the functions  $\mathbf{R}_{\mathbf{x}\mathbf{y}^v}(t, s)$ , with  $v = *, \eta, \eta''$  (respectively,  $v = \eta, \eta''$ ), vanish  $\forall t, s \in \mathbb{Z}$ .

Moreover,  $\mathbf{x}(t)$  and  $\mathbf{y}(t)$  are jointly  $\mathbb{T}_1$ -proper (respectively, jointly  $\mathbb{T}_2$ -proper) if, and only if, they are  $\mathbb{T}_1$ -proper (respectively,  $\mathbb{T}_2$ -proper) and cross  $\mathbb{T}_1$ -proper (respectively, cross  $\mathbb{T}_2$ -proper).

Remark that,  $\mathbb{T}_1$  properness is more restrictive than  $\mathbb{T}_2$  properness. Statistical tests to experimentally check whether a signal is  $\mathbb{T}_k$ -proper,  $k = 1, 2$ , or improper have been proposed in [22] and [23].

It should be highlighted that the different properness properties have direct implications on the optimal linear processing. In general, the optimal linear processing is the widely linear processing which requires to operate on the augmented tessarine vector  $\bar{\mathbf{x}}(t)$ . Nevertheless, in the case of joint  $\mathbb{T}_k$ -properness,  $k = 1, 2$ , the optimal linear processing is reduced to a  $\mathbb{T}_k$  linear processing, with the corresponding decrease in the dimension of the problem. In particular,  $\mathbb{T}_1$  linear processing is based on the tessarine random signal itself, and  $\mathbb{T}_2$  linear processing considers the augmented vector given by the signal and its conjugate [23].

### 3. Problem formulation

Consider the following class of linear discrete stochastic systems with state delays and multiple sensors:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{F}_1(t)\mathbf{x}(t) + \mathbf{F}_2(t)\mathbf{x}^*(t) + \mathbf{F}_3(t)\mathbf{x}^\eta(t) + \mathbf{F}_4(t)\mathbf{x}^{\eta''}(t) + \mathbf{u}(t), \quad t \geq 0 \\ \mathbf{z}^{(i)}(t) &= \mathbf{x}(t) + \mathbf{v}^{(i)}(t), \quad t \geq 0, \quad i = 1, \dots, R \\ \mathbf{y}^{(i)}(t) &= \gamma^{(i)}(t) \star \mathbf{z}^{(i)}(t) + (\mathbf{1}_n - \gamma^{(i)}(t)) \star \mathbf{z}^{(i)}(t-1), \quad t \geq 1, \quad i = 1, \dots, R \end{aligned} \quad (3)$$

where  $R$  is the number of sensors,  $\star$  is the product defined in (2),  $\mathbf{F}_j(t) \in \mathbb{T}^{n \times n}$ ,  $j = 1, 2, 3, 4$ , are deterministic matrices,  $\mathbf{x}(t) \in \mathbb{T}^n$  is the system state to be estimated,  $\mathbf{u}(t) \in \mathbb{T}^n$  is a tessarine noise,  $\mathbf{z}^{(i)}(t) \in \mathbb{T}^n$  is the  $i$ th sensor outputs with tessarine sensor noise  $\mathbf{v}^{(i)}(t) \in \mathbb{T}^n$ ,  $\mathbf{y}^{(i)}(t) \in \mathbb{T}^n$  is the observation of the  $i$ th sensor,  $\gamma^{(i)}(t) = [\gamma_1^{(i)}(t), \dots, \gamma_n^{(i)}(t)]^T \in \mathbb{T}^n$  is a tessarine random vector whose components  $\gamma_j^{(i)}(t) = \gamma_{j,r}^{(i)}(t) + \eta \gamma_{j,\eta}^{(i)}(t) + \eta' \gamma_{j,\eta'}^{(i)}(t) + \eta'' \gamma_{j,\eta''}^{(i)}(t)$ , for  $j = 1, \dots, n$ , are composed of independent Bernoulli random variables  $\gamma_{j,\nu}^{(i)}(t)$ ,  $j = 1, \dots, n$ ,  $\nu = r, \eta, \eta', \eta''$ , with known probabilities  $p_{j,\nu}^{(i)}(t)$ , and whose possible outcomes  $\{0, 1\}$  indicate if the  $\nu$  part of the  $j$ th observation component of the  $i$ th sensor is up-to-date (case  $\gamma_{j,\nu}^{(i)}(t) = 1$ ) or there exists one-step delay (case  $\gamma_{j,\nu}^{(i)}(t) = 0$ ).

The following assumptions for the above system (3) are made:

**Assumption 1.** For a given sensor  $i$ , the Bernoulli variable vector  $\gamma^{(i)}(t)$  is independent of  $\gamma^{(i)}(s)$ , for  $t \neq s$ , and also  $\gamma^{(i)}(t)$  is independent of  $\gamma^{(j)}(t)$ , for any two sensors  $i \neq j$ .

**Assumption 2.** For a given sensor  $i$ ,  $\gamma^{(i)}(t)$  is independent of  $\mathbf{x}(t)$ ,  $\mathbf{u}(t)$  and  $\mathbf{v}^{(j)}(t)$ , for any  $i, j = 1, \dots, R$ .

**Assumption 3.**  $\mathbf{u}(t)$  and  $\mathbf{v}^{(i)}(t)$  are correlated white noises with respective pseudo variances  $\mathbf{Q}(t)$  and  $\mathbf{R}^{(i)}(t)$ . Moreover,  $E[\mathbf{u}(t)\mathbf{v}^{(i)\mathbf{H}}(s)] = \mathbf{S}^{(i)}(t)\delta_{t,s}$ .

**Assumption 4.**  $\mathbf{v}^{(i)}(t)$  is independent of  $\mathbf{v}^{(j)}(t)$ , for any two sensors  $i \neq j$ .

**Assumption 5.** The initial state  $\mathbf{x}(0)$  is independent of the additive noises  $\mathbf{u}(t)$  and  $\mathbf{v}^{(i)}(t)$ , for  $t \geq 0$  and  $i = 1, \dots, R$ .

**Remark 1.** From the hypotheses established on the Bernoulli random variables it follows that, for any  $j_1, j_2 = 1, \dots, n$ ,  $v_1, v_2 = r, \eta, \eta', \eta''$  and  $i_1, i_2 = 1, \dots, R$ ,

$$\begin{aligned} E[\gamma_{j_1, v_1}^{(i_1)}(t) \gamma_{j_2, v_2}^{(i_2)}(t)] &= \begin{cases} p_{j_1, v_1}^{(i_1)}(t), & \text{if } i_1 = i_2, j_1 = j_2, v_1 = v_2 \\ p_{j_1, v_1}^{(i_1)}(t) p_{j_2, v_2}^{(i_2)}(t), & \text{otherwise,} \end{cases} \\ E[(1 - \gamma_{j_1, v_1}^{(i_1)}(t))(1 - \gamma_{j_2, v_2}^{(i_2)}(t))] &= \begin{cases} 1 - p_{j_1, v_1}^{(i_1)}(t), & \text{if } i_1 = i_2, j_1 = j_2, v_1 = v_2 \\ (1 - p_{j_1, v_1}^{(i_1)}(t))(1 - p_{j_2, v_2}^{(i_2)}(t)), & \text{otherwise.} \end{cases} \end{aligned} \quad (4)$$

### 3.1. One-state delay system under $\mathbb{T}_k$ -properness

In this section, a TWL one-state delay system, which exploits the full amount second-order statistics information available, is introduced and analyzed in  $\mathbb{T}_k$ -properness scenarios,  $k = 1, 2$ .

For this purpose, consider the augmented vectors  $\bar{\mathbf{x}}(t)$ ,  $\bar{\mathbf{z}}^{(i)}(t)$ , and  $\bar{\mathbf{y}}^{(i)}(t)$  of  $\mathbf{x}(t)$ ,  $\mathbf{z}^{(i)}(t)$ , and  $\mathbf{y}^{(i)}(t)$ , respectively. Then, by applying Property 2 on the system (3), the following TWL one-state delay model can be defined:

$$\begin{aligned} \bar{\mathbf{x}}(t+1) &= \bar{\Phi}(t) \bar{\mathbf{x}}(t) + \bar{\mathbf{u}}(t), \quad t \geq 0 \\ \bar{\mathbf{z}}^{(i)}(t) &= \bar{\mathbf{x}}(t) + \bar{\mathbf{v}}^{(i)}(t), \quad t \geq 0, \quad i = 1, \dots, R \\ \bar{\mathbf{y}}^{(i)}(t) &= \bar{\mathcal{D}}^{\gamma^{(i)}}(t) \bar{\mathbf{z}}^{(i)}(t) + \bar{\mathcal{D}}^{(1-\gamma^{(i)})}(t) \bar{\mathbf{z}}^{(i)}(t-1), \quad t \geq 1, \quad i = 1, \dots, R \end{aligned} \quad (5)$$

where

$$\bar{\Phi}(t) = \begin{bmatrix} \mathbf{F}_1(t) & \mathbf{F}_2(t) & \mathbf{F}_3(t) & \mathbf{F}_4(t) \\ \mathbf{F}_2^*(t) & \mathbf{F}_1^*(t) & \mathbf{F}_4^*(t) & \mathbf{F}_3^*(t) \\ \mathbf{F}_3^\eta(t) & \mathbf{F}_4^\eta(t) & \mathbf{F}_1^\eta(t) & \mathbf{F}_2^\eta(t) \\ \mathbf{F}_4^{\eta''}(t) & \mathbf{F}_3^{\eta''}(t) & \mathbf{F}_2^{\eta''}(t) & \mathbf{F}_1^{\eta''}(t) \end{bmatrix}.$$

Moreover, from Assumption 3, the *pseudo* correlation matrices associated to the augmented noise vectors  $\bar{\mathbf{u}}(t)$  and  $\bar{\mathbf{v}}^{(i)}(t)$  are given by

$$\begin{aligned} E[\bar{\mathbf{u}}(t) \bar{\mathbf{u}}^H(s)] &= \bar{\mathbf{Q}}(t) \delta_{t,s}; \\ E[\bar{\mathbf{v}}^{(i)}(t) \bar{\mathbf{v}}^{(i)H}(s)] &= \bar{\mathbf{R}}^{(i)}(t) \delta_{t,s}; \\ E[\bar{\mathbf{u}}(t) \bar{\mathbf{v}}^{(i)H}(s)] &= \bar{\mathbf{S}}^{(i)}(t) \delta_{t,s}. \end{aligned}$$

The following result establishes conditions on system (5) which lead to  $\mathbb{T}_k$ -properness properties of the processes involved.

**Proposition 1.** Consider the TWL one-state delay model (5).

1. If  $\mathbf{x}(0)$  and  $\mathbf{u}(t)$  are  $\mathbb{T}_1$ -proper, and  $\bar{\Phi}(t)$  is a block diagonal matrix of the form

$$\bar{\Phi}(t) = \text{diag}(\mathbf{F}_1(t), \mathbf{F}_1^*(t), \mathbf{F}_1^\eta(t), \mathbf{F}_1^{\eta''}(t)),$$

then  $\mathbf{x}(t)$  is  $\mathbb{T}_1$ -proper.

If additionally  $p_{j,r}^{(i)}(t) = p_{j,\eta}^{(i)}(t) = p_{j,\eta'}^{(i)}(t) = p_{j,\eta''}^{(i)}(t) \triangleq p_j^{(i)}(t)$ ,  $\forall t, j, i$ ,  $\mathbf{v}^{(i)}(t)$  is  $\mathbb{T}_1$ -proper, and  $\mathbf{u}(t)$  and  $\mathbf{v}^{(i)}(t)$  are cross  $\mathbb{T}_1$ -proper, then  $\mathbf{x}(t)$  and  $\mathbf{y}^{(i)}(t)$  are jointly  $\mathbb{T}_1$ -proper.

2. If  $\mathbf{x}(0)$  and  $\mathbf{u}(t)$  are  $\mathbb{T}_2$ -proper, and  $\bar{\Phi}(t)$  is a block diagonal matrix of the form

$$\bar{\Phi}(t) = \text{diag}(\Phi_2(t), \Phi_2^\eta(t)), \quad \text{with } \Phi_2(t) = \begin{bmatrix} \mathbf{F}_1(t) & \mathbf{F}_2(t) \\ \mathbf{F}_2^*(t) & \mathbf{F}_1^*(t) \end{bmatrix}, \quad (6)$$



then  $\mathbf{x}(t)$  is  $\mathbb{T}_2$ -proper.

If additionally,  $p_{j,r}^{(i)}(t) = p_{j,\eta}^{(i)}(t)$ ,  $p_{j,\eta'}^{(i)}(t) = p_{j,\eta''}^{(i)}(t)$ ,  $\forall t, j, i$ ,  $\mathbf{v}^{(i)}(t)$  is  $\mathbb{T}_2$ -proper and  $\mathbf{u}(t)$ , and  $\mathbf{v}^{(i)}(t)$  are cross  $\mathbb{T}_2$ -proper, then  $\mathbf{x}(t)$  and  $\mathbf{y}^{(i)}(t)$  are jointly  $\mathbb{T}_2$ -proper.

