

Article

Proper ARMA Modeling and Forecasting in the Generalized Segre's Quaternions Domain

Jesús Navarro-Moreno , Rosa M. Fernández-Alcalá *  and Juan C. Ruiz-Molina 

Department of Statistics and Operations Research, University of Jaén, Paraje Las Lagunillas, 23071 Jaén, Spain; jnavarro@ujaen.es (J.N.-M.); jrcruiz@ujaen.es (J.C.R.-M.)

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

Abstract: The analysis of time series in 4D commutative hypercomplex algebras is introduced. Firstly, generalized Segre's quaternion (GSQ) random variables and signals are studied. Then, two concepts of properness are suggested and statistical tests to check if a GSQ random vector is proper or not are proposed. Further, a method to determine in which specific hypercomplex algebra is most likely to achieve, if possible, the properness properties is given. Next, both the linear estimation and prediction problems are studied in the GSQ domain. Finally, ARMA modeling and forecasting for proper GSQ time series are tackled. Experimental results show the superiority of the proposed approach over its counterpart in the Hamilton quaternion domain.

Keywords: ARMA models; E_{α_0} -properness; generalized Segre's quaternions; H_{α_0} -properness; prediction; proper signal processing

MSC: 37M10



Citation: Navarro-Moreno, J.; Fernández-Alcalá, R.M.; Ruiz-Molina, J.C. Proper ARMA Modeling and Forecasting in the Generalized Segre's Quaternions Domain. *Mathematics* **2022**, *10*, 1083. <https://doi.org/10.3390/math10071083>

Academic Editors: Gordon Morison and Robert C. Atkinson

Received: 16 February 2022

Accepted: 24 March 2022

Published: 28 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Time series theory is a very well-known statistical methodology for the analysis of stochastic processes that has found widespread applications in signal processing, communications engineering, economics, astronomy, meteorology, seismology, oceanography, among other fields [1]. The most extensively used class of stationary process models are the so-called autoregressive moving average (ARMA) processes. The relevance of ARMA models stems from their power to approximate any stationary process arbitrarily well [2,3]. In addition, ARMA models can be cast in a state-space framework and also have enormous significance in prediction problems, since they facilitate the derivation of several efficient prediction algorithms [3].

The traditional settings for developing time series analysis have been the real and complex numbers, while other number systems have demonstrated a high potential to outperform them in some applications. The most paradigmatic example is the system of Hamilton quaternions (HQs), which has proven to be of great value for the description of 3-dimensional (3D) movements, in fields such as graphic design or automated control [4,5]. Actually, HQs have become the most popular and researched system of hypercomplex numbers [6]. In spite of its relevance, the use of this system in time series analysis is scarce [7].

HQs belong to a more general family of algebras, named hypercomplex systems, which extends the real and complex numbers to higher dimensions. Among them, 4D hypercomplex systems are of special relevance since four coordinates are usually considered for describing the physical world [6]. Further, depending on how the product is defined, they can be classified into two types: noncommutative and commutative systems. HQs are of the first type, i.e., they form a normed division algebra, and this property makes them well suited to deal with signal processing benchmark problems, such as the signal estimation problem [8]. Though commutative hypercomplex numbers, which includes the

generalized Segre's quaternions (GSQs), are less well known, their interest is increasing within the signal processing community due to their availability to extend the methodologies developed in the complex and real field to the 4D case [9–21]. These commutative hypercomplex algebras possess zero divisors, which implies that the Euclidean norm is not multiplicative [13]. This drawback, however, has essentially no impact on signal and image processing applications [10].

Interestingly, a recent line of research has discovered that the HQs algebra is not always the best option for processing 4D hypercomplex signals [22–25]. Specifically, some investigations have revealed that properness, a structural property of the hypercomplex signals, determines the performance of the used processing [24,25]. In all of these papers, the analyses have been performed by comparing HQs with tessarines, and several scenarios where the latter outperforms the former have been shown. It should be noted that tessarines are a subset of GSQs obtained by defining a particular product. In general, hypercomplex algebras are characterized by its product, which is usually provided by a multiplication table (Appendix A in [6]). From an intuitive point of view, it is to be expected that more general commutative hypercomplex algebras than tessarines can attain better performance results. On that basis, special attention is paid to such systems in this paper.

As indicated above, properness is a crucial property of hypercomplex signals. This property has two important effects: it determines which type of processing should be used, and also it achieves a significant decrease in the computational complexity involved in comparison to the real processing. Properness has been investigated in detail in the HQ domain [26–30], and for tessarines [24,25]. To our knowledge, however, properness has not been studied for GSQs yet.

Inspired by all the above discussed, the main objective of this paper is to present the time series analysis in the GSQs domain while analyzing the properness properties in this setting. Specifically, the concept of properness is introduced for GSQ signals and their properties studied. Remarkably, two tests to experimentally check whether real data are proper in the GSQs field are proposed and an estimate of the parameter characterizing the product in the hypercomplex algebra is provided. Such an estimate gives the best value of the parameter under which the properness properties could be fulfilled by the signal of interest. It should be highlighted that, unlike the existing approaches, a specific hypercomplex algebra is not fixed in advance in the proposed strategy, but a broad family of commutative hypercomplex algebras is considered. So, the suggested estimate allows us to select the particular algebra in which it is most likely to achieve, if possible, the properness properties. A simulation example illustrates how this approach is effective.

The linear estimation problem is also tackled. Although GSQ numbers are not a normed algebra and thus, the classical projection theorem cannot be applied to obtain the optimum linear estimate, we exploit its metric space feature to prove the existence and uniqueness of the projection and an extension of the innovations algorithm for proper GSQ signals is devised. As could be expected, the proposed algorithms attain a notable reduction in the computational burden. Finally, ARMA modeling for proper GSQ time series is introduced. For that, we first investigate an extension of the Yule–Walker equations and then, the prediction problem is addressed. The particular structure of the ARMA models allows us to improve the innovations algorithm giving rise to a more efficient prediction method. The steady-state properties of the suggested predictor are also studied. To conclude, the superiority in terms of performance of our proposal in comparison with its counterpart in the HQ domain is numerically analyzed.

The rest of this paper is organized as follows. Section 2 presents both the GSQ random variables and signals. Then, Section 3 introduces the concepts of properness in the GSQs domain and provides the statistical inference methods for testing the properness as well as estimating the associated parameter that characterizes the product in the hypercomplex algebra. This section ends with a simulation example. The linear estimation problem for GSQ random variables and the prediction problem for GSQ random signals are analyzed in Section 4. Section 5 focuses on the analysis of time series in the GSQs setting. It deals with

the problems of modeling and forecasting time series data. A final example is included to show the efficiency of the proposed strategy. To improve the readability of the manuscript, all proofs have been deferred to Appendices B–F.

Notation

We designate matrices with boldface uppercase letters, column vectors with lowercase letters, and scalar quantities with lightface lowercase letters. Moreover, the symbols \mathbb{Z} , \mathbb{R} , \mathbb{C} , and \mathbb{S} represent the set of integer, real, complex, and GSQ numbers, respectively. In this regard, $\mathbf{A} \in \mathbb{R}^{p \times q}$ (respectively, $\mathbf{A} \in \mathbb{S}^{p \times q}$) indicates that \mathbf{A} is a real (respectively, a GSQ) matrix of dimension $p \times q$. Similarly $\mathbf{r} \in \mathbb{R}^p$ (respectively, $\mathbf{r} \in \mathbb{S}^p$) means that \mathbf{r} is a p -dimensional real (respectively, a GSQ) vector. In particular, $\mathbf{0}_{p \times q}$ represents the zero matrix of dimension $p \times q$, and \mathbf{I}_p is the identity matrix of dimension p . In addition, A_{ij} denotes the ij element of the matrix \mathbf{A} and $\text{diag}(\cdot)$ stands for a diagonal (or block diagonal) matrix with entries (or block entries) specified in the operator argument.

Furthermore, superscript “T” represents the transpose. “ \otimes ” stands for the Kronecker product. $E[\cdot]$ represents the expectation operator. $\mathcal{R}\{\cdot\}$ denotes the real part of a GSQ, and $|z|$ is either the complex modulus if $z \in \mathbb{C}$ or the absolute value if $z \in \mathbb{R}$. Finally, δ_{nl} denotes the Kronecker delta function.

Unless otherwise stated, all random variables are assumed to have a zero-mean throughout this paper.

A list of symbols that are used in the paper is given in Table 1.

Table 1. List of symbols.

Symbol	Meaning
$\mathbf{x}_r(t)$	Real vector formed with the components of $\mathbf{x} \in \mathbb{S}^p$
$\bar{\mathbf{x}}(t)$	Augmented GSQ signal vector formed by $\mathbf{x} \in \mathbb{S}^p$ and its conjugations
$\Gamma_{\mathbf{a}}(t, s)$	Correlation function of a real random signal vector $\mathbf{a}(t) \in \mathbb{R}^p$
$\Gamma_{\mathbf{ab}}(t, s)$	Correlation function of the real random signal vectors $\mathbf{a}(t) \in \mathbb{R}^p$ and $\mathbf{b}(t) \in \mathbb{R}^q$
$\Gamma_{\mathbf{x}}^{\nu}(\alpha, t, s)$	$E[\mathbf{x}(t)\mathbf{x}^{\text{H}_{3-2\nu}}(s)]$, for $\mathbf{x} \in \mathbb{S}^p$, $\nu = 1, 2$
$\Gamma_{\mathbf{xy}}^{\nu}(\alpha, t, s)$	$E[\mathbf{x}(t)\mathbf{y}^{\text{H}_{3-2\nu}}(s)]$, for $\mathbf{x} \in \mathbb{S}^p$, $\mathbf{y} \in \mathbb{S}^q$, and $\nu = 1, 2$
$\Gamma_{\mathbf{x}}^3(\alpha, t, s)$	$E[\mathbf{x}(t)\mathbf{x}^{\text{H}_{\alpha}}(s)]$, for $\mathbf{x} \in \mathbb{S}^p$
$\Gamma_{\mathbf{xy}}^3(\alpha, t, s)$	$E[\mathbf{x}(t)\mathbf{y}^{\text{H}_{\alpha}}(s)]$, for $\mathbf{x} \in \mathbb{S}^p$ and $\mathbf{y} \in \mathbb{S}^q$
$\hat{\Gamma}_{\mathbf{x}}$	Sample autocorrelation matrix of a random vector \mathbf{x}
\mathcal{C}_{α}	Set of the random vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$, with $\mathbf{x}_i \in \mathbb{S}^p$, $i = 1, \dots, m$
$\bar{\mathcal{C}}_{\alpha}$	Set of the augmented random vectors $\{\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_m\}$, with $\bar{\mathbf{x}}_i \in \mathbb{S}^{4p}$, $i = 1, \dots, m$
$\mathcal{G}_{\mathcal{C}_{\alpha}}$	Set of all finite linear combinations of elements of \mathcal{C}_{α}
$\mathcal{G}_{\bar{\mathcal{C}}_{\alpha}}$	Set of all finite linear combinations of elements of $\bar{\mathcal{C}}_{\alpha}$
\hat{y}^{SWL}	SWL estimator of y respect to \mathcal{C}_{α}
$\hat{y}_{\alpha}^{\text{S}}$	\mathbb{S}_{α} estimator of y respect to \mathcal{C}_{α}
\hat{y}^{QSL}	QSL estimator of y respect to \mathcal{C}_{α}
ϵ^{SWL}	SWL estimation error associated to \hat{y}^{SWL}
$\epsilon_{\alpha}^{\text{S}}$	\mathbb{S}_{α} estimation error associated to $\hat{y}_{\alpha}^{\text{S}}$
ϵ^{QSL}	QSL estimation error associated to \hat{y}^{QSL}

2. Generalized Segre’s Quaternion Random Variables

This section establishes the basic concepts and properties of interest in the GSQ domain. For that, we first introduce the GSQ random variables and then the GSQ random signals are considered.

