Multiple-Model Adaptive Estimation for 3-D and 4-D Signals: A Widely Linear Quaternion Approach

Min Xiang, *Member, IEEE*, Bruno Scalzo Dees, and Danilo P. Mandic, *Fellow, IEEE*

Abstract—Quaternion state estimation techniques have been used in various applications, yet they are only suitable for dynamical systems represented by a single known model. In order to deal with model uncertainty, this paper proposes a class of widely linear quaternion multiple-model adaptive estimation (WL-QMMAE) algorithms based on widely linear quaternion Kalman filters and Bayesian inference. The augmented second-order quaternion statistics is employed to capture complete second-order statistical information in improper quaternion signals. Within the WL-QMMAE framework, a widely linear quaternion interacting multiple-model algorithm is proposed to track time-variant model uncertainty, while a widely linear quaternion static multiple-model algorithm is proposed for time-invariant model uncertainty. A performance analysis of the proposed algorithms shows that, as expected, the WL-QMMAE reduces to semiwidely linear QMMAE for $\mathbb{C}^\eta$-improper signals and further reduces to strictly linear QMMAE for proper signals. Simulation results indicate that for improper signals, the proposed WL-QMMAE algorithms exhibit an enhanced performance over their strictly linear counterparts. The effectiveness of the proposed recursive performance analysis algorithm is also validated.

Index Terms—Interacting multiple-model (IMM) algorithm, multiple-model adaptive estimation (MMAE), static multiple-model (SMM) algorithm, quaternion Kalman filters, quaternion noncircularity, widely linear processing.

I. INTRODUCTION

QUATERNIONS have been traditionally used in aerospace engineering and computer graphics in order to model 3-D rotations and orientations, as their division algebra rectifies numerical problems (accumulation of error and gimbal lock) associated with vector algebras [1]. The recently introduced $\mathbb{HR}$ and generalised $\mathbb{HR}$ calculi [2], [3] and augmented quaternion statistics [4], [5] have triggered a resurgence of research on quaternion-valued signal processing, such as quaternion filters and quaternion neural networks [6]–[12]. This owes to a compact model of mutual information between data channels provided by quaternions and the inherent physically meaningful interpretation for a number of 3-D and 4-D phenomena. These developments have enabled quaternions to find applications in communications, motion tracking, and biomedical signal processing [13]–[16].

The second-order statistical properties of quaternion-valued signals are conventionally characterized by their covariance. However, signal processing techniques that utilize the standard covariance are optimal only for estimating second-order circular (proper) quaternion signals. Advances in quaternion statistics have established that both of the covariance and the three complementary covariances are necessary to characterize the complete augmented second-order statistics of second-order noncircular (improper) quaternion signals [4], [5]. Subsequently, widely linear processing algorithms exploiting the complementary covariances, in addition to the standard covariance, have been proposed, owing to their generic nature. For $\mathbb{C}^\eta$-improper quaternion signals, the widely linear processing reduces to semiwidely linear processing. For example, widely linear state estimation has been used for widely linear quaternion dynamical systems with improper noise [17]–[19], and it reduces to semiwidely linear estimation for semiwidely linear systems with $\mathbb{C}^\eta$-improper noise [20].

The existing quaternion state estimation methods are suitable for dynamical systems represented by a single known model. In reality, there is often considerable uncertainty about the system model. An effective method for circumventing the system uncertainty in the real domain is through multiple-model adaptive estimation (MMAE), which can be seen as an application of the mixture of experts' methodology [21], [22] to the state estimation problem. A real-valued MMAE algorithm typically consists of a bank of filters based on multiple models that represent possible system behavior patterns, referred to as modes, and a fusion algorithm that fuses the state estimates of different filters to form the overall estimate. The fusion algorithm may be based on Bayesian inference [23] or neural network training [24]. The existing MMAE algorithms are divided into two categories: static multiple-model (SMM) algorithms and interacting multiple-model (IMM) algorithms. The former are valid for time-invariant system uncertainty and the latter are effective for time-variant system uncertainty. Both have been extensively studied and used in practical applications, such as target tracking, fault diagnosis, and intelligent control [25]–[28]. The theoretical performance analysis of the MMAE algorithms, especially the IMM, however, is challenging. Recent research on this issue has been focused on recursive approaches, including the hybrid conditional averaging technique for predicting the mean square error (MSE) [29]–[31] and the posterior Cramer–Rao lower bound technique for computing the lower performance bound [32]. These approaches have been shown to be more computationally efficient than

the traditionally used Monte Carlo (MC) simulation method.

For a number of practical engineering problems for which MMAE of 3-D and 4-D signals can be used, such as maneuvering target tracking in the 3-D space, quaternion-valued MMAE (QMMAE) is a compact way to account for mutual information between the dimensions while preserving the inherent physical meaning of the signals. However, QMMAE algorithms for hybrid systems have not yet been investigated in the literature, although real-valued MMAE algorithms have been recently extended to the complex domain [33]. The QMMAE is a nontrivial extension of the real-valued MMAE, since the augmented second-order quaternion statistics must be incorporated to deal with improper quaternion signals. This paper proposes a class of widely linear QMMAE (WL-QMMAE) algorithms based on the widely linear quaternion Kalman filter (WL-QKF). This approach exploits complete second-order statistics of improper quaternion signals. For time-variant system uncertainty, we propose a widely linear quaternion IMM (WL-QIMM) algorithm and a recursive algorithm for the analysis of its performance. For time-invariant system uncertainty, we show that the WL-QIMM can simplify to a widely linear quaternion SMM (WL-QSMM) algorithm. In a similar spirit, we show that for semiwidely linear systems with $\mathbb{C}^\eta$-improper noise, the WL-QMMAE reduces to the semiwidely linear QMMAE (SWL-QMMAE), while for strictly linear systems with $\mathbb{H}$-proper noise, the WL-QMMAE reduces to the strictly linear QMMAE (SL-QMMAE). For rigor, we have also proved the convergence of WL-QSMM.

The rest of this paper is organized as follows. Section II provides an overview of quaternions and quaternion state estimation. Section III presents the WL-QMMAE framework, which yields the WL-QIMM and WL-QSMM algorithms. The performance analysis of these algorithms is provided in Section III. Numerical simulations for the proposed algorithms are presented in Section IV. Section V concludes this paper.

Throughout this paper, we use boldface capital letters to denote matrices, $\mathbf{A}$, boldface lowercase letters for vectors, $\mathbf{a}$, and italic letters for scalar quantities, a. Superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, conjugate, and Hermitian (i.e., transpose and conjugate) operators, respectively. The symbol $E\{\cdot\}$ denotes the statistical expectation operator.

II. BACKGROUND

A. Quaternion Algebra

The quaternion domain $\mathbb{H}$ is a 4-D vector space over the real field $\mathbb{R}$, spanned by the basis $\{1, \imath, \jmath, \kappa\}$. A random quaternion variable $x \in \mathbb{H}$ consists of a real part $\mathfrak{R}[\cdot]$ and an imaginary part $\mathfrak{I}[\cdot]$, which comprises three imaginary components, so that

$$x = \mathfrak{R}[x] + \mathfrak{I}[x] = \mathfrak{R}[x] + \mathfrak{I}_\imath[x]\imath + \mathfrak{I}_\jmath[x]\jmath + \mathfrak{I}_\kappa[x]\kappa \quad (1)$$

where $\mathfrak{R}[x]$, $\mathfrak{I}_\imath[x]$, $\mathfrak{I}_\jmath[x]$, and $\mathfrak{I}_\kappa[x]$ are real variables and $\imath$, $\jmath$, and κ are imaginary units with the properties

$$\imath^2 = \jmath^2 = \kappa^2 = -1, \quad \imath\jmath = -\jmath\imath = \kappa$$
$$\jmath\kappa = -\kappa\jmath = \imath, \quad \kappa\imath = -\imath\kappa = \jmath.$$

The conjugate of x is defined as

$$x^* = \mathfrak{R}[x] - \mathfrak{I}[x] = \mathfrak{R}[x] - \mathfrak{I}_\imath[x]\imath - \mathfrak{I}_\jmath[x]\jmath - \mathfrak{I}_\kappa[x]\kappa.$$

The modulus of x is then

$$|x| = \sqrt{\mathfrak{R}[x]^2 + \mathfrak{I}_\imath[x]^2 + \mathfrak{I}_\jmath[x]^2 + \mathfrak{I}_\kappa[x]^2}$$

and the product of two quaternions $x, y \in \mathbb{H}$ is given by

$$xy = \mathfrak{R}[x]\mathfrak{R}[y] - \mathfrak{I}[x] \cdot \mathfrak{I}[y] + \mathfrak{R}[x]\mathfrak{I}[y]$$
$$+ \mathfrak{R}[y]\mathfrak{I}[x] + \mathfrak{I}[x] \times \mathfrak{I}[y]$$

where "$\cdot$" denotes the scalar product and "$\times$" denotes the vector product. The presence of the vector product causes noncommutativity of the quaternion product, that is, $xy \neq yx$. The quaternion product has the following properties:

$$|xy| = |x||y|, \quad x^{-1} = \frac{x^*}{|x|^2}$$
$$(xy)^{-1} = y^{-1}x^{-1}, \quad (xy)^* = y^*x^*.$$

A quaternion variable x is called a pure quaternion if $\mathfrak{R}[x] = 0$. A quaternion variable x is called a unit quaternion if $|x| = 1$.

Another important notion is the quaternion involution [34], which defines a self-inverse mapping, analogous to the complex conjugate [35]. The general involution of the quaternion variable x is defined as $x^\alpha \triangleq -\alpha x \alpha$, which represents the rotation of the vector part of x by π about a unit pure quaternion α. The quaternion involutions have the property: $(x^\alpha)^\alpha = x$. Accordingly, define $x^{\alpha*} \triangleq (x^\alpha)^* = (x^*)^\alpha$. The three special cases of involutions about the $\imath$, $\jmath$, and κ imaginary axes are given by

$$x^\imath = -\imath x \imath = \mathfrak{R}[x] + \mathfrak{I}_\imath[x]\imath - \mathfrak{I}_\jmath[x]\jmath - \mathfrak{I}_\kappa[x]\kappa$$
$$x^\jmath = -\jmath x \jmath = \mathfrak{R}[x] - \mathfrak{I}_\imath[x]\imath + \mathfrak{I}_\jmath[x]\jmath - \mathfrak{I}_\kappa[x]\kappa$$
$$x^\kappa = -\kappa x \kappa = \mathfrak{R}[x] - \mathfrak{I}_\imath[x]\imath - \mathfrak{I}_\jmath[x]\jmath + \mathfrak{I}_\kappa[x]\kappa. \quad (2)$$

B. Statistics and Estimation of Quaternion Signals

The set of involutions in (2), together with the original quaternion, forms the most frequently used basis for augmented quaternion statistics, which is at the core of the recently proposed widely linear processing methodology [4], [5]. The augmented second-order statistics of a zero-mean random quaternion column vector $\mathbf{x}$ is exploited by the $\imath$-, $\jmath$-, and κ- covariance matrices, $\mathbf{C}_{\mathbf{xx}^\eta} \triangleq E\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^{\eta H}\}$, $\eta \in \{\imath, \jmath, \kappa\}$, which are referred to as complementary covariance matrices, together with the standard Hermitian covariance matrix, $\mathbf{C}_{\mathbf{xx}} \triangleq E\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^H\}$. These covariances, taken together, enable the characterization of the quaternion impropriety (second-order noncircularity) which arises from the degree of correlation and/or power imbalance between the imaginary components relative to the real component. For the nth variable in $\mathbf{x}$, x_n, the impropriety coefficients defined as $\rho_\eta \triangleq |C_{x_n x_n^\eta}/C_{x_n x_n}|$, $\eta = \imath, \jmath, \kappa$, measure the degree of correlation between x_n and each of its involutions [36], [37]. Note $\rho_\eta \in [0, 1]$.