**Proof.** The proof follows straight away from the application of the corresponding conditions on system (5) and the computation of the augmented pseudo correlation matrices  $\mathbf{R}_{\bar{\mathbf{x}}}(t, s)$  and  $\mathbf{R}_{\bar{\mathbf{x}}\bar{\mathbf{y}}^{(i)}}(t, s)$ .  $\square$

**Remark 2.** Note that under  $\mathbb{T}_1$ -properness conditions,  $\bar{\mathbf{\Pi}}^{\gamma^{(i)}}(t) = E[\bar{\mathbf{D}}^{\gamma^{(i)}}(t)]$ ,  $i = 1, \dots, R$ , is a diagonal matrix of the form  $\bar{\mathbf{\Pi}}^{\gamma^{(i)}}(t) = \mathbf{I}_4 \otimes \mathbf{\Pi}_1^{(i)}(t)$ , with  $\mathbf{\Pi}_1^{(i)}(t) = \text{diag}(p_{1,r}^{(i)}(t), \dots, p_{n,r}^{(i)}(t))$ .

Likewise, under  $\mathbb{T}_2$ -properness conditions,  $\bar{\mathbf{\Pi}}^{\gamma^{(i)}}(t) = E[\bar{\mathbf{D}}^{\gamma^{(i)}}(t)]$ ,  $i = 1, \dots, R$ , takes the form of a block diagonal matrix as follows:

$$\bar{\mathbf{\Pi}}^{\gamma^{(i)}}(t) = \text{diag}(\mathbf{\Pi}_2^{(i)}(t), \mathbf{\Pi}_2^{(i)}(t)), \text{ with } \mathbf{\Pi}_2^{(i)}(t) = \frac{1}{2} \begin{bmatrix} \mathbf{\Pi}_a^{(i)}(t) & \mathbf{\Pi}_b^{(i)}(t) \\ \mathbf{\Pi}_b^{(i)}(t) & \mathbf{\Pi}_a^{(i)}(t) \end{bmatrix},$$

where  $\mathbf{\Pi}_a^{(i)}(t) = \text{diag}(p_{1,r}^{(i)}(t) + p_{1,\eta'}^{(i)}(t), \dots, p_{n,r}^{(i)}(t) + p_{n,\eta'}^{(i)}(t))$  and  $\mathbf{\Pi}_b^{(i)}(t) = \text{diag}(p_{1,r}^{(i)}(t) - p_{1,\eta'}^{(i)}(t), \dots, p_{n,r}^{(i)}(t) - p_{n,\eta'}^{(i)}(t))$ .

### 3.2. Compact state-space model

By stacking the observations at each sensor in a global observation vector  $\bar{\mathbf{z}}(t) = [\bar{\mathbf{z}}^{(1)\top}(t), \dots, \bar{\mathbf{z}}^{(R)\top}(t)]^\top$ , the TWL one-state delay system (5) can be rewriting in the following compact form:

$$\begin{aligned} \bar{\mathbf{x}}(t+1) &= \bar{\mathbf{\Phi}}(t)\bar{\mathbf{x}}(t) + \bar{\mathbf{u}}(t), \quad t \geq 0 \\ \bar{\mathbf{z}}(t) &= \bar{\mathbf{C}}\bar{\mathbf{x}}(t) + \bar{\mathbf{v}}(t), \quad t \geq 0 \\ \bar{\mathbf{y}}(t) &= \bar{\mathbf{D}}^{\tilde{\gamma}}(t)\bar{\mathbf{z}}(t) + \bar{\mathbf{D}}^{(1-\tilde{\gamma})}(t)\bar{\mathbf{z}}(t-1), \quad t \geq 1 \end{aligned} \quad (7)$$

where  $\bar{\mathbf{v}}(t)$  and  $\bar{\mathbf{y}}(t)$  denote the stacking vector of  $\bar{\mathbf{v}}^{(i)\top}(t)$  and  $\bar{\mathbf{y}}^{(i)\top}(t)$ , for  $i = 1, \dots, R$ , respectively. Moreover,  $\bar{\mathbf{C}} = \mathbf{1}_R \otimes \mathbf{I}_{4n}$ ,  $\bar{\mathbf{D}}^{\tilde{\gamma}}(t) = \bar{\mathbf{L}} \text{diag}(\tilde{\gamma}^r(t))\bar{\mathbf{L}}^H$  and  $\bar{\mathbf{D}}^{(1-\tilde{\gamma})}(t) = \bar{\mathbf{L}} \text{diag}(\mathbf{1}_{4Rn} - \tilde{\gamma}^r(t))\bar{\mathbf{L}}^H$ , with  $\bar{\mathbf{L}} = \mathbf{I}_R \otimes \mathbf{T}_n$ .

In addition,  $E[\bar{\mathbf{v}}(t)\bar{\mathbf{v}}^H(s)] = \bar{\mathbf{R}}(t)\delta_{t,s}$ , with  $\bar{\mathbf{R}}(t) = \text{diag}(\bar{\mathbf{R}}^{(1)}(t), \dots, \bar{\mathbf{R}}^{(R)}(t))$ , and  $E[\bar{\mathbf{u}}(t)\bar{\mathbf{v}}^H(s)] = \bar{\mathbf{S}}(t)\delta_{t,s}$ , with  $\bar{\mathbf{S}}(t) = [\bar{\mathbf{S}}^{(1)}(t), \dots, \bar{\mathbf{S}}^{(R)}(t)]$ .

In this paper, our aim is to investigate the centralized fusion estimation problem under conditions of  $\mathbb{T}_k$ -properness, with  $k = 1, 2$ . In this sense, the use of  $\mathbb{T}_k$ -properness properties allows us to consider the following observation equation with reduced dimension:

$$\mathbf{y}_k(t) = \bar{\mathbf{D}}_k^{\tilde{\gamma}}(t)\bar{\mathbf{C}}\bar{\mathbf{x}}(t) + \bar{\mathbf{D}}_k^{(1-\tilde{\gamma})}(t)\bar{\mathbf{C}}\bar{\mathbf{x}}(t-1) + \bar{\mathbf{D}}_k^{\tilde{\gamma}}(t)\bar{\mathbf{v}}(t) + \bar{\mathbf{D}}_k^{(1-\tilde{\gamma})}(t)\bar{\mathbf{v}}(t-1), \quad t \geq 1 \quad (8)$$

where  $\bar{\mathbf{x}}(t)$  satisfies the state equation in (7),  $\bar{\mathbf{D}}_k^{\tilde{\gamma}}(t) = \mathbf{L}_k \text{diag}(\tilde{\gamma}^r(t))\bar{\mathbf{L}}^H$  and  $\bar{\mathbf{D}}_k^{(1-\tilde{\gamma})}(t) = \mathbf{L}_k \text{diag}(\mathbf{1}_{4Rn} - \tilde{\gamma}^r(t))\bar{\mathbf{L}}^H$ , with  $\mathbf{L}_k = \mathbf{I}_R \otimes \mathbf{T}_k$  and  $\mathbf{T}_k = \frac{1}{2}\mathbf{B}_k \otimes \mathbf{I}_n$ , where

- $\mathbb{T}_1$ -proper scenario:

$$\begin{aligned} \mathbf{B}_1 &= \begin{bmatrix} 1 & \eta & \eta' & \eta'' \end{bmatrix}; \\ \mathbf{y}_1(t) &\triangleq [\mathbf{y}^{(1)\top}(t), \dots, \mathbf{y}^{(R)\top}(t)]^\top. \end{aligned}$$

- $\mathbb{T}_2$ -proper scenario:

$$\mathcal{B}_2 = \begin{bmatrix} 1 & \eta & \eta' & \eta'' \\ 1 & -\eta & \eta' & -\eta'' \end{bmatrix};$$

$$\mathbf{y}_2(t) \triangleq [\mathbf{y}^{(1)\top}(t), \mathbf{y}^{(1)\mathbf{H}}(t), \dots, \mathbf{y}^{(R)\top}(t), \mathbf{y}^{(R)\mathbf{H}}(t)]^\top.$$

**Remark 3.** Note that under  $\mathbb{T}_k$ -properness conditions,  $\tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) = E[\tilde{\mathcal{D}}_k^{\tilde{\gamma}}(t)]$  is given by  $\tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) = \text{diag}(\tilde{\mathbf{\Pi}}_k^{\gamma^{(1)}}(t), \dots, \tilde{\mathbf{\Pi}}_k^{\gamma^{(R)}}(t))$ , where  $\tilde{\mathbf{\Pi}}_k^{\gamma^{(i)}}(t) = [\mathbf{\Pi}_k^{(i)}(t), \mathbf{0}_{kn \times (4-k)n}]$  with  $\mathbf{\Pi}_k^{(i)}(t)$ ,  $i = 1, \dots, R$ , given in Remark 2.

Similarly,  $\tilde{\mathbf{\Pi}}_k^{(1-\tilde{\gamma})}(t) = E[\tilde{\mathcal{D}}_k^{(1-\tilde{\gamma})}(t)]$  is given by the block diagonal matrix  $\tilde{\mathbf{\Pi}}_k^{(1-\tilde{\gamma})}(t) = \text{diag}(\tilde{\mathbf{\Pi}}_k^{(1-\gamma^{(1)})}(t), \dots, \tilde{\mathbf{\Pi}}_k^{(1-\gamma^{(R)})}(t))$  with  $\tilde{\mathbf{\Pi}}_k^{(1-\gamma^{(i)})}(t) = [\mathbf{I}_{kn} - \mathbf{\Pi}_k^{(i)}(t), \mathbf{0}_{kn \times (4-k)n}]$ .

Accordingly, whereas the optimal linear processing for the estimation of a tessarine signal  $\mathbf{x}(t)$  is the TWL processing based on the set of measurements  $\{\tilde{\mathbf{y}}(1), \dots, \tilde{\mathbf{y}}(t)\}$ , under  $\mathbb{T}_k$ -properness conditions the optimal estimator of  $\mathbf{x}(t) \in \mathbb{T}^n$ ,  $\hat{\mathbf{x}}_k^T(t|s)$ , can be computed by projecting on the set of measurements  $\{\mathbf{y}_k(1), \dots, \mathbf{y}_k(s)\}$ , for  $k = 1, 2$ . Thereby,  $\mathbb{T}_k$  estimators are obtained that have the same performance as TWL estimators but with a lower computational complexity. More importantly, this computational load saving cannot be achieved with the real approach.

Note that tessarine algebra is not a Hilbert space and, as a consequence, neither the existence nor the uniqueness of the projection on a set of tessarines is guaranteed. Nevertheless, this drawback has been overcome in [22] by defining a suitable metric which assures the existence and uniqueness of these projections.

The following property sets the correlations between the noises,  $\tilde{\mathbf{u}}(t)$  and  $\tilde{\mathbf{v}}(t)$ , and both the augmented state  $\tilde{\mathbf{x}}(t)$  and the observations  $\mathbf{y}_k(t)$ .

**Property 3.** Under Assumptions 1-4, the following correlations hold:

1. Correlations between noises and the augmented state:
  - (a)  $E[\tilde{\mathbf{x}}(t+1)\tilde{\mathbf{u}}^H(t)] = \tilde{\mathbf{Q}}(t);$
  - (b)  $E[\tilde{\mathbf{x}}(t)\tilde{\mathbf{u}}^H(s)] = \mathbf{0}_{4n \times 4n},$  for  $t \leq s;$
  - (c)  $E[\tilde{\mathbf{x}}(t+1)\tilde{\mathbf{v}}^H(t)] = \tilde{\mathbf{S}}(t);$
  - (d)  $E[\tilde{\mathbf{x}}(t)\tilde{\mathbf{v}}^H(s)] = \mathbf{0}_{4n \times 4Rn},$  for  $t \leq s.$
2. Correlations between noises and  $\mathbb{T}_k$  observations:
  - (a)  $E[\mathbf{y}_k(t)\tilde{\mathbf{u}}^H(t)] = \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t)\tilde{\mathbf{S}}^H(t);$
  - (b)  $E[\mathbf{y}_k(t+1)\tilde{\mathbf{u}}^H(t)] = \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t+1)\tilde{\mathcal{C}}\tilde{\mathbf{Q}}(t) + \tilde{\mathbf{\Pi}}_k^{(1-\gamma)}(t+1)\tilde{\mathbf{S}}^H(t);$
  - (c)  $E[\mathbf{y}_k(t)\tilde{\mathbf{u}}^H(s)] = \mathbf{0}_{kRn \times 4n},$  for  $t < s;$
  - (d)  $E[\mathbf{y}_k(t)\tilde{\mathbf{v}}^H(t)] = \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t)\tilde{\mathbf{R}}(t);$
  - (e)  $E[\mathbf{y}_k(t+1)\tilde{\mathbf{v}}^H(t)] = \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t+1)\tilde{\mathcal{C}}\tilde{\mathbf{S}}(t) + \tilde{\mathbf{\Pi}}_k^{(1-\gamma)}(t+1)\tilde{\mathbf{R}}(t);$
  - (f)  $E[\mathbf{y}_k(t)\tilde{\mathbf{v}}^H(s)] = \mathbf{0}_{kRn \times 4Rn},$  for  $t < s.$

**Remark 4.** Observe that, under a  $\mathbb{T}_k$ -properness setting, the state equation in (7) is equivalent to the  $\mathbb{T}_k$  state equation:

$$\mathbf{x}_k(t+1) = \mathbf{\Phi}_k(t)\mathbf{x}_k(t) + \mathbf{u}_k(t), \quad t \geq 0 \quad (9)$$

where,

- in a  $\mathbb{T}_1$ -proper scenario,  $\mathbf{x}_1(t) \triangleq \mathbf{x}(t)$ ,  $\mathbf{u}_1(t) \triangleq \mathbf{u}(t)$ , and  $\mathbf{\Phi}_1(t) \triangleq \mathbf{F}_1(t);$
- in a  $\mathbb{T}_2$ -proper scenario,  $\mathbf{x}_2(t) \triangleq [\mathbf{x}^T(t), \mathbf{x}^H(t)]^T$ ,  $\mathbf{u}_2(t) \triangleq [\mathbf{u}^T(t), \mathbf{u}^H(t)]^T$  and  $\mathbf{\Phi}_2(t)$  is as in (6).