A random variable $x \in \mathbb{S}$ can be defined as $x = a + ib + jc + kd$ where $a, b, c, d \in \mathbb{R}$ are random variables and the imaginary units (i, j, k) satisfy the following multiplication rules:

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

where α is a nonzero real number.

As is well-known, the form in which the product is defined is an essential feature of hypercomplex numbers. The product defined in (1) is commutative, associative, and distributive over addition; however, these algebras have divisors of zero. In contrast to GSQs, the product in the HQs is defined in such a way that the algebra becomes noncommutative and does not have divisors of zero [6]. Further, two special algebras in GSQs are of interest: for $\alpha > 0$, \mathbb{S} is called the hyperbolic quaternions system and for $\alpha < 0$, \mathbb{S} is the elliptic quaternions system [6]. In particular, if $\alpha = -1$ then, \mathbb{S} is named the tessarine numbers.

The real vector of $x \in \mathbb{S}$ is defined by $\mathbf{x}_r = [a, b, c, d]^T$. Moreover, we define the following auxiliary GSQ random variables:

$$\begin{aligned} x_\alpha &= a + i\frac{b}{\alpha} + jc + k\frac{d}{\alpha} & x^i &= a + ib - jc - kd \\ x^j &= a - ib + jc - kd & x^k &= a - ib - jc + kd \end{aligned}$$

The triplet (x^i, x^j, x^k) are the so-called principal conjugations and obey the following properties.

Property 1. Let $x, y \in \mathbb{S}$ be random variables, then

1. $(x^\eta)^\eta = x, \quad \eta = i, j, k$
2. $(x^{\eta_1})^{\eta_2} = x^{\eta_3}, \quad \eta_1, \eta_2, \eta_3 = i, j, k, \quad \eta_1 \neq \eta_2 \neq \eta_3$
3. $x^\eta y^\eta = (xy)^\eta, \quad \eta = i, j, k$

Consider the vector random variables $\mathbf{x} \in \mathbb{S}^p$ and $\mathbf{y} \in \mathbb{S}^q$. We define the following functions:

$$\begin{aligned} \Gamma_{\mathbf{xy}}^1(\alpha) &= E[\mathbf{xy}^T] \\ \Gamma_{\mathbf{xy}}^2(\alpha) &= E[\mathbf{xy}^{jT}] \\ \Gamma_{\mathbf{xy}}^3(\alpha) &= E[\mathbf{xy}_\alpha^T] \end{aligned} \tag{2}$$

where $\mathbf{y} = [y_1, \dots, y_q]^T, \mathbf{y}^j = [y_1^j, \dots, y_q^j]^T$, and $\mathbf{y}_\alpha = [y_{1\alpha}, \dots, y_{q\alpha}]^T$; however, for reasons that will become clear, they are denoted by $\Gamma_{\mathbf{xy}}^1(\alpha) \triangleq E[\mathbf{xy}^{H_1}]$, $\Gamma_{\mathbf{xy}}^2(\alpha) \triangleq E[\mathbf{xy}^{H_{-1}}]$, and $\Gamma_{\mathbf{xy}}^3(\alpha) \triangleq E[\mathbf{xy}^{H_\alpha}]$. Likewise, we denote $\Gamma_{\mathbf{xx}}^v(\alpha) \triangleq \Gamma_{\mathbf{x}}^v(\alpha), v = 1, 2, 3$.

Now, we extend some of the above concepts to the case of random signals. Consider a random signal vector $\mathbf{x}(t) = [x_1(t), \dots, x_p(t)]^T \in \mathbb{S}^p, t \in \mathbb{Z}$, where $x_i(t) = a_i(t) + ib_i(t) + jc_i(t) + kd_i(t), i = 1, \dots, p$, with $a_i(t), b_i(t), c_i(t)$, and $d_i(t)$ real random signals. The real vectors associated to $\mathbf{x}(t)$ is denoted by

$$\mathbf{x}_r(t) = [\mathbf{a}^T(t), \mathbf{b}^T(t), \mathbf{c}^T(t), \mathbf{d}^T(t)]^T, \quad t \in \mathbb{Z} \tag{3}$$

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

In a similar way to (2), we define the following functions for the random signal vectors $\mathbf{x}(t) \in \mathbb{S}^p$ and $\mathbf{y}(t) \in \mathbb{S}^q$:

$$\begin{aligned} \Gamma_{\mathbf{x}}^v(\alpha, t, s) &= E[\mathbf{x}(t)\mathbf{x}^{H_{3-2v}}(s)], & \Gamma_{\mathbf{xy}}^v(\alpha, t, s) &= E[\mathbf{x}(t)\mathbf{y}^{H_{3-2v}}(s)], & v &= 1, 2 \\ \Gamma_{\mathbf{x}}^3(\alpha, t, s) &= E[\mathbf{x}(t)\mathbf{x}^{H_\alpha}(s)], & \Gamma_{\mathbf{xy}}^3(\alpha, t, s) &= E[\mathbf{x}(t)\mathbf{y}^{H_\alpha}(s)] \end{aligned}$$

It is easy to check, for $\alpha \neq 0$, that

$$\Gamma_{\mathbf{xy}}^v(\alpha, t, s) = \Gamma_{\mathbf{yx}}^{vH_{3-2v}}(\alpha, s, t), \quad \forall t, s \in \mathbb{Z}, \quad v = 1, 2 \tag{4}$$

We define the augmented vector of $\mathbf{x}(t)$ as

$$\bar{\mathbf{x}}(t) = [\mathbf{x}^T(t), \mathbf{x}^{i^T}(t), \mathbf{x}^{j^T}(t), \mathbf{x}^{k^T}(t)]^T, \quad t \in \mathbb{Z} \tag{5}$$

with $\mathbf{x}^\eta(t) = [x_1^\eta(t), \dots, x_p^\eta(t)]^T$, for $\eta = i, j, k$. The following relationship between the augmented vector and the real vector $\mathbf{x}_r(t)$ given in (3) can be established

$$\bar{\mathbf{x}}(t) = \mathcal{T}_p \mathbf{x}_r(t) \tag{6}$$

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

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

Moreover, from Property 1, the next result can be stated.

Property 2.

$$\Gamma_{\bar{\mathbf{x}}}^v(\alpha, t, s) = \begin{pmatrix} \Gamma_{\mathbf{x}}^v(\alpha, t, s) & \Gamma_{\mathbf{xx}^i}^v(\alpha, t, s) & \Gamma_{\mathbf{xx}^j}^v(\alpha, t, s) & \Gamma_{\mathbf{xx}^k}^v(\alpha, t, s) \\ \Gamma_{\mathbf{xx}^i}^{vi}(\alpha, t, s) & \Gamma_{\mathbf{x}}^{vi}(\alpha, t, s) & \Gamma_{\mathbf{xx}^k}^{vi}(\alpha, t, s) & \Gamma_{\mathbf{xx}^j}^{vi}(\alpha, t, s) \\ \Gamma_{\mathbf{xx}^j}^{vj}(\alpha, t, s) & \Gamma_{\mathbf{xx}^k}^{vj}(\alpha, t, s) & \Gamma_{\mathbf{x}}^{vj}(\alpha, t, s) & \Gamma_{\mathbf{xx}^i}^{vj}(\alpha, t, s) \\ \Gamma_{\mathbf{xx}^k}^{vk}(\alpha, t, s) & \Gamma_{\mathbf{xx}^j}^{vk}(\alpha, t, s) & \Gamma_{\mathbf{x}}^{vk}(\alpha, t, s) & \Gamma_{\mathbf{xx}^i}^{vk}(\alpha, t, s) \end{pmatrix} \tag{7}$$

for $\alpha \neq 0$, $t, s \in \mathbb{Z}$ and $v = 1, 2, 3$. The functions $\Gamma_{\mathbf{x}}^v(\alpha, t, s)$, $\Gamma_{\mathbf{xx}^i}^v(\alpha, t, s)$, $\Gamma_{\mathbf{xx}^j}^v(\alpha, t, s)$ and $\Gamma_{\mathbf{xx}^k}^v(\alpha, t, s)$, $v = 1, 2, 3$, are obtained in Table A1 (see Appendix A) in function of the correlation functions (denoted by Γ) of the components of the real vector (3).

3. Properness Conditions

This section is devoted to analyze the concepts of properness for GSQ random signals. After introducing the notion of properness, a characterization of this property based on the correlations of the real vectors is presented.

Definition 1. A random signal vector $\mathbf{x}(t) \in \mathbb{S}^p$ is said to be H_{α_0} -proper if, and only if, there exists a value $\alpha_0 > 0$ such that the functions $\Gamma_{\mathbf{xx}^\eta}^1(\alpha_0, t, s)$, $\eta = i, j, k$, vanish $\forall t, s \in \mathbb{Z}$.

In a similar manner, two random signal vectors $\mathbf{x}(t) \in \mathbb{S}^p$ and $\mathbf{y}(t) \in \mathbb{S}^q$ are cross H_{α_0} -proper if, and only if, there exists a value $\alpha_0 > 0$ such that the functions $\Gamma_{\mathbf{xy}^\eta}^1(\alpha_0, t, s)$, $\eta = i, j, k$, vanish $\forall t, s \in \mathbb{Z}$. Finally, $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are jointly H_{α_0} -proper if, and only if, are H_{α_0} -proper and cross H_{α_0} -proper.

A similar definition about E_{α_0} -properness can be given by replacing $\alpha_0 > 0$ by $\alpha_0 < 0$, $\Gamma_{\mathbf{xx}^\eta}^1(\alpha_0, t, s)$ by $\Gamma_{\mathbf{xx}^\eta}^2(\alpha_0, t, s)$, and $\Gamma_{\mathbf{xy}^\eta}^1(\alpha_0, t, s)$ by $\Gamma_{\mathbf{xy}^\eta}^2(\alpha_0, t, s)$.

Remark 1. The \mathbb{T}_1 -properness condition introduced in [24] is a particular case of E_{α_0} -properness for $\alpha_0 = -1$.

From Table A1, we have the following characterization result.

Proposition 1. Let $\mathbf{x}(t) \in \mathbb{S}^p$ be a random signal vector. Then

- $\mathbf{x}(t)$ is H_{α_0} -proper if, and only if, the correlation of their real vectors verifies

$$\begin{aligned} \Gamma_{\mathbf{a}}(t, s) &= \Gamma_{\mathbf{c}}(t, s) = \alpha_0 \Gamma_{\mathbf{d}}(t, s), & \Gamma_{\mathbf{b}}(t, s) &= \Gamma_{\mathbf{d}}(t, s) \\ \Gamma_{\mathbf{ab}}(t, s) &= \Gamma_{\mathbf{ba}}(t, s) = \Gamma_{\mathbf{cd}}(t, s) = \Gamma_{\mathbf{dc}}(t, s) \\ \Gamma_{\mathbf{ac}}(t, s) &= \Gamma_{\mathbf{ca}}(t, s) = \alpha_0 \Gamma_{\mathbf{db}}(t, s), & \Gamma_{\mathbf{bd}}(t, s) &= \Gamma_{\mathbf{db}}(t, s) \end{aligned}$$

$$\Gamma_{ad}(t, s) = \Gamma_{bc}(t, s) = \Gamma_{cb}(t, s) = \Gamma_{da}(t, s)$$

with $\alpha_0 > 0$.

