Hypercomplex Widely Linear Estimation Through the Lens of Underpinning Geometry

Tohru Nitta, Masaki Kobayashi, and Danilo P. Mandic, *Fellow, IEEE*

***Abstract*—We provide a rigorous account of the equivalence between the complex-valued widely linear estimation method and the quaternion involution widely linear estimation method with their vector-valued real linear estimation counterparts. This is achieved by an account of degrees of freedom and by providing matrix mappings between a complex variable and an isomorphic bivariate real vector, and a quaternion variable versus a quadri-variate real vector. Furthermore, we show that the parameters in the complex-valued linear estimation method, the complex-valued widely linear estimation method, the quaternion linear estimation method, the quaternion semi-widely linear estimation method, and the quaternion involution widely linear estimation method include distinct geometric structures imposed on complex numbers and quaternions, respectively, whereas the real-valued linear estimation methods do not exhibit any structure. This key difference explains, both in theoretical and practical terms, the advantage of estimation in division algebras (complex, quaternion) over their multivariate real vector counterparts. In addition, we discuss the computational complexities of the estimators of the hypercomplex widely linear estimation methods.**

***Index Terms*—Widely linear estimation, augmented statistics, complex number, quaternion.**

I. Introduction

In RECENT years, there has been an increasing interest in widely linear (WL) estimation methods in the complex or quaternion domains [1], with successful applications in areas such as communication [2], [3], adaptive filters [4], and independent component analysis [5], [6]. The underpinning idea behind the WL estimation methods is to cater for full second-order noncircular statistics in data, which arises through power imbalance or correlation in data channels, through the use of the signal variables and their counterparts [7], [8]. The resulting WL estimation methods are then fully equipped to deal with both second-order circular (proper) and second-order noncircular (improper) signals [9], [10].

Manuscript received August 30, 2018; revised March 26, 2019; accepted May 9, 2019. Date of publication June 11, 2019; date of current version July 3, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kostas D. Berberidis. This work was supported by JSPS KAKENHI Grant Number JP16K00347. *(Corresponding author: Tohru Nitta.)*

T. Nitta is with the Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki 305-8568, Japan (e-mail: tohru-nitta@aist.go.jp).

M. Kobayashi is with the Mathematical Science Center, Yamanashi University, Yamanashi 400-8510, Japan (e-mail: k-masaki@yamanashi.ac.jp).

D. P. Mandic is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: d.mandic@imperial.ac.uk).

Digital Object Identifier 10.1109/TSP.2019.2922151

To this end, Picinbono and Chevalier proved that the estimation error obtained using complex-valued WL estimation is smaller than that obtained using the usual "strictly linear" complex-valued linear estimation method [7]. Nitta proved that the estimation error obtained using the quaternion semi-WL estimation method is smaller than that obtained using the usual quaternion strictly linear estimation method [11], [12]. Furthermore, Xia *et al.* proved the superiority of the quaternion involution WL estimation method over the quaternion semi-WL estimation method with respect to estimation error [13]. Further, Nitta formulated a Clifford-valued WL estimation framework, which is a generalization of the complex-valued and quaternion-valued WL models [11]. However, to date the WL estimation method with more than four dimensions (quaternions) has not been analyzed, a subject of this work.

The equivalence between a WL complex adaptive filter and a dual channel adaptive filter was addressed in [14], in the context of adaptive filtering, and falls within the general framework of this work. Also, in order to reduce the computational complexity, the complex dual channel (CDC) estimation has been proposed which is equivalent to widely linear estimation (Eqs. (12) and (13) in [15]). Furthermore, the CDC estimation has been extended to the quaternion domain [16]. Mizoguchi et al. presented a systematic algebraic translation of the Cayley-Dickson hypercomplex-valued linear systems into a real vector-valued linear model, and pointed out that the complex widely linear model can be treated as the framework (Remark 2 in [17]).

First, we show that the complex-valued WL estimation method is equivalent to a two-dimensional real-valued linear estimation method if we regard a complex number as an arrangement of two real numbers. We further demonstrate that the complex-valued WL estimation method can represent complex-valued data naturally, while it can reduce estimation error as small as the one of the real-valued linear estimation method. This is a key advantage of the complex-valued WL estimation method. Furthermore, we show that the parameters in the complex-valued linear estimation method and the complex-valued WL estimation method both include a geometric structure on complex numbers, whereas the real-valued linear estimation method does not impose any structure. Next, we show that the quaternion involution WL estimation method is equivalent to a four-dimensional real-valued linear estimation method if we regard a quaternion as an arrangement of four real numbers. We demonstrate that the quaternion involution WL estimation method can represent quaternion data naturally, while it can reduce estimation error as small as the one of the real-valued linear estimation method. Furthermore, we show that the parameters in the quaternion linear estimation method, the quaternion semi-WL estimation method and the quaternion involution WL estimation method all include distinct geometric structures imposed on quaternions whereas the real-valued linear estimation method does not have any structure. In addition, we discuss the computational complexities of the estimators of the hypercomplex WL estimation methods. As a result, it is learned that the introduction of the idea of the widely linearity into estimation methods increases the computational complexities of the estimators. This is considered to be one of the disadvantages. However, it seems that this problem can be easily solved by using appropriate parallel computing techniques.

The rest of this paper is organized as follows. Section II analyzes the complex-valued and quaternion WL estimation methods. Section III discusses the research results obtained in the previous section. Finally, Section IV concludes this paper.

II. Analysis of Widely Linear Estimation Methods

In this section, we analyze the several existing WL estimation methods.

A. Complex-Valued Estimation Methods

For completeness, the estimation methods in the complex domain are first addressed.

1) Complex-Valued Linear Mean Square Estimation: We first analyze a framework called the *complex-valued linear mean square estimation*. A true value $y \in \mathbb{C}$ is estimated from an observed value $\boldsymbol{x} \in \mathbb{C}^N$ where y is a complex-valued random variable and $\boldsymbol{x}$ is a complex-valued random vector. Assume an estimated value $\hat{y}_L$, expressed as

$$\hat{y}_L = \boldsymbol{h}^H \boldsymbol{x} \tag{1}$$

where $\boldsymbol{h} \in \mathbb{C}^N$, and H represents complex conjugate transposition. The objective is to find a complex-valued parameter $\boldsymbol{h} \in \mathbb{C}^N$ that minimizes mean square error $E|y - \hat{y}_L|^2$ where $|z| \stackrel{\text{def}}{=} \sqrt{x^2 + y^2}$ for a complex number $z = x + \mathrm{i}y \in \mathbb{C}$ (mean square error is abbreviated to MSE hereafter).

We shall denote by z^R, z^I the real part and the imaginary part of a complex number $z \in \mathbb{C}$, respectively. Here, for any $\boldsymbol{h} = (h_1, \ldots, h_N)^T \in \mathbb{C}^N$ and $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{C}^N$, Equation (1) can be written as

$$\begin{aligned} \hat{y}_L &= \boldsymbol{h}^H \boldsymbol{x} = \sum_{k=1}^{N} h_k^* x_k \\ &= \sum_{k=1}^{N} \left\{ (h_k^R x_k^R + h_k^I x_k^I) + \mathrm{i}(h_k^R x_k^I - h_k^I x_k^R) \right\} \end{aligned} \tag{2}$$

where the operator $(\cdot)^*$ denotes the complex conjugate, i.e., $z^* = x - \mathrm{i}y$ for a complex number $z = x + \mathrm{i}y \in \mathbb{C}$, and T represents the real or complex transposition. Then, it follows from Eq. (2) that

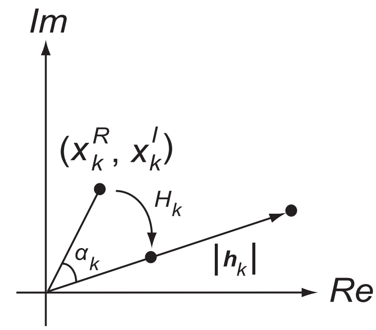

Fig. 1. Geometric interpretation of the operation of an estimator $\hat{y}_L$ within the complex-valued linear mean square estimation framework (Eq. (4)).

$$\begin{aligned} \begin{bmatrix} \hat{y}_L^R \\ \hat{y}_L^I \end{bmatrix} &= \sum_{k=1}^{N} \begin{bmatrix} h_k^R & h_k^I \\ -h_k^I & h_k^R \end{bmatrix} \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix} &(3)\\ &= \sum_{k=1}^{N} |h_k| H_k \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix} &(4) \end{aligned}$$

where for any $1 \leq k \leq N$, $|h_k| \in \mathbb{R}^+ \stackrel{\text{def}}{=} \{u \in \mathbb{R} | u > 0\}$ and

$$H_k \stackrel{\text{def}}{=} \begin{bmatrix} \cos(-\alpha_k) & -\sin(-\alpha_k) \\ \sin(-\alpha_k) & \cos(-\alpha_k) \end{bmatrix} \in SO_2(\mathbb{R}) \tag{5}$$

where $\alpha_k = \tan^{-1}(h_k^I / h_k^R)$. Here, $\mathbb{R}^+$ is the multiplicative group and $SO_2(\mathbb{R})$ is the two-dimensional rotation group. Thus, we obtain the following proposition.