In such cases,  $\mathbf{Q}_k(t) = E[\mathbf{u}_k(t)\mathbf{u}_k^H(t)]$  and  $\mathbf{S}_k(t) = E[\mathbf{u}_k(t)\mathbf{v}_k^H(t)]$ , for  $k = 1, 2$ , where  $\mathbf{v}_1(t) \triangleq \bar{\mathbf{v}}(t)$  and  $\mathbf{v}_2(t) \triangleq [\bar{\mathbf{v}}^T(t), \bar{\mathbf{v}}^H(t)]^T$ , with  $\bar{\mathbf{v}}(t) = [\mathbf{v}^{(1)T}(t), \dots, \mathbf{v}^{(R)T}(t)]^T$ .

Nevertheless, equation (9) cannot be used together with the observation equation (8), since the latter involves the augmented state vector  $\bar{\mathbf{x}}(t)$ .

#### 4. $\mathbb{T}_k$ -proper centralized fusion estimation algorithms

In this section, the  $\mathbb{T}_k$  centralized fusion filter, prediction, and fixed-point smoothing algorithms are designed on the basis of the set of observations  $\{\mathbf{y}_k(1), \dots, \mathbf{y}_k(s)\}$ ,  $k = 1, 2$ , defined in (8).

With this purpose in mind, the observation equation (8) is used to devise filtering, prediction, and smoothing algorithms for the augmented state vector  $\bar{\mathbf{x}}(t)$ . Then, by applying  $\mathbb{T}_k$ -properness properties, the recursive formulas for the filtering, prediction, and smoothing estimators of  $\mathbf{x}_k(t)$  are easily determined. Finally, the desired  $\mathbb{T}_k$  centralized fusion filtering, prediction and fixed-point smoothing estimators are obtained as a subvector of them.

Theorems 1-3 summarize the recursive formulas for the computation of these  $\mathbb{T}_k$  estimators as well as their associated error variances.

##### 4.1. $\mathbb{T}_k$ centralized fusion filter

**Theorem 1.** The optimal  $\mathbb{T}_k$  centralized fusion filter  $\hat{\mathbf{x}}^{T_k}(t|t)$  and one-step predictor  $\hat{\mathbf{x}}^{T_k}(t+1|t)$  for the state  $\mathbf{x}(t)$  are obtained by extracting the first  $n$  fcomponents of the optimal estimator  $\hat{\mathbf{x}}_k(t|t)$  and  $\hat{\mathbf{x}}_k(t+1|t)$ , respectively, which are recursively computed from the expressions

$$\hat{\mathbf{x}}_k(t|t) = \hat{\mathbf{x}}_k(t|t-1) + \mathbf{L}_k(t)\boldsymbol{\varepsilon}_k(t), \quad t \geq 1 \quad (10)$$

$$\hat{\mathbf{x}}_k(t+1|t) = \boldsymbol{\Phi}_k(t)\hat{\mathbf{x}}_k(t|t) + \mathbf{H}_k(t)\boldsymbol{\varepsilon}_k(t), \quad t \geq 1 \quad (11)$$

with  $\hat{\mathbf{x}}_k(0|0) = \mathbf{0}_{kn}$  and  $\hat{\mathbf{x}}_k(1|0) = \mathbf{0}_{kn}$ , and where  $\mathbf{H}_k(t) = \mathbf{S}_k(t)\boldsymbol{\Pi}_k(t)\boldsymbol{\Omega}_k^{-1}(t)$ , with  $\boldsymbol{\Pi}_k(t) = \text{diag}(\boldsymbol{\Pi}_k^{(1)}(t), \dots, \boldsymbol{\Pi}_k^{(R)}(t))$  and  $\boldsymbol{\Pi}_k^{(i)}(t)$ ,  $i = 1, \dots, R$ , defined in Remark 2 for  $k = 1, 2$ .

Moreover,  $\boldsymbol{\varepsilon}_k(t)$  are the innovations calculated as follows

$$\boldsymbol{\varepsilon}_k(t) = \mathbf{y}_k(t) - \boldsymbol{\Pi}_k(t)\mathbf{C}_k\hat{\mathbf{x}}_k(t|t-1) - (\mathbf{I}_m - \boldsymbol{\Pi}_k(t))\mathbf{C}_k\hat{\mathbf{x}}_k(t-1|t-1) - (\mathbf{I}_m - \boldsymbol{\Pi}_k(t))\mathbf{G}_k(t-1)\boldsymbol{\varepsilon}_k(t-1), \quad t \geq 1 \quad (12)$$

with  $m = kRn$ ,  $\boldsymbol{\varepsilon}_k(0) = \mathbf{0}_m$ , and where  $\mathbf{C}_k = \mathbf{1}_R \otimes \mathbf{I}_{kn}$ ,  $\mathbf{G}_k(t) = \mathbf{R}_k(t)\boldsymbol{\Pi}_k(t)\boldsymbol{\Omega}_k^{-1}(t)$ , with  $\mathbf{R}_k(t) = E[\mathbf{v}_k(t)\mathbf{v}_k^H(t)]$ .

In addition,  $\mathbf{L}_k(t) = \boldsymbol{\Theta}_k(t)\boldsymbol{\Omega}_k^{-1}(t)$ , where  $\boldsymbol{\Theta}(t)$  is computed through the equation

$$\begin{aligned} \boldsymbol{\Theta}_k(t) = & \mathbf{P}_k(t|t-1)\mathbf{C}_k^T\boldsymbol{\Pi}_k(t) + \boldsymbol{\Phi}_k(t-1)\mathbf{P}_k(t-1|t-1)\mathbf{C}_k^T(\mathbf{I}_m - \boldsymbol{\Pi}_k(t)) \\ & + \mathbf{S}_k(t-1)(\mathbf{I}_m - \boldsymbol{\Pi}_k(t)) - \mathbf{H}_k(t-1)\boldsymbol{\Theta}_k^H(t-1)\mathbf{C}_k^T(\mathbf{I}_m - \boldsymbol{\Pi}_k(t)) \\ & - \boldsymbol{\Phi}_k(t-1)\boldsymbol{\Theta}_k(t-1)\mathbf{G}_k^H(t-1)(\mathbf{I}_m - \boldsymbol{\Pi}_k(t)) \\ & - \mathbf{H}_k(t-1)\boldsymbol{\Omega}_k(t-1)\mathbf{G}_k^H(t-1)(\mathbf{I}_m - \boldsymbol{\Pi}_k(t)), \quad t > 1 \end{aligned} \quad (13)$$

with  $\boldsymbol{\Theta}_k(1) = \mathbf{P}_k(1|0)\mathbf{C}_k^T\boldsymbol{\Pi}_k(1) + \boldsymbol{\Phi}_k(0)\mathbf{P}_k(0|0)\mathbf{C}_k^T(\mathbf{I}_m - \boldsymbol{\Pi}_k(1)) + \mathbf{S}_k(0)(\mathbf{I}_m - \boldsymbol{\Pi}_k(1))$ , and the innovations covariance matrix  $\boldsymbol{\Omega}_k(t)$  is obtained as

$$\begin{aligned} \boldsymbol{\Omega}_k(t) = & \mathcal{M}_k^1(t) - \mathcal{M}_k^2(t) - \mathcal{M}_k^3(t) + \mathcal{M}_k^4(t) + \boldsymbol{\Pi}_k(t)\mathbf{C}_k\mathbf{P}_k(t|t-1)\mathbf{C}_k^T\boldsymbol{\Pi}_k(t) \\ & + \boldsymbol{\Pi}_k(t)\mathbf{J}_k(t-1)(\mathbf{I}_m - \boldsymbol{\Pi}_k(t)) + (\mathbf{I}_m - \boldsymbol{\Pi}_k(t))\mathbf{J}_k^H(t-1)\boldsymbol{\Pi}_k(t) \\ & + (\mathbf{I}_m - \boldsymbol{\Pi}_k(t))[\mathbf{C}_k\mathbf{P}_k(t-1|t-1)\mathbf{C}_k^T - \mathbf{C}_k\boldsymbol{\Theta}_k(t-1)\mathbf{G}_k^H(t-1) - \mathbf{G}_k(t-1)\boldsymbol{\Theta}_k^H(t-1)\mathbf{C}_k^T \\ & - \mathbf{G}_k(t-1)\boldsymbol{\Omega}_k(t-1)\mathbf{G}_k^H(t-1)](\mathbf{I}_m - \boldsymbol{\Pi}_k(t)), \quad t > 1 \end{aligned} \quad (14)$$

with

$$\begin{aligned}\Omega_k(1) = & \mathcal{M}_k^1(1) - \mathcal{M}_k^2(1) - \mathcal{M}_k^3(1) + \mathcal{M}_k^4(1) + \Pi_k(1)\mathcal{C}_k\mathbf{P}_k(1|0)\mathcal{C}_k^T\Pi_k(1) \\ & + \Pi_k(1)\mathbf{J}_k(0)(\mathbf{I}_m - \Pi_k(1)) + (\mathbf{I}_m - \Pi_k(1))\mathbf{J}_k^H(0)\mathcal{C}_k^T\Pi_k(1) \\ & + (\mathbf{I}_m - \Pi_k(1))\mathcal{C}_k\mathbf{P}_k(0|0)\mathcal{C}_k^T(\mathbf{I}_m - \Pi_k(1)), \quad (15)\end{aligned}$$

where

$$\begin{aligned}\mathbf{J}_k(t) = & \mathcal{C}_k[\Phi_k(t)\mathbf{P}_k(t|t)\mathcal{C}_k^T - \mathbf{H}_k(t)\Theta_k^H(t)\mathcal{C}_k^T + \mathbf{S}_k(t) - \Phi_k(t)\Theta_k(t)\mathbf{G}_k^H(t) \\ & - \mathbf{H}_k(t)\Omega_k(t)\mathbf{G}_k^H(t)],\end{aligned}$$

with  $\mathbf{J}_k(0) = \mathcal{C}_k[\Phi_k(0)\mathbf{P}_k(0|0)\mathcal{C}_k^T + \mathbf{S}_k(0)]$ , and

$$\begin{aligned}\mathcal{M}_k^1(t) &= \mathcal{L}_k\left\{\text{Cov}(\tilde{\gamma}^r(t)) \circ \left(\tilde{\mathcal{L}}^H\tilde{\mathcal{C}}\tilde{\Sigma}(t-1)\tilde{\mathcal{C}}^T\tilde{\mathcal{L}}\right)\right\}\mathcal{L}_k^H, \\ \mathcal{M}_k^2(t) &= \mathcal{L}_k\left\{\text{Cov}(\tilde{\gamma}^r(t)) \circ \left(\tilde{\mathcal{L}}^H\tilde{\mathcal{C}}\tilde{\mathbf{S}}(t)\tilde{\mathcal{L}}\right)\right\}\mathcal{L}_k^H, \\ \mathcal{M}_k^3(t) &= \mathcal{L}_k\left\{\text{Cov}(\tilde{\gamma}^r(t)) \circ \left(\tilde{\mathcal{L}}^H\tilde{\mathbf{S}}^H(t)\tilde{\mathcal{C}}^T\tilde{\mathcal{L}}\right)\right\}\mathcal{L}_k^H, \\ \mathcal{M}_k^4(t) &= \mathcal{L}_k\left\{\Delta^{p^r}(t) \circ \left(\tilde{\mathcal{L}}^H\tilde{\mathbf{R}}(t)\tilde{\mathcal{L}}\right) + \Delta^{1-p^r}(t) \circ \left(\tilde{\mathcal{L}}^H\tilde{\mathbf{R}}(t-1)\tilde{\mathcal{L}}\right)\right\}\mathcal{L}_k^H,\end{aligned}$$

with  $\Delta^{p^r}(t) = E[\tilde{\gamma}^r(t)\tilde{\gamma}^{r^T}(t)]$ ,  $\Delta^{1-p^r}(t) = E[(\mathbf{1}_{4Rn} - \tilde{\gamma}^r(t))(\mathbf{1}_{4Rn} - \tilde{\gamma}^r(t))^T]$ , whose entries are given in (4), and

$$\tilde{\Sigma}(t) = \left[\tilde{\Phi}(t)\tilde{\mathbf{D}}(t)\tilde{\Phi}^H(t) + \tilde{\mathbf{Q}}(t) - \tilde{\Phi}(t)\tilde{\mathbf{D}}(t) - \tilde{\mathbf{D}}(t)\tilde{\Phi}^H(t) + \tilde{\mathbf{D}}(t)\right],$$

where  $\tilde{\mathbf{D}}(t) = \mathbf{R}_{\tilde{\mathbf{x}}}(t, t)$  is recursively computed from

$$\tilde{\mathbf{D}}(t) = \tilde{\Phi}(t-1)\tilde{\mathbf{D}}(t-1)\tilde{\Phi}^H(t-1) + \tilde{\mathbf{Q}}(t-1). \quad (16)$$