2. $\mathbf{x}(t)$ is E_{α_0} -proper if, and only if, the correlation of their real vectors verifies

$$\begin{aligned} \Gamma_{\mathbf{a}}(t, s) &= \Gamma_{\mathbf{c}}(t, s) = -\alpha_0 \Gamma_{\mathbf{d}}(t, s), & \Gamma_{\mathbf{b}}(t, s) &= \Gamma_{\mathbf{d}}(t, s) \\ -\Gamma_{\mathbf{ab}}(t, s) &= \Gamma_{\mathbf{ba}}(t, s) = -\Gamma_{\mathbf{cd}}(t, s) = \Gamma_{\mathbf{dc}}(t, s) \\ \Gamma_{\mathbf{ac}}(t, s) &= \Gamma_{\mathbf{ca}}(t, s) = -\alpha_0 \Gamma_{\mathbf{db}}(t, s), & \Gamma_{\mathbf{bd}}(t, s) &= \Gamma_{\mathbf{db}}(t, s) \\ -\Gamma_{\mathbf{ad}}(t, s) &= \Gamma_{\mathbf{bc}}(t, s) = -\Gamma_{\mathbf{cb}}(t, s) = \Gamma_{\mathbf{da}}(t, s) \end{aligned}$$

with $\alpha_0 < 0$.

From an experimental standpoint, it becomes necessary to develop a method for checking whether a random vector $\mathbf{x} \in \mathbb{S}^p$ is proper and devise an estimate of the associated α_0 . Specifically, based on the information supplied by a random sample, we aim to test if \mathbf{x} is H_{α_0} -proper (or E_{α_0} -proper) and we wish to estimate the suitable value of α_0 under which such properness properties could be satisfied by \mathbf{x} , if possible. To this end, we consider the following statistical hypotheses test:

$$\begin{aligned} H_0 &: \exists \alpha_0 \text{ such that } \mathbf{x} \text{ is } H_{\alpha_0}\text{-proper (respectively, } \mathbf{x} \text{ is } E_{\alpha_0}\text{-proper)} \\ H_1 &: \nexists \alpha_0 \text{ such that } \mathbf{x} \text{ is } H_{\alpha_0}\text{-proper (respectively, } \mathbf{x} \text{ is } E_{\alpha_0}\text{-proper)} \end{aligned} \tag{8}$$

The following result derives two generalized likelihood ratio tests (GLRT), Wilks' statistics, for the problem (8). These statistics are denoted by $\phi_\nu(\mathbf{x}_1, \dots, \mathbf{x}_n)$. The value assigned to the letter "ν" as subscript, superscript or in any other role will indicate when H_{α_0} -properness ($\nu = 1$) or E_{α_0} -properness ($\nu = 2$) should be assumed. This convention is applied henceforth. Notice that, these tests generalize the one given in ([24], Theorem 1) to the GSQs domain. The main difference between them lies in the choice of the appropriate α_0 to achieve properness conditions.

Theorem 1. Given n independent and identically distributed random samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ of a random vector $\mathbf{x} \in \mathbb{S}^p$ such that \mathbf{x}_r follows a Gaussian distribution then, the GLRT statistics for (8) are given by

$$\phi_\nu(\mathbf{x}_1, \dots, \mathbf{x}_n) = -n \left[\ln \left(\det \left(\hat{\Gamma}_{\mathbf{x}}^\nu(\alpha_0) \right)_{\alpha_0} \right) - \ln \left(\det \left(\Gamma_{\mathbb{S}}^\nu(\alpha_0) \right)_{\alpha_0} \right) \right], \quad \nu = 1, 2$$

where $\det(\cdot)_{\alpha_0}$ denotes the determinant under the product associated to α_0 ,

$$\hat{\Gamma}_{\mathbf{x}}^\nu(\alpha_0) = \mathcal{T}_p \hat{\Gamma}_{\mathbf{x}_r} \mathcal{T}_p^{H_3-2\nu}, \quad \nu = 1, 2$$

$\hat{\Gamma}_{\mathbf{x}_r}$ is the sample autocorrelation matrix and

$$\Gamma_{\mathbb{S}}^\nu(\alpha_0) = \text{diag} \left(\hat{\Gamma}_{\mathbf{x}}^\nu(\alpha_0), \hat{\Gamma}_{\mathbf{x}}^{\nu i}(\alpha_0), \hat{\Gamma}_{\mathbf{x}}^{\nu j}(\alpha_0), \hat{\Gamma}_{\mathbf{x}}^{\nu k}(\alpha_0) \right), \quad \nu = 1, 2$$

The parameter α_0 can be estimated as

$$\begin{aligned} \alpha_0 &= \max_{\alpha, \alpha > 0} \left[c_1 - \frac{n}{2} [\ln(\det(\Gamma_{\mathbb{S}}^1(\alpha))_\alpha) + 4p] \right], & \nu &= 1 \\ \alpha_0 &= \max_{\alpha, \alpha < 0} \left[c_1 - \frac{n}{2} [\ln(\det(\Gamma_{\mathbb{S}}^2(\alpha))_\alpha) + 4p] \right], & \nu &= 2 \end{aligned} \tag{9}$$

with $c_1 = -2np \ln(2\pi) + np \ln(16|\alpha|)$.

Further, assuming that H_0 is true, the distributions of the statistics $\phi_1(\mathbf{x}_1, \dots, \mathbf{x}_n)$ and $\phi_2(\mathbf{x}_1, \dots, \mathbf{x}_n)$ tend to a chi-squared distribution with degrees of freedom equal to $6p^2 - 1$ and $2p(3p + 1) - 1$, respectively, as the sample size tends to infinity.

Example

Consider a random vector $\mathbf{x} = [x_1, x_2]^T \in \mathbb{S}^2$ such that the real vectors of x_1 and x_2 given by (3) verify

$$\Gamma_{\mathbf{x}_{ir}} = \begin{pmatrix} |\alpha_0|6 & \lambda & -|\alpha_0|0.4 & \nu \\ \lambda & 6 & \mu & -0.4 \\ -|\alpha_0|0.4 & \mu & |\alpha_0|6 & \lambda \\ \mu & -0.4 & \lambda & 6 \end{pmatrix}, \quad i = 1, 2 \tag{10}$$

and

$$\Gamma_{\mathbf{x}_{1r}, \mathbf{x}_{2r}} = \begin{pmatrix} 2 & (-1)^{\nu+1}0.5 & 1.2 & -(-1)^{\nu+1}0.4 \\ 0.5 & 1 & -0.4 & 0.6 \\ 1.2 & -(-1)^{\nu+1}0.4 & 2 & (-1)^{\nu+1}0.5 \\ -0.4 & 0.6 & 0.5 & 1 \end{pmatrix}$$

The following two cases are studied:

1. Case 1: $\lambda = -0.7$ and $\mu = 0.8$ in an hyperbolic quaternions algebra with $\alpha_0 = 2$. From Proposition 1, we have that the vector \mathbf{x} is H_2 -proper under these conditions. Moreover, from Theorem 1, the GLRT statistic $\phi_1(\mathbf{x}_1, \dots, \mathbf{x}_n)$ converges to a $\chi^2(23)$.
2. Case 2: $\lambda = \mu = 0$ in an elliptic quaternions algebra with $\alpha_0 = -2$. It follows that \mathbf{x} is E_{-2} -proper. Further, $\phi_2(\mathbf{x}_1, \dots, \mathbf{x}_n)$ converges to a $\chi^2(27)$.

On the one hand, the behavior of the cumulative distribution functions (CDFs) of $\phi_\nu(\mathbf{x}_1, \dots, \mathbf{x}_n)$, $\nu = 1, 2$, is analyzed in both cases in terms of the sample size. For that, and through Montecarlo simulations, 2000 values of the statistics $\phi_\nu(\mathbf{x}_1, \dots, \mathbf{x}_n)$, $\nu = 1, 2$, have been generated for $n = 50, 100, 300, 500$. The corresponding empirical CDFs are displayed in Figure 1. At a glance, this figure shows the rapid convergence of the empirical CDFs to the target chi-squared distributions.

On the other hand, we assess the accuracy of the estimations of α_0 depending on the sample size n . The maximization process in (9) can be achieved using the optimization toolbox in Matlab (The Nelder–Mead search algorithm is used). Tables 2 and 3 depict the 95% confidence intervals for α_0 obtained from 2000 simulated runs. As can be expected, the amplitudes of the confidence intervals become narrower as the sample size grows.

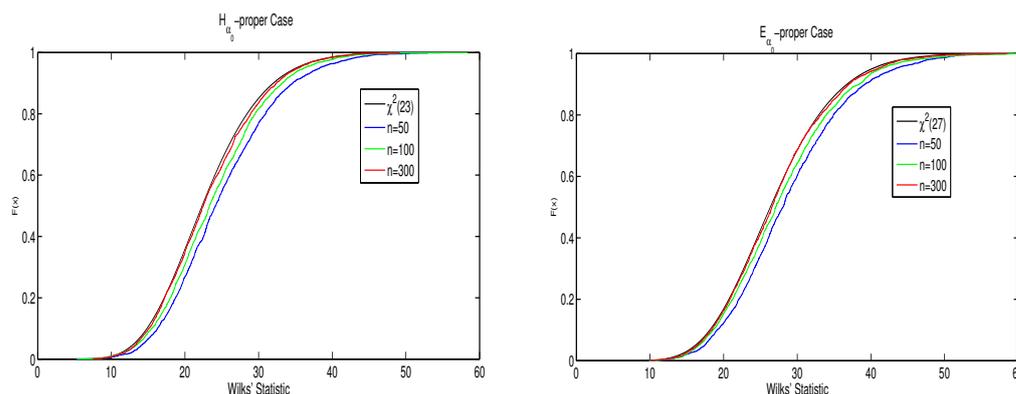


Figure 1. Wilks’ statistic for case 1 (left) and case 2 (right).

Table 2. The 95% confidence intervals for α_0 in case 1.

n	50	100	300	500
C.I.	[1.506, 2.615]	[1.642, 2.433]	[1.788, 2.236]	[1.835, 2.190]

Table 3. The 95% confidence intervals for α_0 in case 2.

n	50	100	300	500
C.I.	$[-2.650, -1.512]$	$[-2.443, -1.633]$	$[-2.233, -1.777]$	$[-2.181, -1.846]$

4. Linear Estimation and Prediction

The objective of this section is to solve the linear estimation problem for proper GSQ random variables. Since it is not possible to define a norm in the set of GSQs then, we cannot apply the projection theorem to obtain the optimum linear estimator. So, we first need to define a suitable metric and thus, we prove the existence and the uniqueness of the projection.

Given the random variables $x, y \in \mathbb{S}$, we define the product

$$\langle x, y \rangle_\alpha = E[xy_\alpha] = \Gamma_{xy}^3(\alpha)$$

where $\Gamma_{xy}^3(\alpha)$ is the scalar version of $\Gamma_{xy}^3(\alpha)$ given in (2).

For any random variables $x, y, u, v \in \mathbb{S}$ and deterministic coefficients $\lambda, \beta \in \mathbb{S}, \gamma, \delta \in \mathbb{R}$, it follows that

- $\langle \lambda x + \beta y, \gamma u + \delta v \rangle_\alpha = \lambda \gamma \langle x, u \rangle_\alpha + \lambda \delta \langle x, v \rangle_\alpha + \beta \gamma \langle y, u \rangle_\alpha + \beta \delta \langle y, v \rangle_\alpha.$
- $\langle x, x \rangle_\alpha = 0 \Leftrightarrow x = 0.$

Thus, we can define a distance as

$$d_\alpha(x, y) = \|x - y\|_\alpha \tag{11}$$

where $\|x\|_\alpha^2 = \mathcal{R}\{\Gamma_x^3(\alpha)\}.$

Consider the set of random vectors $\mathcal{C}_\alpha = \{x_1, \dots, x_m\}$, with $x_i \in \mathbb{S}^p, i = 1, \dots, m$, and denote by $\mathcal{G}_{\mathcal{C}_\alpha}$ the linear span of \mathcal{C}_α , i.e., the set of all finite linear combinations of elements of \mathcal{C}_α . The issue of existence and uniqueness of the projection of a given random variable $y \in \mathbb{S}$ onto the set $\mathcal{G}_{\mathcal{C}_\alpha}$ is now briefly investigated. This projection is denoted by \hat{y}_α .