Proposition 1: An estimator $\hat{y}_L$ is obtained by applying the elements of the group $\mathbb{R}^+ \times SO_2(\mathbb{R})$ ($\{|h_k| H_k\}_{k=1}^N$ in Eq. (4)) to each element of an observed value $\boldsymbol{x}$ in the complex-valued linear mean square estimation (Fig. 1).

2) Complex-Valued Widely Linear Mean Square Estimation: Next, we analyze a framework called the *complex-valued widely linear mean square estimation*. A true value $y \in \mathbb{C}$ is estimated from an observed value $\boldsymbol{x} \in \mathbb{C}^N$ where y is a complex-valued random variable and $\boldsymbol{x}$ a complex-valued random vector. Assume an estimated value, $\hat{y}_{WL}$, expressed as

$$\hat{y}_{WL} = \boldsymbol{h}^H \boldsymbol{x} + \boldsymbol{g}^H \boldsymbol{x}^* \tag{6}$$

where $\boldsymbol{g}, \boldsymbol{h} \in \mathbb{C}^N$. The objective is to find the complex-valued parameter vectors $\boldsymbol{g}, \boldsymbol{h} \in \mathbb{C}^N$ that minimize the MSE $E|y - \hat{y}_{WL}|^2$. It has been proved in [7] that the estimation error obtained using the complex-valued WL mean square estimation is smaller than that obtained using the complex-valued linear mean square estimation analyzed in Section II-A1, that is, $E|y - \hat{y}_L|^2 \geq E|y - \hat{y}_{WL}|^2$, where the equality holds only in exceptional cases. The physical interpretation of this framework of estimation is to add another degree of freedom through a complex conjugate term $\boldsymbol{g}^H \boldsymbol{x}^*$ as an explanatory variable (Equation (6)).

Here, by replacing $[x_k^R \; x_k^I]^T$ with $K[x_k^R \; x_k^I]^T$ in Eq. (3), we obtain

$$\begin{bmatrix} Re[\boldsymbol{g}^H \boldsymbol{x}^*] \\ Im[\boldsymbol{g}^H \boldsymbol{x}^*] \end{bmatrix} = \sum_{k=1}^{N} |g_k| G_k \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix} \tag{7}$$

for any $\boldsymbol{g} = (g_1, \ldots, g_N)^T \in \mathbb{C}^N$ and $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{C}^N$ where the operators $Re[z]$ and $Im[z]$ denote respectively

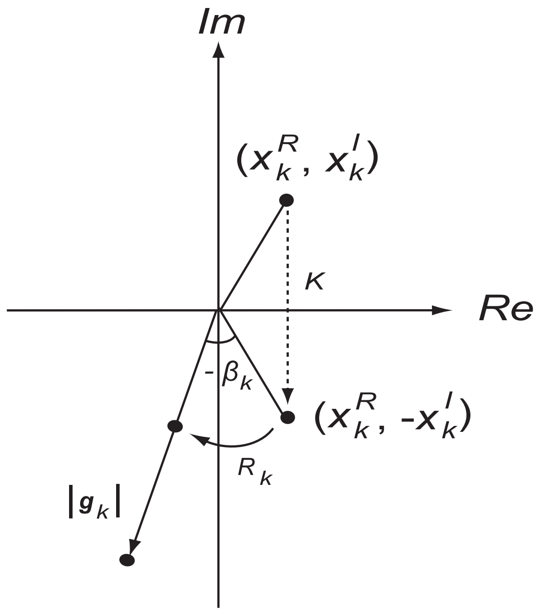

Fig. 2. Geometric interpretation of the operation of an estimator $\hat{y}_{WL}$ within the complex-valued WL mean square estimation framework (see Eq. (7)) .

the real part and the imaginary part of a complex number $z \in \mathbb{C}$, and for any $1 \le k \le N$, $|g_k| \in \mathbb{R}^+$ (the multiplicative group),

$$G_k \stackrel{\text{def}}{=} R_k K \in O_2(\mathbb{R}), \notin SO_2(\mathbb{R}), \tag{8}$$

$$R_k \stackrel{\text{def}}{=} \begin{bmatrix} \cos(-\beta_k) & -\sin(-\beta_k) \\ \sin(-\beta_k) & \cos(-\beta_k) \end{bmatrix} \in SO_2(\mathbb{R}), \tag{9}$$

$$K \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in O_2(\mathbb{R}) \tag{10}$$

where $\beta_k = \tan^{-1}(g_k^I/g_k^R)$. Here, K is a reflection over the real axis in $\mathbb{R}^2$ and $O_2(\mathbb{R})$ is the two-dimensional orthogonal group. So, from Eqs. (4), (6) and (7), we have

$$\begin{bmatrix} \hat{y}_{WL}^R \\ \hat{y}_{WL}^I \end{bmatrix} = \sum_{k=1}^{N} \Big\{ |h_k| H_k + |g_k| G_k \Big\} \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix}. \tag{11}$$

This leads to the following proposition.

Proposition 2: An estimator $\hat{y}_{WL}$ (Eq. (6)) is obtained through linear combinations of an element of the group $\mathbb{R}^+ \times SO_2(\mathbb{R})$ ($|h_k|H_k$ in Eq. (11)) and an element of the group $\mathbb{R}^+ \times (O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$ ($|g_k|G_k$ in Eq. (11)) which are applied to each element of an observed value $\boldsymbol{x}$ in the complex-valued WL mean square estimation (Fig. 2).

Furthermore, it follows from Eqs. (3) and (7) that

$$\begin{bmatrix} \hat{y}_{WL}^R \\ \hat{y}_{WL}^I \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} h_k^R + g_k^R & h_k^I - g_k^I \\ -h_k^I - g_k^I & h_k^R - g_k^R \end{bmatrix} \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix}. \tag{12}$$

Here, for any $p, q \in \mathbb{R}$, let $a \stackrel{\text{def}}{=} 1/2(p+q)$ and $c \stackrel{\text{def}}{=} 1/2(p-q)$. Then, $a + c = p$ and $a - c = q$. So, for any $p, q \in \mathbb{R}$, there exist some $a, c \in \mathbb{R}$ such that $p = a + c$ and $q = a - c$. Hence $h_k^R + g_k^R$ and $h_k^R - g_k^R$ are any real numbers, and the same is said about $h_k^I - g_k^I$ and $-h_k^I - g_k^I$ in Eq. (12). In other words, no interrelation exists among $h_k^R + g_k^R$, $h_k^R - g_k^R$, $h_k^I - g_k^I$ and $-h_k^I - g_k^I$ in Eq. (12), that is,

$$W_k \stackrel{\text{def}}{=} \begin{bmatrix} h_k^R + g_k^R & h_k^I - g_k^I \\ -h_k^I - g_k^I & h_k^R - g_k^R \end{bmatrix} \tag{13}$$

in Eq. (12) is just a general two-dimensional square matrix over real numbers. In other words, any two-dimensional square matrix can be uniquely decomposed into the sum of two two-dimensional square matrices as follows:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{2} \begin{bmatrix} a+d & b-c \\ -(b-c) & a+d \end{bmatrix} + \frac{1}{2} \begin{bmatrix} a-d & b+c \\ b+c & -(a-d) \end{bmatrix}, \tag{14}$$

which was proved in Eqs. (36)–(40) of [18] where it was shown that a rotor Hopfield neural network can be uniquely decomposed into a complex-valued Hopfield neural network and a symmetric complex-valued Hopfield neural network using Eq. (14). Thus, we obtain the following proposition.