Finally, the  $\mathbb{T}_k$  filtering and prediction error pseudo covariance matrices  $\mathbf{P}^{\mathbb{T}_k}(t|t)$  and  $\mathbf{P}^{\mathbb{T}_k}(t+1|t)$ , respectively, are obtained from the filtering and prediction error pseudo covariance matrices  $\mathbf{P}_k(t|t)$  and  $\mathbf{P}_k(t+1|t)$ , calculated from the recursive expressions

$$\mathbf{P}_k(t|t) = \mathbf{P}_k(t|t-1) - \Theta_k(t)\Omega_k^{-1}(t)\Theta_k^H(t), \quad (17)$$

with  $\mathbf{P}_k(0|0) = E[\mathbf{x}_k(0)\mathbf{x}_k^H(0)]$ , and

$$\begin{aligned}\mathbf{P}_k(t+1|t) = & \Phi_k(t)\mathbf{P}_k(t|t)\Phi_k^H(t) - \mathbf{H}_k(t)\Theta_k^H(t)\Phi_k^H(t) \\ & - \Phi_k(t)\Theta_k(t)\mathbf{H}_k^H(t) - \mathbf{H}_k(t)\Omega_k(t)\mathbf{H}_k^H(t) + \mathbf{Q}_k(t), \quad (18)\end{aligned}$$

with  $\mathbf{P}_k(1|0) = \Phi_k(0)\mathbf{P}_k(0|0)\Phi_k^H(0) + \mathbf{Q}_k(0)$ .

**Remark 5.** In the implementation of the above algorithm, the particular structure of  $\tilde{\Sigma}(t)$  under  $\mathbb{T}_k$ -properness conditions should be taken into consideration. In this regard, it is not difficult to check that  $\tilde{\Sigma}(t)$  is a block diagonal matrix of the form

- $\mathbb{T}_1$ -properness:  $\tilde{\Sigma}(t) = \text{diag}\left(\Sigma_1(t), \Sigma_1^*(t), \Sigma_1^\eta(t), \Sigma_1^{\eta''}(t)\right)$ ;
- $\mathbb{T}_2$ -properness:  $\tilde{\Sigma}(t) = \text{diag}\left(\Sigma_2(t), \Sigma_2^\eta(t)\right)$ .

with  $\Sigma_k(t) = \Phi_k(t)\mathbf{D}_k(t)\Phi_k^H(t) + \mathbf{Q}_k(t) - \Phi_k(t)\mathbf{D}_k(t) - \mathbf{D}_k(t)\Phi_k^H(t) + \mathbf{D}_k(t)$ ,  $k = 1, 2$ , where  $\mathbf{D}_k(t) = \mathbf{R}_{\mathbf{x}_k}(t, t)$  is recursively computed from

$$\mathbf{D}_k(t) = \Phi_k(t-1)\mathbf{D}_k(t-1)\Phi_k^H(t-1) + \mathbf{Q}_k(t-1).$$

#### 4.2. $\mathbb{T}_k$ centralized fusion predictor

**Theorem 2.** The optimal  $\mathbb{T}_k$  centralized fusion predictor  $\hat{\mathbf{x}}^{\mathbb{T}_k}(t + \tau|t)$  for the state  $\mathbf{x}(t)$  is obtained by extracting the first  $n$  components of the optimal estimator  $\hat{\mathbf{x}}_k(t + \tau|t)$  which is recursively computed from the expression

$$\hat{\mathbf{x}}_k(t + \tau|t) = \Phi_k(t + \tau - 1)\hat{\mathbf{x}}_k(t + \tau - 1|t), \quad \tau \geq 2 \quad (19)$$

with the initialization the one-step predictor  $\hat{\mathbf{x}}_k(t + 1|t)$  given by (11).

Moreover, the  $\mathbb{T}_k$ -proper prediction error pseudo covariance matrix  $\mathbf{P}^{\mathbb{T}_k}(t + \tau|t)$  is obtained from the prediction error pseudo covariance matrix  $\mathbf{P}_k(t + \tau|t)$ , computed from the recursive expression

$$\mathbf{P}_k(t + \tau|t) = \Phi_k(t + \tau - 1)\mathbf{P}_k(t + \tau - 1|t)\Phi_k^H(t + \tau - 1) + \mathbf{Q}_k(t + \tau - 1), \quad \tau \geq 2 \quad (20)$$

with the initialization the one-step prediction error pseudo covariance matrix given by (18).

#### 4.3. $\mathbb{T}_k$ centralized fusion smoother

**Theorem 3.** The optimal  $\mathbb{T}_k$  centralized fusion fixed-point smoother  $\hat{\mathbf{x}}^{\mathbb{T}_k}(t|s)$ , for a fixed instant  $t < s$ , for the state  $\mathbf{x}(t)$  is obtained by extracting the  $n$  first components of the optimal estimator  $\hat{\mathbf{x}}_k(t|s)$  which is recursively computed from the expressions

$$\hat{\mathbf{x}}_k(t|s) = \hat{\mathbf{x}}_k(t|s - 1) + \mathbf{L}_k(t, s)\mathbf{e}_k(s), \quad t \geq 1 \quad (21)$$

with initial condition  $\hat{\mathbf{x}}_k(t|t)$  given by (10), and where the innovations  $\mathbf{e}_k(s)$  are recursively computed from (12) and  $\mathbf{L}_k(t, s) = \mathbf{\Theta}_k(t, s)\mathbf{\Omega}_k^{-1}(s)$  with  $\mathbf{\Omega}_k^{-1}(s)$  obtained from the recursive expression (14) and

$$\begin{aligned} \mathbf{\Theta}_k(t, s) = & [\mathbf{E}_k(t, s - 1)\Phi_k^H(s - 1) - \mathbf{\Theta}_k(t, s - 1)\mathbf{H}_k^H(s - 1)]\mathbf{C}_k^T\mathbf{\Pi}_k(s) \\ & + [\mathbf{E}_k(t, s - 1)\mathbf{C}_k^T - \mathbf{\Theta}_k(t, s - 1)\mathbf{G}_k^H(s - 1)](\mathbf{I}_m - \mathbf{\Pi}_k(s)), \end{aligned} \quad (22)$$

$$\begin{aligned} \mathbf{E}_k(t, s) = & [\mathbf{E}_k(t, s - 1)\Phi_k^H(s - 1) - \mathbf{\Theta}_k(t, s - 1)\mathbf{H}_k^H(s - 1)](\mathbf{I} - \mathbf{C}_k^T\mathbf{\Pi}_k(s)\mathbf{L}_k^H(s)) \\ & - [\mathbf{E}_k(t, s - 1)\mathbf{C}_k^T - \mathbf{\Theta}_k(t, s - 1)\mathbf{G}_k^H(s - 1)](\mathbf{I}_m - \mathbf{\Pi}_k(s))\mathbf{L}_k^H(s), \end{aligned} \quad (23)$$

with initialization  $\mathbf{\Theta}_k(t, t) = \mathbf{\Theta}_k(t)$  given by (13) and  $\mathbf{E}_k(t, t) = \mathbf{P}_k(t|t)$ .

Furthermore, the  $\mathbb{T}_k$  fixed-point smoothing error pseudo covariance matrix is recursively computed through the expression

$$\mathbf{P}_k(t|s) = \mathbf{P}_k(t|s - 1) - \mathbf{\Theta}_k(t, s)\mathbf{\Omega}_k^{-1}(s)\mathbf{\Theta}_k^H(t, s), \quad (24)$$

with  $\mathbf{P}_k(t|t)$  the filtering error pseudo covariance matrix (17).

As mentioned above, the main advantage of the proposed  $\mathbb{T}_k$  centralized fusion algorithms is that the resulting  $\mathbb{T}_k$  centralized fusion estimators coincide with the optimal TWL counterparts; meanwhile, they lead to computational savings with respect to the one derived from a TWL approach.

## 5. Simulation examples

In this section, the effectiveness of the above  $\mathbb{T}_k$ -proper centralized fusion estimation algorithms is experimentally analyzed. With this aim, the following simulation examples have been chosen to reveal the superiority of the proposed  $\mathbb{T}_k$ -proper estimators over their counterparts in the quaternion domain, when  $\mathbb{T}_k$ -properness conditions are present.

Let us consider the following tessarine system with three sensors:

$$\begin{aligned}x(t+1) &= f_1 x(t) + u(t) \\ z^{(i)}(t) &= x(t) + v^{(i)}(t), \quad i = 1, 2, 3 \\ y^{(i)}(t) &= \gamma^{(i)}(t) \star z^{(i)}(t) + (1 - \gamma^{(i)}(t)) \star z^{(i)}(t-1), \quad i = 1, 2, 3\end{aligned}$$

with  $f_1 = 0.9 - 0.3\eta + 0.02\eta' + 0.1\eta'' \in \mathbb{T}$ . The following assumptions are made on the initial state and additive noises:

1. The initial state  $x_0$  is a tessarine Gaussian variable determined by the following real covariance matrix:

$$E[\mathbf{x}^r(0)\mathbf{x}^{r^T}(0)] = \begin{bmatrix} a & 0 & -2.5 & 0 \\ 0 & 4 & 0 & -2.5 \\ -2.5 & 0 & a & 0 \\ 0 & -2.5 & 0 & 4 \end{bmatrix}. \quad (25)$$

2.  $u(t)$  is a tessarine white Gaussian noise with a real covariance matrix

$$E[\mathbf{u}^r(t)\mathbf{u}^{r^T}(s)] = \begin{bmatrix} 0.9 & 0 & c & 0 \\ 0 & b & 0 & c \\ c & 0 & 0.9 & 0 \\ 0 & c & 0 & b \end{bmatrix} \delta_{t,s}. \quad (26)$$

3. The measurement noises  $v^{(i)}(t)$  of the three sensors are tessarine white Gaussian noises defined as follows:

$$v^{(i)}(t) = \alpha_i u(t) + w^{(i)}(t),$$

where the coefficients  $\alpha_i$  are the constant scalars<sup>1</sup>  $\alpha_1 = 0.5$ ,  $\alpha_2 = 0.8$ , and  $\alpha_3 = 0.4$  and  $w^{(i)}(t)$ ,  $i = 1, 2, 3$ , are  $\mathbb{T}_1$ -proper tessarine white Gaussian noises with mean zeros and real covariance matrices

$$E[\mathbf{w}^{(i)r}(t)\mathbf{w}^{(i)r^T}(s)] = \begin{bmatrix} \beta_i & 0 & 0 & 0 \\ 0 & \beta_i & 0 & 0 \\ 0 & 0 & \beta_i & 0 \\ 0 & 0 & 0 & \beta_i \end{bmatrix} \delta_{t,s},$$

with  $\beta_1 = 4$ ,  $\beta_2 = 8$ , and  $\beta_3 = 25$ , and independent of  $u(t)$ .

Moreover, at every sensor  $i$ , the Bernoulli random variables  $\gamma_v^{(i)}(t)$ ,  $v = r, \eta, \eta', \eta''$ , have the constant probabilities  $P[\gamma_v^{(i)}(t) = 1] = p_v^{(i)}$ , for all  $t \in T$ .

In this framework, a comparative study between tessarine and quaternion approaches is carry out to evaluate the performance of the proposed filtering, prediction and smoothing algorithms under  $\mathbb{T}_1$  and  $\mathbb{T}_2$  properness conditions.

### 5.1. $\mathbb{T}_1$ -Proper case

Consider the values  $a = 4$  in (25) and  $b = 0.9$  and  $c = 0.3$  in (26), and the Bernoulli probabilities

- $p_r^{(1)} = p_\eta^{(1)} = p_{\eta'}^{(1)} = p_{\eta''}^{(1)} = p_1$ ;
- $p_r^{(2)} = p_\eta^{(2)} = p_{\eta'}^{(2)} = p_{\eta''}^{(2)} = p_2$ ;
- $p_r^{(3)} = p_\eta^{(3)} = p_{\eta'}^{(3)} = p_{\eta''}^{(3)} = p_3$ .

<sup>1</sup> If  $\alpha_i = 0$ , then the noises  $u(t)$  and  $v^{(i)}(t)$  are uncorrelated. In the opposite case, when  $\alpha_i$  becomes more different from 0, the correlation between  $u(t)$  and  $v^{(i)}(t)$  is stronger.

Note that, under these conditions, both  $x(t)$  and  $y^{(i)}(t)$ ,  $i = 1, 2, 3$ , are jointly  $\mathbb{T}_1$ -proper.

For the purpose of comparison, the errors variances of both  $\mathbb{T}_1$  and QSL estimators have been computed for different Bernoulli probabilities  $p_i$ ,  $i = 1, 2, 3$ . We denote the QSL error variances by  $\mathbf{P}_{QSL}(t|s)$ .