In a metric space, neither the existence nor the uniqueness of \hat{y}_α is assured; however, following a reasoning similar to [24], the existence and uniqueness of \hat{y}_α can be proved. Further, \hat{y}_α is the projection of y onto the set $\mathcal{G}_{\mathcal{C}_\alpha}$ if, and only if, $\langle y - \hat{y}_\alpha, x \rangle_\alpha = 0, \forall x \in \mathcal{G}_{\mathcal{C}_\alpha}$. As a consequence, if $\{\vartheta_1, \dots, \vartheta_r\}$ is a basis of $\mathcal{G}_{\mathcal{C}_\alpha}$ then,

$$\hat{y}_\alpha = \sum_{i=1}^r h_i \vartheta_i \tag{12}$$

where the deterministic coefficients $h_i \in \mathbb{S}$ are obtained from the system

$$\langle y, \vartheta_j \rangle_\alpha = \sum_{i=1}^r h_i \langle \vartheta_i, \vartheta_j \rangle_\alpha, \quad j = 1, \dots, r$$

Now, we are prepared to introduce the notion of linear estimator in the GSQ domain.

Definition 2. The projection of y onto the set $\mathcal{G}_{\mathcal{C}_\alpha}, \hat{y}_\alpha$, is called the \mathbb{S}_α estimator of y respect to \mathcal{C}_α and, from now on, it is denoted by \hat{y}_α^S . If the information set is formed by augmented vectors defined in (5), i.e., if we consider $\bar{\mathcal{C}}_\alpha = \{\bar{x}_1, \dots, \bar{x}_m\}$ with $\bar{x}_i \in \mathbb{S}^{4p}, i = 1, \dots, m$, then, the projection of y onto the set $\mathcal{G}_{\bar{\mathcal{C}}_\alpha}$ is called SWL estimator of y respect to \mathcal{C}_α and it is denoted by \hat{y}_α^{SWL} .

These concepts are easily extended to the vectorial case. For example, the \mathbb{S}_α estimator of the random vector $y \in \mathbb{S}^q$ respect to \mathcal{C}_α is $\hat{y}_\alpha^S = [\hat{y}_{1\alpha}^S, \dots, \hat{y}_{q\alpha}^S]$, where $\hat{y}_{j\alpha}^S$ is the projection of the component j th of y onto the set $\mathcal{G}_{\mathcal{C}_\alpha}$.

Theorem 2. Consider a random variable $y \in \mathbb{S}$ and the set of random vectors $C_\alpha = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$, with $\mathbf{x}_i \in \mathbb{S}^p, i = 1, \dots, m$. Then,

1. The SWL estimator of y respect to C_α is obtained as

$$\hat{y}_\alpha^{SWL} = \sum_{i=1}^m \left(\mathbf{h}_{1i}^T(\alpha)\mathbf{x}_i + \mathbf{h}_{2i}^T(\alpha)\mathbf{x}_i^i + \mathbf{h}_{3i}^T(\alpha)\mathbf{x}_i^j + \mathbf{h}_{4i}^T(\alpha)\mathbf{x}_i^k \right)$$

where the deterministic vectors $\mathbf{h}_{ji}(\alpha) \in \mathbb{S}^p$ are computed through the equation

$$\begin{bmatrix} \mathbf{h}_{11}^T(\alpha), \dots, \mathbf{h}_{1m}^T(\alpha), \mathbf{h}_{21}^T(\alpha), \dots, \mathbf{h}_{2m}^T(\alpha), \\ \mathbf{h}_{31}^T(\alpha), \dots, \mathbf{h}_{3m}^T(\alpha), \mathbf{h}_{41}^T(\alpha), \dots, \mathbf{h}_{4m}^T(\alpha) \end{bmatrix} = \mathbf{\Gamma}_{\mathbf{y}\bar{\mathbf{z}}}^3(\alpha)\mathbf{\Gamma}_{\bar{\mathbf{z}}}^{3^{-1}}(\alpha) \tag{13}$$

with $\bar{\mathbf{z}} = [\mathbf{x}_1^T, \dots, \mathbf{x}_m^T, \mathbf{x}_1^{iT}, \dots, \mathbf{x}_m^{iT}, \mathbf{x}_1^{jT}, \dots, \mathbf{x}_m^{jT}, \mathbf{x}_1^{kT}, \dots, \mathbf{x}_m^{kT}]^T$.

Moreover, this estimator and its associated error are independent of the value of α , i.e., $\hat{y}^{SWL} = \hat{y}_\alpha^{SWL}$ and $\epsilon^{SWL} = \epsilon_\alpha^{SWL}, \forall \alpha \neq 0$, where $\epsilon_\alpha^{SWL} = \|y - \hat{y}_\alpha^{SWL}\|_\alpha^2$.

2. The \mathbb{S}_α estimator of y respect to C_α is calculated as

$$\hat{y}_\alpha^S = \sum_{i=1}^m \mathbf{g}_i^T(\alpha)\mathbf{x}_i$$

where the deterministic vectors $\mathbf{g}_i(\alpha) \in \mathbb{S}^p$ are computed through the equation

$$[\mathbf{g}_1^T(\alpha), \dots, \mathbf{g}_m^T(\alpha)] = \mathbf{\Gamma}_{\mathbf{y}\mathbf{z}}^3(\alpha)\mathbf{\Gamma}_{\mathbf{z}}^{3^{-1}}(\alpha) \tag{14}$$

with

$$\mathbf{z} = [\mathbf{x}_1^T, \dots, \mathbf{x}_m^T]^T \tag{15}$$

3. $\epsilon^{SWL} \leq \epsilon_\alpha^S = \|y - \hat{y}_\alpha^S\|_\alpha^2, \forall \alpha \neq 0$.
4. If $\mathbf{x}_1, \dots, \mathbf{x}_m$ are jointly H_{α_0} -proper (respectively, E_{α_0} -proper) and y is cross H_{α_0} -proper (respectively, E_{α_0} -proper) with each element of $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ then, $\hat{y}^{SWL} = \hat{y}_{\alpha_0}^S$.

Remark 2. Theorem 2 shows that whatever hypercomplex algebra you choose provides the same SWL estimator. Likewise, the computational burden of the SWL estimator can be notably reduced whenever hyperbolic or elliptic conditions of properness are fulfilled (compare (13) with (14)). Effectively, while the computational complexity of SWL estimators is of order $\mathcal{O}(64p^3m^3)$, this is of order $\mathcal{O}(p^3m^3)$ under hyperbolic or elliptic properness. More importantly, this reduction in computational complexity cannot be achieved in the real numbers domain.

Now, we are concerned with the prediction problem for random signals. Consider a random signal vector $\mathbf{x}(t) \in \mathbb{S}^p$. Our aim is to obtain the \mathbb{S}_α one-stage predictor of $\mathbf{x}(t + 1)$ from the information contained in the set $C_\alpha^t = \{\mathbf{x}(1), \dots, \mathbf{x}(t)\}$, denoted by $\hat{\mathbf{x}}_\alpha^S(t + 1|t)$. Analogously, the SWL one-stage predictor is denoted by $\hat{\mathbf{x}}^{SWL}(t + 1|t)$.

Note that a similar result to Theorem 2 can be stated for random signals, but we do not include it here for the sake of brevity. It should be remarked, however, that a computational problem remains still unsolved. Although the estimation problem is indeed simplified due to the reduction in dimension implied by the properness conditions, we have to cope with the inverse in (14) as the number of observations t grows. The following result overcomes this problem for proper signals.

Theorem 3 (The innovations algorithm). Consider a H_{α_0} -proper (respectively, E_{α_0} -proper) random signal vector $\mathbf{x}(t) \in \mathbb{S}^p$. If $\mathbf{\Gamma}_\mathbf{x}^v(\alpha_0, t, s)$ is nonsingular for $v = 1, 2$ then, $\hat{\mathbf{x}}^{SWL}(t + 1|t) = \hat{\mathbf{x}}_{\alpha_0}^S(t + 1|t)$ and can be recursively obtained in the following way:

$$\hat{\mathbf{x}}_{\alpha_0}^S(t + 1|t) = \sum_{i=1}^t \mathbf{\Lambda}(t, i) (\mathbf{x}(t + 1 - i) - \hat{\mathbf{x}}_{\alpha_0}^S(t + 1 - i|t - i)) \quad t \geq 1 \tag{16}$$

with $\hat{\mathbf{x}}_{\alpha_0}^S(1|0) = \mathbf{0}_{p \times 1}$ and where

$$\mathbf{\Lambda}(t, t - k) = \left(\mathbf{\Gamma}_x^v(\alpha_0, t + 1, k + 1) - \sum_{i=0}^{k-1} \mathbf{\Lambda}(t, t - i) \mathbf{\Pi}(i) \mathbf{\Lambda}^{H_{3-2v}}(k, k - i) \right) \mathbf{\Pi}^{-1}(k) \tag{17}$$

for $k = 1, \dots, t - 1$, with $\mathbf{\Lambda}(t, t) = \mathbf{\Gamma}_x^v(\alpha_0, t + 1, 1) \mathbf{\Pi}^{-1}(0)$ and

$$\mathbf{\Pi}(t) = \mathbf{\Gamma}_x^v(\alpha_0, t + 1, t + 1) - \sum_{i=0}^{t-1} \mathbf{\Lambda}(t, t - i) \mathbf{\Pi}(i) \mathbf{\Lambda}^{H_{3-2v}}(t, t - i) \tag{18}$$

with the initial condition $\mathbf{\Pi}(0) = \mathbf{\Gamma}_x^v(\alpha_0, 1, 1)$. Moreover, the error associated to the component $x_j(t + 1)$ is given by

$$\epsilon_j^{SWL}(t + 1|t) = \frac{2|\alpha_0| + 2}{4|\alpha_0|} \mathcal{R}\{\mathbf{\Pi}_{jj}(t)\} \tag{19}$$

5. Proper ARMA Models

This section constitutes the core of the paper. Starting from the definition of proper ARMA models in the GSQs domain we then solve the one-stage prediction problem.

5.1. ARMA Modeling

Next, we adapt the proper ARMA signal concept to the GSQs domain and extend the Yule–Walker equations to this kind of signals. To do so, we first extend the concept of stationarity to the GSQs domain in the proper signal scenario.

Definition 3. A H_{α_0} -proper random signal vector $\mathbf{x}(t) \in \mathbb{S}^p$ is said to be stationary if $\mathbf{\Gamma}_x^1(\alpha_0, t + h, t) = \mathbf{\Gamma}_x^1(\alpha_0, s + h, s), \forall t, s, h \in \mathbb{Z}$, i.e., $\mathbf{\Gamma}_x^1$ only depends on the lag h for a given α_0 . In this case, we denote $\mathbf{\Gamma}_x^1(\alpha_0, h) = \mathbf{\Gamma}_x^1(\alpha_0, t + h, t)$. The concept of a E_{α_0} -proper stationary signal is defined analogously by replacing $\mathbf{\Gamma}_x^1$ by $\mathbf{\Gamma}_x^2$.