Proposition 3: If the observed value $\boldsymbol{x}$, the true value y and the estimated value $\hat{y}_{WL}$ are regarded as real, that is, $\boldsymbol{x} \in \mathbb{R}^{2N}$, $y \in \mathbb{R}^2$ and $\hat{y}_{WL} \in \mathbb{R}^2$, then the complex-valued WL mean square estimation is equivalent to two-dimensional real-valued linear MSE.

B. Quaternion Estimation Methods

Next, we analyze the estimation methods in the quaternion domain.

1) Quaternions: Before analyzing the quaternion estimation methods, we shall first briefly review the algebra of quaternions. A quaternion is defined over $\mathbb{R}^4$ and comprises three imaginary units: $\mathsf{i}, \mathsf{j}, \mathsf{k}$, such that

$$\mathsf{ij} = -\mathsf{ji} = \mathsf{k},\ \mathsf{jk} = -\mathsf{kj} = \mathsf{i},\ \mathsf{ki} = -\mathsf{ik} = \mathsf{j},$$
$$\mathsf{i}^2 = \mathsf{j}^2 = \mathsf{k}^2 = \mathsf{ijk} = -1. \tag{15}$$

Every quaternion, q, can be written explicitly as

$$q = a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k} \in \mathbb{H}, \quad a, b, c, d \in \mathbb{R} \tag{16}$$

where $\mathbb{H}$ denotes the set of quaternions. Observe that the commutativity does not hold, that is, $pq \neq qp$ for any $p, q \in \mathbb{H}$. A quaternion conjugate is defined as

$$q^* = a - b\mathsf{i} - c\mathsf{j} - d\mathsf{k}, \tag{17}$$

and the norm by

$$|q| = \sqrt{qq^*}. \tag{18}$$

The quaternion involution is defined as

$$q^{\mathsf{i}} = -\mathsf{i}q\mathsf{i} = a + b\mathsf{i} - c\mathsf{j} - d\mathsf{k}, \tag{19}$$

$$q^{\mathsf{j}} = -\mathsf{j}q\mathsf{j} = a - b\mathsf{i} + c\mathsf{j} - d\mathsf{k}, \tag{20}$$

$$q^{\mathsf{k}} = -\mathsf{k}q\mathsf{k} = a - b\mathsf{i} - c\mathsf{j} + d\mathsf{k}. \tag{21}$$

Note that the quaternion involution (Eqs. (19)–(21) is a special case of the nonstandard involution, which was first given in [19] (Definition 2.4.5) and further developed and discussed in [20], [21]. For a quaternion $q = a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k} \in \mathbb{H}$, its quaternion conjugates are

$$q^{\mathsf{i}^*} = a - b\mathsf{i} + c\mathsf{j} + d\mathsf{k}, \tag{22}$$

$$q^{\mathsf{j}^*} = a + b\mathsf{i} - c\mathsf{j} + d\mathsf{k}, \tag{23}$$

$$q^{\mathsf{k}^*} = a + b\mathsf{i} + c\mathsf{j} - d\mathsf{k}. \tag{24}$$

From now on, we shall denote by q^R, q^I, q^J, q^K the real part a and the three imaginary parts b, c, d of a quaternion $q = a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k} \in \mathbb{H}$, respectively.

2) Quaternion Linear Mean Square Estimation: We first analyze a framework called the *quaternion linear mean square estimation*. A true value $y \in \mathbb{H}$ is estimated from an observed value $\boldsymbol{x} \in \mathbb{H}^N$ where y is a quaternion random variable and $\boldsymbol{x}$ is a quaternion random vector. Assume that an estimated value $\hat{y}_L$ expressed in a strictly linear form as

$$\hat{y}_L = \boldsymbol{h}^H \boldsymbol{x} \tag{25}$$

where $\boldsymbol{h} \in \mathbb{H}^N$, and H represents quaternion conjugate transposition. The objective is to find a quaternion-valued parameter $\boldsymbol{h} \in \mathbb{H}^N$ that minimizes the MSE $E|y - \hat{y}_L|^2$.

Proposition 4: An estimator $\hat{y}_L$ (Eq. (25)) is obtained through the elements of the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ which are applied to each element of an observed value $\boldsymbol{x}$ in the quaternion linear mean square estimation, where $\mathbb{R}^+$ is the multiplicative group and $SO_4(\mathbb{R})$ is the four-dimensional rotation group.

Proof: For any $\boldsymbol{h} = (h_1, \ldots, h_N)^T \in \mathbb{H}^N$ and $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{H}^N$, Equation (25) can be written as

$$\begin{aligned}\hat{y}_L &= \sum_{k=1}^{N} h_k^* x_k \\ &= \sum_{k=1}^{N} \Big\{ (h_k^R x_k^R + h_k^I x_k^I + h_k^J x_k^J + h_k^K x_k^K) \\ &\quad + \mathsf{i}(h_k^R x_k^I - h_k^I x_k^R - h_k^J x_k^K + h_k^K x_k^J) \\ &\quad + \mathsf{j}(h_k^R x_k^J + h_k^I x_k^K - h_k^J x_k^R - h_k^K x_k^I) \\ &\quad + \mathsf{k}(h_k^R x_k^K - h_k^I x_k^J + h_k^J x_k^I - h_k^K x_k^R) \Big\}.\end{aligned} \tag{26}$$

Then, it follows from Eq. (26) that

$$\begin{bmatrix} \hat{y}_L^R \\ \hat{y}_L^I \\ \hat{y}_L^J \\ \hat{y}_L^K \end{bmatrix} = \sum_{k=1}^{N} |h_k| H_k \begin{bmatrix} x_k^R \\ x_k^I \\ x_k^J \\ x_k^K \end{bmatrix} \tag{27}$$

where $|h_k| \in \mathbb{R}^+$ and

$$H_k \overset{\text{def}}{=} \frac{1}{|h_k|} \begin{bmatrix} h_k^R & h_k^I & h_k^J & h_k^K \\ -h_k^I & h_k^R & h_k^K & -h_k^J \\ -h_k^J & -h_k^K & h_k^R & h_k^I \\ -h_k^K & h_k^J & -h_k^I & h_k^R \end{bmatrix} \in SO_4(\mathbb{R}) \tag{28}$$

for any $1 \le k \le N$. Thus, we have found that an estimator $\hat{y}_L$ (Eq. (25)) is obtained by applying the elements of the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ ($\{|h_k| H_k\}_{k=1}^N$ in Eq. (27)) to each element of an observed value $\boldsymbol{x}$ in the quaternion linear MSE. ∎

Here, we define a measure called the *degree of freedom of parameters* of quaternion estimation methods for comparing the three quaternion estimation methods: the quaternion linear mean square estimation, the quaternion semi-WL mean square estimation, and the quaternion involution WL mean square estimation.

Definition 1: By the *degree of freedom of parameters*, we refer to the number of parameters required when the quaternion-valued estimator is expressed by four-dimensional real-valued vectors.

Proposition 5: The degree of freedom of parameters of the quaternion linear mean square estimation is $4N$ where N is the dimension of the observed value $\boldsymbol{x} \in \mathbb{H}^N$.

Proof: It follows from Eq. (28) that the matrix $H_k \in SO_4(\mathbb{R})$ is determined by the four parameters $h_k^R, h_k^I, h_k^J, h_k^K$. Therefore, the degree of freedom of parameters is $4N$. ∎

3) Quaternion Semi-Widely Linear Mean Square Estimation: Next, we analyze a framework called the *quaternion semi-widely linear mean square estimation* [8], [11], [12]. A true value $y \in \mathbb{H}$ is estimated from an observed value $\boldsymbol{x} \in \mathbb{H}^N$ where y is a quaternion random variable and $\boldsymbol{x}$ a quaternion random vector. Assume an estimated value, $\hat{y}_{SWL}$, expressed as

$$\hat{y}_{SWL} = \boldsymbol{h}^H \boldsymbol{x} + \boldsymbol{g}^H \boldsymbol{x}^* \tag{29}$$

where $\boldsymbol{g}, \boldsymbol{h} \in \mathbb{H}^N$. The objective is to find a quaternion-valued parameter $\boldsymbol{g}, \boldsymbol{h} \in \mathbb{H}^N$ that minimizes the MSE $E|y - \hat{y}_{SWL}|^2$.

Proposition 6: A linear estimator $\hat{y}_{SWL}$ (Eq. (29)) is obtained through linear combinations of an element of the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ and an element of the group $\mathbb{R}^+ \times (O_4(\mathbb{R}) \setminus SO_4(\mathbb{R}))$ which are applied to each element of an observed value $\boldsymbol{x}$ in the quaternion semi-WL mean square estimation where $\mathbb{R}^+$ is the multiplicative group, $SO_4(\mathbb{R})$ is the four-dimensional rotation group, and $O_4(\mathbb{R})$ is the four-dimensional orthogonal group.