Benefiting from the information contained in the complementary covariances, widely linear processing with the augmented signal vector

$$\mathbf{x}^a \triangleq [\mathbf{x}^T, \mathbf{x}^{\iota T}, \mathbf{x}^{J T}, \mathbf{x}^{\kappa T}]^T \tag{3}$$

and semiwidely linear processing with the semiaugmented signal vector $\mathbf{x}^b \triangleq [\mathbf{x}^T, \mathbf{x}^{\eta T}]^T$, where $\eta \in \{\iota, J, \kappa\}$, achieve a better performance for improper quaternion signals compared to traditional strictly linear processing with $\mathbf{x}$ [38]. The covariance matrices of $\mathbf{x}^a$ and $\mathbf{x}^b$ can be represented by

$$\mathbf{C}_{\mathbf{x}^a \mathbf{x}^a} \triangleq E\{(\mathbf{x}^a - E\{\mathbf{x}^a\})(\mathbf{x}^a - E\{\mathbf{x}^a\})^H\}$$
$$= \begin{bmatrix} \mathbf{C}_{\mathbf{xx}} & \mathbf{C}_{\mathbf{xx}^\iota} & \mathbf{C}_{\mathbf{xx}^J} & \mathbf{C}_{\mathbf{xx}^\kappa} \\ \mathbf{C}_{\mathbf{xx}^\iota}^\iota & \mathbf{C}_{\mathbf{xx}}^\iota & \mathbf{C}_{\mathbf{xx}^\kappa}^\iota & \mathbf{C}_{\mathbf{xx}^J}^\iota \\ \mathbf{C}_{\mathbf{xx}^J}^J & \mathbf{C}_{\mathbf{xx}^\kappa}^J & \mathbf{C}_{\mathbf{xx}}^J & \mathbf{C}_{\mathbf{xx}^\iota}^J \\ \mathbf{C}_{\mathbf{xx}^\kappa}^\kappa & \mathbf{C}_{\mathbf{xx}^J}^\kappa & \mathbf{C}_{\mathbf{xx}^\iota}^\kappa & \mathbf{C}_{\mathbf{xx}}^\kappa \end{bmatrix}$$
$$\mathbf{C}_{\mathbf{x}^b \mathbf{x}^b} \triangleq E\{(\mathbf{x}^b - E\{\mathbf{x}^b\})(\mathbf{x}^b - E\{\mathbf{x}^b\})^H\}$$
$$= \begin{bmatrix} \mathbf{C}_{\mathbf{xx}} & \mathbf{C}_{\mathbf{xx}^\eta} \\ \mathbf{C}_{\mathbf{xx}^\eta}^\eta & \mathbf{C}_{\mathbf{xx}}^\eta. \end{bmatrix}$$

The four degrees of freedom in the quaternion domain allow for different notions of properness: $\mathbb{H}$-properness, $\mathbb{R}^\eta$-properness, and $\mathbb{C}^\eta$-properness.

Definition 1 (Properness of a Random Quaternion Vector): A random quaternion vector $\mathbf{x}$ is $\mathbb{H}$-proper if it is uncorrelated with its involutions $\mathbf{x}^\iota$, $\mathbf{x}^J$ and $\mathbf{x}^\kappa$, so that $\mathbf{C}_{\mathbf{xx}^\iota} = \mathbf{C}_{\mathbf{xx}^J} = \mathbf{C}_{\mathbf{xx}^\kappa} = \mathbf{0}$; $\mathbf{x}$ is $\mathbb{R}^\eta$-proper if it is only uncorrelated with the involution $\mathbf{x}^\eta$, so that only $\mathbf{C}_{\mathbf{xx}^\eta}$ among the three complementary covariances vanishes; $\mathbf{x}$ is $\mathbb{C}^\eta$-improper if it is only correlated with the involution $\mathbf{x}^\eta$, so that all complementary covariances except $\mathbf{C}_{\mathbf{xx}^\eta}$ vanish; otherwise, $\mathbf{x}$ is generally improper.

Based on the above augmented second-order statistics, a generic quaternion multivariate Gaussian distribution (QMGD) for improper random vectors can be expressed as $\mathbf{x}^a \sim \mathcal{N}(\boldsymbol{\mu}^a, \mathbf{C}_{\mathbf{x}^a \mathbf{x}^a})$, where $\boldsymbol{\mu}^a \triangleq E\{\mathbf{x}^a\}$, with a probability density function given by [5]

$$f(\mathbf{x}^a | \boldsymbol{\mu}^a, \mathbf{C}^a) = \frac{4^N \exp\left\{-\frac{1}{2}(\mathbf{x}^a - \boldsymbol{\mu}^a)^H \mathbf{C}_{\mathbf{x}^a \mathbf{x}^a}^{-1}(\mathbf{x}^a - \boldsymbol{\mu}^a)\right\}}{\pi^{2N} \sqrt{|\mathbf{C}_{\mathbf{x}^a \mathbf{x}^a}|}} \tag{4}$$

where $|\mathbf{C}_{\mathbf{x}^a \mathbf{x}^a}|$ is the determinant of $\mathbf{C}_{\mathbf{x}^a \mathbf{x}^a}$.

The augmented form of a quaternion random vector $\mathbf{x}$ can be written as $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^{\eta T}, \mathbf{x}^{\beta T}, \mathbf{x}^{\zeta T}]^T$ for distinct $\eta, \beta, \zeta \in \{\iota, J, \kappa\}$. Note that the ordering does not change the probability density in (4). If $\mathbf{x}$ is $\mathbb{C}^\eta$-improper, then $\mathbf{C}_{\mathbf{x}^a \mathbf{x}^a} = \text{bdiag}(\mathbf{C}_{\mathbf{x}^b \mathbf{x}^b}, \mathbf{C}_{\mathbf{x}^b \mathbf{x}^b}^\beta)$, and (4) becomes

$$f(\mathbf{x}^b | \boldsymbol{\mu}^b, \mathbf{C}^b) = \frac{4^N \exp\left\{-(\mathbf{x}^b - \boldsymbol{\mu}^b)^H \mathbf{C}_{\mathbf{x}^b \mathbf{x}^b}^{-1}(\mathbf{x}^b - \boldsymbol{\mu}^b)\right\}}{\pi^{2N} |\mathbf{C}_{\mathbf{x}^b \mathbf{x}^b}|} \tag{5}$$

where $\boldsymbol{\mu}^b \triangleq E\{\mathbf{x}^b\}$ and the distribution can be expressed as $\mathbf{x}^b \sim \mathcal{N}(\boldsymbol{\mu}^b, \mathbf{C}_{\mathbf{x}^b \mathbf{x}^b})$.

If $\mathbf{x}$ is $\mathbb{H}$-proper, then $\mathbf{C}_{\mathbf{x}^a \mathbf{x}^a} = \text{bdiag}(\mathbf{C}_{\mathbf{xx}}, \mathbf{C}_{\mathbf{xx}}^\iota, \mathbf{C}_{\mathbf{xx}}^J, \mathbf{C}_{\mathbf{xx}}^\kappa)$, and (4) becomes

$$f(\mathbf{x} | \boldsymbol{\mu}, \mathbf{C}) = \frac{4^N \exp\{-2(\mathbf{x} - \boldsymbol{\mu})^H \mathbf{C}_{\mathbf{xx}}^{-1}(\mathbf{x} - \boldsymbol{\mu})\}}{\pi^{2N} |\mathbf{C}_{\mathbf{xx}}|^2} \tag{6}$$

where $\boldsymbol{\mu} \triangleq E\{\mathbf{x}\}$ and the distribution can be expressed as $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C}_{\mathbf{xx}})$.

To deal with improper quaternion signals, the WL-QKF based on the widely linear state-space model

$$\mathbf{x}_n^a = \mathbf{F}_n^a \mathbf{x}_{n-1}^a + \mathbf{w}_n^a$$
$$\mathbf{z}_n^a = \mathbf{H}_n^a \mathbf{x}_n^a + \mathbf{v}_n^a \tag{7}$$

has been proposed [17]. It is optimal for widely linear quaternion dynamical systems with general (proper or improper) noise [17]. In (7), n denotes the time instant, and $\mathbf{x}_n^a$, $\mathbf{z}_n^a$, $\mathbf{w}_n^a$, and $\mathbf{v}_n^a$ are the augmented forms of the state variable vector $\mathbf{x}_n$, the observation vector $\mathbf{z}_n$, the state noise vector $\mathbf{w}_n$, and the observation noise vector $\mathbf{v}_n$, as in (3), while the state transition and observation matrices, $\mathbf{F}_n^a$ and $\mathbf{H}_n^a$, have the following structure:

$$\mathbf{F}_n^a = \begin{bmatrix} \mathbf{F}_n & \mathbf{F}_{\iota,n} & \mathbf{F}_{J,n} & \mathbf{F}_{\kappa,n} \\ \mathbf{F}_{\iota,n}^\iota & \mathbf{F}_n^\iota & \mathbf{F}_{\kappa,n}^\iota & \mathbf{F}_{J,n}^\iota \\ \mathbf{F}_{J,n}^J & \mathbf{F}_{\kappa,n}^J & \mathbf{F}_n^J & \mathbf{F}_{\iota,n}^J \\ \mathbf{F}_{\kappa,n}^\kappa & \mathbf{F}_{J,n}^\kappa & \mathbf{F}_{\iota,n}^\kappa & \mathbf{F}_n^\kappa \end{bmatrix}$$

$$\mathbf{H}_n^a = \begin{bmatrix} \mathbf{H}_n & \mathbf{H}_{\iota,n} & \mathbf{H}_{J,n} & \mathbf{H}_{\kappa,n} \\ \mathbf{H}_{\iota,n}^\iota & \mathbf{H}_n^\iota & \mathbf{H}_{\kappa,n}^\iota & \mathbf{H}_{J,n}^\iota \\ \mathbf{H}_{J,n}^J & \mathbf{H}_{\kappa,n}^J & \mathbf{H}_n^J & \mathbf{H}_{\iota,n}^J \\ \mathbf{H}_{\kappa,n}^\kappa & \mathbf{H}_{J,n}^\kappa & \mathbf{H}_{\iota,n}^\kappa & \mathbf{H}_n^\kappa \end{bmatrix}$$

where $\mathbf{F}_n, \mathbf{F}_{\iota,n}, \mathbf{F}_{J,n}, \mathbf{F}_{\kappa,n}, \mathbf{H}_n, \mathbf{H}_{\iota,n}, \mathbf{H}_{J,n}$, and $\mathbf{H}_{\kappa,n}$ are quaternion matrices.

Observe that under the condition of $\mathbb{C}^\eta$-improperness, the widely linear model in (7) reduces to the semiwidely linear one given by

$$\mathbf{x}_n^b = \mathbf{F}_n^b \mathbf{x}_{n-1}^b + \mathbf{w}_n^b$$
$$\mathbf{z}_n^b = \mathbf{H}_n^b \mathbf{x}_n^b + \mathbf{v}_n^b \tag{8}$$

where the state transition and observation matrices, $\mathbf{F}_n^b$ and $\mathbf{H}_n^b$, have the following structure:

$$\mathbf{F}_n^b = \begin{bmatrix} \mathbf{F}_n & \mathbf{F}_{\eta,n} \\ \mathbf{F}_{\eta,n}^\eta & \mathbf{F}_n^\eta \end{bmatrix}, \quad \mathbf{H}_n^b = \begin{bmatrix} \mathbf{H}_n & \mathbf{H}_{\eta,n} \\ \mathbf{H}_{\eta,n}^\eta & \mathbf{H}_n^\eta. \end{bmatrix}$$

The semiwidely linear quaternion state estimation based on the semiwidely linear model in (8) is therefore optimal when the true system model is semiwidely linear with $\mathbb{C}^\eta$-improper $\mathbf{w}_n$ and $\mathbf{v}_n$ [39].

Furthermore, under the condition of $\mathbb{H}$-properness, the widely linear model in (7) reduces to the strictly linear one given by

$$\mathbf{x}_n = \mathbf{F}_n \mathbf{x}_{n-1} + \mathbf{w}_n$$
$$\mathbf{z}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{v}_n. \tag{9}$$

The strictly linear quaternion state estimation based on the strictly linear model in (9) is optimal only when the true system model is strictly linear and $\mathbf{w}_n$ and $\mathbf{v}_n$ are $\mathbb{H}$-proper [40]. This conclusion is an extension of the analysis result for complex Kalman filters [41] to the quaternion domain.