Definition 4. A random signal vector $\mathbf{x}(t) \in \mathbb{S}^p$ is said to be a H_{α_0} -proper ARMA signal of orders p and q (ARMA(p, q)) if it is H_{α_0} -proper, stationary, and can be modeled by the following system:

$$\mathbf{x}(t) = \sum_{i=1}^p \mathbf{F}(i) \mathbf{x}(t - i) + \sum_{i=0}^q \mathbf{G}(i) \mathbf{v}(t - i) \tag{20}$$

where $\mathbf{G}(0) = \mathbf{I}_p, \mathbf{F}(i), \mathbf{G}(i) \in \mathbb{S}^{p \times p}$ are deterministic matrices and $\mathbf{v}(t)$ is a H_{α_0} -proper white noise verifying

$$\mathbf{\Gamma}_v^1(\alpha_0, t, s) = \mathbf{\Omega} \delta_{ts} \tag{21}$$

with $\mathbf{\Omega} \in \mathbb{S}^{p \times p}$ a deterministic matrix.

The concept of a E_{α_0} -proper ARMA(p, q) signal is defined analogously by replacing H_{α_0} -proper by E_{α_0} -proper, and $\mathbf{\Gamma}_v^1$ by $\mathbf{\Gamma}_v^2$ in (21).

The following result extends the Yule–Walker equations to the case of proper ARMA signals.

Proposition 2. Let $\mathbf{x}(t) \in \mathbb{S}^p$ be a H_{α_0} -proper (or E_{α_0} -proper) ARMA(p, q) signal. Consider the matrix-valued polynomial $\mathfrak{F}(z) = \mathbf{I}_{4p} - \sum_{i=1}^p \bar{\mathbf{F}}(i) z^i$ with $\bar{\mathbf{F}}(i) = \text{diag}(\mathbf{F}(i), \mathbf{F}^i(i), \mathbf{F}^j(i), \mathbf{F}^k(i)), i = 1, \dots, p$. If $\det(\mathfrak{F}(z))_{\alpha_0} \neq 0, \forall z \in \mathbb{C}$ such that $|z| \leq 1$ then, $\mathbf{\Gamma}_x^v(\alpha_0, h), v = 1, 2$, obeys the equation

$$\mathbf{\Gamma}_x^v(\alpha_0, h) - \sum_{i=1}^p \mathbf{F}(i) \mathbf{\Gamma}_x^v(\alpha_0, h - i) = \sum_{h \leq i \leq q} \mathbf{G}(i) \mathbf{\Omega} \mathbf{\Lambda}^{H_{3-2v}}(i - h), \quad h \geq 0 \tag{22}$$

where the matrices $\mathbf{L}(t)$ are determined by

$$\begin{aligned} \mathbf{L}(0) &= \mathbf{I}_p, & \mathbf{L}(t) &= \mathbf{0}_{p \times p}, \quad t < 0 \\ \mathbf{L}(t) &= \sum_{i=1}^t \mathbf{F}(i)\mathbf{L}(t-i) + \mathbf{G}(t), & t &\geq 1 \end{aligned}$$

with $\mathbf{F}(t) = \mathbf{0}_{p \times p}, t > p$ and $\mathbf{G}(t) = \mathbf{0}_{p \times p}, t > q$.

5.2. One-Stage Prediction

We aim to improve the recursive algorithm for $\hat{\mathbf{x}}_{\alpha_0}^S(t+1|t)$ given in Theorem 3 by making use of the particular structure of proper ARMA(p, q) models.

Theorem 4. Let $\mathbf{x}(t) \in \mathbb{S}^p$ be a H_{α_0} -proper (or E_{α_0} -proper) ARMA(p, q) signal. If $\det(\mathfrak{F}(z))_{\alpha_0} \neq 0, \forall z \in \mathbb{C}$ such that $|z| \leq 1$ then, $\hat{\mathbf{x}}^{\mathcal{S}WL}(t+1|t) = \hat{\mathbf{x}}_{\alpha_0}^S(t+1|t)$ and this predictor can be calculated through the following recursive procedure:

$$\hat{\mathbf{x}}_{\alpha_0}^S(t+1|t) = \begin{cases} \sum_{i=1}^t \Lambda(t,i)(\mathbf{x}(t+1-i) - \hat{\mathbf{x}}_{\alpha_0}^S(t+1-i|t-i)), & 1 \leq t < m \\ \sum_{i=1}^p \mathbf{F}(i)\mathbf{x}(t+1-i) + \sum_{i=1}^q \Lambda(t,i)(\mathbf{x}(t+1-i) - \hat{\mathbf{x}}_{\alpha_0}^S(t+1-i|t-i)), & t \geq m \end{cases} \quad (23)$$

with $\hat{\mathbf{x}}_{\alpha_0}^S(1|0) = \mathbf{0}_{p \times 1}, m = \max(p, q)$ and the matrices $\Lambda(t, i)$ are calculated by replacing $\Gamma_{\mathbf{x}}^v$ in (17) and (18) by the following function:

$$\mathbf{Y}^v(t, s) = \begin{cases} \Gamma_{\mathbf{x}}^v(\alpha_0, t-s), & 1 \leq t \leq s \leq m \\ \Gamma_{\mathbf{x}}^v(\alpha_0, t-s) - \sum_{i=1}^p \Gamma_{\mathbf{x}}^v(\alpha_0, t+i-s)\mathbf{F}^{\#_{3-2v}}(i), & 1 \leq t \leq m \leq s \leq 2m \\ \mathbf{0}_{p \times p}, & 1 \leq t \leq m, s > 2m \\ \sum_{i=0}^q \mathbf{G}(i)\Omega\mathbf{G}^{\#_{3-2v}}(i+s-t), & m < t \leq s \leq t+q \\ \mathbf{0}_{p \times p}, & m < t, t+q < s \\ \mathbf{Y}^{v\#_{3-2v}}(s, t), & s < t \end{cases} \quad (24)$$

with $\mathbf{G}(i) = \mathbf{0}_{p \times p}$ for $i > q, \Omega$ given in (21), and $\Gamma_{\mathbf{x}}^v(\alpha_0, t-s)$ obeys (22). Moreover, the error associated to (23) coincides with (19).

Remark 3. The advantage of working with this algorithm is that $\mathbf{Y}^v(t, s) = \mathbf{0}_{p \times p}$ for $|s-t| > q$ and thus, unlike the innovations algorithm given in Theorem 3, only $\Lambda(t, i)$, for $i = 1, \dots, q$, have to be computed in each recursion.

Corollary 1. For an ARMA(1, 1) signal it follows that

$$\mathbf{Y}^v(t, s) = \begin{cases} \Gamma_{\mathbf{x}}^v(\alpha_0, 0), & t = s = 1 \\ \Omega\mathbf{G}^{\#_{3-2v}}(1), & 1 \leq t, s = t+1 \\ \Omega + \mathbf{G}(1)\Omega\mathbf{G}^{\#_{3-2v}}(1), & 1 < t = s \\ \mathbf{0}_{p \times p}, & 1 \leq t, t+1 < s \\ \mathbf{Y}^{v\#_{3-2v}}(s, t), & s < t \end{cases}, \quad v = 1, 2$$

and the one-stage predictor is computed through

$$\hat{\mathbf{x}}_{\alpha_0}^S(t+1|t) = \mathbf{F}(1)\mathbf{x}(t) + \Lambda(t, 1)(\mathbf{x}(t) - \hat{\mathbf{x}}_{\alpha_0}^S(t|t-1)), \quad t \geq 1 \quad (25)$$

where

$$\begin{aligned} \Lambda(t, 1) &= \mathbf{G}(1)\mathbf{\Omega}\mathbf{\Pi}^{-1}(t - 1) \\ \mathbf{\Pi}(t) &= \mathbf{\Omega} + \mathbf{G}(1)\mathbf{\Omega}\mathbf{G}^{H_{3-2v}}(1) - \Lambda(t, 1)\mathbf{\Pi}(t - 1)\Lambda^{H_{3-2v}}(t, 1) \end{aligned} \tag{26}$$

and $\mathbf{\Pi}(0) = \mathbf{\Gamma}_x^v(\alpha_0, 0)$ is computed solving the following system of matrix equations:

$$\begin{aligned} \mathbf{\Gamma}_x^v(\alpha_0, 0) - \mathbf{F}(1)\mathbf{\Gamma}_x^{vH_{3-2v}}(\alpha_0, -1) &= \mathbf{\Omega} + \mathbf{G}(1)\mathbf{\Omega}(\mathbf{F}^{H_{3-2v}}(1) + \mathbf{G}^{H_{3-2v}}(1)) \\ \mathbf{\Gamma}_x^v(\alpha_0, 1) - \mathbf{F}(1)\mathbf{\Gamma}_x^v(\alpha_0, 0) &= \mathbf{G}(1)\mathbf{\Omega} \end{aligned}$$

To conclude, a steady-state performance analysis for the one-stage predictor (23) and its associated error is fulfilled.

Corollary 2. Under conditions of Theorem 4, if $\mathfrak{G}(z)$ given in (A4) satisfies $\det(\mathfrak{G}(z))_{\alpha_0} \neq 0$, $\forall z \in \mathbb{C}$ such that $|z| \leq 1$, then,

$$\Lambda(t, i) \xrightarrow{t \uparrow \infty} \mathbf{G}(i), \quad \epsilon_j^{SWL}(t + 1|t) \xrightarrow{t \uparrow \infty} \epsilon_j \tag{27}$$

with

$$\epsilon_j = \frac{2|\alpha_0| + 2}{4|\alpha_0|} \mathcal{R}\{\Omega_{jj}\} \tag{28}$$

being Ω_{jj} the j th diagonal entry of $\mathbf{\Omega}$ given in (21).

5.3. Example

By way of illustration, we consider a proper ARMA(1,1) model. With this model, we first carry out a performance comparative analysis between the \mathbb{S}_{α_0} one-stage predictor (25), and its counterpart in the HQ domain, $\hat{x}_{\alpha_0}^{QSL}(t + 1|t)$. Our aim is to show the superiority in terms of performance of the proposed approach respect to the usual HQ methodology under different initial conditions. Finally, a steady-state performance analysis is carried out.

Let the ARMA(1,1) system be

$$x(t) = fx(t - 1) + v(t) + gv(t - 1)$$

with $f = 0.3 - 0.05j + 0.1k$, $g = 0.1 + 0.2j + 0.4k$ and $v(t)$ a white noise whose real vector (3) verifies $\mathbf{\Gamma}_{v_r}(t, s) = \mathbf{\Omega}\delta_{ts}$ where $\mathbf{\Omega}$ is the matrix given in (10).

Two scenarios are contemplated: in the first one, we consider that $\lambda = -0.7$ and $\mu = 0.8$ in $\mathbf{\Omega}$, and three different hyperbolic quaternion algebras are studied ($\alpha_0 = 1, 2, 3$). In the second one, we set the value of the parameters λ and μ to be both 0 and three elliptic quaternion algebras are analyzed ($\alpha_0 = -1, -2, -3$). These scenarios present an interesting characteristic: from Proposition 1, $x(t)$ is either H_{α_0} -proper (in scenario 1) or E_{α_0} -proper (in scenario 2); however, the degree of improperness of $x(t)$ in the HQ system increases as $|\alpha_0|$ grows.

Note that the conditions of Proposition 2 and Corollary 2 are fulfilled and hence, $\hat{x}_{\alpha_0}^S(t + 1|t)$ can be obtained through the recursions (25) and (26) and the steady-state results in (27) hold.