Proof: For any $\boldsymbol{g} = (g_1, \ldots, g_N)^T \in \mathbb{H}^N$ and $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{H}^N$, the second term of Eq. (29) can be written as

$$\begin{aligned}\boldsymbol{g}^H \boldsymbol{x}^* &= \sum_{k=1}^{N} g_k^* x_k^* \\ &= \sum_{k=1}^{N} \Big\{ (g_k^R x_k^R - g_k^I x_k^I - g_k^J x_k^J - g_k^K x_k^K) \\ &\quad + \mathsf{i}(-g_k^R x_k^I - g_k^I x_k^R + g_k^J x_k^K - g_k^K x_k^J) \\ &\quad + \mathsf{j}(-g_k^R x_k^J - g_k^I x_k^K - g_k^J x_k^R + g_k^K x_k^I) \\ &\quad + \mathsf{k}(-g_k^R x_k^K + g_k^I x_k^J - g_k^J x_k^I - g_k^K x_k^R) \Big\}.\end{aligned} \tag{30}$$

Then, it follows from Eqs. (27) and (30) that

$$\begin{bmatrix} \hat{y}_{SWL}^R \\ \hat{y}_{SWL}^I \\ \hat{y}_{SWL}^J \\ \hat{y}_{SWL}^K \end{bmatrix} = \sum_{k=1}^{N} \Big\{ |h_k| H_k + |g_k| G_k \Big\} \begin{bmatrix} x_k^R \\ x_k^I \\ x_k^J \\ x_k^K \end{bmatrix} \tag{31}$$

where $|g_k| \in \mathbb{R}^+$ and

$$\begin{aligned} G_k &\overset{\text{def}}{=} \frac{1}{|g_k|} \begin{bmatrix} g_k^R & -g_k^I & -g_k^J & -g_k^K \\ -g_k^I & -g_k^R & -g_k^K & g_k^J \\ -g_k^J & g_k^K & -g_k^R & -g_k^I \\ -g_k^K & -g_k^J & g_k^I & -g_k^R \end{bmatrix} \\ &= R_k K^{(\mathsf{r})} \in O_4(\mathbb{R}) \setminus SO_4(\mathbb{R}), \end{aligned} \tag{32}$$

$$R_k \overset{\text{def}}{=} G_k K^{(\mathsf{r})} \in SO_4(\mathbb{R}), \tag{33}$$

$$K^{(\mathsf{r})} \overset{\text{def}}{=} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \in O_4(\mathbb{R}) \tag{34}$$

for any $1 \le k \le N$. Here, $K^{(\mathsf{r})}$ is a reflection over the real axis in $\mathbb{R}^4$. We have thus found that an estimator, $\hat{y}_{SWL}$ in Eq. (29), is obtained through linear combinations of an element of the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ ($|h_k|H_k$ in Eq. (31)) and an element of the group $\mathbb{R}^+ \times (O_4(\mathbb{R}) \setminus SO_4(\mathbb{R}))$ ($|g_k|G_k$ in Eq. (31)) which are applied to each element of an observed value $\boldsymbol{x}$ in the quaternion semi-WL mean square estimation. ■

For rigour, we next investigate the degree of freedom of parameters of quaternion semi-WL mean square estimation method (see Definition 1).

Proposition 7: The degree of freedom of parameters of the quaternion semi-WL mean square estimation is $8N$ where N is the dimension of the observed value $\boldsymbol{x} \in \mathbb{H}^N$.

Proof: The statement of the proposition is obvious because $2N$ quaternions are used. Actually, Eq. (31) can be rewritten as follows.

$$\begin{bmatrix} \hat{y}^R_{SWL} \\ \hat{y}^I_{SWL} \\ \hat{y}^J_{SWL} \\ \hat{y}^K_{SWL} \end{bmatrix} = \sum_{k=1}^{N} \left\{ \begin{bmatrix} \alpha_k & \beta_k & \gamma_k & \delta_k \\ -\beta_k & \alpha_k & \delta_k & -\gamma_k \\ -\gamma_k & -\delta_k & \alpha_k & \beta_k \\ -\delta_k & \gamma_k & -\beta_k & \alpha_k \end{bmatrix} \right.$$

$$\left. + \begin{bmatrix} \varepsilon_k & -\zeta_k & -\eta_k & -\theta_k \\ -\zeta_k & -\varepsilon_k & -\theta_k & \eta_k \\ -\eta_k & \theta_k & -\varepsilon_k & -\zeta_k \\ -\theta_k & -\eta_k & \zeta_k & -\varepsilon_k \end{bmatrix} \right\} \begin{bmatrix} x^R_k \\ x^I_k \\ x^J_k \\ x^K_k \end{bmatrix}$$

$$= \sum_{k=1}^{N} \left\{ \begin{bmatrix} \alpha_k+\varepsilon_k & \beta_k-\zeta_k & \gamma_k-\eta_k & \delta_k-\theta_k \\ -\beta_k-\zeta_k & \alpha_k-\varepsilon_k & \delta_k-\theta_k & -\gamma_k+\eta_k \\ -\gamma_k-\eta_k & -\delta_k+\theta_k & \alpha_k-\varepsilon_k & \beta_k-\zeta_k \\ -\delta_k-\theta_k & \gamma_k-\eta_k & -\beta_k+\zeta_k & \alpha_k-\varepsilon_k \end{bmatrix} \right\} \begin{bmatrix} x^R_k \\ x^I_k \\ x^J_k \\ x^K_k \end{bmatrix}$$

$$= \sum_{k=1}^{N} \left\{ \begin{bmatrix} A_k & -F_k & -H_k & -C_k \\ E_k & B_k & -C_k & H_k \\ G_k & C_k & B_k & -F_k \\ D_k & -H_k & F_k & B_k \end{bmatrix} \right\} \begin{bmatrix} x^R_k \\ x^I_k \\ x^J_k \\ x^K_k \end{bmatrix} \tag{35}$$

where the elements of the matrices are simplified for the sake of simplicity as $\alpha_k = h^R_k, \beta_k = h^I_k, \gamma_k = h^J_k, \delta_k = h^K_k, \varepsilon_k = g^R_k$, $\zeta_k = g^I_k, \eta_k = g^J_k, \theta_k = g^K_k$, and where $A_k = \alpha_k + \varepsilon_k, B_k = \alpha_k - \varepsilon_k, C_k = -\delta_k + \theta_k, D_k = -\delta_k - \theta_k, E_k = -\beta_k - \zeta_k, F_k = -\beta_k + \zeta_k, G_k = -\gamma_k - \eta_k, H_k = -\gamma_k + \eta_k$. Note here that $A_k, B_k, C_k, D_k, E_k, F_k, G_k, H_k$ are independent parameters each other because $s+t$ and $s-t$ are generally arbitrary real numbers for any $s, t \in \mathbb{R}$. Therefore, the degree of freedom of parameters is $8N$. ■

4) Quaternion Involution Widely Linear Mean Square Estimation: Finally, we analyze the framework called the *quaternion involution widely linear mean square estimation* [22], [23]. A true value $y \in \mathbb{H}$ is estimated from an observed value $\boldsymbol{x} \in \mathbb{H}^N$ where y is a quaternion random variable and $\boldsymbol{x}$ is a quaternion random vector. Assume an estimated value, $\hat{y}_{IWL}$, expressed as

$$\hat{y}_{IWL} = \boldsymbol{h}^H \boldsymbol{x} + \boldsymbol{g}^H \boldsymbol{x}^{\mathsf{i}} + \boldsymbol{u}^H \boldsymbol{x}^{\mathsf{j}} + \boldsymbol{v}^H \boldsymbol{x}^{\mathsf{k}} \tag{36}$$

where $\boldsymbol{g}, \boldsymbol{h}, \boldsymbol{u}, \boldsymbol{v} \in \mathbb{H}^N$ (see Eqs. (19)–(21) for $\boldsymbol{x}^{\mathsf{i}}, \boldsymbol{x}^{\mathsf{j}}, \boldsymbol{x}^{\mathsf{k}}$). The objective is to find a quaternion-valued parameter $\boldsymbol{g}, \boldsymbol{h}, \boldsymbol{u}, \boldsymbol{v} \in \mathbb{H}^N$ that minimizes the MSE $E|y - \hat{y}_{IWL}|^2$.