III. QUATERNION MMAE

If a quaternion dynamical system is represented by the widely linear model in (7), but the state transition matrix, the observation matrix, and the properties of the noises are uncertain over time, a quaternion Kalman filter based on a single assumed model will suffer a great performance loss. In such cases, we may employ L potential models as candidates for the unknown system model. We can assume that the system model is always included in the L candidate models, each of which is represented by

$$\mathbf{x}_n^a = \mathbf{F}_n^{a[l]}\mathbf{x}_{n-1}^a + \mathbf{w}_n^{a[l]}$$
$$\mathbf{z}_n^a = \mathbf{H}_n^{a[l]}\mathbf{x}_n^a + \mathbf{v}_n^{a[l]} \qquad (10)$$

where $l \in \{1, 2, \ldots, L\}$ is the model index, $\mathbf{w}_n^{a[l]}$ and $\mathbf{v}_n^{a[l]}$ are zero-mean improper state and observation noise, and $\mathbf{F}_n^{a[l]}$ and $\mathbf{H}_n^{a[l]}$ are state transition and observation matrices corresponding to model l. We refer to the system behavior pattern as a mode. For example, if the system can be represented by model l at time n, we say that the system is in mode l at time n. Upon implementing L WL-QKFs, we can obtain their augmented state estimates and error covariances given by

$$\hat{\mathbf{x}}_{n|n}^{a[l]} = E\{\mathbf{x}_n^a | \mathbf{Z}_n^a, \tau_n^{[l]}\}$$
$$\mathbf{P}_{n|n}^{a[l]} = E\{(\mathbf{x}_n^a - \hat{\mathbf{x}}_{n|n}^{a[l]})(\mathbf{x}_n^a - \hat{\mathbf{x}}_{n|n}^{a[l]})^H | \mathbf{Z}_n^a, \tau_n^{[l]}\}.$$

By fusing the L estimates, the overall augmented estimate can be obtained as

$$E\{\mathbf{x}_n^a | \mathbf{Z}_n^a\} = \sum_{l=1}^{L} p(\tau_n^{[l]} | \mathbf{Z}_n^a)\hat{\mathbf{x}}_{n|n}^{a[l]} \qquad (11)$$

while the overall augmented error covariance becomes

$$E\{(\mathbf{x}_n^a - E\{\mathbf{x}_n^a | \mathbf{Z}_n^a\})(\mathbf{x}_n^a - E\{\mathbf{x}_n^a | \mathbf{Z}_n^a\})^H | \mathbf{Z}_n^a, \tau_n^{[l]}\}$$
$$= \sum_{l=1}^{L} p(\tau_n^{[l]} | \mathbf{Z}_n^a)[\mathbf{P}_{n|n}^{a[l]} + (\hat{\mathbf{x}}_{n|n}^{a[l]} - \hat{\mathbf{x}}_{n|n}^a)(\hat{\mathbf{x}}_{n|n}^{a[l]} - \hat{\mathbf{x}}_{n|n}^a)^H]$$
$$= \sum_{l=1}^{L} p(\tau_n^{[l]} | \mathbf{Z}_n^a)[\mathbf{P}_{n|n}^{a[l]} + \hat{\mathbf{x}}_{n|n}^{a[l]}(\hat{\mathbf{x}}_{n|n}^{a[l]})^H] - \hat{\mathbf{x}}_{n|n}^a(\hat{\mathbf{x}}_{n|n}^a)^H \qquad (12)$$

where $\mathbf{Z}_n^a \triangleq [\mathbf{z}_1^{aT}, \ldots, \mathbf{z}_n^{aT}]^T = [\mathbf{Z}_{n-1}^{aT}, \mathbf{z}_n^{aT}]^T$ is the set of measurements and $\tau_n^{[l]}$ is the event that the system is in mode l at time instant n. Using Bayes' law, we next obtain

$$p(\tau_n^{[l]} | \mathbf{Z}_n^a)$$
$$= \frac{p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[l]})p(\tau_n^{[l]} | \mathbf{Z}_{n-1}^a)}{\sum_{i=1}^{L} p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[i]})p(\tau_n^{[i]} | \mathbf{Z}_{n-1}^a)}$$
$$= \frac{p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[l]})\sum_{i=1}^{L} p(\tau_{n-1}^{[i]} | \mathbf{Z}_{n-1}^a)p(\tau_n^{[l]} | \tau_{n-1}^{[i]})}{\sum_{i=1}^{L} p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[i]})\sum_{j=1}^{L} p(\tau_{n-1}^{[j]} | \mathbf{Z}_{n-1}^a)p(\tau_n^{[i]} | \tau_{n-1}^{[j]})}$$
$$\qquad (13)$$

where $p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[l]})$ is the likelihood function of $\tau_n^{[l]}$. Note that $p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[l]}) = p(\mathbf{r}_n^{a[l]} | \tau_n^{[l]})$, where

$$\mathbf{r}_n^{a[l]} \triangleq \mathbf{z}_n^a - \mathbf{H}_n^{[l]}\hat{\mathbf{x}}_{n|n-1}^{a[l]} \qquad (14)$$

is referred to as the augmented residual vector. Following the principle commonly used for real-valued and complex-valued MMAE, we assume the augmented residual vector to be normally distributed, with $\mathbf{r}_n^{a[l]} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_n^{a[l]})$, where $\mathbf{S}_n^{a[l]}$ is the augmented covariance matrix of $\mathbf{r}_n^{a[l]}$. Denote the likelihood function $p(\mathbf{z}_n^a | \mathbf{Z}_{n-1}^a, \tau_n^{[l]})$ by $\mathcal{L}_n^{[l]}$. $\mathcal{L}_n^{[l]}$ can be computed as

$$\mathcal{L}_n^{[l]} = f(\mathbf{r}_n^{a[l]} | \mathbf{0}, \mathbf{S}_n^{a[l]}). \qquad (15)$$

Note that $\mathbf{S}_n^{a[l]}$ is an intermediate result in the WL-QKF algorithm corresponding to model l. The estimated mode at time n is therefore given by $\arg\max_l p(\tau_n^{[l]} | \mathbf{Z}_n^a)$. The above constitutes the main framework of the proposed WL-QMMAE. This framework is flexible and therefore can incorporate other approaches for nonlinear estimation. For example, the parameters within the selected models may be a function of the order of the signal subspace or the number of noncircular signals in the subspace [42]. The fusion in (11) and (12) is a data-driven approach and can be viewed as a more direct and adaptive way of implementing the multistep generalized likelihood ratio test with the threshold being unit [42]. The conditional probabilities of the various hypotheses modeled in the filters can be based on criteria other than the likelihood. For instance, in order to select the model that best fits the observation data while avoiding overfitting, the likelihood function can be replaced with $\exp(-\text{ITC}_l)$, where $\text{ITC}_l = -\ln[f(\mathbf{r}_n^{a[l]} | \mathbf{0}, \mathbf{S}_n^{a[l]})] + \alpha O^{[l]}$ is the information-theoretic criterion of model l, with $\alpha O^{[l]}$ a penalty term that penalizes complicated models to avoid overfitting, α the penalty factor dependent on the chosen criterion, and $O^{[l]}$ the order of model l [43].

A. WL-QIMM

To understand the WL-QMMAE, the probability $p(\tau_n^{[l]} | \mathbf{Z}_{n-1}^a)$ in (13) needs to be explained further. For systems with a time-varying mode, we make the standard assumption that the mode transition is governed by a first-order Markov chain, that is

$$p(\tau_n^{[l]} | \tau_{n-1}^{[i]}) = \pi^{[il]} \quad \forall i, l \in \{1, 2, \ldots, L\} \qquad (16)$$

where $\pi^{[il]}$ is the Markov transition probability from mode i to mode l. Upon incorporating this assumption and imposing a reinitialization step at the beginning of the estimation to couple the L WL-QKFs, the WL-MMAE framework reduces to a practical WL-QIMM algorithm, an extension of the real-valued IMM [44]. An iteration of this recursive algorithm at time n can be described as follows and can be shown in the schematic in Fig. 1.

Step 1 (Reinitialization): For $l = 1, \ldots, L$, $i = 1, \ldots, L$, the probability of the system is in mode i at time $(n-1)$ given that the system is in mode l at time n, which is computed as

$$\gamma_n^{[li]} \triangleq p(\tau_{n-1}^{[i]} | \tau_n^{[l]}, \mathbf{Z}_{n-1}^a) = \frac{\pi^{[il]}\mu_{n-1}^{[i]}}{\sum_{j=1}^{L} \pi^{[jl]}\mu_{n-1}^{[j]}}$$

where $\mu_{n-1}^{[i]}$ termed mode probability is the estimate of $p(\tau_{n-1}^{[l]} | \mathbf{Z}_{n-1}^a)$ in the previous iteration. For $l = 1, \ldots, L$,
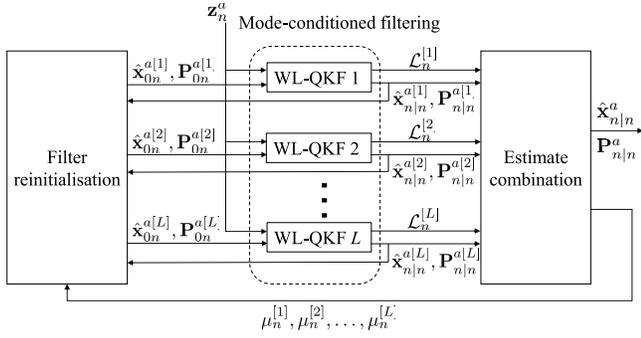
Fig. 1. Block diagram of the WL-QIMM algorithm.

compute the augmented initial state of filter l based on Bayes' law as

$$\mathbf{x}_{0n}^{a[l]} = E\{\mathbf{x}_{n-1}^a | \tau_n^{[l]}, \mathbf{Z}_{n-1}^a\} = \sum_{i=1}^{L} \hat{\mathbf{x}}_{n-1|n-1}^{a[i]} \gamma_n^{[li]} \quad (17)$$

and the augmented initial error covariance of filter l as

$$\mathbf{P}_{0n}^{a[l]}$$
$$= \sum_{i=1}^{L} \gamma_n^{[li]} \big[\mathbf{P}_{n-1|n-1}^{a[i]} + \big(\mathbf{x}_{0n}^{[l]} - \hat{\mathbf{x}}_{n-1|n-1}^{a[i]}\big)\big(\mathbf{x}_{0n}^{[l]} - \hat{\mathbf{x}}_{n-1|n-1}^{a[i]}\big)^H\big]$$
$$= \sum_{i=1}^{L} \gamma_n^{[li]} \big[\mathbf{P}_{n-1|n-1}^{a[i]} + \hat{\mathbf{x}}_{n-1|n-1}^{a[i]} \big(\hat{\mathbf{x}}_{n-1|n-1}^{a[i]}\big)^H\big] - \mathbf{x}_{0n}^{a[l]}\big(\mathbf{x}_{0n}^{a[l]}\big)^H.$$
$$(18)$$

Step 2 (Mode-Conditioned Kalman Filtering): Implement L WL-QKFs with the initial conditions $\mathbf{x}_{0n}^{a[l]}$ and $\mathbf{P}_{0n}^{a[l]}$. Specifically, for $l = 1, \ldots, L$, compute the predicted augmented state

$$\hat{\mathbf{x}}_{n|n-1}^{a[l]} = \mathbf{F}_n^{a[l]} \mathbf{x}_{0n}^{a[l]}$$

predicted augmented error covariance

$$\mathbf{P}_{n|n-1}^{a[l]} = \mathbf{F}_n^{a[l]} \mathbf{P}_{0n}^{a[l]} \big(\mathbf{F}_n^{a[l]}\big)^H + \mathbf{Q}_n^{a[l]}$$

augmented residual

$$\mathbf{r}_n^{a[l]} = \mathbf{z}_n^a - \mathbf{H}_n^{[l]} \hat{\mathbf{x}}_{n|n-1}^{a[l]}$$

augmented residual covariance

$$\mathbf{S}_n^{a[l]} = \mathbf{H}_n^{a[l]} \mathbf{P}_{n|n-1}^{a[l]} \big(\mathbf{H}_n^{a[l]}\big)^H + \mathbf{R}_n^{a[l]}$$

filter gain

$$\mathbf{K}_n^{a[l]} = \mathbf{P}_{n|n-1}^{a[l]} \big(\mathbf{H}_n^{a[l]}\big)^H \big(\mathbf{S}_n^{a[l]}\big)^{-1}$$

updated augmented state

$$\hat{\mathbf{x}}_{n|n}^{a[l]} = \hat{\mathbf{x}}_{n|n-1}^{a[l]} + \mathbf{K}_n^{a[l]} \mathbf{r}_n^{a[l]}$$

updated augmented error covariance

$$\mathbf{P}_{n|n}^{a[l]} = \mathbf{P}_{n|n-1}^{a[l]} - \mathbf{K}_n^{a[l]} \mathbf{H}_n^{a[l]} \mathbf{P}_{n|n-1}^{a[l]}$$

likelihood for mode l as in (15).