The errors associated to $\hat{x}_{\alpha_0}^S(t + 1|t)$ and $\hat{x}_{\alpha_0}^{QSL}(t + 1|t)$ depend on the value α_0 and thus, they are denoted by $\epsilon_{\alpha_0}^{SWL}(t + 1|t)$ and $\epsilon_{\alpha_0}^{QSL}(t + 1|t)$, respectively. To assess the performance of both approaches, the following measure is computed:

$$DP_{\alpha_0}(t) = \epsilon_{\alpha_0}^{QSL}(t + 1|t) - \epsilon_{\alpha_0}^{SWL}(t + 1|t) \tag{29}$$

Figure 2 depicts the measure $DP_{\alpha_0}(t)$ in (29), for $t = 1, \dots, 100$, in the above two scenarios. As can be seen, the \mathbb{S}_{α_0} one-stage predictor outperforms the HQ one-stage predictor in both scenarios. Furthermore, the measure becomes greater as $|\alpha_0|$ grows, i.e., as the degree of HQ improperness grows; being $DP_{\alpha_0}(t)$ slightly greater for the E_{α_0} -proper case.

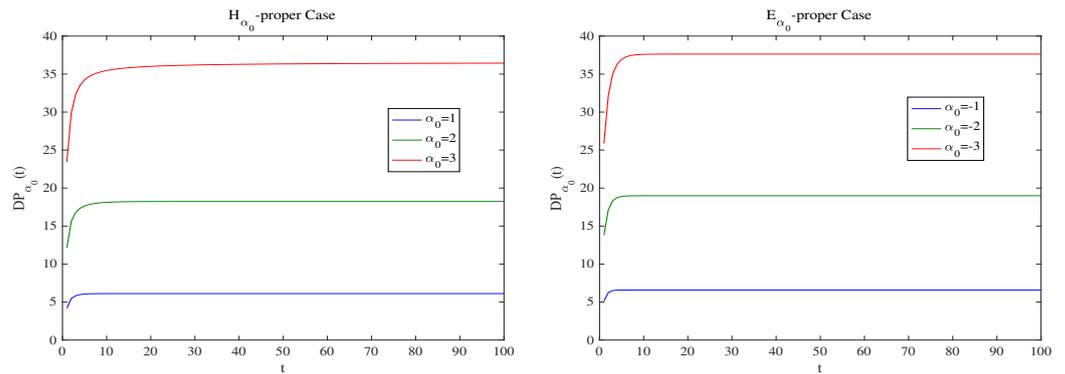


Figure 2. $DP_{\alpha_0}(t)$ in scenarios 1 (left) and 2 (right).

Now, we analyze the steady-state performance of both the estimate (25) and its error. The errors (28) computed for different values of α_0 appear in Table 4. Further, Figure 3 shows as $\epsilon_{\alpha_0}^{SWL}(t+1|t)$ converges toward their respective steady-state errors when t grows. These plots indicate that the convergence is rapid.

Finally, Figure 4 displays the convergence performance of the components of $\Lambda(t, 1) = \Lambda_1(t, 1) + \Lambda_2(t, 1)i + \Lambda_3(t, 1)j + \Lambda_4(t, 1)k$ given in (26). As could be expected from (27), we also obtain here a rapid convergence toward the target values and, in general, this convergence is slightly slower as $|\alpha_0|$ grows as well as in the H_{α_0} -proper case.

Table 4. Steady-state errors of the one-stage predictor for the different values of α_0 .

$ \alpha_0 $	1	2	3
ϵ	24	36	48

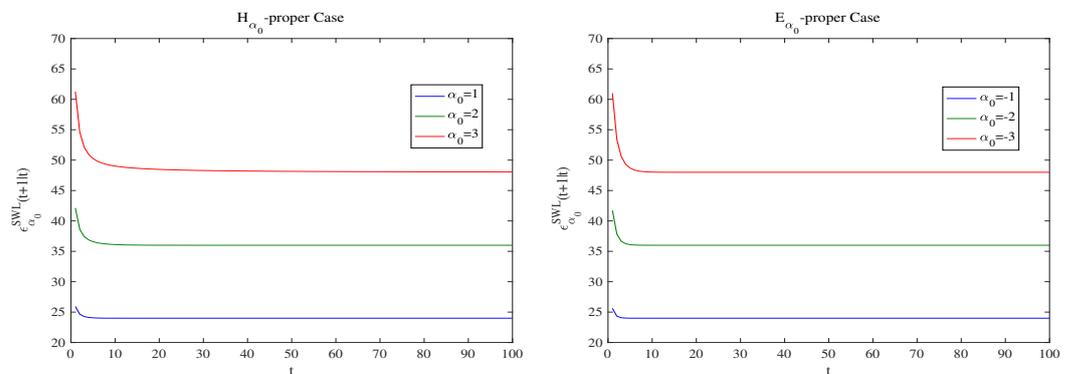


Figure 3. $\epsilon_{\alpha_0}^{SWL}(t+1|t)$ for $\alpha_0 = -3, -2, -1, 1, 2, 3$.

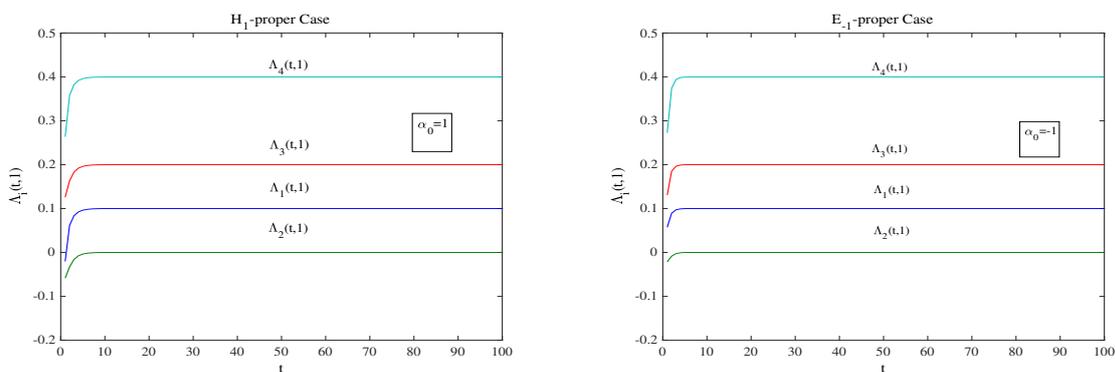


Figure 4. Cont.

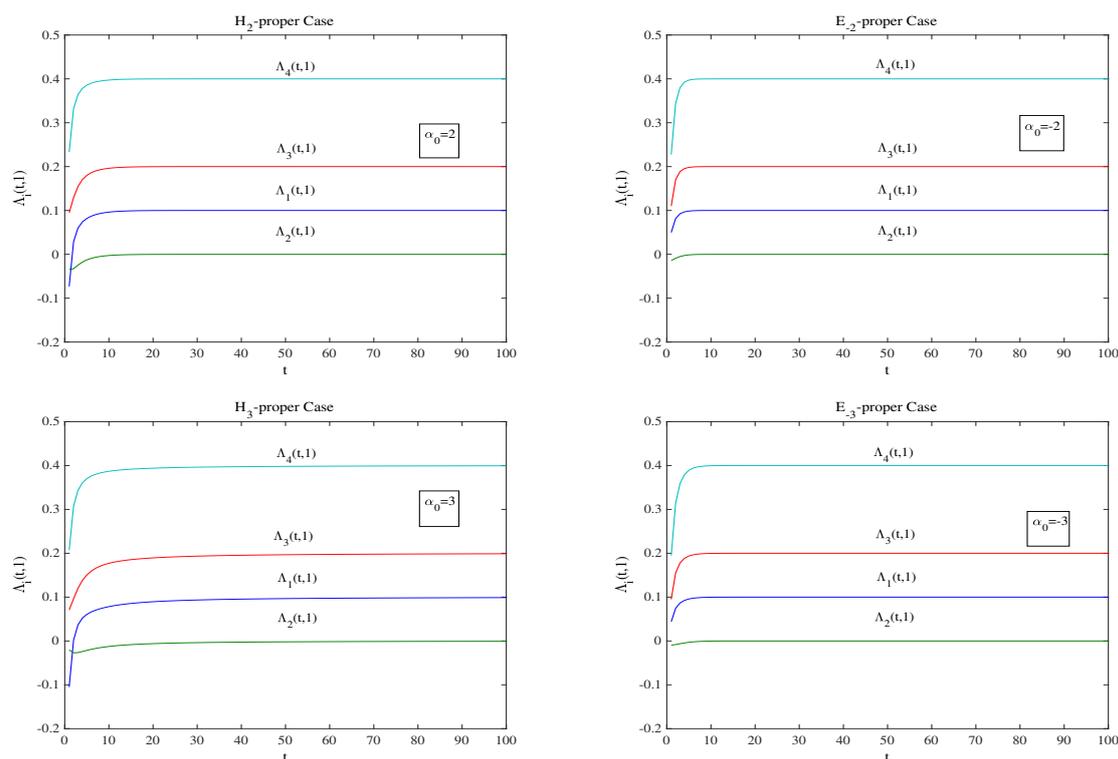


Figure 4. Values of the components $\Lambda_i(t, 1), i = 1, \dots, 4$, for $\alpha_0 = -3, -2, -1, 1, 2, 3$.

6. Discussion

Recent investigations have shown that HQs are not always the best choice for processing 4D hypercomplex signals. These results suggest distinct strategies depending on whether properness conditions are complied with. As the main characteristic, they offer a significant computational saving in relation to the real-valued processing. Unlike earlier research, the approach described in this paper has the additional advantage of the flexibility in the choice of the hypercomplex algebra. In our opinion, this research topic opens a broad range of possibilities for studying new applications in signal processing. This paper precisely proposes the time series analysis as a possible field of application of the proper GSQ signal processing. Other potential areas include dynamic linear models, filtering and smoothing, signal detection, etc.

Author Contributions: Conceptualization, J.C.R.-M.; Formal analysis, J.N.-M.; Methodology, J.N.-M.; Investigation, J.N.-M., R.M.F.-A. and J.C.R.-M.; Visualization, J.N.-M., R.M.F.-A. and J.C.R.-M.; Writing—original draft preparation, J.N.-M.; Writing—review and editing, J.N.-M., R.M.F.-A. and J.C.R.-M.; Funding acquisition, J.N.-M. and R.M.F.-A.; Project administration, J.N.-M. and R.M.F.-A.; Software, J.N.-M.; Supervision, J.N.-M., R.M.F.-A. and J.C.R.-M.; Validation, J.N.-M., R.M.F.-A. and J.C.R.-M. All authors 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.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Table A1. Expressions for $\Gamma_{\mathbf{x}}^{\nu}(\alpha, t, s)$ and $\Gamma_{\mathbf{xx}^k}^{\nu}(\alpha, t, s)$, $\nu = 1, 2, 3$, in function of the correlation functions of the components of $\mathbf{x}_r(t)$ (3).