Proposition 8: An estimator $\hat{y}_{IWL}$ (Eq. (36)) is obtained through linear combinations of four different types of elements of the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ which are applied to each element of an observed value $\boldsymbol{x}$ in the quaternion involution WL mean square estimation where $\mathbb{R}^+$ is the multiplicative group, and $SO_4(\mathbb{R})$ is the four-dimensional rotation group.

Proof: For any $\boldsymbol{g} = (g_1, \ldots, g_N)^T, \boldsymbol{u} = (u_1, \ldots, u_N)^T, \boldsymbol{v} = (v_1, \ldots, v_N)^T \in \mathbb{H}^N$ and $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{H}^N$, the second, third and fourth terms of Eq. (36) can be respectively written as

$$\begin{aligned} \boldsymbol{g}^H \boldsymbol{x}^{\mathsf{i}} &= \sum_{k=1}^{N} g^*_k x^{\mathsf{i}}_k \\ &= \sum_{k=1}^{N} \Big\{ (g^R_k x^R_k + g^I_k x^I_k - g^J_k x^J_k - g^K_k x^K_k) \\ &\quad + \mathsf{i}(g^R_k x^I_k - g^I_k x^R_k + g^J_k x^K_k - g^K_k x^J_k) \\ &\quad - \mathsf{j}(g^R_k x^J_k + g^I_k x^K_k + g^J_k x^R_k + g^K_k x^I_k) \\ &\quad + \mathsf{k}(-g^R_k x^K_k + g^I_k x^J_k + g^J_k x^I_k - g^K_k x^R_k) \Big\}, \end{aligned} \tag{37}$$

$$\begin{aligned} \boldsymbol{u}^H \boldsymbol{x}^{\mathsf{j}} &= \sum_{k=1}^{N} u^*_k x^{\mathsf{j}}_k \\ &= \sum_{k=1}^{N} \Big\{ (u^R_k x^R_k - u^I_k x^I_k + u^J_k x^J_k - u^K_k x^K_k) \\ &\quad + \mathsf{i}(-u^R_k x^I_k - u^I_k x^R_k + u^J_k x^K_k + u^K_k x^J_k) \\ &\quad + \mathsf{j}(u^R_k x^J_k - u^I_k x^K_k - u^J_k x^R_k + u^K_k x^I_k) \\ &\quad - \mathsf{k}(u^R_k x^K_k + u^I_k x^J_k + u^J_k x^I_k + u^K_k x^R_k) \Big\}, \end{aligned} \tag{38}$$

$$\begin{aligned} \boldsymbol{v}^H \boldsymbol{x}^{\mathsf{k}} &= \sum_{k=1}^{N} v^*_k x^{\mathsf{k}}_k \\ &= \sum_{k=1}^{N} \Big\{ (v^R_k x^R_k - v^I_k x^I_k - v^J_k x^J_k + v^K_k x^K_k) \\ &\quad - \mathsf{i}(v^R_k x^I_k + v^I_k x^R_k + v^J_k x^K_k + v^K_k x^J_k) \\ &\quad + \mathsf{j}(-v^R_k x^J_k + v^I_k x^K_k - v^J_k x^R_k + v^K_k x^I_k) \\ &\quad + \mathsf{k}(v^R_k x^K_k + v^I_k x^J_k - v^J_k x^I_k - v^K_k x^R_k) \Big\}. \end{aligned} \tag{39}$$

Then, it follows from Eqs. (27), (37), (38) and (39) that

$$\begin{bmatrix} \hat{y}^R_{IWL} \\ \hat{y}^I_{IWL} \\ \hat{y}^J_{IWL} \\ \hat{y}^K_{IWL} \end{bmatrix} = \sum_{k=1}^{N} \Big\{ |h_k| H_k + |g_k| G^{(\mathsf{i})}_k + |u_k| G^{(\mathsf{j})}_k + |v_k| G^{(\mathsf{k})}_k \Big\} \begin{bmatrix} x^R_k \\ x^I_k \\ x^J_k \\ x^K_k \end{bmatrix} \tag{40}$$

where $|h_k|, |g_k|, |u_k|, |v_k| \in \mathbb{R}^+$ and,

$$G_k^{(\mathsf{i})} \stackrel{\text{def}}{=} \frac{1}{|g_k|}\begin{bmatrix} g_k^R & g_k^I & -g_k^J & -g_k^K \\ -g_k^I & g_k^R & -g_k^K & g_k^J \\ -g_k^J & -g_k^K & -g_k^R & -g_k^I \\ -g_k^K & g_k^J & g_k^I & -g_k^R \end{bmatrix} = R_k^{(\mathsf{i})} K^{(\mathsf{i})},$$

$$G_k^{(\mathsf{j})} \stackrel{\text{def}}{=} \frac{1}{|u_k|}\begin{bmatrix} u_k^R & -u_k^I & u_k^J & -u_k^K \\ -u_k^I & -u_k^R & u_k^K & u_k^J \\ -u_k^J & u_k^K & u_k^R & -u_k^I \\ -u_k^K & -u_k^J & -u_k^I & -u_k^R \end{bmatrix} = R_k^{(\mathsf{j})} K^{(\mathsf{j})},$$

$$G_k^{(\mathsf{k})} \stackrel{\text{def}}{=} \frac{1}{|v_k|}\begin{bmatrix} v_k^R & -v_k^I & -v_k^J & v_k^K \\ -v_k^I & -v_k^R & -v_k^K & -v_k^J \\ -v_k^J & v_k^K & -v_k^R & v_k^I \\ -v_k^K & -v_k^J & v_k^I & v_k^R \end{bmatrix} = R_k^{(\mathsf{k})} K^{(\mathsf{k})},$$

$$R_k^{(\mathsf{i})} \stackrel{\text{def}}{=} G_k^{(\mathsf{i})} K^{(\mathsf{i})}, R_k^{(\mathsf{j})} \stackrel{\text{def}}{=} G_k^{(\mathsf{j})} K^{(\mathsf{j})}, R_k^{(\mathsf{k})} \stackrel{\text{def}}{=} G_k^{(\mathsf{k})} K^{(\mathsf{k})},$$

$$K^{(\mathsf{i})} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

$$K^{(\mathsf{j})} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

$$K^{(\mathsf{k})} \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \in SO_4(\mathbb{R}) \tag{41}$$

for any $1 \leq k \leq N$. Here, $K^{(\mathsf{i})}$, $K^{(\mathsf{j})}$, $K^{(\mathsf{k})}$ are the reflections over the planes spanned by $\{1, \mathsf{i}\}$, $\{1, \mathsf{j}\}$, $\{1, \mathsf{k}\}$ in $\mathbb{R}^4$, respectively. We have thus found that an estimator, $\hat{y}_{IWL}$ in Eq. (36), is obtained through linear combinations of the four elements of the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ ($|h_k|H_k, |g_k|G_k^{(\mathsf{i})}, |u_k|G_k^{(\mathsf{j})}, |v_k|G_k^{(\mathsf{k})}$ in Eq. (40)) which are applied to each element of an observed value $\boldsymbol{x}$ in the quaternion involution WL mean square estimation. ■

We next clarify the relationship of the quaternion involution WL mean square estimation method with a four-dimensional real-valued linear mean square estimation method, and the degree of freedom of parameters of the quaternion involution WL mean square estimation method (see Definition 1).

The following lemma is a four-dimensional generalization of the unique decomposition of a two-dimensional square matrix proved in Eqs. (36)–(40) of [18].

Lemma 1: Any four-dimensional square matrix $W = (w_{ij})_{(1\leq i,j\leq 4)}$ can be decomposed into the sum of four-dimensional square matrices uniquely as follows:

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$
$$= \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 & -b_3 & -b_4 \\ -b_2 & b_1 & -b_4 & b_3 \\ -b_3 & -b_4 & -b_1 & -b_2 \\ -b_4 & b_3 & b_2 & -b_1 \end{bmatrix}$$
$$+ \begin{bmatrix} c_1 & -c_2 & c_3 & -c_4 \\ -c_2 & -c_1 & c_4 & c_3 \\ -c_3 & c_4 & c_1 & -c_2 \\ -c_4 & -c_3 & -c_2 & -c_1 \end{bmatrix} + \begin{bmatrix} d_1 & -d_2 & -d_3 & d_4 \\ -d_2 & -d_1 & -d_4 & -d_3 \\ -d_3 & d_4 & -d_1 & d_2 \\ -d_4 & -d_3 & d_2 & d_1 \end{bmatrix} \tag{42}$$

where $a_k, b_k, c_k, d_k \in \mathbb{R}$ $(1 \leq k \leq 4)$.