Step 3 (Combination): For $l = 1, \ldots, L$, compute the estimate of $p(\tau_n^{[l]} | \mathbf{Z}_n^a)$, denoted by $\mu_n^{[l]}$, as

$$\mu_n^{[l]} = \frac{\mathcal{L}_n^{[l]} \sum_{i=1}^{L} \pi^{[il]} \mu_{n-1}^{[i]}}{\sum_{i=1}^{L} \mathcal{L}_n^{[i]} \sum_{j=1}^{L} \pi^{[ji]} \mu_{n-1}^{[j]}} \quad (19)$$

according to (13) and (16). Based on (11), compute the overall augmented state estimate as a weighted sum of the L augmented state estimates from the L WL-QKFs, that is

$$\hat{\mathbf{x}}_{n|n}^a = \sum_{l=1}^{L} \mu_n^{[l]} \hat{\mathbf{x}}_{n|n}^{a[l]}. \quad (20)$$

Based on (12), compute the overall augmented error covariance as

$$\mathbf{P}_{n|n}^a = \sum_{l=1}^{L} \mu_n^{[l]} \big[\mathbf{P}_{n|n}^{a[l]} + \big(\hat{\mathbf{x}}_{n|n}^{a[l]} - \hat{\mathbf{x}}_{n|n}^a\big)\big(\hat{\mathbf{x}}_{n|n}^{a[l]} - \hat{\mathbf{x}}_{n|n}^a\big)^H\big]$$
$$= \sum_{l=1}^{L} \mu_n^{[l]} \big[\mathbf{P}_{n|n}^{a[l]} + \hat{\mathbf{x}}_{n|n}^{a[l]} \big(\hat{\mathbf{x}}_{n|n}^{a[l]}\big)^H\big] - \hat{\mathbf{x}}_{n|n}^a \big(\hat{\mathbf{x}}_{n|n}^a\big)^H. \quad (21)$$

B. Performance Analysis of WL-QIMM

The stability condition of the real-valued IMM algorithm proved in [45] is readily extended to the WL-QIMM, that is, if the system is uniformly controllable and uniformly observable, the covariances of the mode-conditioned WL-QKFs are uniformly bounded from above and below. Furthermore, to efficiently evaluate the estimation performance of WL-QIMM, this section extends the hybrid conditional averaging algorithms for the performance analysis of real-valued IMM [29], [30] to the quaternion domain. The resulting recursive algorithm, without recourse to multiple trials, is more computationally efficient than the commonly used MC simulation method. Note that owing to the widely linear nature of the WL-QIMM algorithm, the performance analysis algorithm must also be in a widely linear form, since the augmented second-order quaternion statistics must be incorporated.

The performance of WL-QIMM depends on the operating scenario, so that we must analyze the average performance of WL-QIMM in a definite operating scenario $\mathcal{S}$ in which the system can be represented by a widely linear quaternion state model in the form of (7). Compared with (10), the model in (7) is general in the sense that $\mathbf{F}_n^a$, $\mathbf{H}_n^a$, $\mathbf{w}_n^a$, and $\mathbf{v}_n^a$ may or may not belong to the candidate model sets used by the L WL-QKFs. We define $\bar{\mathbf{x}} \triangleq E\{\mathbf{x} | \mathcal{S}\}$ to denote the mean of the quaternion random vector $\mathbf{x}$ in a scenario $\mathcal{S}$, while $\mathrm{Cov}(\mathbf{x}, \mathbf{y}) \triangleq E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})^H | \mathcal{S}\}$ denotes the cross covariance of two quaternion random vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{S}$. Based on these definitions, we can derive a recursive algorithm that estimates the MSE of WL-QIMM without generating observation data in multiple trials. This is achieved by the conditional expectation operation mentioned earlier, through which the randomness of the error covariance due to the noise is averaged out. This recursive algorithm is similar to the performance analysis algorithm for real-valued IMM [30], which originated from [29], but is different in two respects: 1) the augmented forms of the quaternion variables are employed and 2) the generic QMGD with the probability density function in (4) is incorporated into the evaluation of the average likelihood functions and the average mode probabilities. Next, we describe an iteration of the proposed performance analysis algorithm at time n, which is shown in Fig. 2, but for space concerns, omit the derivation,
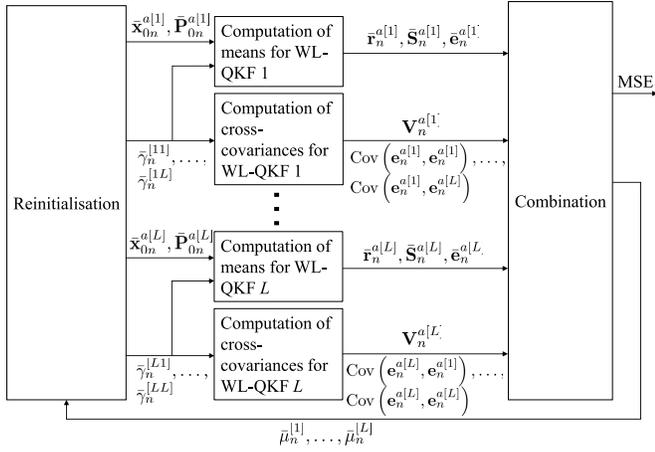
Fig. 2. Block diagram of the performance analysis algorithm for WL-QIMM.

owing to its generic extension from the real-valued algorithm. For more details, we refer the reader to [29] and [30].

Step 1 (Reinitialization): Compute the means of the initial conditions of the mode-conditioned WL-QKFs. For $l = 1, \ldots, L$, $i = 1, \ldots, L$, compute the mean of the mixing probability as

$$\bar{\gamma}_n^{[li]} \approx \frac{\pi^{[il]} \bar{\mu}_{n-1}^{[i]}}{\sum_{j=1}^{L} \pi^{[jl]} \bar{\mu}_{n-1}^{[j]}}. \tag{22}$$

For $l = 1, \ldots, L$, compute the mean of the initial condition of filter l as

$$\bar{\mathbf{x}}_{0n}^{[l]} = \sum_{i=1}^{L} \bar{\gamma}_n^{[li]} \bar{\mathbf{x}}_{n-1|n-1}^{[i]}$$

$$\bar{\mathbf{P}}_{0n}^{a[l]} = \sum_{i=1}^{L} \bar{\gamma}_n^{[li]} \big[\bar{\mathbf{P}}_{n-1|n-1}^{a[i]} + \big(\bar{\mathbf{x}}_{0n}^{a[l]} - \bar{\mathbf{x}}_{n-1|n-1}^{a[i]} \big)$$
$$\times \big(\bar{\mathbf{x}}_{0n}^{a[l]} - \bar{\mathbf{x}}_{n-1|n-1}^{a[i]} \big)^H \big]$$

$$= \sum_{i=1}^{L} \bar{\gamma}_n^{[li]} \big[\bar{\mathbf{P}}_{n-1|n-1}^{a[i]} + \bar{\mathbf{x}}_{n-1|n-1}^{a[i]} \big(\bar{\mathbf{x}}_{n-1|n-1}^{a[i]} \big)^H \big]$$
$$- \bar{\mathbf{x}}_{0n}^{a[l]} \big(\bar{\mathbf{x}}_{0n}^{a[l]} \big)^H.$$

Step 2 (Mode-Conditioned Kalman Filtering): Compute the means of residuals and estimates of the mode-conditioned WL-QKFs. For $l = 1, \ldots, L$, define the estimation error of filter l as $\mathbf{e}_n^{[l]} \triangleq \mathbf{x}_n - \hat{\mathbf{x}}_{n|n}^{[l]}$, and the error of the initial state of filter l as $\mathbf{e}_{0n}^{[l]} \triangleq \mathbf{x}_{n-1} - \mathbf{x}_{0n}^{[l]}$. Their means are related by

$$\bar{\mathbf{e}}_{0n}^{[l]} \approx \sum_{i=1}^{L} \bar{\gamma}_n^{[li]} \bar{\mathbf{e}}_{n-1}^{[i]}.$$

For $l = 1, \ldots, L$, define the model error as

$$\Delta^{a[l]} \triangleq \mathbf{H}_n^a \mathbf{F}_n^a - \mathbf{H}_n^{a[l]} \mathbf{F}_n^{a[l]}$$

and compute

$$\bar{\mathbf{r}}_n^{a[l]} = \mathbf{H}_n^{a[l]} \mathbf{F}_n^{a[l]} \bar{\mathbf{e}}_{0n}^{a[l]} + \Delta^{a[l]} \bar{\mathbf{x}}_{n-1}^a$$

$$\bar{\mathbf{P}}_{n|n-1}^{a[l]} = \mathbf{F}_n^{a[l]} \bar{\mathbf{P}}_{0n}^{a[l]} \big(\mathbf{F}_n^{a[l]} \big)^H + \mathbf{Q}_n^{a[l]}$$

$$\bar{\mathbf{S}}_n^{a[l]} = \mathbf{H}_n^{a[l]} \bar{\mathbf{P}}_{n|n-1}^{a[l]} \big(\mathbf{H}_n^{a[l]} \big)^H + \mathbf{R}_n^{a[l]}$$

$$\bar{\mathbf{K}}_n^{a[l]} = \bar{\mathbf{P}}_{n|n-1}^{a[l]} \big(\mathbf{H}_n^{a[l]} \big)^H \big(\bar{\mathbf{S}}_n^{a[l]} \big)^{-1}$$

$$\bar{\mathbf{x}}_{n|n}^{[l]} = \mathbf{F}_n^{a[l]} \mathbf{x}_{0n}^{a[l]} + \bar{\mathbf{K}}_n^{a[l]} \bar{\mathbf{r}}_n^{a[l]}$$

$$\bar{\mathbf{P}}_{n|n}^{a[l]} = \bar{\mathbf{P}}_{n|n-1}^{a[l]} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}_n^{a[l]} \bar{\mathbf{P}}_{n|n-1}^{a[l]}$$

$$\bar{\mathbf{e}}_n^{a[l]} = \big(\mathbf{I} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}_n^{a[l]} \big) \mathbf{F}_n^{a[l]} \bar{\mathbf{e}}_{0n}^{a[l]}$$
$$+ \big(\mathbf{F}_n^a - \mathbf{F}_n^{a[l]} - \bar{\mathbf{K}}_n^{a[l]} \Delta^{a[l]} \big) \bar{\mathbf{x}}_{n-1}^a$$

where $\mathbf{I}$ is the identity matrix. Update $\bar{\mathbf{x}}_n^a = \mathbf{F}_n^a \bar{\mathbf{x}}_{n-1}^a$.