$\Gamma_{\mathbf{x}}^3(\alpha, t, s) = \Gamma_{\mathbf{a}}(t, s) + \Gamma_{\mathbf{b}}(t, s) + \Gamma_{\mathbf{c}}(t, s) + \Gamma_{\mathbf{d}}(t, s)$ $+i\left(\frac{1}{\alpha}\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) + \frac{1}{\alpha}\Gamma_{\mathbf{cd}}(t, s) + \Gamma_{\mathbf{dc}}(t, s)\right)$ $+j(\Gamma_{\mathbf{ac}}(t, s) + \Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) + \Gamma_{\mathbf{db}}(t, s))$ $+k\left(\frac{1}{\alpha}\Gamma_{\mathbf{ad}}(t, s) + \Gamma_{\mathbf{bc}}(t, s) + \frac{1}{\alpha}\Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s)\right)$	$\Gamma_{\mathbf{x}}^1(\alpha, t, s) = \Gamma_{\mathbf{xx}^i}^2(\alpha, t, s) =$ $\Gamma_{\mathbf{a}}(t, s) + \alpha\Gamma_{\mathbf{b}}(t, s) + \Gamma_{\mathbf{c}}(t, s) + \alpha\Gamma_{\mathbf{d}}(t, s)$ $+i(\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) + \Gamma_{\mathbf{cd}}(t, s) + \Gamma_{\mathbf{dc}}(t, s))$ $+j(\Gamma_{\mathbf{ac}}(t, s) + \alpha\Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) + \alpha\Gamma_{\mathbf{db}}(t, s))$ $+k(\Gamma_{\mathbf{ad}}(t, s) + \Gamma_{\mathbf{bc}}(t, s) + \Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s))$
$\Gamma_{\mathbf{xx}^i}^3(\alpha, t, s) = \Gamma_{\mathbf{a}}(t, s) + \Gamma_{\mathbf{b}}(t, s) - \Gamma_{\mathbf{c}}(t, s) - \Gamma_{\mathbf{d}}(t, s)$ $+i\left(\frac{1}{\alpha}\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) - \frac{1}{\alpha}\Gamma_{\mathbf{cd}}(t, s) - \Gamma_{\mathbf{dc}}(t, s)\right)$ $+j(-\Gamma_{\mathbf{ac}}(t, s) - \Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) + \Gamma_{\mathbf{db}}(t, s))$ $+k\left(-\frac{1}{\alpha}\Gamma_{\mathbf{ad}}(t, s) - \Gamma_{\mathbf{bc}}(t, s) + \frac{1}{\alpha}\Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s)\right)$	$\Gamma_{\mathbf{xx}^i}^1(\alpha, t, s) = \Gamma_{\mathbf{xx}^k}^2(\alpha, t, s) =$ $\Gamma_{\mathbf{a}}(t, s) + \alpha\Gamma_{\mathbf{b}}(t, s) - \Gamma_{\mathbf{c}}(t, s) - \alpha\Gamma_{\mathbf{d}}(t, s)$ $+i(\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) - \Gamma_{\mathbf{cd}}(t, s) - \Gamma_{\mathbf{dc}}(t, s))$ $+j(-\Gamma_{\mathbf{ac}}(t, s) - \alpha\Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) + \alpha\Gamma_{\mathbf{db}}(t, s))$ $+k(-\Gamma_{\mathbf{ad}}(t, s) - \Gamma_{\mathbf{bc}}(t, s) + \Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s))$
$\Gamma_{\mathbf{xx}^j}^3(\alpha, t, s) = \Gamma_{\mathbf{a}}(t, s) - \Gamma_{\mathbf{b}}(t, s) + \Gamma_{\mathbf{c}}(t, s) - \Gamma_{\mathbf{d}}(t, s)$ $+i\left(-\frac{1}{\alpha}\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) - \frac{1}{\alpha}\Gamma_{\mathbf{cd}}(t, s) + \Gamma_{\mathbf{dc}}(t, s)\right)$ $+j(\Gamma_{\mathbf{ac}}(t, s) - \Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) - \Gamma_{\mathbf{db}}(t, s))$ $+k\left(-\frac{1}{\alpha}\Gamma_{\mathbf{ad}}(t, s) + \Gamma_{\mathbf{bc}}(t, s) - \frac{1}{\alpha}\Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s)\right)$	$\Gamma_{\mathbf{xx}^j}^1(\alpha, t, s) = \Gamma_{\mathbf{x}}^2(\alpha, t, s) =$ $\Gamma_{\mathbf{a}}(t, s) - \alpha\Gamma_{\mathbf{b}}(t, s) + \Gamma_{\mathbf{c}}(t, s) - \alpha\Gamma_{\mathbf{d}}(t, s)$ $+i(-\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) - \Gamma_{\mathbf{cd}}(t, s) + \Gamma_{\mathbf{dc}}(t, s))$ $+j(\Gamma_{\mathbf{ac}}(t, s) - \alpha\Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) - \alpha\Gamma_{\mathbf{db}}(t, s))$ $+k(-\Gamma_{\mathbf{ad}}(t, s) + \Gamma_{\mathbf{bc}}(t, s) - \Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s))$
$\Gamma_{\mathbf{xx}^k}^3(\alpha, t, s) = \Gamma_{\mathbf{a}}(t, s) - \Gamma_{\mathbf{b}}(t, s) - \Gamma_{\mathbf{c}}(t, s) + \Gamma_{\mathbf{d}}(t, s)$ $+i\left(-\frac{1}{\alpha}\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) + \frac{1}{\alpha}\Gamma_{\mathbf{cd}}(t, s) - \Gamma_{\mathbf{dc}}(t, s)\right)$ $+j(-\Gamma_{\mathbf{ac}}(t, s) + \Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) - \Gamma_{\mathbf{db}}(t, s))$ $+k\left(\frac{1}{\alpha}\Gamma_{\mathbf{ad}}(t, s) - \Gamma_{\mathbf{bc}}(t, s) - \frac{1}{\alpha}\Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s)\right)$	$\Gamma_{\mathbf{xx}^k}^1(\alpha, t, s) = \Gamma_{\mathbf{xx}^i}^2(\alpha, t, s) =$ $\Gamma_{\mathbf{a}}(t, s) - \alpha\Gamma_{\mathbf{b}}(t, s) - \Gamma_{\mathbf{c}}(t, s) + \alpha\Gamma_{\mathbf{d}}(t, s)$ $+i(-\Gamma_{\mathbf{ab}}(t, s) + \Gamma_{\mathbf{ba}}(t, s) + \Gamma_{\mathbf{cd}}(t, s) - \Gamma_{\mathbf{dc}}(t, s))$ $+j(-\Gamma_{\mathbf{ac}}(t, s) + \alpha\Gamma_{\mathbf{bd}}(t, s) + \Gamma_{\mathbf{ca}}(t, s) - \alpha\Gamma_{\mathbf{db}}(t, s))$ $+k(\Gamma_{\mathbf{ad}}(t, s) - \Gamma_{\mathbf{bc}}(t, s) - \Gamma_{\mathbf{cb}}(t, s) + \Gamma_{\mathbf{da}}(t, s))$

Appendix B

Proof of Theorem 1. The proof is similar to that of Theorem 1 in [24]. \square

Appendix C

Proof of Theorem 2. Point 1 is obtained from (12) and the fact that the SWL and the real processing are isomorphic. Again, point 2 follows from (12). The proof of point 3 is immediate due to $\mathcal{G}_{C_{\alpha}} \subseteq \mathcal{G}_{\bar{C}_{\alpha}}, \forall \alpha \neq 0$.

Next, point 4 is demonstrated under the H_{α_0} -properness condition. The proof of the E_{α_0} -proper case is similar. From (2) and (6) we obtain

$$\begin{aligned} \Gamma_{y\bar{z}}^3(\alpha_0) &= [1, i, j, k]\Gamma_{y, z_r} \mathcal{T}_p^{H_{\alpha_0}}, & \Gamma_{\bar{z}}^3(\alpha_0) &= \mathcal{T}_p \Gamma_{z_r} \mathcal{T}_p^{H_{\alpha_0}} \\ \Gamma_{y\bar{z}}^1(\alpha_0) &= [1, i, j, k]\Gamma_{y, z_r} \mathcal{T}_p^T, & \Gamma_{\bar{z}}^1(\alpha_0) &= \mathcal{T}_p \Gamma_{z_r} \mathcal{T}_p^T \end{aligned} \tag{A1}$$

with \mathbf{z} given in (15). Then, since $\Gamma_{y\bar{z}}^3(\alpha_0)\Gamma_{\bar{z}}^{3^{-1}}(\alpha_0) = \Gamma_{y\bar{z}}^1(\alpha_0)\Gamma_{\bar{z}}^{1^{-1}}(\alpha_0)$, the proof follows from (A1), Definition 1 and (7). \square

Appendix D

Proof of Theorem 3. We carry out the demonstration for the H_{α_0} -proper case. The E_{α_0} -proper case is analogous. Firstly, since $\mathbf{x}(t)$ is H_{α_0} -proper and applying Theorem 2 we have that $\hat{\mathbf{x}}^{\text{SWL}}(t+1|t) = \hat{\mathbf{x}}_{\alpha_0}^{\text{S}}(t+1|t)$. Now, consider the next product respect to α_0

$$\langle x, y \rangle_1 = E[xy] = \Gamma_{xy}^1(\alpha_0), \quad x, y \in \mathbb{S}$$

and define the distance

$$d_1(x, y) = \|x - y\|_1 \tag{A2}$$

where $\|x\|_1^2 = \mathcal{R}\{\Gamma_x^1(\alpha_0)\}$. Following a similar procedure to the one used with (11), the existence and uniqueness of the projection of $x(t + 1)$ onto the set $\mathcal{G}_{C_{\alpha_0}^t}$ associated to the metric (A2) is guaranteed. Further, this projection coincides with $\hat{x}_{\alpha_0}^S(t + 1|t)$, which is obtained from (11).

Define the innovations as $\tilde{x}_{\alpha_0}^S(i) = x(i) - \hat{x}_{\alpha_0}^S(i|i - 1)$. Thus, the set $\{\tilde{x}_{\alpha_0}^S(1), \dots, \tilde{x}_{\alpha_0}^S(t)\}$ spans $\mathcal{G}_{C_{\alpha_0}^t}$ and, from (4), it follows that

$$\langle \tilde{x}_{\alpha_0}^S(i), \tilde{x}_{\alpha_0}^S(j) \rangle_1 = 0, \quad i \neq j$$

Then, (16) holds. Now, denoting $\Pi(k) = \Gamma_{\tilde{x}_{\alpha_0}^S}^1(\alpha_0, k + 1, k + 1)$, we obtain

$$\Gamma_{\tilde{x}_{\alpha_0}^S}^1(\alpha_0, t + 1, k + 1) = \Gamma_{\tilde{x}_{\alpha_0}^S \tilde{x}_{\alpha_0}^S}^1(\alpha_0, t + 1, k + 1) = \Lambda(t, t - k)\Pi(k), \quad 0 \leq k < t \quad (A3)$$

By using (16), Equation (A3) can be rewritten as

$$\Lambda(t, t - k)\Pi(k) = \Gamma_x^1(\alpha_0, t + 1, k + 1) - \sum_{i=0}^{k-1} \Gamma_{\tilde{x}_{\alpha_0}^S}^1(\alpha_0, t + 1, i + 1)\Lambda^T(k, k - i)$$

Taking into account that $\Gamma_x^1(\alpha_0, t, s)$ is nonsingular and hence, $\Pi(k)$ is also nonsingular, we obtain (17).