Proof: By solving Eq. (42), a_k, b_k, c_k, d_k $(1 \leq k \leq 4)$ are expressed using w_{ij} $(1 \leq i, j \leq 4)$ as follows

$$a_1 = \frac{1}{4}(w_{11} + w_{22} + w_{33} + w_{44}), \tag{43}$$

$$a_2 = -\frac{1}{4}(w_{12} - w_{21} + w_{34} - w_{43}), \tag{44}$$

$$a_3 = -\frac{1}{4}(w_{13} - w_{24} - w_{31} + w_{42}), \tag{45}$$

$$a_4 = -\frac{1}{4}(w_{14} + w_{23} - w_{32} - w_{41}), \tag{46}$$

$$b_1 = \frac{1}{4}(w_{11} + w_{22} - w_{33} - w_{44}), \tag{47}$$

$$b_2 = \frac{1}{4}(w_{12} - w_{21} - w_{34} + w_{43}), \tag{48}$$

$$b_3 = -\frac{1}{4}(w_{13} - w_{24} + w_{31} - w_{42}), \tag{49}$$

$$b_4 = -\frac{1}{4}(w_{14} + w_{23} + w_{32} + w_{41}), \tag{50}$$

$$c_1 = \frac{1}{4}(w_{11} - w_{22} + w_{33} - w_{44}), \tag{51}$$

$$c_2 = -\frac{1}{4}(w_{12} + w_{21} + w_{34} + w_{43}), \tag{52}$$

$$c_3 = \frac{1}{4}(w_{13} + w_{24} - w_{31} - w_{42}), \tag{53}$$

$$c_4 = -\frac{1}{4}(w_{14} - w_{23} - w_{32} + w_{41}), \tag{54}$$

$$d_1 = \frac{1}{4}(w_{11} - w_{22} - w_{33} + w_{44}), \tag{55}$$

$$d_2 = -\frac{1}{4}(w_{12} + w_{21} - w_{34} - w_{43}), \tag{56}$$

$$d_3 = -\frac{1}{4}(w_{13} + w_{24} + w_{31} + w_{42}), \tag{57}$$

$$d_4 = \frac{1}{4}(w_{14} - w_{23} + w_{32} - w_{41}). \tag{58}$$

This completes the proof. ■

Proposition 9: (i) If the observed value $\boldsymbol{x} \in \mathbb{H}^N$, the true value $y \in \mathbb{H}$ and the estimated value $\hat{y}_{IWL} \in \mathbb{H}$ are regarded as real, that is, $\boldsymbol{x} \in \mathbb{R}^{4N}$, $y \in \mathbb{R}^4$ and $\hat{y}_{IWL} \in \mathbb{R}^4$, then the quaternion involution WL mean square estimation is equivalent to a four-dimensional real-valued linear mean square estimation. (ii) The degree of freedom of parameters of the quaternion involution WL mean square estimation is $16N$ where N is the dimension of the observed value $\boldsymbol{x} \in \mathbb{H}^N$.

Proof: It follows from Lemma 1 that the N matrices in the right-hand side of Eq. (40) can be regarded as just N four-dimensional square matrices over real numbers, and their elements are all mutually independent. Therefore, the degree of freedom of parameters is $16N$. This completes the proof. ■

III. Discussion

In this section, we discuss the research opportunities based on the WL estimation methods described in Section II.

First, we compare the complex-valued linear mean square estimation with the complex-valued WL mean square estimation. From Eq. (3), the complex-valued linear mean square estimation can be rewritten as follows

$$\begin{bmatrix} \hat{y}_L^R \\ \hat{y}_L^I \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} a_k & b_k \\ -b_k & a_k \end{bmatrix} \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix} \tag{59}$$

where $a_k, b_k \in \mathbb{R}$ $(1 \leq k \leq N)$. The complex-valued WL mean square estimation can be rewritten, as follows from Eq. (12), as

$$\begin{bmatrix} \hat{y}_{WL}^R \\ \hat{y}_{WL}^I \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} c_k & d_k \\ e_k & f_k \end{bmatrix} \begin{bmatrix} x_k^R \\ x_k^I \end{bmatrix} \tag{60}$$

where $c_k, d_k, e_k, f_k \in \mathbb{R}$ $(1 \leq k \leq N)$.

Remark 1: The output $\hat{y}_{WL}^R$ is estimated using $2N$ parameters $\{c_k, d_k\}_{k=1}^N$, and the $\hat{y}_{WL}^I$ is estimated using $2N$ parameters $\{e_k, f_k\}_{k=1}^N$ from observed values $\{x_k^R, x_k^I\}_{k=1}^N$, respectively. Accordingly, $\hat{y}_{WL}^R$ and $\hat{y}_{WL}^I$ are determined independently. So, the complex-valued WL mean square estimation has sufficient degrees of freedom.

Remark 2: In the complex-valued linear mean square estimation (Eq. (59)), both $\hat{y}_L^R$ and $\hat{y}_L^I$ are estimated respectively using $2N$ parameters $\{a_k, b_k\}_{k=1}^N$ from observed values $\{x_k^R, x_k^I\}_{k=1}^N$. Therefore, a strong restriction is posed on $\hat{y}_L^R$ and $\hat{y}_L^I$ through the parameters $\{a_k, b_k\}_{k=1}^N$. Consequently, for minimizing estimation error, complex-valued WL mean square estimation is intuitively advantageous over the complex-valued linear mean square estimation in terms of the restriction of parameters. This is another aspect of Picinbono's mathematical result [7], taken from another perspective.

In the complex-valued WL mean square estimation, complex-valued data can be represented naturally, while estimation error can be reduced as small as the one of the real-valued linear mean square estimation; this is an advantage of the complex-valued WL mean square estimation.

Remark 3: As seen in Proposition 2, parameters in complex-valued WL mean square estimation impose the geometric structure on complex numbers (rotation and reflection). If this feature is utilized as intended, then this is another advantage of the complex-valued WL mean square estimation (see the last paragraphs in this section).

TABLE I
Comparison of the Real-Valued Linear Mean Square Estimation (MSE) Method and the Two Complex-Valued Mean Square Estimation Methods. $\varepsilon_2 \leq \varepsilon_1$ Holds True [7]

	Representation for Complex-Valued Data	Estimation Error
Real-Valued Linear MSE	Not Natural	ε_2
Complex-Valued Linear MSE	Natural	ε_1
Complex-Valued WL MSE	Natural	ε_2

Table I summaries the complex-valued linear mean square estimation, the complex-valued WL mean square estimation, and the real-valued mean square estimation methods.

In addition, we discuss the computational complexities of the estimation methods. Various types of algorithms can be considered that minimizes the mean square error $E|y - \hat{y}|^2$, and the computational complexities of estimation methods depend on the algorithms used. Thus, it is difficult to compare the computational complexities of estimation methods from a general point of view. Then, we here adopt the computational complexity of estimator itself as an evaluation criterion because it does not depend on the algorithm used. A computational complexity consists of time and space complexities. Here, time complexity means the sum of four operations for real numbers, and space complexity the sum of parameters where a complex-valued parameter is counted as two because it consists of a real part and an imaginary part. Table II shows that the computational complexity of the estimator of the complex-valued WL method is twice that of the complex-valued linear method, and is four times that of the real-valued linear method.