Step 3 (Cross-Covariance Analysis): Compute the cross covariances of augmented residuals and augmented estimation errors of the mode-conditioned WL-QKFs. For $l = 1, \ldots, L$, $i = 1, \ldots, L$, compute the cross covariances of augmented residuals as

$$\text{Cov}\big(\mathbf{r}_n^{a[l]}, \mathbf{r}_n^{a[i]} \big)$$
$$= \mathbf{H}_n^a \mathbf{Q}_n^a \mathbf{H}_n^{aH} + \Delta^{a[l]} \text{Cov}\big(\mathbf{x}_{n-1}^a, \mathbf{x}_{n-1}^a \big) \big(\Delta^{a[i]} \big)^H$$
$$+ \mathbf{H}_n^{a[l]} \mathbf{F}_n^{a[l]} \text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{e}_{0n}^{a[i]} \big) \big(\mathbf{H}_n^{a[i]} \mathbf{F}_n^{a[i]} \big)^H$$
$$+ \mathbf{H}_n^{a[l]} \mathbf{F}_n^{a[l]} \text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{x}_{n-1}^a \big) \big(\Delta^{a[i]} \big)^H$$
$$+ \Delta^{a[l]} \text{Cov}\big(\mathbf{e}_{0n}^{a[i]}, \mathbf{x}_{n-1}^a \big)^H \big(\mathbf{H}_n^{a[i]} \mathbf{F}_n^{a[i]} \big)^H + \mathbf{R}_n^a$$

and the cross covariances of augmented estimation errors as

$$\text{Cov}\big(\mathbf{e}_n^{a[l]}, \mathbf{e}_n^{a[i]} \big)$$
$$= \big(\mathbf{I} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}_n^{a[l]} \big) \mathbf{F}^{a[l]} \text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{e}_{0n}^{a[i]} \big)$$
$$\times \big(\mathbf{F}^{a[i]} - \bar{\mathbf{K}}_n^{a[i]} \mathbf{H}^{a[i]} \mathbf{F}^{a[i]} \big)^H$$
$$+ \big(\mathbf{F}^a - \mathbf{F}^{a[l]} - \bar{\mathbf{K}}_n^{a[l]} \Delta^{a[l]} \big) \text{Cov}\big(\mathbf{x}_{n-1}^a, \mathbf{x}_{n-1}^a \big)$$
$$\times \big(\mathbf{F}^a - \mathbf{F}^{a[i]} - \bar{\mathbf{K}}_n^{a[i]} \Delta^{a[i]} \big)^H$$
$$+ \big(\mathbf{F}^{a[l]} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}^{a[l]} \mathbf{F}^{a[l]} \big) \text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{x}_{n-1}^a \big)$$
$$\times \big(\mathbf{F}^a - \mathbf{F}^{a[i]} - \bar{\mathbf{K}}_n^{a[i]} \Delta^{a[i]} \big)^H$$
$$+ \big(\mathbf{F}^a - \mathbf{F}^{a[l]} - \bar{\mathbf{K}}_n^{a[l]} \Delta^{a[l]} \big) \text{Cov}\big(\mathbf{e}_{0n}^{a[i]}, \mathbf{x}_{n-1}^a \big)^H$$
$$\times \big(\mathbf{F}^{a[i]} - \bar{\mathbf{K}}_n^{a[i]} \mathbf{H}^{a[i]} \mathbf{F}^{a[i]} \big)^H$$
$$+ \big(\mathbf{I} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}^a \big) \mathbf{Q}_n^a \big(\mathbf{I} - \bar{\mathbf{K}}_n^{a[i]} \mathbf{H}^a \big)^H + \bar{\mathbf{K}}_n^{a[l]} \mathbf{R}_n^a \big(\bar{\mathbf{K}}_n^{a[i]} \big)^H$$

where

$$\text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{x}_{n-1}^a \big) \approx \sum_{i=1}^{L} \bar{\gamma}_n^{[li]} \text{Cov}\big(\mathbf{e}_{n-1}^{a[i]}, \mathbf{x}_{n-1}^a \big) \tag{23}$$

$$\text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{e}_{0n}^{a[i]} \big) \approx \sum_{k=1}^{L} \sum_{j=1}^{L} \bar{\gamma}_n^{[lk]} \bar{\gamma}_n^{[ij]} \text{Cov}\big(\mathbf{e}_{n-1}^{a[k]}, \mathbf{e}_{n-1}^{a[j]} \big). \tag{24}$$

For $l = 1, \ldots, L$, update

$$\text{Cov}\big(\mathbf{e}_n^{a[l]}, \mathbf{x}_n^a \big) = \big(\mathbf{I} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}_n^{a[l]} \big) \mathbf{F}_n^{a[l]} \text{Cov}\big(\mathbf{e}_{0n}^{a[l]}, \mathbf{x}_{n-1}^a \big) \mathbf{F}_n^{aH}$$
$$+ \big(\mathbf{F}_n^a - \mathbf{F}_n^{a[l]} - \bar{\mathbf{K}}_n^{a[l]} \Delta^{a[l]} \big) \text{Cov}\big(\mathbf{x}_{n-1}^a, \mathbf{x}_{n-1}^a \big)$$
$$\times \mathbf{F}_n^{aH} + \big(\mathbf{I} - \bar{\mathbf{K}}_n^{a[l]} \mathbf{H}_n^a \big) \mathbf{Q}_n^a$$

$$\text{Cov}\big(\mathbf{x}_n^a, \mathbf{x}_n^a \big) = \mathbf{F}_n^a \text{Cov}\big(\mathbf{x}_{n-1}^a, \mathbf{x}_{n-1}^a \big) \mathbf{F}_n^{aH} + \mathbf{Q}_n^a.$$

Step 4 (Combination): For $l = 1, \ldots, L$, employing the fact that $\mathbf{r}_n^{a[l]} \sim \mathcal{N}(\bar{\mathbf{r}}_n^{a[l]}, \mathbf{V}_n^{a[l]})$, where $\mathbf{V}_n^{a[l]} = \mathrm{Cov}(\mathbf{r}_n^{a[l]}, \mathbf{r}_n^{a[l]})$, we can compute the mean of likelihood function as

$$
\begin{aligned}
\bar{\mathcal{L}}_n^{[l]} &= \int_{\mathbb{H}^N} \mathcal{L}_n^{[l]} f\left(\mathbf{r}_n^{a[l]} \big| \bar{\mathbf{r}}_n^{a[l]}, \mathbf{V}_n^{a[l]}\right) d\mathbf{r}_n^{[l]} \\
&= \int_{\mathbb{H}^N} f\left(\mathbf{r}_n^{a[l]} \big| \mathbf{0}, \bar{\mathbf{S}}_n^{a[l]}\right) f\left(\mathbf{r}_n^{a[l]} \big| \bar{\mathbf{r}}_n^{a[l]}, \mathbf{V}_n^{a[l]}\right) d\mathbf{r}_n^{[l]} \\
&= \left(\frac{2}{\pi}\right)^{2N} \frac{\left|\left[(\bar{\mathbf{S}}_n^{a[l]})^{-1} + (\mathbf{V}_n^{a[l]})^{-1}\right]^{-1}\right|^{0.5}}{\left|\bar{\mathbf{S}}_n^{a[l]}\right|^{0.5} \left|\mathbf{V}_n^{a[l]}\right|^{0.5}} \\
&\quad \times \exp\left\{-0.5(\bar{\mathbf{r}}_n^{a[l]})^H (\mathbf{V}_n^{a[l]} + \bar{\mathbf{S}}_n^{a[l]})^{-1} \bar{\mathbf{r}}_n^{a[l]}\right\}
\end{aligned} \tag{25}
$$

and the mean of the mode probability as

$$
\begin{aligned}
\bar{\mu}_n^{[l]} &= \int_{\mathbb{H}^N} \mu_n^{[l]} f\left(\mathbf{r}_n^{a[l]} \big| \bar{\mathbf{r}}_n^{a[l]}, \mathbf{V}_n^{a[l]}\right) d\mathbf{r}_n^{[l]} \\
&= \int_{\mathbb{H}^N} \frac{\sum_{i=1}^L \pi^{[il]} \mu_{n-1}^{[i]} f\left(\mathbf{r}_n^a \big| \mathbf{0}, \bar{\mathbf{S}}_n^{a[l]}\right) f\left(\mathbf{r}_n^a \big| \bar{\mathbf{r}}_n^{a[l]}, \mathbf{V}_n^{a[l]}\right)}{\sum_{i=1}^L \sum_{j=1}^L \pi^{[ji]} \mu_{n-1}^{[j]} f\left(\mathbf{r}_n^a \big| \mathbf{0}, \bar{\mathbf{S}}_n^{a[i]}\right)} d\mathbf{r}_n \\
&= \bar{\mathcal{L}}_n^{[l]} \sum_{i=1}^L \pi^{[il]} \bar{\mu}_{n-1}^{[i]} \times \int_{\mathbb{H}^N} \\
&\quad \times \frac{f\left(\mathbf{r}_n^a \big| [\mathbf{V}_n^{a[l]}(\bar{\mathbf{S}}_n^{a[l]})^{-1} + \mathbf{I}]^{-1} \bar{\mathbf{r}}_n^{a[l]}, [(\bar{\mathbf{S}}_n^{a[l]})^{-1} + (\mathbf{V}_n^{a[l]})^{-1}]^{-1}\right)}{\sum_{i=1}^L \sum_{j=1}^L \pi^{[ji]} \bar{\mu}_{n-1}^{[j]} f\left(\mathbf{r}_n^a \big| \mathbf{0}, \bar{\mathbf{S}}_n^{a[i]}\right)} d\mathbf{r}_n
\end{aligned} \tag{26}
$$

Then, compute the covariance of the overall augmented estimation error as

$$
\begin{aligned}
&E\left\{(\mathbf{x}_n^a - \hat{\mathbf{x}}_{n|n}^a)(\mathbf{x}_n^a - \hat{\mathbf{x}}_{n|n}^a)^H \big| \mathcal{S}\right\} \\
&\approx \sum_{l=1}^L \sum_{i=1}^L \bar{\mu}_n^{[l]} \bar{\mu}_n^{[i]} \left[\mathrm{Cov}(\mathbf{e}_n^{a[l]}, \mathbf{e}_n^{a[i]}) + \bar{\mathbf{e}}_n^{a[l]} (\bar{\mathbf{e}}_n^{a[i]})^H\right]
\end{aligned} \tag{27}
$$

and the MSE of the overall estimate is given by

$$
\begin{aligned}
&E\{(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n})^H (\mathbf{x}_n - \hat{\mathbf{x}}_{n|n}) | \mathcal{S}\} \\
&\qquad = \frac{1}{4} \mathrm{Tr}\left(E\{(\mathbf{x}_n^a - \hat{\mathbf{x}}_{n|n}^a)(\mathbf{x}_n^a - \hat{\mathbf{x}}_{n|n}^a)^H | \mathcal{S}\}\right).
\end{aligned} \tag{28}
$$

Remark 1: In order to derive a closed-form solution, we have applied the approximations in (22)–(24) and (27) based on the zero correlation assumption between the involved variables.

C. WL-QSMM

If the system mode is time invariant, then

$$
\pi^{[il]} = \begin{cases} 1 & l = i \\ 0 & l \neq i \end{cases}
$$

so that (19) simplifies to

$$
\mu_n^{[l]} = \frac{\mathcal{L}_n^{[l]} \mu_{n-1}^{[l]}}{\sum_{i=1}^L \mathcal{L}_n^{[i]} \mu_{n-1}^{[i]}} \tag{29}
$$

and the reinitialization step can be proven to be unnecessary. In this case, the WL-QIMM algorithm reduces to the WL-QSMM algorithm, for which each iteration includes the following two steps.

Step 1 (Mode-Conditioned Kalman Filtering): Implement L WL-QKFs with the initial states and covariances being the estimates obtained in the previous iteration. This step is the same as Step 2 of the WL-QIMM except that $\mathbf{x}_{0n}^{[l]}$ and $\mathbf{P}_{0n}^{a[l]}$ in (17) and (18) are replaced with $\hat{\mathbf{x}}_{n-1|n-1}^{a[l]}$ and $\mathbf{P}_{n-1|n-1}^{a[l]}$ for $l = 1, \ldots, L$.

Step 2 (Combination): Compute the mode probabilities according to (15) and (29). The overall augmented estimate and overall augmented error covariance are given by (20) and (21), respectively.

The performance analysis algorithm proposed in Section III-B can also be applied to the WL-QSMM algorithm by replacing $\bar{\mathbf{e}}_{0n}^{a[l]}$, $\bar{\mathbf{x}}_{0n}^{[l]}$, and $\bar{\mathbf{P}}_{0n}^{a[l]}$ with $\bar{\mathbf{e}}_{n-1}^{a[l]}$, $\bar{\mathbf{x}}_{n-1|n-1}^{a[l]}$, and $\bar{\mathbf{P}}_{n-1|n-1}^{a[l]}$, for $l = 1, \ldots, L$.