By way of illustration, we consider the same delay probabilities in the three sensors ( $p^{(1)} = p^{(2)} = p^{(3)} = p$ ) and, as a performance measure, we compute the difference between the  $\mathbb{T}_1$  and QSL error variances associated to the filter,  $DE_1(t|t) = \mathbf{P}_{QSL}(t|t) - \mathbf{P}_1(t|t)$ , the 3-step predictor,  $DE_1(t+3|t) = \mathbf{P}_{QSL}(t+3|t) - \mathbf{P}_1(t+3|t)$ , and the fixed-point smoother at  $t = 20$ ,  $DE_1(20|t) = \mathbf{P}_{QSL}(20|t) - \mathbf{P}_1(20|t)$ , for  $t > 20$ .

In Fig. 1, these differences are displayed for the Bernoulli probabilities  $p = 0, 0.2, 0.4, 0.6, 0.8, 1$ . In these figures, the superiority in performance of  $\mathbb{T}_1$  estimators over QSL estimators is confirmed since  $DE_1 > 0$  in every case. Additionally, in the filtering and prediction problem, it is observed that this superiority is higher when the Bernoulli probabilities are smaller, i.e., when the delay probabilities are greater. On the other hand, in the case of the fixed-point smoothing problem, the difference between the associated  $\mathbb{T}_1$  and QSL error variances for any Bernoulli probability  $p$  is similar to the one obtained for Bernoulli probability  $1 - p$ , and the benefit of the  $\mathbb{T}_1$  algorithm over the QSL one is higher when Bernoulli probabilities approach to 0.5.

Next, we analyze in detail the benefits of our  $\mathbb{T}_1$  estimation algorithms in terms of the Bernoulli probabilities of the three sensors  $p$  and for different values of  $c$  in (26). With this aim, the means of the difference between the  $\mathbb{T}_1$  and QSL filtering, prediction and fixed-point smoothing error variances have been computed as

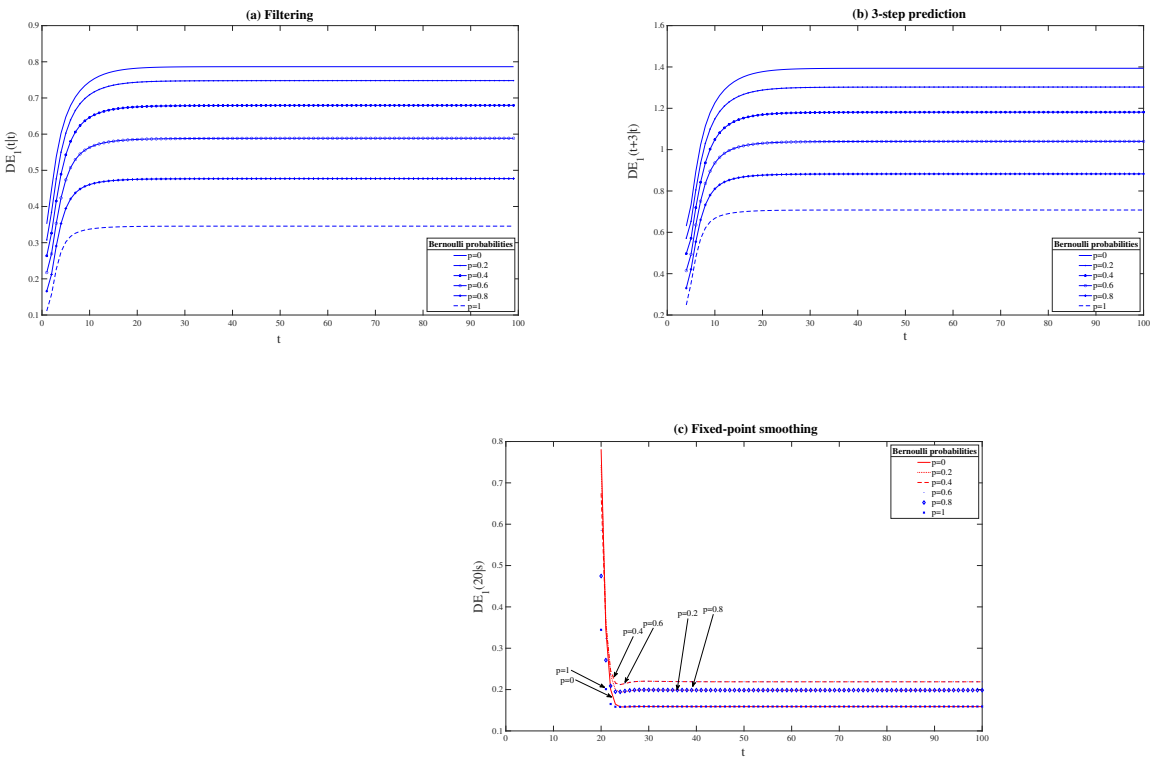
- Filtering problem:  $MDE_1^p(t|t) = \frac{1}{100} \sum_{t=1}^{100} DE_1^p(t|t)$ ;
- 3-step prediction problem:  $MDE_1^p(t+3|t) = \frac{1}{97} \sum_{t=1}^{97} DE_1^p(t+3|t)$ ;
- Fixed-point smoothing problem:  $MDE_1^p(20|t) = \frac{1}{80} \sum_{t=1}^{80} DE_1^p(20|t)$ .

for  $p$  varying from 0 to 1 and the values of  $c = 0, 0.1$  and  $0.2$ , where  $DE_1^p(t|t)$ ,  $DE_1^p(t+3|t)$  and  $DE_1^p(20|t)$  denote the difference between the  $\mathbb{T}_1$  and QSL filtering, 3-step prediction, and fixed-point smoothing error variances, respectively, for a value of the Bernoulli probability  $p$ . Note that in the case  $c = 0$ , the noise  $u(t)$  is, besides being  $\mathbb{T}_1$ -proper,  $\mathbb{Q}$ -proper, and a higher value of  $c$  means that the noise  $u(t)$  moves further away from the  $\mathbb{Q}$ -properness condition. The results of this analysis are depicted in Fig. 2 where, on the one hand, we can clearly observe how the best performance of  $\mathbb{T}_1$  filtering and prediction estimators is achieved when the Bernoulli probability takes the value 0, i.e., when only one-step delay exists in the measurements. However, in the fixed-point smoothing problem  $\mathbb{T}_1$  is more advantageous when the Bernoulli probability  $p$  tends to 0.5. On the other hand, in every case, the superiority of our  $\mathbb{T}_1$  estimation algorithms is more evident as the parameter  $c$  in (26) grows, i.e., the noise  $u(t)$  is further away from the  $\mathbb{Q}$ -properness condition.

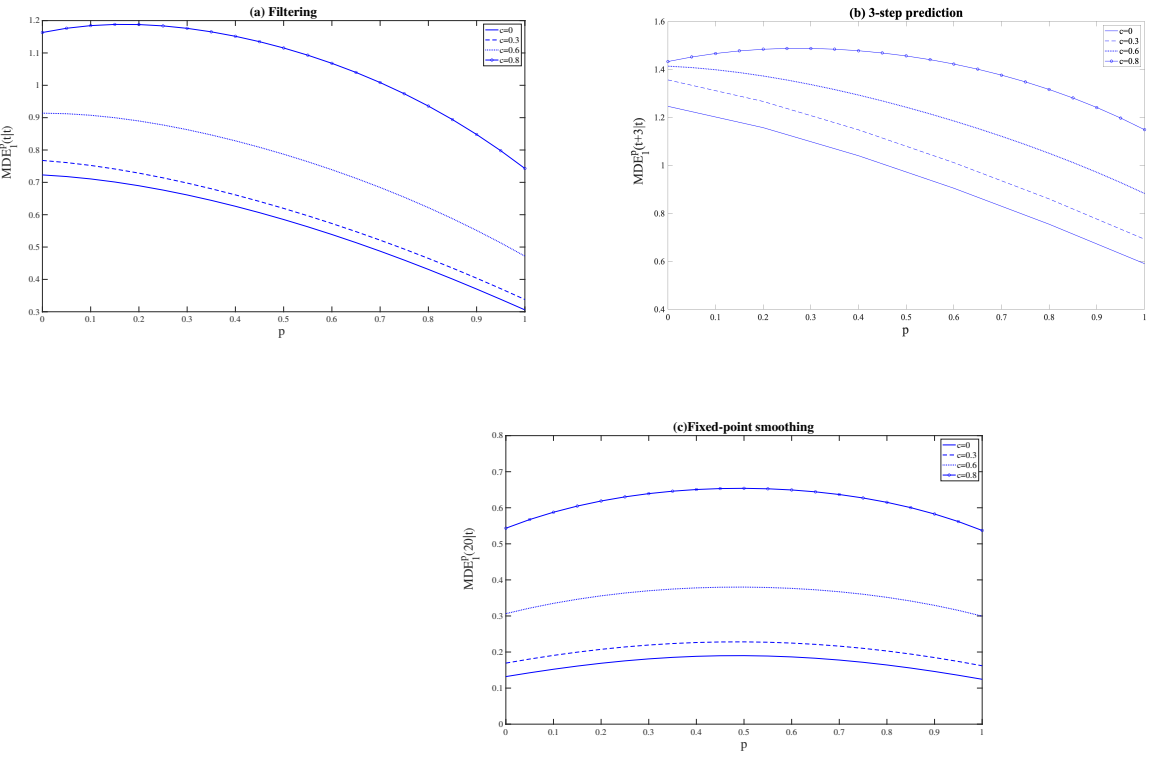
## 5.2. $\mathbb{T}_2$ -Proper case

Consider the values  $a = 6$  in (25),  $b = 0.3$  in (26), and the Bernoulli probabilities for the three sensors as in Section 5.1. Note that, under these conditions, both  $x(t)$  and  $y^{(i)}(t)$ ,  $i = 1, 2, 3$ , are jointly  $\mathbb{T}_2$ -proper.

Thus, we are interested in comparing the behavior of  $\mathbb{T}_2$  centralized fusion estimators with their counterparts in the quaternion domain, i.e., the quaternion semi-widely linear (QSWL) estimators. For this purpose, the  $\mathbb{T}_2$  and QSWL error variances,  $\mathbf{P}_2(t|s)$  and  $\mathbf{P}_{QSWL}(t|s)$ , respectively, have been computed by considering different Bernoulli probabilities for the three sensors.



**Figure 1.** Difference between QSL and  $T_1$  error variances for the problem of (a) filtering, (b) 3-step prediction and (c) fixed-point smoothing.



**Figure 2.** Mean of the difference between QSL and  $T_1$  error variances for the problem of (a) filtering, (b) 3-step prediction and (c) fixed-point smoothing.



Specifically, we consider the filtering, the 3-step prediction, and the fixed-point smoothing problems at  $t = 20$ , and, as a measure of comparison, we use the difference between both QSWL and  $\mathbb{T}_2$  error variances, which are defined as  $DE_2(t|t) = \mathbf{P}_{QSWL}(t|t) - \mathbf{P}_2(t|t)$  (filtering),  $DE_2(t+3|t) = \mathbf{P}_{QSWL}(t+3|t) - \mathbf{P}_2(t+3|t)$  (3-step prediction), and  $DE_2(20|t) = \mathbf{P}_{QSWL}(20|t) - \mathbf{P}_2(20|t)$  (fixed-point smoothing).

Figures 3-4 compare the difference between QSWL and  $\mathbb{T}_2$  centralized estimation error variances for different Bernoulli probabilities  $p_1$ ,  $p_2$  and  $p_3$ . Specifically, Fig. 3 analyzes the filtering and 3-step prediction error variance differences  $DE_2(t|t)$  and  $DE_2(t+3|t)$  for the following cases:

1. Case 1: for values of  $p_1 = 0.1, 0.5, 0.9$  in three situations:  $p_2 = 0.9$  and  $p_3 = 0.1$ ,  $p_2 = 0.1$  and  $p_3 = 0.9$ , and  $p_2 = p_3 = 0.5$ ;
2. Case 2: for values of  $p_2 = 0.1, 0.5, 0.9$  in three situations:  $p_1 = 0.9$  and  $p_3 = 0.1$ ,  $p_1 = 0.1$  and  $p_3 = 0.9$ , and  $p_1 = p_3 = 0.5$ ;
3. Case 3: for values of  $p_3 = 0.1, 0.5, 0.9$  in three situations:  $p_1 = 0.9$  and  $p_2 = 0.1$ ,  $p_1 = 0.1$  and  $p_2 = 0.9$ , and  $p_1 = p_2 = 0.5$ .

It should be highlighted that similar results are obtained with any other combination of Bernoulli probabilities  $p_i$ ,  $i = 1, 2, 3$ .

From these figures, we can reaffirm that  $\mathbb{T}_2$  processing is a better approach than the QSWL processing in terms of performance ( $DE_2 > 0$ ), and also this fact is more evident when the probabilities of the Bernoulli variables decreases (that is, the delays probabilities increase).