Next, we provide a comprehensive insight into the three quaternion linear mean square estimation methods: the quaternion linear mean square estimation, the quaternion semi-WL mean square estimation, and the quaternion involution WL mean square estimation. It follows from Eqs. (27), (35) and the proof of Proposition 9 that the vector representation of an estimated value $\hat{y}$ has the following form:

$$\begin{bmatrix} \hat{y}^R \\ \hat{y}^I \\ \hat{y}^J \\ \hat{y}^K \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} a_{11}^{(k)} & a_{12}^{(k)} & a_{13}^{(k)} & a_{14}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} & a_{23}^{(k)} & a_{24}^{(k)} \\ a_{31}^{(k)} & a_{32}^{(k)} & a_{33}^{(k)} & a_{34}^{(k)} \\ a_{41}^{(k)} & a_{42}^{(k)} & a_{43}^{(k)} & a_{44}^{(k)} \end{bmatrix} \begin{bmatrix} x_k^R \\ x_k^I \\ x_k^J \\ x_k^K \end{bmatrix} \tag{61}$$

in each of the three quaternion estimation methods where $a_{lm}^{(k)} \in \mathbb{R}$ $(1 \leq l, m \leq 4)$. Here, a different constraint is imposed on the 16 elements $\{a_{lm}^{(k)}\}$ in the matrix in Eq. (61) for each of the three quaternion estimations, which causes the differences among the degrees of freedom of parameters of the three quaternion estimations (Table III; Propositions 5, 7 and 9). It is intuitively obvious that an estimation method with a large freedom of parameters is more advantageous for the purpose of minimizing estimation error. Actually, it has been proved in [13] that the inequality $\alpha \leq \beta \leq \gamma$ holds true where α, β, γ are respectively the

TABLE II
COMPARISON OF COMPUTATIONAL COMPLEXITIES OF ESTIMATORS OF THE REAL-VALUED LINEAR MEAN SQUARE ESTIMATION (MSE) METHOD AND THE TWO COMPLEX-VALUED MEAN SQUARE ESTIMATION METHODS. N IS THE DIMENSION OF THE OBSERVED VALUE

	Time Complexity			Space Complexity
	$\times, \div$	$+, -$	Sum	
Real-Valued Linear MSE	$2N$	N	$3N$	N
Complex-Valued Linear MSE (Eq. (1))	$4N$	$2N$	$6N$	$2N$
Complex-Valued WL MSE (Eq. (6))	$8N$	$4N$	$12N$	$4N$

TABLE III
COMPARISON OF FREEDOM OF PARAMETERS OF THE THREE QUATERNION MEAN SQUARE ESTIMATION (MSE) METHODS. N IS THE DIMENSION OF THE OBSERVED VALUE

	Freedom of Parameters
Quaternion Linear MSE	$4N$
Quaternion Semi-WL MSE	$8N$
Quaternion Involution WL MSE	$16N$

TABLE IV
COMPARISON OF COMPUTATIONAL COMPLEXITIES OF ESTIMATORS OF THE THREE QUATERNION MEAN SQUARE ESTIMATION (MSE) METHODS. N IS THE DIMENSION OF THE OBSERVED VALUE

	Time Complexity			Space Complexity
	$\times, \div$	$+, -$	Sum	
Quaternion Linear MSE (Eq. (25))	$16N$	$12N$	$28N$	$4N$
Quaternion Semi-WL MSE (Eq. (29))	$32N$	$24N$	$56N$	$8N$
Quaternion Involution WL MSE (Eq. (36))	$64N$	$48N$	$112N$	$16N$

estimation errors of the quaternion involution WL mean square estimation, the quaternion semi-WL mean square estimation and the quaternion linear mean square estimation.

Remark 4: As we have seen in Proposition 9, the quaternion involution WL mean square estimation is equivalent to a four-dimensional real-valued linear mean square estimation. As a matter of fact, the number of parameters of the quaternion involution WL mean square estimation is the same as that of a four-dimensional real-valued linear mean square estimation (see Table III).

From a theoretical point of view, Proposition 9 (i) immediately means that the quaternion involution WL mean square estimation can reduce estimation error as small as the one of the real-valued linear estimation method. However, from a practical point of view, it depends on algorithms used.

We here discuss the computational complexities of the estimators of the three quaternion estimation methods described in this paper where a quaternion-valued parameter is counted as four because it consists of a real part and the three imaginary parts. Table IV shows that the computational complexity of the estimator of the quaternion involution WL method is twice that of the quaternion Semi-WL method, and is four times that of the quaternion linear method.

Remark 5: The computational complexities of the estimators of the complex-valued linear method, and the complex-valued WL method are twice, four times that of the real-valued linear method, respectively (Table II). The computational complexities of the estimators of the quaternion Semi-WL and the quaternion involution WL methods are twice, four times that of the quaternion linear method, respectively (Table IV).

As described above, the introduction of the idea of the widely linearity into estimation methods increases the computational complexities of the estimators. This is considered to be one of the disadvantages. However, this problem will be easily solved by using appropriate parallel computing techniques.

Remark 6: As seen in Proposition 8, the parameters of the quaternion involution WL mean square estimation impose a geometric structure on the group $\mathbb{R}^+ \times SO_4(\mathbb{R})$ (rotation and three reflections over planes), whereas the real-valued linear mean square estimation does not have any structure. This is a key distinguishing point between the estimation in the division algebra of quaternions and vector algebra of $\mathbb{R}^4$.

In the estimation methods using quaternions described above, quaternion data can be represented and processed naturally. This is an advantage of the quaternion estimation methods. On the other hand, the parameters in the three quaternion estimation methods described above include the geometric structure on the similarity transformation, rotation and reflections (Propositions 4, 6 and 8). If this feature is utilized as intended, then it is another advantage of the quaternion mean square estimation methods. We will give an example below.

Finally, we suggest a possibility of applications of the WL estimation methods described in Section II to neural networks using their geometric structures.

It was shown in [24], [25] that a complex-valued neural network whose weights and thresholds are all complex numbers, has the ability to learn the transformation of geometric figures, e.g. rotation, similarity transformation and parallel displacement of straight lines, circles, etc., which originates from complex number operations, especially the multiplication of complex numbers: zw where z, w are complex numbers. The ability to learn transformations has been applied to complement the optical flow (2D velocity vector field on an image) [26] and to the generation of fractal images [27]. Also, Isokawa et al. presented the ability of a quaternion-valued neural network to generalize affine transformation in three-dimensional space such as translation, dilatation and rotation [28]. The ability to learn transformation originates from quaternion operations, especially the multiplication of quaternions: zw where z, w are quaternions. Furthermore, it was shown in [29] that the hyperbolic neural network has an ability to learn hyperbolic rotation as its inherent property, which originates from the multiplication of hyperbolic numbers: zw where z, w are hyperbolic numbers. As described above, extending neural networks to higher dimensions creates such new functions according to algebras used and expand their application fields as a result.

An attempt to introduce the idea of widely linearity into complex-valued neural networks has already been performed in [30] where additional extended input neurons are created for complex conjugate of input signals. Such complex-valued neural network models combined with the idea of widely linearity might have the ability to learn 2D geometric transformation that originates from the widely linearity introduced, especially the multiplication of a complex number and a conjugate of a complex number: zw^* where z, w are complex numbers. The same goes for quaternion neural networks defined in the quaternion domain. Complex-valued neural networks using Eq. (6) such as [30] might have the ability to learn 2D geometric transformation including the reflection over the real axis in $\mathbb{R}^2$ (see Eq. (8) and Fig. 2). Quaternion neural networks using Eq. (29), Eq. (36) might have the ability to learn 4D geometric transformation including the reflection over the real axis in $\mathbb{R}^4$ (see Eq. (32)), the reflections over the planes spanned by $\{1, \mathsf{i}\}$, $\{1, \mathsf{j}\}$, $\{1, \mathsf{k}\}$ in $\mathbb{R}^4$ (see Eq. (41)), respectively.

IV. Conclusions

We have analyzed the hypercomplex WL estimation methods with the aim to examine their associated degrees of freedom against the corresponding estimation errors and underlying geometric structures. It has been shown that the complex-valued WL estimation method and the quaternion involution WL estimation method are equivalent to the two-dimensional or four-dimensional real-valued linear estimation method, respectively, if we regard a complex number as an arrangement of two real numbers and treat a quaternion as an arrangement of four real numbers. The complex-valued WL estimation method and the quaternion involution WL estimation method have been shown to possess the ability to represent complex-valued data and quaternion-valued data naturally, while they can reduce estimation errors to be as small as the one of the real-valued linear estimation method. Moreover, we have shown that the parameters in the complex-valued linear estimation method, the complex-valued WL estimation method, the quaternion linear estimation method, the quaternion semi-WL estimation method and the quaternion involution WL estimation method impose on distinct geometric structures on the corresponding complex numbers and quaternions, whereas the real-valued linear estimation method does not exhibit any structure. This is a key distinguishing different point among the estimation in division algebras of $\mathbb{C}$ and $\mathbb{H}$, compared to estimation in the vector algebras of $\mathbb{R}^2$ and $\mathbb{R}^4$. Although the introduction of the idea of the widely linearity into estimation methods increases the computational complexities of the estimators, it seems that this problem can be easily solved by using appropriate parallel computing techniques.