We next prove that the WL-QSMM algorithm converges to the candidate model closest to the true system model in the sense of the Kullback–Leibler divergence. The Kullback–Leibler divergence between $\mathcal{N}(E\{\mathbf{r}_n^a\}, \mathbf{S}_n^a)$ and $\mathcal{N}(E\{\mathbf{r}_n^{a[l]}\}, \mathbf{S}_n^{a[l]})$, where $\mathbf{r}_n^a$ is the augmented residual corresponding to the filter that exactly matches the true system model and $\mathbf{S}_n^a$ is its covariance matrix, is defined by

$$
\begin{aligned}
K_l &\triangleq \frac{1}{2} \ln \frac{|\mathbf{S}_n^{a[l]}|}{|\mathbf{S}_n^a|} + \frac{1}{2} \mathrm{Tr}\left(\Re[(\mathbf{S}_n^{a[l]})^{-1} \mathbf{S}_n^a] - \mathbf{I}_{4N}\right) \\
&\quad + \frac{1}{2}\left(E\{\mathbf{r}_n^{a[l]}\} - E\{\mathbf{r}_n^a\}\right)^H (\mathbf{S}_n^{a[l]})^{-1}\left(E\{\mathbf{r}_n^{a[l]}\} - E\{\mathbf{r}_n^a\}\right) \\
&= \frac{1}{2} \ln \frac{|\mathbf{S}_n^{a[l]}|}{|\mathbf{S}_n^a|} + \frac{1}{2} \mathrm{Tr}\left(\Re[(\mathbf{S}_n^{a[l]})^{-1} E\{\mathbf{r}_n^{a[l]}(\mathbf{r}_n^{a[l]})^H\}] - \mathbf{I}\right).
\end{aligned} \tag{30}
$$

The second equality in (30) is based on $E\{\mathbf{r}_n^a\} = \mathbf{0}$ and

$$
E\{\mathbf{r}_n^{a[l]}(\mathbf{r}_n^{a[l]})^H\} - E\{\mathbf{r}_n^{a[l]}\}\left(E\{\mathbf{r}_n^{a[l]}\}\right)^H = \mathbf{S}_n^a
$$

which is a straightforward extension of the analysis result for real-valued mismatched Kalman filters [46] to the quaternion domain. Let $t = \arg\min_{l \in \{1,2,\ldots,L\}} K_l$, which is the index of the candidate model closest to the true system model. For $l \neq t$, define $\Lambda_n^{[l]} \triangleq \mu_n^{[l]}/\mu_n^{[t]}$ to obtain

$$
\begin{aligned}
\frac{\Lambda_n^{[l]}}{\Lambda_{n-1}^{[l]}} &= \frac{\mathcal{L}_n^{[l]}}{\mathcal{L}_n^{[t]}} = \sqrt{\frac{|\mathbf{S}_n^{a[t]}|}{|\mathbf{S}_n^{a[l]}|}} \\
&\quad \times \exp\left\{\frac{1}{2}\left[(\mathbf{r}_n^{a[t]})^H (\mathbf{S}_n^{a[t]})^{-1} \mathbf{r}_n^{a[t]}\right.\right. \\
&\qquad \left.\left. - (\mathbf{r}_n^{a[l]})^H (\mathbf{S}_n^{a[l]})^{-1} \mathbf{r}_n^{a[l]}\right]\right\} = \sqrt{\frac{|\mathbf{S}_n^{a[t]}|}{|\mathbf{S}_n^{a[l]}|}} \\
&\quad \times \exp\left\{\frac{1}{2} \mathrm{Tr}\left(\Re[(\mathbf{S}_n^{a[t]})^{-1} \mathbf{r}_n^{a[t]}(\mathbf{r}_n^{a[t]})^H - \mathbf{S}_n^{a[l]})^{-1} \mathbf{r}_n^{a[l]}(\mathbf{r}_n^{a[l]})^H]\right)\right\}
\end{aligned} \tag{31}
$$

Applying the statistical expectation operator and the natural logarithm to (31) yields

$$
\ln\left(E\{\Lambda_n^{[l]}/\Lambda_{n-1}^{[l]}\}\right) = K_t - K_l < 0
$$

whereby $E\{\Lambda_n^{[l]}/\Lambda_{n-1}^{[l]}\} < 1$, which indicates that $\mu_n^{[l]}$ decreases to 0 and hence $\mu_n^{[t]}$ increases to 1 when $n \to \infty$. This proves the convergence of the WL-QSMM algorithm to model t.

Remark 2: Since the Kullback–Leibler divergence in (30) is nonnegative and is vanishing only when the filter model exactly matches the true system, the algorithm will converge to the true system model if the true model belongs to the candidate model set.

D. SWL-QMMAE and SL-QMMAE

For semiwidely linear quaternion systems with $\mathbb{C}^\eta$-improper quaternion Gaussian noise, the above WL-QMMAE algorithms reduce to corresponding SWL-QMMAE algorithms by replacing the augmented forms of variables with the semiaugmented forms. Based on the probability density function in (5), the likelihood functions in the SWL-QIMM and SWL-QSMM algorithms are given by

$$\mathcal{L}_n^{[l]} = f\left(\mathbf{r}_n^{b[l]} \big| \mathbf{0}, \mathbf{S}_n^{b[l]}\right).$$

For strictly linear quaternion systems with $\mathbb{H}$-proper quaternion Gaussian noise, the WL-QMMAE algorithms reduce to corresponding SL-QMMAE algorithms by replacing the augmented forms of variables with the original forms. Since the QMGD is associated with the probability density function in (6) under this condition, the likelihood functions in the SL-QIMM and SL-QSMM algorithms become

$$\mathcal{L}_n^{[l]} = f\left(\mathbf{r}_n^{[l]} \big| \mathbf{0}, \mathbf{S}_n^{[l]}\right).$$

The SL-QMMAE and SWL-QMMAE algorithms can be seen as reduced-order WL-QMMAE algorithms [47]. Therefore, they are suboptimal for widely linear quaternion systems and general improper quaternion noise, and their performance disadvantages increase with the degree of widely linear nature of the system and/or the degree of noise impropriety.

IV. SIMULATIONS

A. QIMM Algorithms for Filtering Quaternion Signals

This section presents simulation results for evaluating the SL-QIMM, SWL-QIMM, and WL-QIMM algorithms and the corresponding performance analysis algorithms. Since the linear nature (strictly, semiwidely, or widely linear) of the involved quaternion system essentially affects the estimation performance, we employed a strictly linear quaternion hybrid system, with noise at different levels of impropriety, to study the impact of impropriety on the estimation performance. The considered system, which was used in [33], can be represented by the following 1-D state-space model:

$$x_n = b^{[l]} x_{n-1} + w_n^{[l]}$$
$$z_n = c^{[l]} x_n + v_n^{[l]}$$

where $l \in \{1, 2\}$, $b^{[1]} = 0.95$, $b^{[2]} = 0.85$, $c^{[1]} = 1$, $c^{[2]} = 0.4$, and $w_n^{[1]}$, $w_n^{[2]}$, $v_n^{[1]}$, $v_n^{[2]}$ are proper or improper quaternion Gaussian noise sequences with the power of 2, 0, 0, and 0 dB. The system mode was 1 from $n = 1$ to 100 and from $n = 200$ to 300, and was 2 from $n = 100$ to 200. In each experiment, 500 MC simulation runs of the three considered QIMM algorithms were implemented and the estimation performance was evaluated through the averaged MSE. The initial mode probabilities were set to
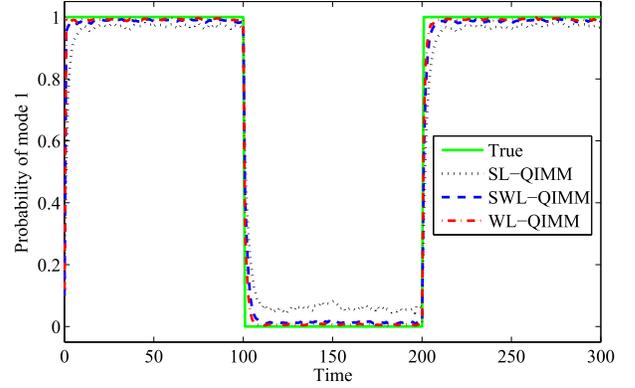
(a)
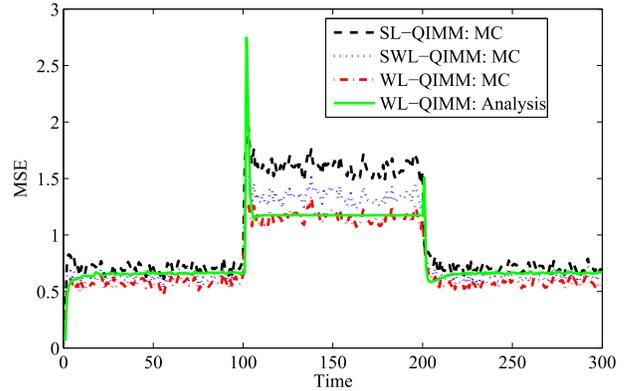
(b)

Fig. 3. Learning curves of the three QIMM algorithms when the state and observation noise are improper. (a) Probability of the time-varying mode 1. (b) MSE of the algorithms considered.

$\mu_0^{[1]} = 0.1$ and $\mu_0^{[2]} = 0.9$. The proposed performance analysis algorithms were also implemented for comparison with the MC results.

Fig. 3 shows the simulation results of the three QIMM algorithms for improper state noise with impropriety coefficients $\rho_\iota = \rho_J = \rho_\kappa = 0.2$ and improper observation noise with $\rho_\iota = \rho_J = \rho_\kappa = 0.9$. Fig. 3(a) shows the average probabilities of the hypothesizing mode 1 over 500 MC runs. Observe that the WL-QIMM has converged faster to the true mode than the other two algorithms, while the SL-QIMM has exhibited the slowest convergence. As shown in Fig. 3(b), the WL-QIMM has achieved the smallest MSE, while the SL-QSMM has attained the largest MSE; the MSE computed from the proposed performance analysis algorithm for WL-QIMM has matched the MSE averaged over the MC runs with 12% difference on average. The error arose from the approximations in (22)–(24), (27), and the integral calculation in (26). The 500 MC runs of WL-QIMM took 79 min, while the proposed performance analysis algorithm took only 2 min on a standard PC. Obviously, this computation efficiency advantage of the performance analysis algorithm will increase with the data size.

Fig. 4 shows the simulation results of the three QIMM algorithms for $\mathbb{C}^\iota$-improper state and observation noise.
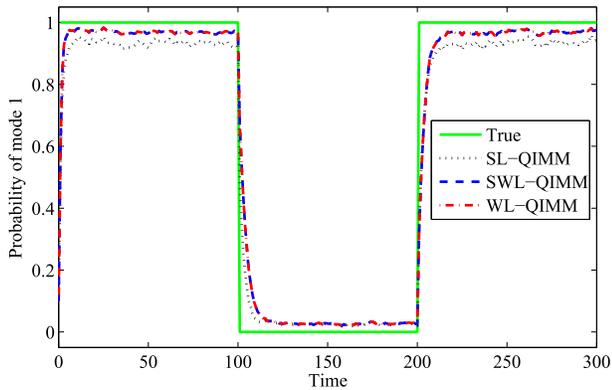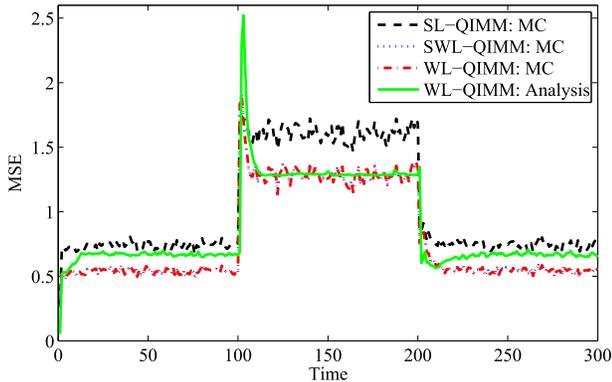
Fig. 4. Learning curves of the three QIMM algorithms when the state and observation noise are $\mathbb{C}^i$-improper. (a) Probability of the time-varying mode 1. (b) MSE of the algorithms considered.

Fig. 5. Learning curves of the three QIMM algorithms when the state and observation noise are $\mathbb{H}$-proper. (a) Probability of the time-varying mode 1. (b) MSE of the algorithms considered.

Fig. 4(a) shows the average probabilities of the hypothesizing mode 1 over 500 MC runs, and Fig. 4(b) shows the MSEs of the three QIMM algorithms. As shown in Fig. 4, the WL-QIMM and SWL-QIMM have exhibited the same performance, achieving a lower MSE than the SL-QIMM, and the MSE computed from the performance analysis algorithm for WL-QIMM has matched the MSE averaged over the MC runs with 16% difference on average.

Fig. 5 shows the simulation results of the three QIMM algorithms for $\mathbb{H}$-proper state and observation noise. Fig. 5(a) shows the average probabilities of the hypothesizing mode 1 over the 500 MC simulation runs, and Fig. 5(b) shows the MSEs of the three QIMM algorithms. As shown in Fig. 5, the three QIMM algorithms had the same performance for $\mathbb{H}$-proper noise, and the MSE computed from the performance analysis algorithm for WL-QIMM has matched the MSE averaged over the MC runs with 14% difference on average.