The differences between both QSWL and  $\mathbb{T}_2$  error variances for the fixed-point smoothing problem are illustrated in Fig. 4. Note that, since the behavior of the differences between QSWL and  $\mathbb{T}_2$  fixed-point smoothing errors is similar for Bernoulli probabilities values  $p_i$  and  $1 - p_i$ , these differences are analyzed in the following cases:

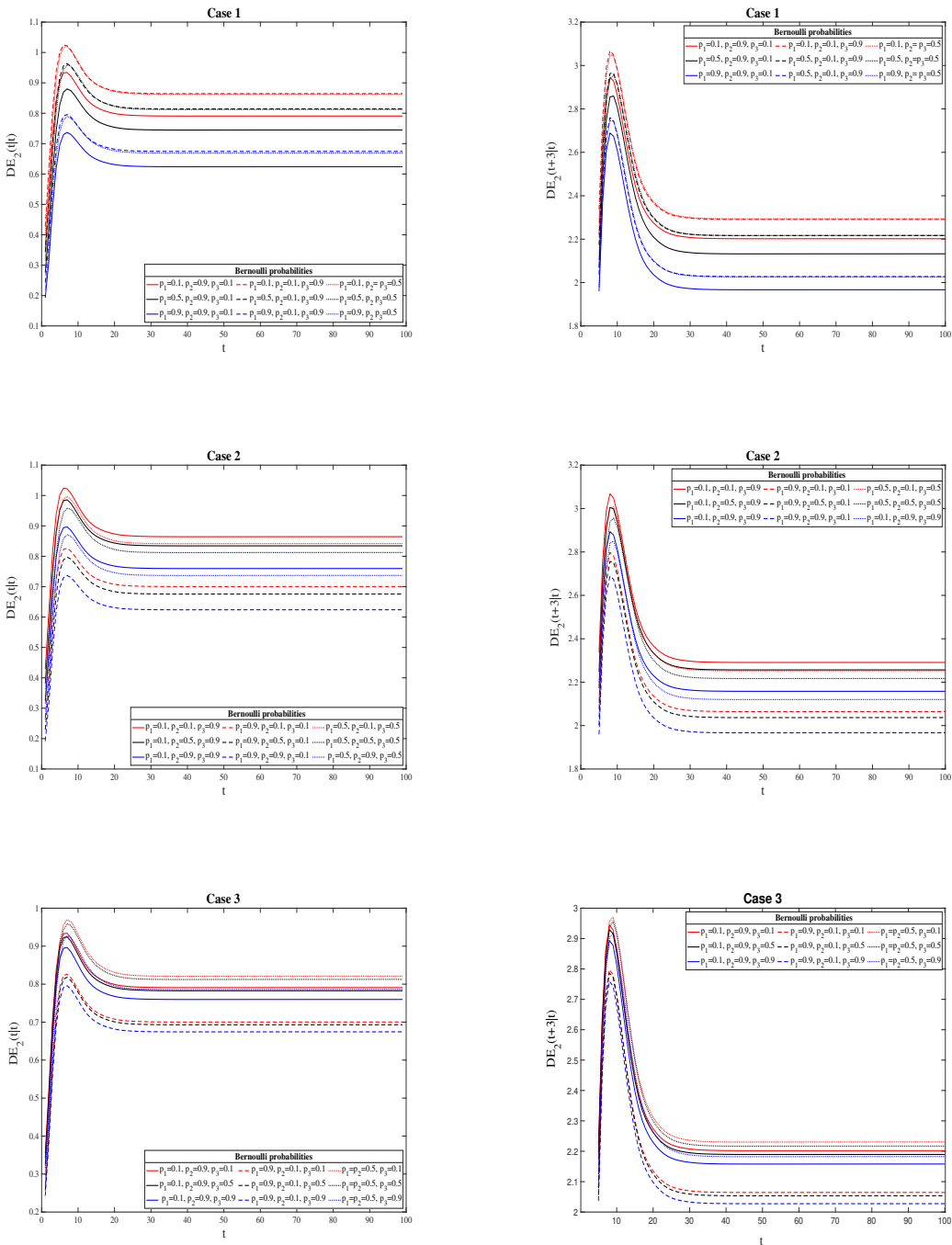
1. Case 4: for values of  $p_1 = 0.1, 0.3, 0.5$  in three situations:  $p_2 = 0.1$  and  $p_3 = 0.3$ ,  $p_2 = 0.3$  and  $p_3 = 0.1$ , and  $p_2 = p_3 = 0.3$ .
2. Case 5: for values of  $p_2 = 0.1, 0.3, 0.5$  in three situations:  $p_1 = 0.1$  and  $p_3 = 0.3$ ,  $p_1 = 0.3$  and  $p_3 = 0.1$ , and  $p_1 = p_3 = 0.3$ .
3. Case 6: for values of  $p_3 = 0.1, 0.3, 0.5$  in three situations:  $p_1 = 0.1$  and  $p_2 = 0.3$ ,  $p_1 = 0.3$  and  $p_2 = 0.1$ , and  $p_1 = p_2 = 0.3$ .

In every situation the better behavior of  $\mathbb{T}_2$  processing over the QSWL processing is verified and also this superiority increases when the Bernoulli probabilities tends to 0.5, i.e., when there is a similar chance of receiving updated and delayed information.

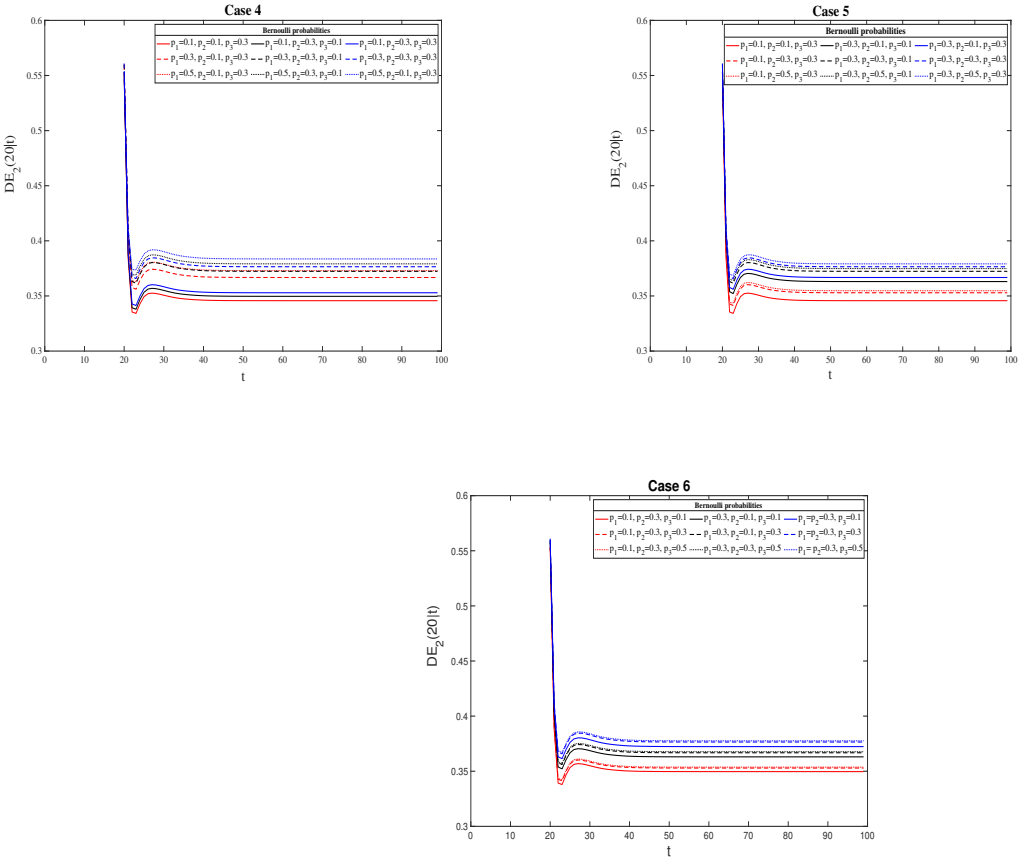
## 6. Discussion

From among the different sensor fusion methods, it is the centralized fusion techniques that provide the optimal estimators from measurements of all sensors. Nevertheless, to avoid the computational burden involved in these estimates, especially in systems with a large number of sensors, suboptimum estimation algorithms have been traditionally designed by using a decentralized fusion approach. This paper has overcome the above computational difficulties without renouncing to obtain the optimal solution, by considering hypercomplex algebras. Quaternions and, more recently, tessarines are the most common 4D hypercomplex algebra used in signal processing. Neither of these algebras always performs better than the other and the choice of the most suitable one is conditioned by the characteristics of the signal.

Due to the commutativity and reduced computational complexity, the tessarine algebra makes it particularly interesting for our purposes. Thus, under conditions of  $\mathbb{T}_k$ -properness, filtering, prediction, and fixed-point smoothing algorithms of reduced dimension have been devised for the estimation of a vectorial tessarine signal based



**Figure 3.** Difference between QSWL and  $\mathbb{T}_2$  error variances for the problem of filtering (left column) and 3-step prediction (right column) for Cases 1-3.



**Figure 4.** Difference between QSWL and  $\mathbb{T}_2$  error variances for the fixed-point smoothing problem for Cases 4-6.

on one-step randomly delayed observations coming from multiple sensors stochastic systems with different delay rates and correlated noises. The reduction of the dimension of the problem under  $\mathbb{T}_k$ -properness scenarios makes it possible for these algorithms to facilitate the computation of the optimal estimates with a lower computational cost in comparison with the real processing approach. It should be highlighted that this computational saving cannot be attained in the real field.

The good performance of the algorithms proposed has been experimentally illustrated by means of two simulation examples, where the better behavior of the proposed  $\mathbb{T}_k$  estimates over their counterparts in the quaternion domain under  $\mathbb{T}_k$ -properness conditions has been evidenced.

In future research, we will set out to explore the design of decentralized fusion estimation algorithms for hypercomplex signals and investigate the use of new hypercomplex algebras in this field.

**Author Contributions:** All authors have contributed equally to the work. Moreover, all of them have read and agreed to the published version of the manuscript.

**Funding:** This work has been supported in part by I+D+i Project with reference number 1256911, under Programa Operativo FEDER Andalucía 2014–2020, Junta de Andalucía, and Project EI\_FQM2\_2021 of ‘Plan de Apoyo a la Investigación 2021–2022 of the University of Jaén.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Appendix A Proof of Theorem 1

The proof is based on the innovation technique. Consider the one-state delay model:

$$\begin{aligned}\bar{\mathbf{x}}(t+1) &= \bar{\Phi}(t)\bar{\mathbf{x}}(t) + \bar{\mathbf{u}}(t), \quad t \geq 0 \\ \mathbf{y}_k(t) &= \tilde{\mathcal{D}}_k^\gamma(t)\bar{\mathcal{C}}\bar{\mathbf{x}}(t) + \tilde{\mathcal{D}}_k^{(1-\gamma)}(t)\bar{\mathcal{C}}\bar{\mathbf{x}}(t-1) + \tilde{\mathcal{D}}_k^\gamma(t)\bar{\mathbf{v}}(t) + \tilde{\mathcal{D}}_k^{(1-\gamma)}(t)\bar{\mathbf{v}}(t-1), \quad t \geq 1\end{aligned}\quad (\text{A1})$$

and define the innovations as  $\boldsymbol{\varepsilon}_k(t) = \mathbf{y}_k(t) - \hat{\mathbf{y}}_k(t|t-1)$ .

In order to simplify the proof of Theorem 1, the following results have been previously established.

### Appendix A.1 Preliminary results

The following property, stated without proof, about the correlations between the innovations  $\boldsymbol{\varepsilon}_k(t)$  and the augmented state  $\bar{\mathbf{x}}(t)$  and the noises  $\bar{\mathbf{u}}(t)$  and  $\bar{\mathbf{v}}_k(t)$ , will be useful in the proof of Theorem 1.

**Property A1.** *Given the system (A1), and under the Assumptions 1-4, the following correlations hold:*

1.  $E[\bar{\mathbf{u}}(t)\boldsymbol{\varepsilon}_k^H(t)] = \bar{\mathbf{S}}(t)\bar{\mathbf{\Pi}}_k^{\gamma H}(t).$
2.  $E[\bar{\mathbf{u}}(t)\boldsymbol{\varepsilon}_k^H(s)] = \mathbf{0}_{4n \times m},$  for  $t > s.$
3.  $E[\bar{\mathbf{v}}(t)\boldsymbol{\varepsilon}_k^H(t)] = \bar{\mathbf{R}}(t)\bar{\mathbf{\Pi}}_k^{\gamma H}(t).$
4.  $E[\bar{\mathbf{v}}(t)\boldsymbol{\varepsilon}_k^H(s)] = \mathbf{0}_{4Rn \times m},$  for  $t > s.$

Moreover, the following results will be of interest in the derivation of the formulas given in Theorem 1.