In our future studies, we will proceed with discovery of inherent properties of geometric structures of the hypercomplex-valued WL estimation method, and will consider various formulations of the WL methods based on the Clifford algebra [31] and the Caley-Dickson number system including octonions [32], [33], that could utilize the second-order statistics and their analyses. A quaternion involution estimation model using another nonstandard quaternion involution given in [21] could also be considered.

Acknowledgment

The authors would like to give special thanks to the members of the Cooperative Research Project of the Research Institute of Electrical Communication, Tohoku University for the stimulated discussion, and the anonymous reviewers for valuable comments.

References

[1] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[2] H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1512–1523, Sep. 2003.
[3] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 746–756, Mar. 2004.
[4] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ, USA: Wiley, 2009.
[5] J. Via, D. P. Palomar, L. Vielva, and I. Santamaria, "Quaternion ICA from second-order statistics," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1586–1600, Apr. 2011.
[6] S. Javidi, C. C. Took, and D. P. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 1967–1978, Dec. 2011.
[7] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
[8] C. C. Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009.
[9] F. Neeser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[10] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Process.*, vol. 91, no. 2, pp. 214–224, 2011.
[11] T. Nitta, "Widely linear processing of hypercomplex signals," in *Proc. Int. Conf. Neural Inf. Process. (Lecture Notes Comput. Sci.)*, 2011, pp. 519–525.
[12] T. Nitta, "A theoretical foundation for the widely linear processing of quaternion-valued data," *Appl. Math.*, vol. 4, no. 12, pp. 1616–1620, 2013.
[13] Y. Xia, C. Jahanchahi, T. Nitta, and D. P. Mandic, "Performance bounds of quaternion estimators," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3287–3292, Dec. 2015.
[14] D. P. Mandic, S. Still, and S. C. Douglas, "Duality between widely linear and dual channel adaptive filtering," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, vol. 4, 2009, pp. 1729–1732.
[15] C. Jahanchahi, S. Kanna, and D. P. Mandic, "Complex dual channel estimation: cost effective widely linear adaptive filtering," *Signal Process.*, vol. 104, pp. 33–42, 2014.
[16] M. Xiang, C. C. Took, and D. P. Mandic, "Cost-effective quaternion minimum mean square error estimation: From widely linear to four-channel processing," *Signal Process.*, vol. 136, pp. 81–91, 2017.
[17] T. Mizoguchi and I. Yamada, "An algebraic translation of Cayley-Dickson linear systems and its applications to online learning," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1438–1453, Mar. 2014.
[18] M. Kobayashi, "Decomposition of rotor Hopfield neural networks using complex numbers," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1366–1370, Apr. 2018.
[19] L. Rodman, *Topics in Quaternion Linear Algebra*. Princeton, NJ, USA: Princeton Univ. Press, 2014.
[20] Z. H. He, "Structure, properties and applications of some simultaneous decompositions for quaternion matrices involving ϕ-skew-Hermicity," *Adv. Appl. Clifford Algebras*, vol. 29, no. 6, 2019. [Online] Available: https://doi.org/10.1007/s00006-018-0921-4
[21] Z. H. He, J. Liu, and T. Y. Tam, "The general ϕ-Hermitian solution to mixed pairs of quaternion matrix Sylvester equations," *Electron. J. Linear Algebra*, vol. 32, pp. 475–499, 2017.

[22] J. Via, D. Ramirez, and I. Santamaria, “Properness and widely linear processing of quaternion random vectors,” *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010.

[23] C. C. Took and D. P. Mandic, “A quaternion widely linear adaptive filter,” *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4427–4431, Aug. 2010.

[24] T. Nitta and T. Furuya, “A complex back-propagation learning,” (in Japanese), *Trans. Inf. Process. Soc. Jpn.*, vol. 32, no. 10, pp. 1319–1329, 1991.

[25] T. Nitta, “An extension of the back-propagation algorithm to complex numbers,” *Neural Netw.*, vol. 10, no. 8, pp. 1392–1415, 1997.

[26] A. Watanabe, N. Yazawa, A. Miyauchi, and M. Miyauchi, “A method to interpret 3D motions using neural networks,” *IEICE Trans. Fund. Electr. Comm. Comput. Sci.*, vol. E77-A, no. 8, pp. 1363–1370, 1994.

[27] M. Miura and E. Aiyoshi, “Approximation and designing of fractal images by complex neural networks,” (in Japanese), *IEEJ Trans. Electron. Inf. Syst.*, vol. 123, no. 8, pp. 1465–1472, 2003.

[28] T. Isokawa, N. Matsui, and H. Nishimura, “Quaternionic neural networks: fundamental properties and applications,” in *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*, T. Nitta, Ed. New York, NY, USA: Information Science Reference, 2009, ch. XVI, pp. 411–439.

[29] T. Nitta and Y. Kuroe, “Hyperbolic gradient operator and hyperbolic back-propagation learning algorithms,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1689–1702, May 2018.

[30] D. Xu, H. Zhang, and D. P. Mandic, “Convergence analysis of an augmented algorithm for fully complex-valued neural networks,” *Neural Netw.*, vol. 69, pp. 44–50, 2015.

[31] K. Gürlebeck, K. Habetha, and W. Spröβig, *Holomorphic Functions in the Plane and N-Dimensional Space*. Cambridge, MA, USA: Birkhäuser, 2008.

[32] I. Kantor and A. Solodovnikov, *Hypercomplex Numbers: An Elementary Introduction to Algebras*. New York, NY, USA: Springer-Verlag, 1989.

[33] J. C. Baez, “The octonions,” *Bull. Amer. Math. Soc.*, vol. 39, no. 2, pp. 145–205, 2001.

Tohru Nitta received the B.S. degree in mathematics and the M.E. and Ph.D. degrees in information engineering from the University of Tsukuba, Tsukuba, Japan, in 1983, 1985, and 1995, respectively. From 1985 to 1990, he was with NEC Corporation and engaged in research on expert systems. He joined the Electrotechnical Laboratory, Agency of Industrial Science and Technology, Ministry of International Trade and Industry, Japan in 1990. Currently, he is a Chief Senior Researcher with the National Institute of Advanced Industrial Science and Technology (AIST) (former Electrotechnical Laboratory), Japan. He was also with the Department of Mathematics, Graduate School of Science, Osaka University, Japan, as an Associate Professor from 2000 to 2006, and as a Professor from 2006 to 2008 (additional post). He has been working in the areas of game theory, stochastic processes, optimization, expert systems, and hypercomplex neural networks such as complex- and quaternion-valued neural networks. From 2010 to 2014, Prof. Nitta was a Board Member of the Japanese Neural Network Society (JNNS). He is also a member of the task force on complex neural networks and a member of the Young Researcher Award Committee, Japan chapter (both within IEEE CIS).

Masaki Kobayashi is a Professor with the University of Yamanashi, Kofu, Japan. He received the B.S., M.S., and D.S. degrees in mathematics from Nagoya University, Nagoya, Japan, in 1989, 1991, and 1996, respectively. He became a Research Associate and an Associate Professor with the University of Yamanashi in 1993 and 2006, respectively. He has been a Professor since 2014.

Danilo P. Mandic (M’99–SM’03–F’12) received the Ph.D. degree in nonlinear adaptive signal processing from Imperial College London, London, U.K., in 1999. He is a Professor of Signal Processing with Imperial College London, London, U.K., where he has been involved in nonlinear adaptive signal processing and nonlinear dynamics. He is a Deputy Director of the Financial Signal Processing Laboratory, Imperial College London. He has been a Guest Professor with Katholieke Universiteit Leuven, Leuven, Belgium; the Tokyo University of Agriculture and Technology, Tokyo, Japan; and Westminster University, London, U.K., and a Frontier Researcher with RIKEN, Wako, Japan. He has two research monographs titled *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (West Sussex, U.K.: Wiley, 2001) and *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models* (West Sussex, U.K.: Wiley, 2009), an edited book titled *Signal Processing Techniques for Knowledge Extraction and Information Fusion* (New York, NY, USA: Springer, 2008), and more than 200 publications on signal and image processing. He has produced award winning papers and products resulting from his collaboration with the industry. Prof. Mandic has been a member of the IEEE Technical Committee on Signal Processing Theory and Methods, and an Associate Editor of the IEEE SIGNAL PROCESSING MAGAZINE, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, IEEE TRANSACTIONS ON SIGNAL PROCESSING, and IEEE TRANSACTIONS ON NEURAL NETWORKS.