Fig. 6 shows the average MSEs of the three QIMM algorithms over the whole simulation period. Fig. 6(a) shows the results for $\mathbb{H}$-proper observation noise and improper state noise with varying impropriety coefficients. For simplicity, the three impropriety coefficients of the state noise were set to be equal, that is, $\rho_\iota = \rho_\jmath = \rho_\kappa = \rho$. Fig. 6(b) shows the results for $\mathbb{H}$-proper state noise and improper observation noise with

various impropriety coefficients. The three impropriety coefficients of the observation noise were also equal. Fig. 6 shows that for the $\mathbb{H}$-proper noise, the three QIMM algorithms have exhibited the same MSE, while for the improper noise, the WL-QIMM has achieved the smallest MSE and the SL-QIMM has attained the largest MSE. As the impropriety degree of the noise is increased, the MSE difference between the three algorithms is also increased. Interestingly, the MSE of SL-QIMM was increased with the impropriety coefficient of state noise but it was unaffected by the impropriety coefficient of observation noise, while the MSEs of SWL-QIMM and WL-QIMM were decreased with the impropriety coefficients of state and observation noise. The MSE estimated from the performance analysis algorithm for WL-QIMM shows the same trend as the MC simulation result with about 10% difference in value.

B. QSMM Algorithms for Filtering Quaternion Signals

This section presents the simulation results for evaluating the SL-QSMM, SWL-QSMM, and WL-QSMM algorithms and the performance analysis algorithm for WL-QSMM. The simulation settings were the same as that for the QIMM algorithms in Section IV-A, except that the system mode was constantly 1. Fig. 7 shows the simulation results of
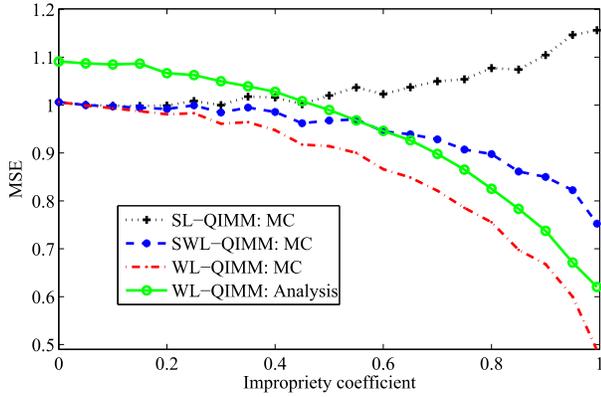
(a)
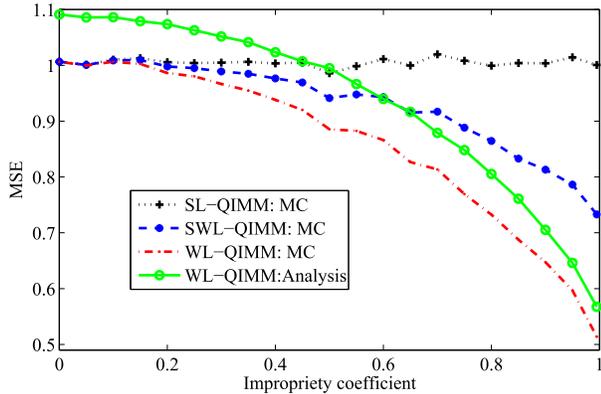
(b)

Fig. 6. MC simulated MSEs of the three QIMM algorithms and the MSE computed from the proposed performance analysis algorithm (analysis) for WL-QIMM, as a function of the impropriety coefficient of noise. (a) Improper state noise and $\mathbb{H}$-proper observation noise. (b) $\mathbb{H}$-proper state noise and improper observation noise.

(a)

(b)

Fig. 7. Learning curves of the three QSMM algorithms when the state and observation noise are improper. (a) Probability of the time-varying mode 1. (b) MSE of the algorithms considered.

the three QSMM algorithms for improper state noise with $\rho_\iota = \rho_J = \rho_\kappa = 0.2$ and improper observation noise with $\rho_\iota = \rho_J = \rho_\kappa = 0.9$. Fig. 7(a) shows the average probabilities of the hypothesizing mode 1 over 500 MC runs. Observe that the WL-QSMM has converged faster to the true mode than the other two algorithms, and the SL-QSMM has exhibited the slowest convergence. As shown in Fig. 7(b), the WL-QSMM has achieved the smallest MSE, while the SL-QSMM has attained the largest MSE; the MSE computed from the performance analysis algorithm for WL-QSMM has matched the MSE averaged over the MC runs with 5% difference. The performance analysis algorithm for the WL-QSMM was more accurate than the analysis algorithm for WL-QIMM simulated in Section IV-A because of the less complicated mechanism of the QSMM algorithm compared with the QIMM algorithm.

Fig. 8 shows the average MSEs of the three QSMM algorithms over the whole simulation period. Fig. 8(a) shows the results for $\mathbb{H}$-proper observation noise and improper state noise over a range of impropriety coefficients. For simplicity, the three impropriety coefficients of the state noise were set to be equal. Fig. 8(b) shows the results for $\mathbb{H}$-proper state noise and improper observation noise with various impropriety coefficients. The three impropriety coefficients of the observation noise were also equal. Fig. 8 shows that for $\mathbb{H}$-proper noise, the three QSMM algorithms have exhibited the same MSE, while for the improper noise, the WL-QSMM has achieved the smallest MSE and the SL-QSMM has attained the largest MSE. As the impropriety degree of the noise is increased, the MSE difference between the three algorithms is increased. It is also noticeable that the MSEs of WL-QSMM and SWL-QSMM was decreased with the impropriety coefficients of state and observation noise, and the MSE of SL-QSMM was unaffected by the impropriety coefficients. The MSE estimated from the performance analysis algorithm for WL-QSMM has matched the MC simulation result.

C. QIMM Algorithms for 3-D Target Tracking

The effectiveness of the proposed WL-QIMM algorithm is next illustrated in the context of target tracking in a 3-D space. The task is to track the position and velocity of a moving target using the measurements of its position. Denote the position by (X_n, Y_n, Z_n), velocity by $(\dot{X}_n, \dot{Y}_n, \dot{Z}_n)$, and acceleration by $(\ddot{X}_n, \ddot{Y}_n, \ddot{Z}_n)$. Consider two quaternion-valued kinematic models. The first one is the constant velocity model given
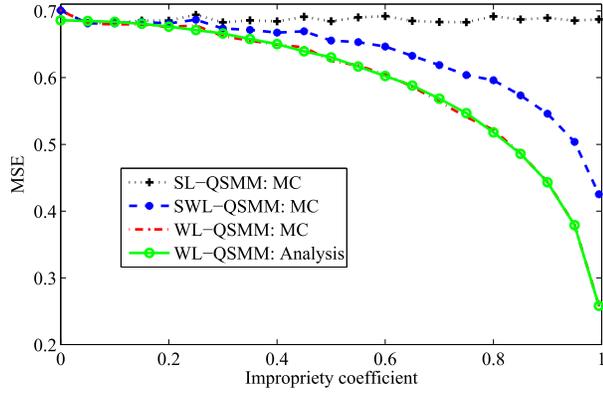
(a)
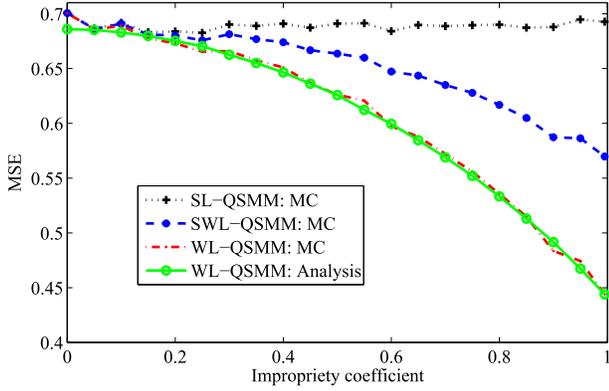
(b)

Fig. 8.   MC simulated MSEs of the three QSMM algorithms and the MSE computed from the proposed performance analysis algorithm (analysis) for WL-QSMM, as a function of the impropriety coefficient of noise. (a) Improper state noise and $\mathbb{H}$-proper observation noise. (b) $\mathbb{H}$-proper state noise and improper observation noise.

by [25]

$$\mathbf{x}_n = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \mathbf{x}_{n-1} + \mathbf{w}_n$$

with the measurement given by

$$z_n = [\, 1 \quad 0\,]\mathbf{x}_n + v_n$$

where $\mathbf{x}_n = [\iota X_n + J Y_n + \kappa Z_n; \iota \dot{X}_n + J \dot{Y}_n + \kappa \dot{Z}_n]$, Δt is the sampling interval, $\mathbf{w}_n$ is a zero-mean pure quaternion Gaussian noise vector, and v_n is zero-mean pure quaternion Gaussian noise. The covariance matrices of state noise are

$$\mathbf{C_{ww}} = \begin{bmatrix} \dfrac{1}{3}\Delta t^3 & \dfrac{1}{2}\Delta t^2 \\ \dfrac{1}{2}\Delta t^2 & \Delta t \end{bmatrix} \sigma_w^2$$

$$\mathbf{C_{ww^\iota}} = \rho_{w\iota}\mathbf{C_{ww}}, \quad \mathbf{C_{ww^J}} = \rho_{wJ}\mathbf{C_{ww}}, \quad \mathbf{C_{ww^\kappa}} = \rho_{w\kappa}\mathbf{C_{ww}}$$

where σ_w^2 indicates the power of state noise, and $\rho_{w\iota}, \rho_{wJ}, \rho_{wk} \in \mathbb{R}$ are the impropriety coefficients satisfying $\rho_{w\iota} + \rho_{wJ} + \rho_{wk} = -1$. The covariances of observation noise are $C_{vv} = \sigma_v^2$, $C_{vv^\iota} = \rho_{v\iota}C_{vv}$, $C_{vv^J} = \rho_{vJ}C_{vv}$, and $C_{vv^\kappa} = \rho_{v\kappa}C_{vv}$, where σ_v^2 indicates the power of observation noise, and $\rho_{v\iota}, \rho_{vJ}, \rho_{vk} \in \mathbb{R}$ are the impropriety coefficients satisfying $\rho_{v\iota} + \rho_{vJ} + \rho_{vk} = -1$.
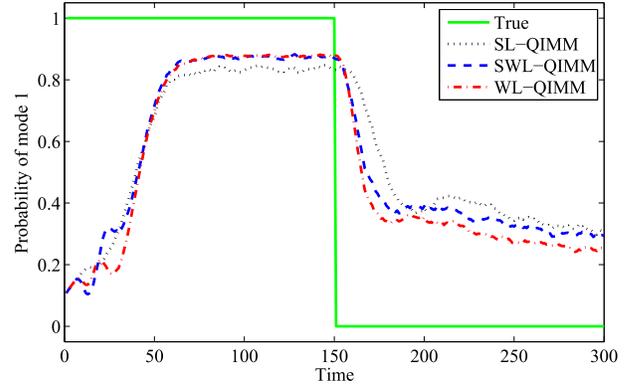
(a)
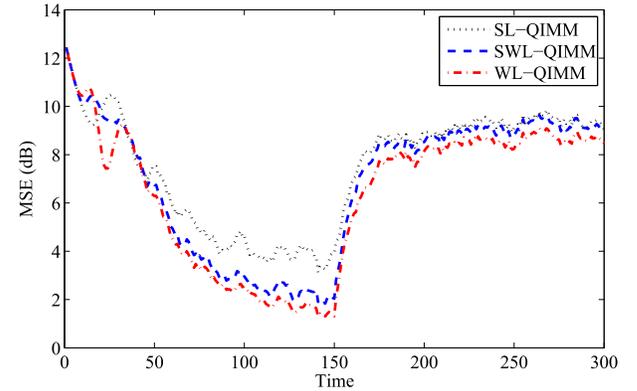
(b)

Fig. 9.   Learning curves of the three QIMM algorithms for target tracking in a 3-D space. (a) Probability of the time-varying mode 1. (b) MSE of the position and velocity estimation.