**Lemma A1.** *Denote  $\Delta\tilde{\mathcal{D}}_k^\gamma(t) = \tilde{\mathcal{D}}_k^\gamma(t) - \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t)$  and  $\Delta\tilde{\mathcal{D}}_k^{(1-\gamma)}(t) = \tilde{\mathcal{D}}_k^{(1-\gamma)}(t) - \tilde{\mathbf{\Pi}}_k^{(1-\gamma)}(t)$ . For any tessarine random vectors  $\boldsymbol{\alpha}_1(t), \boldsymbol{\alpha}_2(t) \in \mathbb{T}^{4Rn}$  and  $\boldsymbol{\beta}(t) \in \mathbb{T}^q$ , for any dimension  $q$ , the following relations hold:*

1.  $E[\Delta \tilde{\mathcal{D}}_k^\gamma(t) \alpha_1(t) \alpha_2^H(s) \Delta \tilde{\mathcal{D}}_k^{\gamma^H}(t)] = \mathcal{L}_k \left\{ \text{Cov}(\tilde{\gamma}^r(t)) \circ \left( \tilde{\mathcal{L}}^H E[\alpha_1(t) \alpha_2^H(s)] \tilde{\mathcal{L}} \right) \right\} \mathcal{L}_k^H.$
2.  $E[\Delta \tilde{\mathcal{D}}_k^\gamma(t) \alpha_1(t) \alpha_2^H(s) \tilde{\mathcal{D}}_k^{\gamma^H}(t)] = \mathcal{L}_k \left\{ \text{Cov}(\tilde{\gamma}^r(t)) \circ \left( \tilde{\mathcal{L}}^H E[\alpha_1(t) \alpha_2^H(s)] \tilde{\mathcal{L}} \right) \right\} \mathcal{L}_k^H.$
3.  $E[\Delta \tilde{\mathcal{D}}_k^\gamma(t) \alpha_1(t) \alpha_2^H(s) \tilde{\mathcal{D}}_k^{(1-\gamma)^H}(t)] = -\mathcal{L}_k \left\{ \text{Cov}(\tilde{\gamma}^r(t)) \circ \left( \tilde{\mathcal{L}}^H E[\alpha_1(t) \alpha_2^H(s)] \tilde{\mathcal{L}} \right) \right\} \mathcal{L}_k^H.$
4.  $E[\Delta \tilde{\mathcal{D}}_k^\gamma(t) \alpha_1(t) \beta^H(s)] = \mathbf{0}_{m \times q}.$
5.  $E[\tilde{\mathcal{D}}_k^\gamma(t) \alpha_1(t) \alpha_2^H(s) \tilde{\mathcal{D}}_k^{\gamma^H}(t)] = \mathcal{L}_k \left\{ E[\tilde{\gamma}^r(t) \tilde{\gamma}^{r^T}(t)] \circ \left( \tilde{\mathcal{L}}^H E[\alpha_1(t) \alpha_2^H(s)] \tilde{\mathcal{L}} \right) \right\} \mathcal{L}_k^H.$
6.  $E[\tilde{\mathcal{D}}_k^\gamma(t) \alpha_1(t) \alpha_2^H(s) \tilde{\mathcal{D}}_k^{(1-\gamma)^H}(t)] =$   
 $\mathcal{L}_k \left\{ E[\tilde{\gamma}^r(t) (\mathbf{1}_{4Rn} - \tilde{\gamma}^r(t))^T] \circ \left( \tilde{\mathcal{L}}^H E[\alpha_1(t) \alpha_2^H(s)] \tilde{\mathcal{L}} \right) \right\} \mathcal{L}_k^H.$
7.  $E[\tilde{\mathcal{D}}_k^{(1-\gamma)}(t) \alpha_1(t) \alpha_2^H(s) \tilde{\mathcal{D}}_k^{(1-\gamma)^H}(t)] =$   
 $\mathcal{L}_k \left\{ E[(\mathbf{I}_m - \tilde{\gamma}^r(t)) (\mathbf{I}_m - \tilde{\gamma}^r(t))^T] \circ \left( \tilde{\mathcal{L}}^H E[\alpha_1(t) \alpha_2^H(s)] \tilde{\mathcal{L}} \right) \right\} \mathcal{L}_k^H.$

**Proof.** The proof is immediate from (1) and taking into account that  $\tilde{\mathcal{D}}_k^\gamma(t) = \mathcal{L}_k \text{diag}(\tilde{\gamma}^r(t)) \tilde{\mathcal{L}}^H$  and  $\tilde{\mathcal{D}}_k^{(1-\gamma)}(t) = \mathcal{L}_k \text{diag}(\mathbf{1}_{4Rn} - \tilde{\gamma}^r(t)) \tilde{\mathcal{L}}^H$ .  $\square$

#### Appendix A.2 Expressions in Theorem 1

Although tessarine algebra is not a Hilbert space, the existence and uniqueness of the projection of an element on the set of measurements  $\{\mathbf{y}_k(1), \dots, \mathbf{y}_k(s)\}$ , for  $k = 1, 2$ , is guaranteed ([22]). Now, from Theorem 3 of [22], we obtain

$$\hat{\mathbf{x}}(t|t) = \hat{\mathbf{x}}(t|t-1) + \tilde{\mathbf{L}}_k(t) \varepsilon_k(t), \quad (\text{A2})$$

with  $\tilde{\mathbf{L}}_k(t) = \tilde{\mathbf{\Theta}}_k(t) \mathbf{\Omega}_k^{-1}(t)$ , where  $\tilde{\mathbf{\Theta}}_k(t) = E[\tilde{\mathbf{x}}(t) \varepsilon_k^H(t)]$  and  $\mathbf{\Omega}_k(t) = E[\varepsilon_k(t) \varepsilon_k^H(t)]$ . Then, by applying  $\mathbb{T}_k$ -properness conditions, (10) is directly devised.

Taking projections on both sides of the state and observation equations in (A1) onto the linear space spanned by  $\{\varepsilon_k(1), \dots, \varepsilon_k(t-1)\}$ , and using Property A1, we have

$$\hat{\mathbf{x}}(t+1|t) = \tilde{\mathbf{\Phi}}(t) \hat{\mathbf{x}}(t|t) + \tilde{\mathbf{H}}_k(t) \varepsilon_k(t) \quad (\text{A3})$$

$$\hat{\mathbf{y}}_k(t|t-1) = \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) \tilde{\mathcal{C}} \hat{\mathbf{x}}(t|t-1) + \tilde{\mathbf{\Pi}}_k^{(1-\gamma)}(t) [\tilde{\mathcal{C}} \hat{\mathbf{x}}(t-1|t-1) + \hat{\mathbf{v}}(t-1|t-1)] \quad (\text{A4})$$

where  $\tilde{\mathbf{H}}_k(t) = \tilde{\mathbf{S}}(t) \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) \mathbf{\Omega}_k^{-1}(t)$  and  $\hat{\mathbf{v}}(t|t) = \tilde{\mathbf{G}}_k(t) \varepsilon_k(t)$ , with  $\tilde{\mathbf{G}}_k(t) = \tilde{\mathbf{R}}(t) \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) \mathbf{\Omega}_k^{-1}(t)$ .

Then, (11) follows from (A3) and the  $\mathbb{T}_k$ -properness conditions on  $\tilde{\mathbf{\Phi}}(t)$  and  $\tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t)$  established in Proposition 1 and Remark 3. Likewise, (12) is easily obtained from (A4).

Consider now the gain matrix  $\tilde{\mathbf{L}}_k(t) = \tilde{\mathbf{\Theta}}_k(t) \mathbf{\Omega}_k^{-1}(t)$  in (A2). Denote the prediction error and its covariance matrix as  $\tilde{\mathbf{e}}(t|t-1) = \tilde{\mathbf{x}}(t) - \hat{\mathbf{x}}(t|t-1)$  and  $\tilde{\mathbf{P}}(t|t-1) = E[\tilde{\mathbf{e}}(t|t-1) \tilde{\mathbf{e}}^H(t|t-1)]$ , respectively. Then, by applying (A1), (A4),  $\tilde{\mathbf{e}}(t|t-1) \perp \hat{\mathbf{x}}(t|t-1)$ , Property 3 and Property A1, we have

$$\begin{aligned} \tilde{\mathbf{\Theta}}_k(t) &= \tilde{\mathbf{P}}(t|t-1) \tilde{\mathcal{C}}^T \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) + \tilde{\mathbf{\Phi}}(t-1) \tilde{\mathbf{P}}(t-1|t-1) \tilde{\mathcal{C}}^T \tilde{\mathbf{\Pi}}_k^{(1-\gamma)^H}(t) + \tilde{\mathbf{S}}(t-1) \tilde{\mathbf{\Pi}}_k^{(1-\gamma)^H}(t) \\ &\quad - \tilde{\mathbf{H}}_k(t-1) \tilde{\mathbf{\Theta}}_k^H(t-1) \tilde{\mathcal{C}}^T \tilde{\mathbf{\Pi}}_k^{(1-\gamma)^H}(t) - \tilde{\mathbf{\Phi}}(t-1) \tilde{\mathbf{\Theta}}_k(t-1) \tilde{\mathbf{G}}_k^H(t-1) \tilde{\mathbf{\Pi}}_k^{(1-\gamma)^H}(t) \\ &\quad - \tilde{\mathbf{H}}_k(t-1) \mathbf{\Omega}_k(t-1) \tilde{\mathbf{G}}_k^H(t-1) \tilde{\mathbf{\Pi}}_k^{(1-\gamma)^H}(t), \quad t > 1 \quad (\text{A5}) \end{aligned}$$

and thus, the recursive expression (13) is directly obtained from (A5), by applying the  $\mathbb{T}_k$ -properness conditions on  $\tilde{\mathbf{\Phi}}(t)$ , and denoting by  $\mathbf{P}_k(t|t-1)$  the first  $m \times m$  submatrix of  $\tilde{\mathbf{P}}(t|t-1)$ .

Next, we devise the expression for the innovation covariance matrix (14). For this purpose, the innovations are rewritten in the following form

$$\begin{aligned} \varepsilon_k(t) = & \Delta \tilde{\mathcal{D}}_k^\gamma(t) \tilde{\mathcal{C}} \bar{\mathbf{x}}(t) + \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(t) \tilde{\mathcal{C}} \bar{\varepsilon}(t|t-1) - \Delta \tilde{\mathcal{D}}_k^\gamma(t) \tilde{\mathcal{C}} \bar{\mathbf{x}}(t-1) + \tilde{\mathbf{\Pi}}_k^{(1-\gamma)}(t) \tilde{\mathcal{C}} \bar{\varepsilon}(t-1|t-1) \\ & + \tilde{\mathcal{D}}_k^\gamma(t) \tilde{\mathbf{v}}(t) + \tilde{\mathcal{D}}_k^{(1-\gamma)}(t) \tilde{\mathbf{v}}(t-1) - \tilde{\mathbf{\Pi}}_k^{(1-\gamma)}(t) \tilde{\mathbf{G}}_k(t-1) \varepsilon_k(t-1). \end{aligned} \quad (\text{A6})$$

From (A3), the prediction error  $\bar{\varepsilon}(t+1|t)$  can be expressed as

$$\bar{\varepsilon}(t+1|t) = \tilde{\Phi}(t) \bar{\varepsilon}(t|t) + \tilde{\mathbf{u}}(t) - \tilde{\mathbf{H}}_k(t) \varepsilon_k(t). \quad (\text{A7})$$

As a consequence, from (A7), using Property 3 and Property A1, and taking into account that  $\bar{\varepsilon}(t|t) \perp \varepsilon_k(t)$ , we have

$$E[\tilde{\mathbf{u}}(t) \bar{\varepsilon}^H(t|t)] = -\tilde{\mathbf{H}}_k(t) \tilde{\Theta}_k^H(t), \quad (\text{A8})$$

$$E[\bar{\varepsilon}(t+1|t) \bar{\varepsilon}^H(t|t)] = \tilde{\Phi}(t) \tilde{\mathbf{P}}(t|t) - \tilde{\mathbf{H}}_k(t) \tilde{\Theta}_k^H(t), \quad (\text{A9})$$

$$E[\bar{\varepsilon}(t|t) \tilde{\mathbf{v}}^H(t)] = -\tilde{\Theta}_k(t) \tilde{\mathbf{G}}_k^H(t), \quad (\text{A10})$$

$$E[\bar{\varepsilon}(t|t) \tilde{\mathbf{v}}^H(t+1)] = \mathbf{0}_{4n \times 4Rn}, \quad (\text{A11})$$

$$E[\bar{\varepsilon}(t+1|t) \tilde{\mathbf{v}}^H(t)] = \tilde{\mathbf{S}}(t) - \tilde{\Phi}(t) \tilde{\Theta}_k(t) \tilde{\mathbf{G}}_k^H(t) - \tilde{\mathbf{H}}_k(t) \tilde{\Omega}_k(t) \tilde{\mathbf{G}}_k^H(t), \quad (\text{A12})$$

$$E[\bar{\varepsilon}(t+1|t) \tilde{\mathbf{v}}^H(t+1)] = \mathbf{0}_{4n \times 4Rn}. \quad (\text{A13})$$

Then, the expression (14) for the innovation covariance matrix is obtained from (A6), by using Lemma A1, Property 3, Property A1, (A9)-(A13),  $\bar{\varepsilon}(t+1|t) \perp \varepsilon_k(t)$ , and by applying  $\mathbb{T}_k$  properness conditions. Furthermore, the recursion of  $\tilde{\mathbf{D}}(t) = E[\bar{\mathbf{x}}(t) \bar{\mathbf{x}}^H(t)]$  given in (16) is a direct consequence of the augmented state equation in system (A1). In a similar way, Eq. (15) follows.