Now, since $\Pi(t) = \Gamma_x^1(\alpha_0, t + 1, t + 1) - \Gamma_{\tilde{x}_{\alpha_0}^S}^1(\alpha_0, t + 1, t + 1)$ and from (16), we obtain (18). Finally, taking into account that $\tilde{x}^S(t + 1)$ is H_{α_0} -proper and

$$\|\tilde{x}_j^S(t + 1)\|_{\alpha_0}^2 = \frac{2\alpha_0 + 2}{4\alpha_0} \mathcal{R}\{\Gamma_{\tilde{x}_j^S}^1(\alpha_0, t + 1, t + 1)\}$$

then (19) follows. \square

Appendix E

Proof of Proposition 2. From (20), we have

$$\mathfrak{F}(B)\bar{x}(t) = \mathfrak{G}(B)\bar{v}(t) \quad (A4)$$

where B denotes the backward shift operator, $\mathfrak{G}(z) = \mathbf{I}_{4p} + \sum_{i=1}^q \bar{\mathbf{G}}(i)z^i$ and $\bar{\mathbf{G}}(i) = \text{diag}(\mathbf{G}(i), \mathbf{G}^1(i), \mathbf{G}^j(i), \mathbf{G}^k(i))$, $i = 1, \dots, q$. Thus, post-multiplying (A4) by $\bar{x}^{H_{3-2\nu}}(t - h)$, taking expectations and taking into account the H_{α_0} -properness (respectively, E_{α_0} -properness), (22) holds. The recursive expression for the matrices $\mathbf{L}(t)$ can be derived as follows. Taking into account that $\det(\mathfrak{F}(z))_{\alpha_0} \neq 0 \forall z \in \mathbb{C}$ such that $|z| \leq 1$, we have that (Theorem 11.3.1 in [3]).

$$\bar{x}(t) = \mathfrak{F}^{-1}(B)\mathfrak{G}(B)\bar{v}(t) = \sum_{j=0}^{\infty} \bar{\mathbf{L}}(j)\bar{v}(t - j)$$

By comparing coefficients of z^j in the power series $\sum_{j=0}^{\infty} \bar{\mathbf{L}}(j)z^j$ after multiplying through by $\mathfrak{F}(z)$ and from the block diagonality of $\bar{\mathbf{F}}(i)$ and $\bar{\mathbf{G}}(i)$, the result follows. \square

Appendix F

Proof of Theorem 4. The proof is given for the H_{α_0} -proper case and the E_{α_0} -proper case is similarly obtained. Since $x(t)$ is H_{α_0} -proper and from Theorem 2 we have that $\hat{x}_{\alpha_0}^{SWL}(t + 1|t) = \hat{x}_{\alpha_0}^S(t + 1|t)$.

Now, consider the transformed signal

$$\begin{aligned} \mathbf{w}(t) &= \mathbf{x}(t), & 1 \leq t \leq m \\ \mathbf{w}(t) &= \mathbf{x}(t) - \sum_{i=1}^p \mathbf{F}(i)\mathbf{x}(t-i), & t > m \end{aligned}$$

It is trivial that $\mathcal{G}_{\mathcal{C}_{\alpha_0}^t} = \mathcal{G}_{\mathcal{D}_{\alpha_0}^t}$ with $\mathcal{C}_{\alpha_0}^t = \{\mathbf{x}(1), \dots, \mathbf{x}(t)\}$ and $\mathcal{D}_{\alpha_0}^t = \{\mathbf{w}(1), \dots, \mathbf{w}(t)\}$. Then, defining $\hat{\mathbf{w}}_{\alpha_0}(1|0) = \mathbf{0}_{p \times 1}$ and denoting by $\hat{\mathbf{w}}_{\alpha_0}(t+1|t)$ the \mathbb{S}_{α_0} one-stage predictor of $\mathbf{w}(t+1)$ respect to $\mathcal{D}_{\alpha_0}^t$ under (A2), we have

$$\begin{aligned} \hat{\mathbf{x}}_{\alpha_0}^S(t+1|t) &= \hat{\mathbf{w}}_{\alpha_0}(t+1|t), & 0 \leq t < m \\ \hat{\mathbf{x}}_{\alpha_0}^S(t+1|t) &= \sum_{i=1}^p \mathbf{F}(i)\mathbf{x}(t-i) + \hat{\mathbf{w}}_{\alpha_0}(t+1|t), & t \geq m \end{aligned} \tag{A5}$$

Thus, the problem of computing $\hat{\mathbf{x}}_{\alpha_0}^S(t+1|t)$ is equivalent to calculating $\hat{\mathbf{w}}_{\alpha_0}(t+1|t)$. It is not difficult to prove that $\Gamma_{\mathbf{w}}^1(\alpha_0, t, s) = \mathbf{Y}^1(t, s)$ given in (24). Since $\mathbf{Y}^1(t, s) = \mathbf{0}_{p \times p}$ if $t > m$ and $|t-s| > q$, then, from Theorem 3, we obtain $\Lambda(t, i) = \mathbf{0}_{p \times p}$ for $t \geq m$ and $i > q$, and hence,

$$\hat{\mathbf{w}}_{\alpha_0}(t+1|t) = \begin{cases} \sum_{i=1}^t \Lambda(t, i)(\mathbf{w}(t+1-i) - \hat{\mathbf{w}}_{\alpha_0}(t+1-i|t-i)), & 1 \leq t < m \\ \sum_{i=1}^q \Lambda(t, i)(\mathbf{w}(t+1-i) - \hat{\mathbf{w}}_{\alpha_0}(t+1-i|t-i)), & t \geq m \end{cases} \tag{A6}$$

with $\hat{\mathbf{w}}_{\alpha_0}(1|0) = \mathbf{0}_{p \times 1}$ and the matrices $\Lambda(t, i)$ calculated from (17) and (18) for $\mathbf{Y}^1(t, s)$.

Finally, from (A5), (A6) and the following equality

$$\mathbf{x}(t+1) - \hat{\mathbf{x}}_{\alpha_0}^S(t+1|t) = \mathbf{w}(t+1) - \hat{\mathbf{w}}_{\alpha_0}(t+1|t), \quad t \geq 0$$

(23) follows. \square

References

1. Pollock, D.S.G. *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*; Academic Press: San Diego, CA, USA, 1999.
2. Manolakis, D.; Bosowski, N.; Ingle, V.K. Count time-series analysis: A signal processing perspective. *IEEE Signal Process. Mag.* **2019**, *36*, 64–81. [CrossRef]
3. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*, 2nd ed.; Springer: New York, NY, USA, 1991.
4. Kuipers, J.B. *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*; Princeton University Press: Princeton, NJ, USA, 2002.
5. Hanson, A.J. *Visualizing Quaternions*; The Morgan Kaufmann Series in Interactive 3D Technology; Elsevier: Amsterdam, The Netherlands, 2005.
6. Catoni, F.; Boccaletti, D.; Cannata, R.; Catoni, V.; Nichelatti, E.; Zampetti, P. *The Mathematics of Minkowski Space-Time: With an Introduction to Commutative Hypercomplex Numbers*; Birkhäuser Verlag: Basel, Switzerland, 2008.
7. Ginzberg, P.; Walden, A.T. Quaternion VAR modelling and estimation. *IEEE Trans. Signal Process.* **2013**, *61*, 154–158. [CrossRef]
8. Tobar, F.; Mandic, D.P. Quaternion reproducing kernel Hilbert spaces: Existence and uniqueness conditions. *IEEE Trans. Inform. Theory* **2014**, *60*, 5736–5749. [CrossRef]
9. Schütte, H.D.; Wenzel, J. Hypercomplex numbers in digital signal processing. In Proceedings of the International Symposium on Circuits and Systems, New Orleans, LA, USA, 1–3 May 1990; pp. 1557–1560. Available online: <https://ieeexplore.ieee.org/document/112431> (accessed on 15 February 2022)
10. Pei, S.H.; Chang, J.H.; Ding, J.J. Commutative reduced biquaternions and their Fourier transform for signal and image processing applications. *IEEE Trans. Signal Process.* **2004**, *52*, 2012–2031. [CrossRef]
11. Labunets, V. Clifford algebras as unified language for image processing and pattern recognition. In *Computational Noncommutative Algebra and Applications*; NATO Science Series II: Mathematics, Physics and Chemistry; Byrnes, J., Ed.; Springer: Dordrecht, The Netherlands, 2004; Volume 136, pp. 197–225.
12. Alfsmann, D. On families of 2N-dimensional hypercomplex algebras suitable for digital signal processing. In Proceedings of the 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, 4–8 September 2006; pp. 1–4.

13. Alfsmann, D.; Göckler, H.G.; Sangwine, S.J.; Ell, T.A. Hypercomplex algebras in digital signal processing: Benefits and drawbacks. In Proceedings of the 15th European Signal Processing Conference, Poznan, Poland, 3–7 September 2007; pp. 1322–1326.
14. Pei, S.H.; Chang, J.H.; Ding, J.J.; Chen, M.Y. Eigenvalues and singular value decompositions of reduced biquaternion matrices. *IEEE Trans. Circuits Syst. I* **2008**, *525*, 2673–2685.
15. Guo, L.; Zhu, M.; Ge, X. Reduced biquaternion canonical transform, convolution and correlation. *Signal Process.* **2011**, *91*, 2147–2153. [[CrossRef](#)]
16. Hahn, S.L.; Snopce, K.M. *Complex and Hypercomplex Analytic Signals: Theory and Applications*; Artech House: Norwood, MA, USA, 2016.
17. Kamal, A.T.; El-Melegy, M.T. Color image processing using reduced biquaternions with applications to face recognition in a PCA framework. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 3039–3046.
18. Esposito, A.; Faundez-Zanuy, M.; Morabito, F.C.; Pasero, E. *Neural Advances in Processing Nonlinear Dynamic Signals*; Springer: Cham, Switzerland, 2018.
19. Kobayashi, M. Twin-multistate commutative quaternion Hopfield neural networks. *Neurocomputing* **2018**, *320*, 150–156. [[CrossRef](#)]
20. Zanetti de Castro, F.; Eduardo Valle, M. A broad class of discrete-time hypercomplex-valued Hopfield neural networks. *Neural Netw.* **2020**, *122*, 54–67. [[CrossRef](#)]
21. Takahashi, K. Comparison of high-dimensional neural networks using hypercomplex numbers in a robot manipulator control. *Artif. Life Robot.* **2021**, *26*, 367–377. [[CrossRef](#)]
22. Ortolani, F.; Scarpiniti, M.; Comminiello, D.; Uncini, A. On the influence of microphone array geometry on the behavior of hypercomplex adaptive filters. In Proceedings of the 5th IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS), Kiev, Ukraine, 29–31 August 2017; pp. 37–42.
23. Ortolani, F.; Scarpiniti, M.; Comminiello, D.; Uncini, A. On 4-dimensional hypercomplex algebras in adaptive signal processing. In *Neural Advances in Processing Nonlinear Dynamic Signals*; Springer: Berlin, Germany, 2017; pp. 131–140.
24. Navarro-Moreno, J.; Fernández-Alcalá, R.M.; López-Jiménez, J.D.; Ruiz-Molina, J.C. Tessarine signal processing under the \mathbb{T} -properness condition. *J. Frankl. Inst.* **2020**, *357*, 10100–10126. [[CrossRef](#)]
25. Navarro-Moreno, J.; Ruiz-Molina, J.C. Wide-sense Markov signals on the tessarine domain. A study under properness conditions. *Signal Process.* **2021**, *183C*, 108022. [[CrossRef](#)]
26. Took, C.C.; Mandic, D.P. Augmented second-order statistics of quaternion random signals. *Signal Process.* **2011**, *91*, 214–224. [[CrossRef](#)]
27. Navarro-Moreno, J.; Fernández-Alcalá, R.M.; Ruiz-Molina, J.C. Semi-widely simulation and estimation of continuous-time \mathbb{C}^n -proper quaternion random signals. *IEEE Trans. Signal Process.* **2015**, *63*, 4999–5012. [[CrossRef](#)]
28. Navarro-Moreno, J.; Ruiz-Molina, J.C. Semi-widely linear estimation of \mathbb{C}^n -proper quaternion random signal vectors under Gaussian and stationary conditions. *Signal Process.* **2016**, *119*, 56–66. [[CrossRef](#)]
29. Nitta, T.; Kobayashi, M.; Mandic, D.P. Hypercomplex widely linear estimation through the lens of underpinning geometry. *IEEE Trans. Signal Process.* **2019**, *67*, 3985–3994. [[CrossRef](#)]
30. Grassucci, E.; Comminiello, D.; Uncini, A. An information-theoretic perspective on proper quaternion variational autoencoders. *Entropy* **2021**, *23*, 856. [[CrossRef](#)] [[PubMed](#)]