The second model is the constant acceleration model given by [25]

$$\mathbf{x}_n = \begin{bmatrix} 1 & \Delta t & \dfrac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{n-1} + \mathbf{w}_n$$

with the measurement given by

$$z_n = [\, 1 \quad 0 \quad 0\,]\mathbf{x}_n + v_n$$

where

$$\mathbf{x}_n = \begin{bmatrix} \iota X_n + J Y_n + \kappa Z_n \\ \iota \dot{X}_n + J \dot{Y}_n + \kappa \dot{Z}_n \\ \iota \ddot{X}_n + J \ddot{Y}_n + \kappa \ddot{Z}_n \end{bmatrix}.$$

The pure quaternion Gaussian noises $\mathbf{w}_n$ and v_n in this model have the same properties as those in the constant velocity model, except for

$$\mathbf{C_{ww}} = \begin{bmatrix} \dfrac{1}{20}\Delta t^5 & \dfrac{1}{8}\Delta t^4 & \dfrac{1}{6}\Delta t^3 \\ \dfrac{1}{8}\Delta t^4 & \dfrac{1}{3}\Delta t^3 & \dfrac{1}{2}\Delta t^2 \\ \dfrac{1}{6}\Delta t^3 & \dfrac{1}{2}\Delta t^2 & dt \end{bmatrix} \sigma_w^2.$$

In our simulation, $\Delta t = 0.1\,s$, and the target initially had a position of $(0, 0, 0)$, velocity of $(1, 2, 0.5)$, and vanishing acceleration. The target has followed the constant velocity model with $\sigma_w^2 = 0.01$ (mode 1) for 150 s before following the constant acceleration model with $\sigma_w^2 = 1$ (mode 2) for another 150 s. During the whole 300 s, $\rho_{w\imath} = \rho_{w\jmath} = \rho_{wk} = -1/3$, $\rho_{v\imath} = \rho_{v\jmath} = -25/27$, $\rho_{vk} = 23/27$, and $\sigma_v^2 = 1$. In the estimation algorithms, the initial estimate of the state vector was set to $[2\imath + 2.1\jmath + 2.2\kappa, 0.1\imath + 0.2\jmath + 0.3\kappa, 0.05\imath + 0.05\jmath + 0.05\kappa]$, and the initial mode probabilities were $\mu_0^{[1]} = 0.1$ and $\mu_0^{[2]} = 0.9$. Fig. 9 shows the MC simulated performance of the three considered QIMM algorithms. As shown in Fig. 9(a), the probability of mode 1 is increased to 0.9 in the first 150 s and decreased toward 0.2 in the second 150 s. Fig. 9(b) shows that the WL-QIMM has achieved a lower MSE of the position and velocity estimation compared with the SL-QIMM and SWL-QIMM.

V. Conclusion

We have proposed a class of WL-QMMAE algorithms which employ the augmented second-order quaternion statistics to cater for both proper and improper quaternion Gaussian signals. The WL-QIMM algorithm for tracking time-variant system mode uncertainty and the WL-QSMM algorithm for tracking time-invariant system mode uncertainty have been proposed based on the Bayesian inference. A recursive algorithm has been derived to efficiently assess the performance of WL-QIMM, and a convergence proof of WL-QSMM has been provided. We have shown that, as expected, the WL-QMMAE reduces to the semiwidely linear form for $\mathbb{C}^\eta$-improper signals and further reduces to the strictly linear form for $\mathbb{H}$-proper signals. The effectiveness of the proposed algorithms has been verified by numerical simulations over 3-D and 4-D signal processing case studies. Although the presented algorithms in this paper use Kalman filters for linear state estimation, the proposed WL-QMMAE framework is flexible and therefore can incorporate other types of filters and neural networks for nonlinear estimation [48]. The proposed algorithms can also be modified to the widely linear quaternion counterpart of the Gaussian sum filtering algorithms for non-Gaussian signals [49], [50].

References

[1] J. B. Kuipers, *Quaternions and Rotation Sequences: A Primer With Applications to Orbits, Aerospace and Virtual Reality*. Princeton, NJ, USA: Princeton Univ. Press, 1999.

[2] D. P. Mandic, C. Jahanchahi, and C. C. Took, "A quaternion gradient operator and its applications," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 47–50, Jan. 2011.

[3] D. Xu, C. Jahanchahi, C. C. Took, and D. P. Mandic, "Enabling quaternion derivatives: The generalized HR calculus," *Roy. Soc. Open Sci.*, vol. 2, no. 8, p. 150255, 2015.

[4] J. Vía, D. Ramírez, and I. Santamaría, "Properness and widely linear processing of quaternion random vectors," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010.

[5] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Process.*, vol. 91, no. 2, pp. 214–224, Feb. 2011.

[6] 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.

[7] 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.

[8] T. K. Paul and T. Ogunfunmi, "A kernel adaptive algorithm for quaternion-valued inputs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2422–2439, Oct. 2015.

[9] F. Shang and A. Hirose, "Quaternion neural-network-based PolSAR land classification in poincare-sphere-parameter space," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5693–5703, Sep. 2014.

[10] Y. Xia, C. Jahanchahi, and D. P. Mandic, "Quaternion-valued echo state networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 4, pp. 663–673, Apr. 2015.

[11] M. E. Valle and F. Z. de Castro, "On the dynamics of hopfield neural networks on unit quaternions," *IEEE Trans. Neural Netw. Learn. Syst.*, doi: 10.1109/TNNLS.2017.2691462.

[12] T. Minemoto, T. Isokawa, H. Nishimura, and N. Matsui, "Feed forward neural network with random quaternionic neurons," *Signal Process.*, vol. 136, pp. 59–68, Jul. 2017.

[13] S. Miron, N. Le Bihan, and J. I. Mars, "Quaternion-MUSIC for vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006.

[14] J.-F. Gu and K. Wu, "Quaternion modulation for dual-polarized antennas," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 286–289, Feb. 2017.

[15] H. Fourati, N. Manamanni, L. Afilal, and Y. Handrich, "Complementary observer for body segments motion capturing by inertial and magnetic sensors," *IEEE/ASME Trans. Mechatronics*, vol. 19, no. 1, pp. 149–157, Feb. 2014.

[16] K. Adhikari, S. Tatinati, W. T. Ang, K. C. Veluvolu, and K. Nazarpour, "A quaternion weighted Fourier linear combiner for modeling physiological tremor," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 11, pp. 2336–2346, Nov. 2016.

[17] C. Jahanchahi and D. P. Mandic, "A class of quaternion Kalman filters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 533–544, Mar. 2014.

[18] J. Navarro-Moreno, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "A quaternion widely linear model for nonlinear Gaussian estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6414–6424, Dec. 2014.

[19] J. D. Jiménez-López, R. M. Fernández-Alcalá, J. Navarro-Moreno, and J. C. Ruiz-Molina, "Widely linear estimation of quaternion signals with intermittent observations," *Signal Process.*, vol. 136, pp. 92–101, Jul. 2017.

[20] J. Navarro-Moreno, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Semi-widely simulation and estimation of continuous-time $\mathbb{C}^\eta$-proper quaternion random signals," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4999–5012, Sep. 2015.

[21] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixtures of local experts," *Neural Comput.*, vol. 3, no. 1, pp. 79–87, 2012.

[22] S. E. Yuksel, J. N. Wilson, and P. D. Gader, "Twenty years of mixture of experts," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 8, pp. 1177–1193, Aug. 2012.

[23] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.

[24] M. Gwak, K. Jo, and M. Sunwoo, "Neural-network multiple models filter (NMM)-based position estimation system for autonomous vehicles," *Int. J. Autom. Technol.*, vol. 14, no. 2, pp. 265–274, Apr. 2013.

[25] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation: Theory, Algorithms and Software*. Hoboken, NJ, USA: Wiley, 2001.

[26] L. Chen and K. S. Narendra, "Nonlinear adaptive control using neural networks and multiple models," *Automatica*, vol. 37, no. 8, pp. 1245–1255, Aug. 2001.

[27] L. Chen and K. S. Narendra, "Intelligent control using multiple neural networks," *Int. J. Adapt. Control Signal Process.*, vol. 17, no. 6, pp. 417–430, 2003.

[28] Y. Fu and T. Chai, "Nonlinear multivariable adaptive control using multiple models and neural networks," *Automatica*, vol. 43, no. 6, pp. 1101–1110, Jun. 2007.

[29] X. R. Li and Y. Bar-Shalom, "Performance prediction of the interacting multiple model algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 3, pp. 755–771, Jul. 1993.

[30] C. E. Seah and I. Hwang, "Algorithm for performance analysis of the IMM algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 1114–1124, Apr. 2011.

[31] R. W. Osborne and W. D. Blair, "Update to the hybrid conditional averaging performance prediction of the IMM algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 4, pp. 2967–2974, Oct. 2011.

[32] C. Hue, J.-P. Le Cadre, and P. Perez, "Posterior Cramer–Rao bounds for multi-target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 1, pp. 37–49, Jan. 2006.

[33] A. Mohammadi and K. N. Plataniotis, "Improper complex-valued multiple-model adaptive estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1528–1542, Mar. 2015.

[34] T. A. Ell and S. J. Sangwine, "Quaternion involutions and anti-involutions," *Comput. Math. Appl.*, vol. 53, no. 1, pp. 137–143, Jan. 2007.

[35] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ, USA: Wiley, 2009.

[36] S. P. Talebi, S. Kanna, and D. P. Mandic, "Real-time estimation of quaternion impropriety," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 557–561.

[37] J. Vía, D. Palomar, L. Vielva, and I. Santamaría, "Quaternion ICA from second-order statistics," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1586–1600, Apr. 2011.

[38] 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.

[39] J. D. J. López, R. M. Fernández-Alcalá, J. Navarro-Moreno, and J. C. Ruiz-Molina, "A semi-widely linear filtering algorithm for C^η-proper quaternion signals based on randomly modeled observations," in *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–5.

[40] C. Jahanchahi, "Quaternion valued adaptive signal processing," Ph.D. dissertation, Dept. Elect. Electron. Eng., Imperial College London, London, U.K., 2013.

[41] D. H. Dini and D. P. Mandic, "Class of widely linear complex Kalman filters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 5, pp. 775–786, May 2012.

[42] T. Hasija, C. Lameiro, and P. J. Schreier, "Determining the dimension of the improper signal subspace in complex-valued data," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1606–1610, Nov. 2017.

[43] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 2, pp. 387–392, Apr. 1985.

[44] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, vol. AC-33, no. 8, pp. 780–783, Aug. 1988.

[45] I. Hwang, C. E. Seah, and S. Lee, "A study on stability of the interacting multiple model algorithm," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 901–906, Feb. 2017.

[46] P. D. Hanlon and P. S. Maybeck, "Characterization of Kalman filter residuals in the presence of mismodeling," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 1, pp. 114–131, Jan. 2000.

[47] B. Friedland, "On the properties of reduced order Kalman filters," *IEEE Trans. Autom. Control*, vol. 34, no. 3, pp. 321–324, Mar. 1989.

[48] F. A. Tobar, S.-Y. Kung, and D. P. Mandic, "Multikernel least mean square algorithm," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 265–277, Feb. 2014.

[49] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Autom. Control*, vol. AC-17, no. 4, pp. 439–448, Aug. 1972.

[50] A. Mohammadi and K. N. Plataniotis, "Complex-valued Gaussian sum filter for nonlinear filtering of non-Gaussian/non-circular noise," *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 440–444, Apr. 2015.
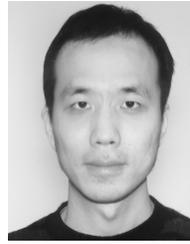
Min Xiang (M'18) received the Ph.D. degree in electrical engineering from Imperial College London, London, U.K., in 2018.

He is currently a Research Associate with Imperial College London. His current research interests include quaternion-valued statistical signal processing and adaptive filtering.

Bruno Scalzo Dees received the M.Eng. degree in aeronautical engineering from Imperial College London, London, U.K., where he is currently pursuing the Ph.D. degree with the Department of Electrical and Electronic Engineering.

His current research interests include statistical signal processing, dynamical systems, and tensor decompositions.

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 has been a Guest Professor with Katholieke Universiteit Leuven, Leuven, Belgium, the Tokyo University of Agriculture and Technology, Tokyo, Japan, Westminster University, London, and a Frontier Researcher with RIKEN, Wako, Japan. He is currently a Professor of signal processing with Imperial College London, where he is involved in nonlinear adaptive signal processing and nonlinear dynamics. He is also the Deputy Director of the Financial Signal Processing Laboratory, Imperial College London. He has two research monographs *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 *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.

Dr. Mandic has been a member of the IEEE Technical Committee on Signal Processing Theory and Methods. He has produced award winning papers and products resulting from his collaboration with the industry. He has been an Associate Editor of the *IEEE Signal Processing Magazine*, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON NEURAL NETWORKS.