In the following step, consider the filtering error covariance matrix  $\tilde{\mathbf{P}}(t|t) = E[\bar{\varepsilon}(t|t) \bar{\varepsilon}^H(t|t)]$  with  $\bar{\varepsilon}(t|t) = \bar{\mathbf{x}}(t) - \hat{\mathbf{x}}(t|t)$ . From (A2), we directly obtain that  $\tilde{\mathbf{P}}(t|t) = \tilde{\mathbf{P}}(t|t-1) - \tilde{\Theta}_k(t) \Omega_k^{-1}(t) \tilde{\Theta}_k^H(t)$  and thus (17) holds by virtue of  $\mathbb{T}_k$  properness conditions.

Finally, from (A7), and taking into consideration that  $\bar{\varepsilon}(t|t) \perp \varepsilon_k(t)$ , (A8), and Property A1, we have

$$\begin{aligned} \tilde{\mathbf{P}}(t+1|t) = & \tilde{\Phi}(t) \tilde{\mathbf{P}}(t|t) \tilde{\Phi}^H(t) - \tilde{\mathbf{H}}_k(t) \tilde{\Theta}_k^H(t) \tilde{\Phi}^H(t) \\ & - \tilde{\Phi}(t) \tilde{\Theta}_k(t) \tilde{\mathbf{H}}_k^H(t) - \tilde{\mathbf{H}}_k(t) \Omega_k(t) \tilde{\mathbf{H}}_k^H(t) + \tilde{\mathbf{Q}}(t). \end{aligned}$$

From  $\mathbb{T}_k$  properness conditions (18) follows.

## Appendix B Proof of Theorem 2

From the projection of  $\mathbf{x}(t+\tau)$  onto the linear space spanned by  $\{\varepsilon_k(1), \dots, \varepsilon_k(t)\}$ , we have

$$\hat{\mathbf{x}}(t+\tau|t) = \Phi(t+\tau-1) \hat{\mathbf{x}}(t+\tau-1|t), \tau \geq 2$$

Then, from  $\mathbb{T}_k$  properness conditions, (19) holds.

Finally, from (19), it is clear that the prediction error covariance matrix  $\mathbf{P}_k(t+\tau|t)$  satisfies the recursive expression (20).

## Appendix C Proof of Theorem 3

By projecting the state  $\mathbf{x}(t)$  onto the linear space spanned by  $\{\varepsilon_k(1), \dots, \varepsilon_k(s)\}$ , we have

$$\hat{\mathbf{x}}(t|s) = \hat{\mathbf{x}}(t|s-1) + \tilde{\mathbf{L}}_k(t,s) \varepsilon_k(s), \quad t < s \quad (\text{A14})$$

with  $\tilde{\mathbf{L}}_k(t,s) = \tilde{\Theta}_k(t,s) \Omega_k^{-1}(s)$ , where  $\tilde{\Theta}_k(t,s) = E[\bar{\mathbf{x}}(t) \varepsilon_k^H(s)]$ . Then, (21) is directly derived from (A14), by applying  $\mathbb{T}_k$  properness conditions.



Consider the matrix  $\tilde{\theta}_k(t, s)$ . From (12) and (8), we have

$$\begin{aligned}\tilde{\theta}_k(t, s) = E[\bar{\mathbf{x}}(t)\bar{\mathbf{e}}^H(s|s-1)]\tilde{\mathbf{C}}^T\tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s) + E[\bar{\mathbf{x}}(t)\bar{\mathbf{e}}^H(s-1|s-1)]\tilde{\mathbf{C}}^T(I_m - \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s)) \\ - \tilde{\theta}_k(t, s-1)\tilde{\mathbf{G}}_k^H(s-1)(I_m - \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s))\end{aligned}$$

Let us define the matrix  $\bar{\mathbf{E}}(t, s) = E[\bar{\mathbf{x}}(t)\bar{\mathbf{e}}^H(s|s)]$ . Thus, from (A1) and (A3), it follows that

$$\begin{aligned}\tilde{\theta}_k(t, s) = [\bar{\mathbf{E}}(t, s-1)\tilde{\mathbf{\Phi}}^H(s-1) - \tilde{\theta}_k(t, s-1)\tilde{\mathbf{H}}_k^H(s-1)]\tilde{\mathbf{C}}^T\tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s) \\ + [\bar{\mathbf{E}}(t, s-1)\tilde{\mathbf{C}}^T - \tilde{\theta}_k(t, s-1)\tilde{\mathbf{G}}_k^H(s-1)](I_m - \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s))\end{aligned}\quad (\text{A15})$$

Then, (22) follows from  $\mathbb{T}$  proper conditions.

In a similar way, from (A1), (A2), (A3) and (A15),  $\bar{\mathbf{E}}(t, s)$  is of the form

$$\begin{aligned}\bar{\mathbf{E}}(t, s) = [\bar{\mathbf{E}}(t, s-1)\tilde{\mathbf{\Phi}}^H(s-1) - \tilde{\theta}_k(t, s-1)\tilde{\mathbf{H}}_k^H(s-1)](I_m - \tilde{\mathbf{C}}^T\tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s)\mathbf{L}^H(s)) \\ - [\bar{\mathbf{E}}(t, s-1)\tilde{\mathbf{C}}^T - \tilde{\theta}_k(t, s-1)\tilde{\mathbf{G}}_k^H(s-1)](I_m - \tilde{\mathbf{\Pi}}_k^{\tilde{\gamma}}(s))\mathbf{L}^H(s)\end{aligned}\quad (\text{A16})$$

where  $\bar{\mathbf{E}}(t, t) = \bar{\mathbf{P}}(t|t)$ . Then, (23) follows from  $\mathbb{T}_k$  proper conditions.

Finally, (24) can be easily derived from (21).

## References

1. Li, W.L.; Jia, Y.M.; Du, J.P.; Fu J.C. State estimation for nonlinearly coupled complex networks with application to multi-target tracking. *Neurocomputing* **2018**, *275*, 1884–1892.
2. Lee, J.S.; McBride, J. Extended object tracking via positive and negative information fusion. *IEEE Trans. Signal Process.* **2019**, *67* (7), 1812–1823.
3. Kurkin, A.A.; Tyugin, D.Y.; Kuzin, V.D.; Chernov, A.G.; Makarov, V.S.; Beresnev, P.O.; Filatov, V.I.; Zeziulin, D.V. Autonomous mobile robotic system for environment monitoring in a coastal zone. *Procedia Comput. Sci.* **2017**, *103*, 459–465.
4. Gingras, D. An overview of positioning and data fusion techniques applied to land vehicle navigation systems. In *Automotive Informatics and Communicative Systems*; Information Science Reference, 2009; 219–246.
5. Gao, B.; Hu, G.; Gao, S.; Zhong, Y.; Gu, C.; Beresnev, P.O.; Filatov, V.I.; Zeziulin, D.V. Multi-sensor optimal data fusion for INS/GNSS/CNS integration based on unscented Kalman filter. *Int. J. Control Autom. Syst.* **2018**, *16* (1), 129–140.
6. Din, S.; Ahmad, A.; Paul, A.; Rathore, M.M.U.; Gwanggil, J. A clusterbased data fusion technique to analyze big data in wireless multi-sensor system. *IEEE Access* **2017**, *5*, 5069–5083.
7. Liggins, M. E.; Hall, D. L.; Llinas, J. *Handbook of Multisensor Data Fusion: Theory and Practice*, ser. The Electrical Engineering and Applied Signal Processing Series.; CRC PressINC, 2009.
8. Hounkpevi, F.O.; Yaz, E.E. Minimum variance generalized state estimators for multiple sensors with different delay rates. *Signal Process.* **2007**, *87*, 602–613.
9. Ma, J.; Sun, S. Centralized fusion estimators for multisensor systems with random sensor delays, multiple packet dropouts and uncertain observations. *IEEE Sens. J.* **2013**, *13* (4), 1228–1235.
10. Chen, D.; Xu, L. Optimal filtering with finite-step autocorrelated process noises, random one-step sensor delay and missing measurements. *Commun. Nonlinear Sci. Numer. Simul.* **2016**, *32*, 211–224.
11. Sun, S.; Ma, J. Linear estimation for networked control systems with random transmission delays and packet dropouts. *Inf. Sci.* **2014**, *269*, 349–365.
12. Li, N.; Sun, S.; Ma, J. Multi-sensor distributed fusion filtering for networked systems with different delay and loss rates. *Digital Signal Process.* **2014**, *34*, 29–38.
13. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. Optimal fusion estimation with multi-step random delays and losses in transmission. *Sensors* **2017**, *17* (5), 1151.
14. Tian, T.; Sun, S.; Li, N. Multi-sensor information fusion estimators for stochastic uncertain systems with correlated noises. *Inf. Fusion* **2016**, *27*, 126–137.
15. Abu Bakr, M.; Lee, S. Distributed multisensor data fusion under unknown correlation and data inconsistency. *Sensors* **2017**, *17*, 2472.
16. Liu, W.Q.; Wang, X.M.; Deng, Z.L. Robust centralized and weighted measurement fusion kalman estimators for uncertain multisensor systems with linearly correlated white noises. *Inf. Fusion.* **2017**, *35*, 11–25.
17. Caballero-Águila, R.; Hermoso-Carazo, A.; Linares-Pérez, J. Centralized fusion approach to the Estimation problem with multi-packet processing under uncertainty in outputs and transmissions. *Sensors* **2018**, *18*, 2697.

18. Shen, Q.; Liu, J.; Zhou, X.; Qin, W.; Wang, L.; Wang, Q. Centralized fusion methods for multi-sensor system with bounded disturbances. *IEEE Access* **2019**, *7*, 141612–141626.
19. Alfsmann, D. ; Göckler, H.G. ; Sangwine, S.J.; Ell, T.A. Hypercomplex Algebras in Digital Signal Processing: Benefits and Drawbacks. In Proc. 15th Eur. Signal Process. Conf. (EUSIPCO 2007), 2007, pp.1322–1326.
20. Ortolani, F.; Scarpiniti, M.; Comminiello, D.; Uncini, A. 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.
21. Ortolani, F.; Comminiello, D.; Scarpiniti, M.; Uncini, A. On 4-dimensional hypercomplex algebras in adaptive signal processing. In *Neural Advances in Processing Nonlinear Dynamic Signals*, 2017, pp. 131–140.
22. Navarro-Moreno, J.; Fernández Alcalá, R.M.; Jiménez López, J.D.; Ruiz-Molina, J.C. Tessarine signal processing under the T-properness condition. *J. Frankl. Inst.* **2020**, *357* (14), 10099–10125.
23. Navarro-Moreno, J.; Ruiz-Molina, J.C. Wide-sense Markov signals on the tessarine domain. A study under properness conditions. *Signal Process.* **2021**, *183*, 108022.
24. Sabatelli, S.; Sechi, F.; Fanucci, L.; Rocchi, A. A sensor fusion algorithm for an integrated angular position estimation with inertial measurement units. Design, Automation and Test in Europe (DATE 2011), Grenoble, France, March 14-18, 2011, 273–276.
25. Tannous, H.; Istrate, D.; Benlarbi-Delai, A.; Sarrazin, J.; Gamet, D.; Ho Ba Tho, M.C.; Dao, T.T. A new multi-sensor fusion scheme to improve the accuracy of knee flexion kinematics for functional rehabilitation movements, *J. Sensors* **2016**, *16* (11), 1914.
26. Talebi, S.; Kanna, S.; Mandic, D. A distributed quaternion Kalman filter with applications to smart grid and target tracking, *IEEE Trans. Signal Inf. Process. Networks* **2016**, *2*, 477–488.
27. Talebi, S.P.; Werner, S.; Mandic, D.P. Quaternion-valued distributed filtering and control, *IEEE Trans. Autom. Control.* **2020**, *65* (10), 4246–4256.
28. Wu, J.; Zhou, Z.; Fourati, H.; Li, R.; Liu, M. Generalized linear quaternion complementary filter for attitude estimation from multi-sensor observations: An optimization approach. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 1330–1343.
29. Navarro-Moreno, J.; Fernández-Alcalá, R.M.; Jiménez López, J.D.; Ruiz-Molina, J.C. Widely linear estimation for multisensor quaternion systems with mixed uncertainties in the observations, *J. Frankl. Inst.* **2019**, *356*, 3115–3138.
30. Vía, J.; Ramírez, D.; Santamaría, I. Properness and widely linear processing of quaternion random vectors, *IEEE Trans. Inform. Theory* **2010**, *56* (7), 3502–3515.
31. Jiménez-López, J.D.; Fernández-Alcalá, R.M.; Navarro-Moreno, J.; Ruiz-Molina, J.C. Widely linear estimation of quaternion signals with intermittent observations, *Signal Process.* **2017**, *136*, 92–101.
32. Fernández Alcalá, R.M.; Navarro-Moreno, J.; Jiménez López, J.D.; Ruiz-Molina, J.C. Semi-widely linear estimation algorithms of quaternion signals with missing observations and correlated noises. *J. Frankl. Inst.* **2020**, *357*, 3075–3096.
33. Nitta, T. ; Kobayashi, M.; Mandic, D.P. Hypercomplex widely linear estimation through the lens of underpinning geometry, *IEEE Trans. Signal Process.* **2019** *67* (15), 3985–3994.